

DOI: 10.21122/2220-9506-2025-16-4-379-385

Прецизионное детектирование пространственно-разнесённых объектов на аэрокосмических снимках для решения обратных задач

И.В. Саечников, В.В. Скакун, Э.А. Чернявская

Белорусский государственный университет,
пр-т Независимости, 4, г. Минск 220030, Беларусь

Поступила 10.10.2025

Принята к печати 21.11.2025

Обработка данных дистанционного зондирования, основанная на методах глубокого обучения, находит широкое применение в различных междисциплинарных задачах. Современные подходы преимущественно сосредоточены на разработке методов глубокого обучения, предназначенных для обнаружения объектов в густонаселённых городских условиях, которые плохо подходят для обнаружения пространственно-разнесённых объектов, особенно мелкозернистых. Характерным примером является обнаружение морских судов на открытой водной поверхности. К проблемам в этой области относятся неоднородный шумовой фон, вызванный меняющимися условиями освещения, внешние артефакты, размытие, вызванное волнами, вид сверху, дающий некачественные карты характеристик объектов, и необходимость сканирования больших территорий для поиска объектов. Для решения этих проблем в данной статье предложен новый гибридный подход к глубокому обучению для обнаружения пространственно-разнесённых объектов. Предлагаемый метод включает в себя ансамбль классификаторов на основе DenseNet_1 и ResNet_1 в сочетании с детектором YOLOv11+ и полуконтролируемой стратегией обучения. Кроме того, для прецизионного детектирования мелкозернистых объектов предлагается улучшенная версия модели YOLOv11+fg с адаптированными модулями C3K2, UpSample и SPFF. Метод детектирования был разработан, предварительно обучен и валидирован с использованием набора спутниковых изображений Airbus SPOT, а затем дополнительно дообучен на базе датасета ShipRSImageNet для задачи обнаружения мелкозернистых объектов. Результаты демонстрируют эффективность крупномасштабного сканирования в дистанционном зондировании и распознавания пространственно-разнесённых объектов, а также превосходят по точности показатели обнаружения мелкозернистых объектов оригинальной моделью.

Ключевые слова: гибридное глубокое обучение, детектирование объекта, ансамблевый метод, полуконтролируемое обучение, обратная задача

Адрес для переписки:

Саечников И.В.
Белорусский государственный университет,
пр-т Независимости, 4, г. Минск 220030, Беларусь
e-mail: saetchnikovivan@gmail.com

Address for correspondence:

Saetchnikov I.V.
Belarusian State University,
Nesavisimosty Ave., 4, Minsk 220030, Belarus
e-mail: saetchnikovivan@gmail.com

Для цитирования:

И.В. Саечников, В.В. Скакун, Э.А. Чернявская.
Прецизионное детектирование пространственно-разнесённых
объектов на аэрокосмических снимках для решения обратных задач.
Приборы и методы измерений.
2025. Т. 16. № 4. С. 379–385.
DOI: 10.21122/2220-9506-2025-16-4-379-385

For citation:

Saetchnikov IV, Skakun VV, Tcherniavskaja EA.
Detection of Spatially Distributed Fine-Grained Objects in Aerospace
Images for Inverse Problem Solving.
Devices and Methods of Measurements.
2025;16(4):379–385. (In Russ.).
DOI: 10.21122/2220-9506-2025-16-4-379-385

DOI: 10.21122/2220-9506-2025-16-4-379-385

Detection of Spatially Distributed Fine-Grained Objects in Aerospace Images for Inverse Problem Solving

I.V. Saetchnikov, V.V. Skakun, E.A. Tcherniavskaia

Belarusian State University,
Nesavisimosty Ave., 4, Minsk 220030, Belarus

Received 10.10.2025

Accepted for publication 21.11.2025

Abstract

Remote sensing data processing based on deep learning methods has been widely used in several cross-domain tasks. Current approaches predominantly focus on developing deep learning methods dedicated to object detection in densely populated urban environments, which are not well suited for sparsely distributed objects, especially fine-grained ones. A representative example is maritime vessel detection on open water surfaces. The challenges in this domain include heterogeneous noise backgrounds caused by varying lighting conditions, external artifacts, wave-induced blurs, top-view perspectives that yield poor feature maps of the objects, and the need to scan large areas for object searching. To address these challenges, this paper proposes a novel hybrid deep learning approach for sparse distributed fine-grained object detection. The proposed method includes an ensemble of classifiers based on DenseNet_1 and ResNet_1, combined with a YOLOv11+ detector and a semi-supervised learning strategy. Furthermore, for fine-grained object detection, an improved version of the YOLOv11+fg model is proposed, featuring adapted C3K2, UpSample, and SPFF modules. The approach was developed, pretrained, and validated using the Airbus SPOT Satellite Imagery dataset, and further trained with a subset of ShipRSImageNet for fine-grained detection. Results demonstrate the effectiveness of large-scale scanning in remote sensing and recognize sparsely distributed objects, outperforming the fine-grained detection score compared to the original model.

Keywords: hybrid deep learning, object detection, ensemble, semi-supervised, inverse problem

Адрес для переписки:

Саечников И.В.
Белорусский государственный университет,
пр-т Независимости, 4, г. Минск 220030, Беларусь
e-mail: saetchnikovivan@gmail.com

Address for correspondence:

Saetchnikov I.V.
Belarusian State University,
Nesavisimosty Ave., 4, Minsk 220030, Belarus
e-mail: saetchnikovivan@gmail.com

Для цитирования:

И.В. Саечников, В.В. Скакун, Э.А. Чернявская.
Прецизионное детектирование пространственно-разнесённых
объектов на аэрокосмических снимках для решения обратных задач.
Приборы и методы измерений.
2025. Т. 16. № 4. С. 379–385.
DOI: 10.21122/2220-9506-2025-16-4-379-385

For citation:

Saetchnikov IV, Skakun VV, Tcherniavskaia EA.
Detection of Spatially Distributed Fine-Grained Objects in Aerospace
Images for Inverse Problem Solving.
Devices and Methods of Measurements.
2025;16(4):379–385. (In Russ.).
DOI: 10.21122/2220-9506-2025-16-4-379-385

Введение

Достижения в области прикладных методов машинного обучения значительно усовершенствовали подходы к построению карты признаков на сложных гетерогенных данных, что способствует развитию множества приложений, включая робототехнику и анализ геопространственных данных. В частности, результаты предыдущих исследований показали высокую эффективность обнаружения плотно сгруппированных объектов на аэрофотоснимках, сделанных в городской местности [1–2]. Однако более актуальной и на данный момент менее изученной является задача детектирования пространственно-разнесённых объектов с низкой визуальной различимостью внутри одного класса на гетерогенном фоне [3–5]. Например, обнаружение частных самолетов на необорудованных взлётно-посадочных полосах или обнаружение кораблей в открытом море.

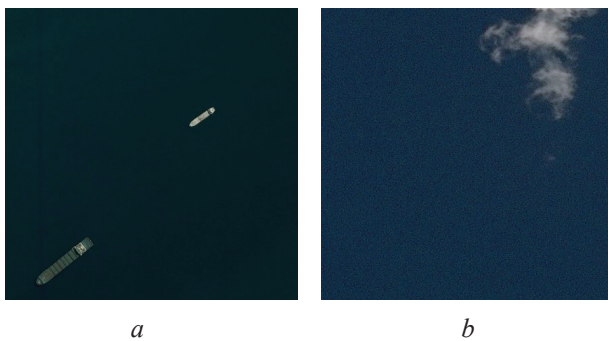


Рисунок 1 – Примеры изображений исследуемых объектов: *a* – однородное изображение с двумя объектами исследования; *b* – гетерогенное изображение без объектов исследования

Figure 1 – Examples of images of the observed objects: *a* – homogeneous image with two objects of interest; *b* – heterogeneous image without objects of interest

Задача обнаружения морских объектов на спутниковых снимках в данных условиях дополнительно осложняется спутниковым ракурсом сверху, который обеспечивает ограниченное разнообразие визуальных признаков, что затрудняет распознавание судов, особенно принадлежащих к разным подклассам [6–7]. Кроме того, гетерогенный фон, включающий волны, облака, отражения и разнообразную текстуру воды, дополнительно усложняет процесс обнаружения (рисунок 1*a*) по сравнению с однородным фоном, представленным на рисунке 1*b*.

Несмотря на то, что основная часть современных методик анализа аэрокосмических изображений фокусировались на детектировании объектов в городских условиях, были предложены ряд комбинированных подходов применительно к детектированию пространственно-разнесённых объектов. В работе [8] рассмотрена задача обнаружения морских судов как проблема с высоким уровнем дисбаланса по размеру объектов и предложена пирамидальная модель сети. В статье [9] предложена комбинированная модель eYOLO, основанная на методе взвешенного слияния блоков, объединяющем YOLOv4 и YOLOv5. Однако, несмотря на значительный прогресс в данной области, всё ещё сохраняются ряд актуальных аспектов для изучения:

1. Современные одноэтапные модели обнаружения объектов менее эффективны при анализе изображений с низкой плотностью объектов без существенных модификаций, что требует дополнительной предварительной обработки.

2. Построение карт признаков для прецизионного детектирования ограничено спецификой аэрокосмических изображений, что приводит к увеличению числа ложноположительных ошибок.

3. Применение существующих методов в широких географических масштабах ограничено из-за неоднородной и зашумлённой среды.

Таким образом, в данной статье предлагается гибридная модель глубокого обучения для распознавания пространственно-разнесённых объектов в многопараметрических наборах данных с акцентом на решении обратной задачи прецизионного детектирования объектов и их анализа.

Методология

Для обучения и валидации модели использовались два датасета: датасет на основе спутниковых снимков Airbus SPOT для детектирования и датасет ShipRSImageNet для прецизионного детектирования объектов с определением подкласса морского объекта. [10]

В качестве первого набора данных использовалась аннотированная часть датасета спутниковых изображений Airbus SPOT, включающая оптические изображения со спутников Pléiades, SPOT, Vision-1 и DMC,

а также радиолокационные изображения с аппаратов TanDEM-X, PAZ и TanDEM-X. Изначально набор данных содержал изображения с разрешением 1,5 м размером 768 на 768 в формате JPEG. Данные изображения были фрагментами, извлечёнными из сверхбольших данных дистанционного зондирования. При этом несмотря на однородный фон по типу (например, преимущественно водная поверхность), фон заметно отличался как по цвету, так и по текстуре из-за различных местоположений съёмки, времени суток и погодных условий. Также были проведена балансировка обучающего датасета: количество изображений, не содержащих исследуемые объекты, была уменьшена с 85 % изображений до 15 %. Первоначальная разметка в виде сегментационных масок была преобразована в ориентированные ограничительные рамки для задачи детектирования.

Для подготовки данных к обучению классификатора (в частности, для обработки на графическом процессоре) из исходных изображений были извлечены фрагменты размером 224×224 пикселя с перекрытием в 48 пикселей. Для повышения способности модели обучаться на инвариантах и улучшения обобщающей способности применялась аугментация данных – горизонтальные и вертикальные перевороты, перспективные трансформации, а также mixup аугментация.

Для задачи прецизионного детектирования изначальный набор данных дополнялся изображениями части датасета ShipRSImageNet, содержащего первоначально 3435 изображений с различных сенсоров, спутниковых платформ, локаций и сезонов. Изображения размером 930×930 пикселей включали 14 подклассов гражданских и коммерческих морских судов разного масштаба, ориентации и соотношения сторон.

Детектирование пространственно-разнесённых объектов на аэрокосмических снимках проводилось гибридной моделью глубокого обучения в два этапа (рисунок 2). На первом этапе проводилось сканирование и классификация входных фрагментов для выявления потенциальных объектов для детектирования. В дальнейшем успешные кандидаты с высокой вероятностью наличия объектов на изображениях детектировались. Таким образом, данный двухэтапный процесс исключал необходимость использования тяжёлых моделей детектирования для поиска объектов на изначально пустых изображениях. Кроме того,

он снижал количество ложных срабатываний, вызванных внешними артефактами, такими как блики волн, морской мусор и др.

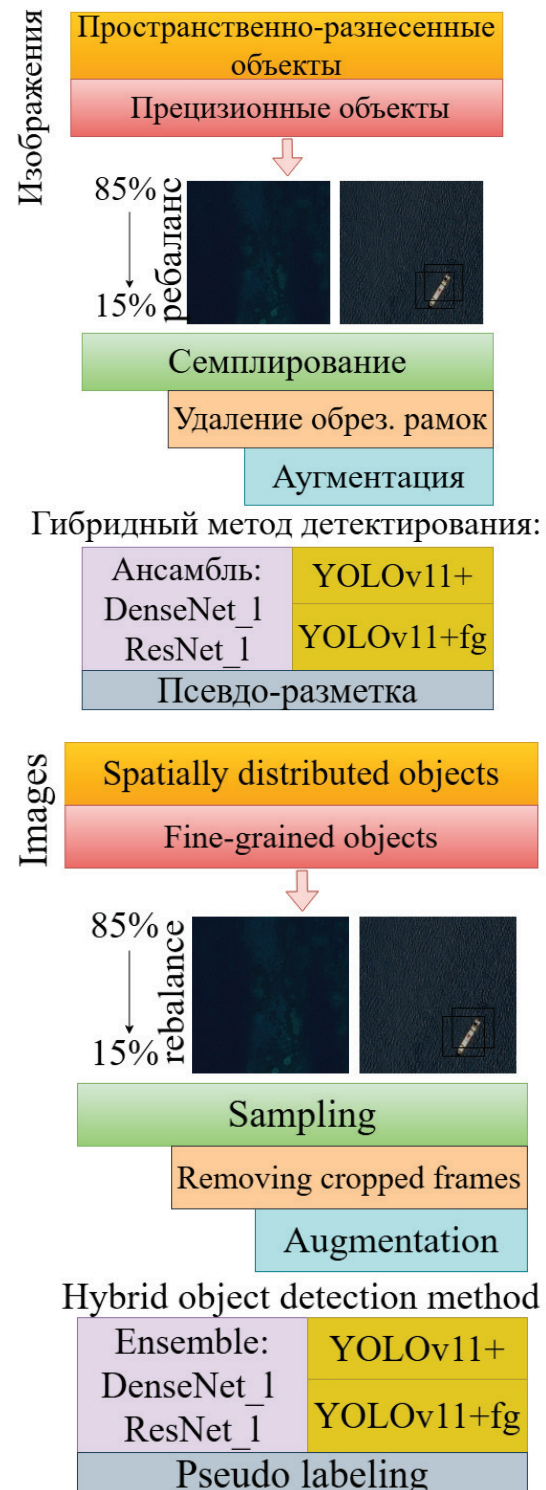


Рисунок 2 – Схема гибридного метода глубокого обучения для детектирования пространственно-разнесённых объектов на аэрокосмических снимках

Figure 2 – A hybrid deep learning approach for detecting spatially distributed objects on aerospace images

Первый этап базировался на ансамбле классификаторов, основанных на архитектурах DenseNet [11] и ResNet [12]. Архитектура сети ResNet позволяет фиксировать иерархические связи благодаря skip connections, тогда как сеть на основе DenseNet позволяет сохранять мелкие детали благодаря техники многослойного соединения. Были предложены адаптированные версии: в DenseNet_1 первая свёртка использовала меньший шаг, равный 1, и менее агрессивный размер ядра 3×3 , после чего применялся слой average pooling вместо max pooling. Это увеличение размера выборки на каждом последующем слое позволяло получать более мелкую пространственную информацию. Аналогичная модификация была применена к архитектуре ResNet_1. Эти модели обучались отдельно, после обучения полученные результаты усреднялись для прогнозирования вероятности присутствия объекта на изображении с целью отбора образцов для дальнейшей обработки на втором этапе.

Детектирование морских судов выполнялось на более легковесной модели архитектуры YOLOv11+ [13], используя ориентированные ограничительные рамки и адаптивные параметры max_channels, depth_multiple и width_multiple. Также в процессе обучения детектора применялся подход progressive resizing: обучение проходило в три цикла с постепенным увеличением разрешения изображения. Использовался оптимизатор AdamW со скоростью обучения 0,001 с последующим постепенным уменьшением скорости по косинусной стратегии с рестартами. Также в метод была интегрирована стратегия полуконтролируемого обучения с использованием псевдомаркировки, при котором 15 % обучающего датасета разбавляли псевдометками, сгенерированными обученной моделью на тестовых семплах.

После детектирования пространственно-рассеянных объектов была предложена последующая модификация архитектуры YOLOv11+fg с упором на задачу тонкого детектирования и определения подкласса объектов. В контексте задачи прецизионного детектирования были оптимизированы модули C3K2, UpSample и SPFF, так как они манипулировали многоуровневыми картами признаков. В частности, блок C3K2 был дополнен слоями spatial и channel attention,

а также агрегацией признаков, что усиливало релевантные каналы. Также в рамках модуля upsample использовались деформируемые адаптивные свертки, что позволяло точнее учитывать локальные геометрические вариации и контекстные особенности объектов на изображении. Также вместо классической метрики порога использовалась метрика WIoU.

Обучение модели проводилось на следующей аппаратной конфигурации: 14-ядерный процессор Intel Xeon W-2275 и 128 ГБ оперативной памяти, а также видеокарта Nvidia Quadro RTX 6000 с 24 ГБ оперативной памяти. Обработка данных осуществлялась с использованием Python, а также ряда библиотек глубокого обучения (Pytorch, TensorFlow, Keras, Rasterio, Abluminations и MMRotate).

Результаты

Оценка эффективности работы предложенного метода детектирования пространственно-разнесённых объектов на аэрокосмических снимках на примере морских судов включала количественное сравнение спрогнозированных ограничивающих рамок и аннотированных. В качестве основной метрики порога использовалась метрика пересечения по объединению (IoU).

Учитывая разнообразие размеров наблюдаемых объектов, мы применили оценку по метрике F2, усреднённую по набору пороговых значений в диапазоне от 0,5 до 0,95 с шагом 0,05.

Данный подход позволял избежать смещения к определённому пороговому значению. При каждом пороговом уровне оценка F2 рассчитывалась на базе значений истинно положительных (TP), ложноположительных (FP) и ложноотрицательных (FN) результатов. Истинно положительный результат регистрировался, когда предсказанный объект совпадал с истинным объектом со значением по IoU, превышающим пороговое значение. Ложноотрицательный результат возникал, когда истинный объект оставался необнаруженным, в то время как ложноположительный результат возникал, когда фрагмент фона определялся как исследуемый объект. Представленный результат по метрике F2 рассчитывался как усреднение значений этой метрики по всем пороговым значениям IoU.

Датасет не разделялся на обучающий и тестовый: валидация проводилась на основе стратегии перекрёстной проверки. Набор данных был разделён на фолды, и в течение циклов обученная модель проходила проверку по каждому из них, затем результаты усреднялись для гарантированного использования всего набора данных в валидации, что повышало репрезентативность финальных результатов. Достигнутые результаты распознавания объектов составили 0,84 до расширенной аугментации, 0,86 после её применения и 0,92 после применения полуконтролируемого обучения через псевдомаркировку.

Для валидации модели прецизионного детектирования YOLOv11+fg использовалась основная метрика mAP (50) в связи с увеличением числа детектируемых классов до 16 по сравнению с предыдущим случаем, где было всего 2 класса. Модель была предобучена на предыдущем AirbusSPOT и дообучена для решения

мультиклассовой задачи прецизионного детектирования. Увеличение точности детектирования по метрике mAP составило 2,5 % по сравнению с точностью базовой архитектуры YOLOv11n.

Результаты работы гибридного метода детектирования пространственно-разнесённых объектов и её визуализация представлены на рисунке 3.

Таким образом была продемонстрирована устойчивость по точности детектирования пространственно-разнесённых объектов на примере морских судов с использованием предложенного метода. Это касается как вариаций степени гетерогенности фона, как показано на рисунках 3*a* и 3*b*, так и детектирования на гомогенном фоне (рисунок 3*f*). Кроме того, предложенная модель устойчива к детектированию объектов малого размера, что иллюстрируется рисунками 3*c* и 3*d* и рядом расположенных морских судов без сливаний (рисунок 3*e*)

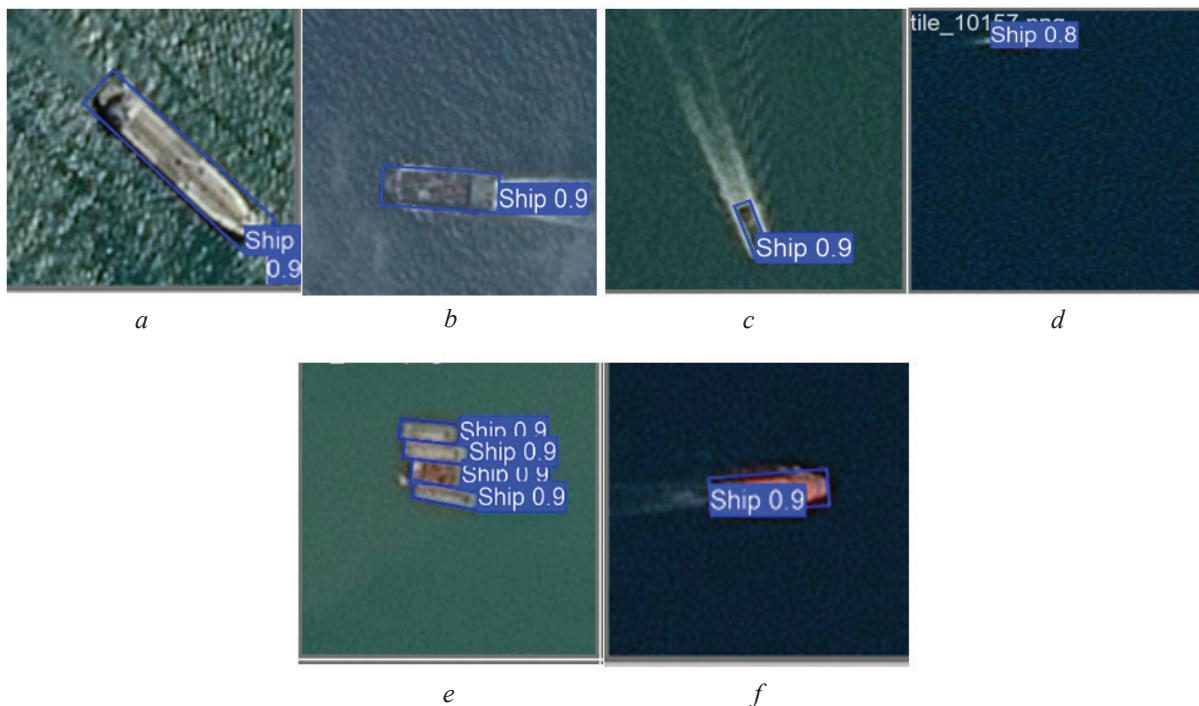


Рисунок 3 – Результаты работы гибридного метода детектирования пространственно-разнесённых морских судов в различных условиях: *a, b* – детектированные суда на гетерогенном фоне; *c, d* – детектированные суда малого размера; *e* – детектированные рядом расположенные суда; *f* – детектированное судно на гомогенном фоне

Figure 3 – Results of the spatially distributed ships detection in various conditions: *a, b* – detected ships on a heterogeneous background; *c, d* – detected small vessels; *e* – detected vessels located closely; *f* – detected vessel on a homogeneous background

В заключение следует отметить, что предлагаемый метод может быть интегрирован в решение ряда других

обратных задач, таких как отслеживание объектов [14] и определение траектории их движения.

Заключение

Данная работа акцентируется на задаче прецизионного детектирования пространственно-разнесённых объектов на примере объектов исследования типа морских кораблей.

В работе предложен двухэтапный метод детектирования пространственно-разнесённых объектов, реализуемый на базе ансамбля классификаторов DenseNet и ResNet в сочетании с детектором YOLOv11n+ и стратегией полуконтролируемого обучения. Метод обучен и валидирован на спутниковом наборе данных Airbus SPOT. Достигнутая точность распознавания морских судов на базе метрики F2 составила 0,92. Полученные результаты демонстрируют эффективность метода для задач крупномасштабного сканирования аэрокосмических изображений с целью обнаружения пространственно-разнесённых объектов, что обусловлено эффективной предобработкой данных и использованием ансамбля классификаторов. Предлагаемый подход обладает потенциалом расширения на другие классы пространственно-разнесённых объектов. Для последующей задачи прецизионного детектирования предложена модификация детектора YOLOv11+fg. Предложенный детектор, дообученный и провалидированный на части датасета ShipRSImageNet, продемонстрировал улучшение на 2,5 % по метрике mAP по сравнению с базовой архитектурой детектора версии nano. Обсуждены возможности применения предложенного гибридного метода к ряду других обратных задач.

References

1. He Lh. [et al.]. Research on object detection and recognition in remote sensing images based on YOLOv11. *Nature: Scientific Reports*. 2025;(15):14032. **DOI:** 10.1038/s41598-025-96314-x
2. Saetchnikov I, Tcherniavskaia E, Skakun V. Object Detection for Unmanned Aerial Vehicle Camera via Convolutional Neural Networks. *IEEE Journal on Miniaturization for Air and Space Systems*. 2021;2(2):98-103. **DOI:** 10.1109/JMASS.2020.3040976
3. Zhu XX. [et al.]. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*. 2017;5(4):8-36. **DOI:** 10.1109/MGRS.2017.2762307

4. Cheng G, Zhou P, Han J. Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*. 2016;54(12):7405-7415. **DOI:** 10.1109/TGRS.2016.2601622
5. Ding J. [et al.]. Object Detection in Aerial Images: A Large-Scale Benchmark and Challenges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2022;44(11):7778-7796. **DOI:** 10.1109/TPAMI.2021.3117983
6. Guo H, Yang X, Wang N, Gao X. A CenterNet++ model for ship detection in SAR images. *Pattern Recognition*. 2021;112(1):107787-107787. **DOI:** 10.1016/j.patcog.2020.107787
7. Gupta H. [et al.]. Ship detection using ensemble deep learning techniques from synthetic aperture radar imagery. *Scientific Reports*. 2024;14(1). **DOI:** 10.1038/s41598-024-80239-y
8. Liu J, Yang D, Hu F. Multiscale Object Detection in Remote Sensing Images Combined with Multi-Receptive-Field Features and Relation-Connected Attention. *Remote Sensing*. 2022;14(2):427-447. **DOI:** 10.3390/rs14020427
9. Fan X. [et al.]. A Small-Ship Object Detection Method for Satellite Remote Sensing Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2024;(17):11886-11898. **DOI:** 10.1109/JSTARS.2024.3419786
10. Zhang Z, Zhang L, Wang Y, Feng P, He R. ShipRSImageNet: A Large-Scale Fine-Grained Dataset for Ship Detection in High-Resolution Optical Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2021;(14):8458-8472. **DOI:** 10.1109/JSTARS.2021.3104230
11. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017;2261-2269. arXiv:1608.06993v5
12. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA. 2016;770-778 pp. arXiv:1512.03385
13. Redmon J, Divvala S, Girshick R, Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA. 2016;779-788 pp. **DOI:** 10.1109/CVPR.2016.91
14. Saetchnikov I, Skakun V, Tcherniavskaia E. Deep Neural Network-Based Dynamical Object Recognition and Robust Multiobject Tracking Technique for Onboard Unmanned Aerial Vehicle's Computer Vision-Based Systems. *IEEE Journal on Miniaturization for Air and Space Systems*. 2023;4(3):250-256. **DOI:** 10.1109/JMASS.2023.3274929