РАКЕТНО-КОСМИЧЕСКОЕ ПРИБОРОСТРОЕНИЕ И ИНФОРМАЦИОННЫЕ СИСТЕМЫ
2024, том 11, выпуск 3, с. 55–63

КОСМИЧЕСКИЕ НАВИГАЦИОННЫЕ СИСТЕМЫ И ПРИБОРЫ. РАДИОЛОКАЦИЯ И РАДИОНАВИГАЦИЯ

УДК 004.8 EDN GAUWKC

Метод выделения надводных объектов со спутниковых снимков и аэрофотоснимков с использованием нейронных сетей

Д. В. Горбунов, *orcid: 0000-0002-4646-2636, darkwen3@mail.ru*
МГТУ им. Н.Э. Баумана, Москва, Российская Федерация
Г. А. Миронов, *orcid: 0009-0007-1859-7238, migregal@mail.ru*
МГТУ им. Н.Э. Баумана, Москва, Российская Федерация
К. Л. Тассов, *spin: 1141-6778, ktassov@bmstu.ru*
МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация. В последнее время задачи выделения надводных объектов становятся все более актуальными. Под задачей выделения подразумевается поиск местоположения объектов на изображении. Одно из возможных решений — это нейронные сети, способные как убрать шумы, так и провести выделение объектов. Находящиеся в открытом доступе статьи опираются на использование одной нейронной сети, из-за чего предлагаемые методы обладают относительно большой погрешностью. В связи с этим предложен метод выделения надводных объектов, основанный на ансамбле нейронных сетей. В рамках метода предложен алгоритм, включающий в себя этапы предобработки изображений и выделения соответствующих объектов на спутниковых снимках. Метод может служить в системах контроля различных акваторий как вспомогательный, выделять надводные объекты для дальнейшей обработки и получения необходимой информации. В статье рассмотрены предобработка изображений и выделение надводных объектов, приведены точность для тренировочного и валидационного наборов, а также результаты, полученные в ходе программной реализации метода, структура используемой нейронной сети и разработанного ансамбля.

Ключевые слова: выделение надводных объектов, сверточные нейронные сети, ансамбли нейронных сетей

Для цитирования: Горбунов Д. В., Миронов Г. А., Тассов К. Л. Метод выделения надводных объектов со спутниковых снимков и аэрофотоснимков с использованием нейронных сетей. *Ракетно-космическое приборостроение и информационные системы*. 2024. Т. 11. № 3. С. 55–63.

©Горбунов Д. В., Миронов Г. А., Тассов К. Л., 2024

РАКЕТНО-КОСМИЧЕСКОЕ ПРИБОРОСТРОЕНИЕ И ИНФОРМАЦИОННЫЕ СИСТЕМЫ
2024, том 11, выпуск 3, с. 55–63

КОСМИЧЕСКИЕ НАВИГАЦИОННЫЕ СИСТЕМЫ И ПРИБОРЫ. РАДИОЛОКАЦИЯ И РАДИОНАВИГАЦИЯ

Neural Network-Based Method of Detecting Surface Objects from Satellite and Aerial Images

D. V. Gorbunov, *orcid: 0000-0002-4646-2636, darkwen3@mail.ru*
Bauman Moscow State Technical University, Moscow, Russian Federation
G. A. Mironov, *orcid: 0009-0007-1859-7238; migregal@mail.ru*
Bauman Moscow State Technical University, Moscow, Russian Federation
K. L. Tassov, *spin: 1141-6778, ktassov@bmstu.ru*
Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract. Nowadays the tasks of surface object detection have become more and more relevant. The task of object detection is finding the location of objects in the image. One of the possible solutions is neural networks capable of removing noises and detecting objects. The publicly available papers rely on the use of a single neural network. Hence, the proposed methods have a relatively large margin of an error. In this regard, a method for detecting surface objects based on an ensemble of neural networks is proposed. The method proposes an algorithm that includes the steps of pre-processing images and detecting the corresponding objects on satellite images. The method can be applied in control systems of various water areas as an auxiliary one. Moreover, it can detect surface objects for further processing and obtaining the necessary information. The paper discusses the preprocessing of images and detection of surface objects. It also provides the accuracy for training and validation sets, as well as the results obtained during the software implementation of the method, the structure of the neural network used and the ensemble developed.

Keywords: detection of surface objects, convolutional neural networks, ensembles of neural networks

For citation: Gorbunov D. V., Mironov G. A., Tassov K. L. Neural Network-Based Method of Detecting Surface Objects from Satellite and Aerial Images. *Rocket-Space Device Engineering and Information Systems*. 2024. Vol. 11. No. 3. P. 55–63. (in Russian)

Введение

В настоящее время все больше внимания привлекает тема охраны и анализа акваторий. Помимо этого, особый интерес представляет задача автоматизации отслеживания малых надводных объектов в акваториях стран, например в акватории Черного моря.

Методы выделения различных надводных объектов имеют особое значение в текущем контексте, так как позволяют решить множество проблем: избежание столкновений судов, осуществление автономного плавания [1], обеспечение безопасности, целостности акваторий и прочие.

Существующие методы выделения надводных объектов основываются на использовании одной нейронной сети, в связи с чем обладают сравнительно большой погрешностью результатов, либо радара, который очень плохо выделяет как близко расположенные к берегу, так и малые надводные объекты [2]. Кроме того, большинство доступных методов предназначено для выделения объектов только с аэрофотоснимков, либо только со спутниковых снимков.

В статье рассмотрен метод выделения надводных объектов с использованием ансамбля нейронных сетей. Представленный метод лишен указанных выше недостатков: использование ансамбля позволило повысить точность результатов, работать с аэрофотоснимками и спутниковыми снимками, а также выделять близко расположенные к берегу и малые надводные объекты.

Постановка задачи

Рассмотрим задачу выделения надводных объектов.

На вход алгоритму подается файл, содержащий соответствующий снимок в формате PNG. Изображение должно иметь разрешение не менее 640×640 пикселей. При этом для корректной работы метода необходимо, чтобы надводные объекты имели размер не менее 25×25 пикселей.

Результатом работы алгоритма является множество обрамляющих окон, в каждом из которых находится надводный объект. Под обрамляющим окном понимается прямоугольная область изображения, чьи стороны расположены перпендикулярно соответствующим сторонам изображения. Окно задается координатами левого нижнего угла и длинами соответствующих сторон.

Метод должен быть устойчив к шумам и помехам, инвариантен к масштабу.

Метод выделения надводных объектов

Предлагаемый метод заключается в использовании ансамбля нейронных сетей, результаты работы которых объединяются посредством детерминированного алгоритма голосования. Ввиду использования нейронных сетей алгоритм требует предварительного обучения на некотором наборе данных.

1. Описание разработанного ансамбля нейронных сетей

«Ансамбль — это набор из нескольких алгоритмов или методов, которые называются слабыми экспертами. Результат работы ансамбля рассчитывается на основании результатов работы составляющих его слабых экспертов путем объединения с помощью некоторого алгоритма, который может быть представлен в разных формах: дерево решений, нейронная сеть и другие. Ансамблевые методы привлекательны в основном потому, что они способствуют усилению слабых экспертов, то есть повышается точность работы метода» [3, с. 35].

«Для повышения точности работы ансамбля прибегают к использованию различных алгоритмов, к которым относятся стекинг, бустинг и бэггинг, каждый из которых обладает определенными преимуществами и недостатками» [4, с. 5].

В предложенном методе в качестве слабых экспертов выступают нейронные сети одинаковой архитектуры, тренировка которых происходила на разных обучающих выборках. На рис. 1 схематично изображен разработанный ансамбль, построенный с использованием бэггинга.

2. Описание используемых слабых экспертов

Один из основных типов нейронных сетей, используемых при решении задач выделения, — сверточные нейронные сети.

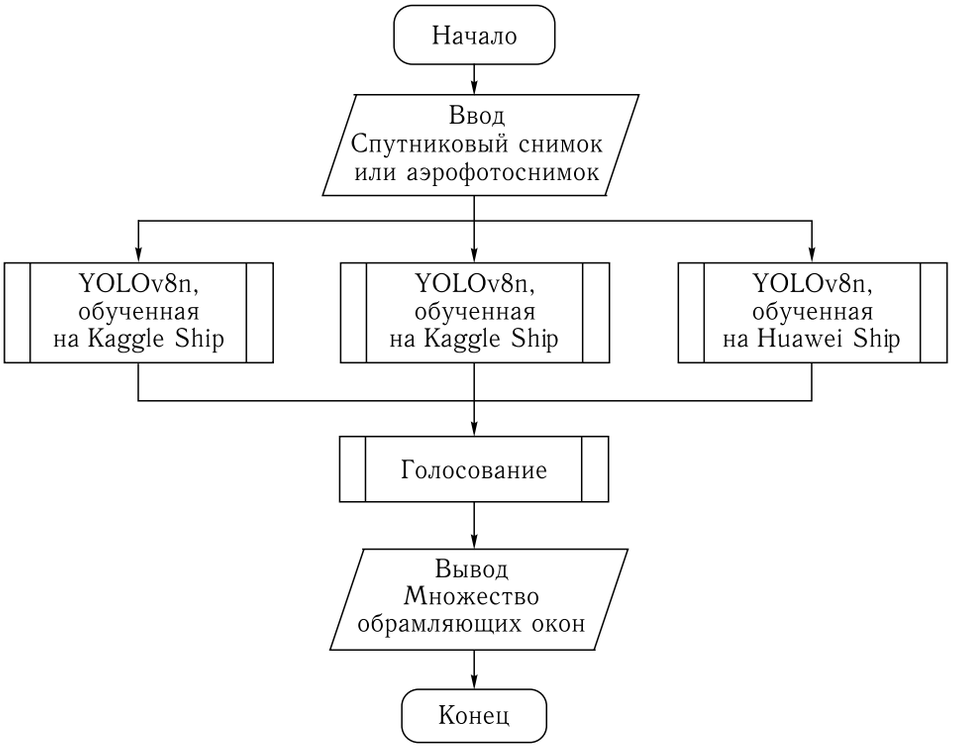

Рис. 1. Схема алгоритма работы предложенного метода
Fig. 1. The scheme of the operation algorithm of the proposed method

Попытки применить нейронные сети семейства YOLO — одного из наиболее активно развивающихся семейств сверточных нейронных сетей — к решению задач выделения надводных объектов не прекращаются [5, с. 2]. Так, при использовании модифицированной версии YOLOv3 tiny была достигнута средняя точность выделения более 61 % [6]. Однако серьезным недостатком данного семейства нейронных сетей является снижение точности работы алгоритма при выделении относительно больших и малых объектов [7].

В разработанном ансамбле в качестве слабых экспертов выступают несколько нейронных сетей последней версии архитектуры — YOLOv8, демонстрирующей наиболее высокую среднюю точность результатов в семействе [5, с. 18]. Так как ансамбль может быть обучен на меньшем размере выборки, чем аналогичная нейронная сеть, используется «младшая» версия сети — YOLOv8 nano — с минимальным числом параметров.

YOLOv8 имеет схожую с предыдущими версиями архитектуру. Тем не менее среди внесенных изменений стоит отметить одновременное использование как пирамиды признаков (англ. FPN), так и агрегацию путей (англ. PAN).

FPN — сверточная нейронная сеть, построенная в виде пирамиды, служащая для объединения достоинств карт признаков нижних и верхних уровней сети, первые имеют высокое разрешение, но низкую семантическую, обобщающую способность; вторые — наоборот [6].

PAN направлена на решение другой проблемы: в процессе выделения карт признаков часть информации о низкоуровневых признаках неизбежно теряется ввиду свойств нейронных сетей: чем больше слоев включает в себя нейронная сеть, тем сложнее получить корректные данные с нижних слоев на очередном слое. Как показывает практика, наличие этой информации непосредственно на более поздним слоях сети может заметно повысить точность результатов выделения [9].

3. Алгоритм голосования

Для принятия того или иного решения в ансамбле проводится голосование по результатам работы слабых экспертов, на основании которого

выбираются наиболее правдоподобные из представленных вариантов.

Результаты голосования могут быть определены среди прочего посредством консенсуса, простого большинства или же взвешивания. В представленном методе использован последний алгоритм, подразумевающий сопоставление каждому из слабых экспертов веса, описывающего достоверность результатов работы конкретного слабого эксперта.

Суть предлагаемого алгоритма сводится к следующему: при объединении результатов необходимо выбирать наиболее достоверный результат из возможных или же среднее среди наиболее достоверных, если существует неоднозначность.

Говоря формально, алгоритм принимает на вход множество взвешенных обрамляющих окон, включающее в себя результаты работы всех слабых экспертов ансамбля, а в качестве результата возвращает множество обрамляющих окон, соответствующих выделяемым надводным объектам.

Перейдем к более детальному рассмотрению алгоритма. Пусть дано множество предсказаний

$$X = (x_{ij}), \;\; i = \overline{1..n}, \; j = \overline{1..m},$$

где x_{ij} — достоверность выделения j-го объекта i-м экспертом. Тогда результат работы алгоритма можно представить следующим образом:

$$\left\{x : x_{ij} w_i = \max(xw), \;\; i = \overline{1..n}, \; j = \overline{1..m}\right\},$$

где $w = \{w_i\}$ — вектор весов, сопоставленных i-м экспертом.

Определения значений элементов данного вектора является нетривиальной задачей. В зависимости от задачи могут быть использованы такие алгоритмы, как OOB (Out of Bag Error), усреднение по байесовской модели, Random Forest, а также различные методы, учитывающие величину ошибки каждого из экспертов в процессе валидации результатов обучения. В текущей работе использовался алгоритм BEM (Basic ensemble method), предлагающий использование равных весов для всех слабых экспертов. Например, для N слабых экспертов имеем:

$$w_1 = \ldots = w_i = \ldots = \frac{1}{N}.$$

В общем случае обрамляющие окна, полученные в результате работы слабых экспертов, могут пересекаться или же полностью дублироваться, имея при этом одинаковый вес.

Предположим, что в результате работы слабых экспертов в множестве предсказаний X имеется подмножество пересекающихся окон Y.

Пусть для наглядности обрамляющее окно b задается координатами левого нижнего (b_{x1}, b_{y1}) и правого верхнего (b_{x2}, b_{y2}) углов. Тогда центр обрамляющего окна вычисляется следующим образом:

$$c(x) = \left(\frac{(b_{x1} + b_{x1})}{2}, \frac{(b_{y1} + b_{y1})}{2}\right).$$

Обозначим Y^k, $k \in [1, n]$, подмножество из k произвольных предсказаний из Y.

Обозначим также C_k, $k \in [1, n]$, геометрический центр следующего множества:

$$\left\{c : c(y_i), \;\; y_i \in Y^k, \;\; i = \overline{1..k}\right\},$$

где k — число пересекающихся предсказаний, имеющих одинаковый вес.

Наконец, обозначим d евклидово расстояние между двумя точками изображения. Тогда решением в случае пересекающихся предсказаний будем считать пересечение следующего множества предсказаний:

$$\left\{y : y \in Y^l, \;\; d(C_1, C_n) \to \min, \;\; l \in [1, n]\right\}.$$

Соответствующая схема алгоритма представлена на рис. 2.

Данные для обучения

Рассматриваются снимки из следующих наборов данных:

- Kaggle Ships In Google Earth;
- ShipRSImageNet [10];
- Huawei Ship.

Отдельная выборка состоит из списка пар изображение — множество размеченных объектов. Под разметкой объекта подразумевается положение обрамляющего окна в относительных координатах изображения [7].

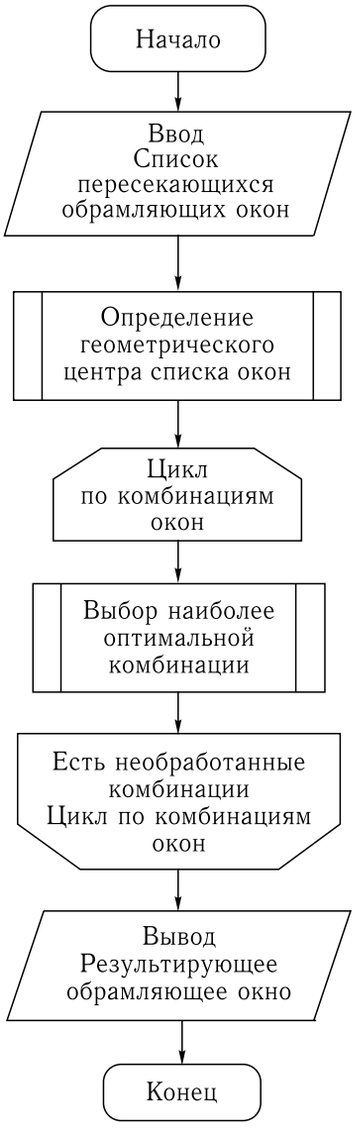

Рис. 2. Схема алгоритма голосования для пересекающихся предсказаний
Fig. 2. The scheme of the voting algorithm for overlapping predictions

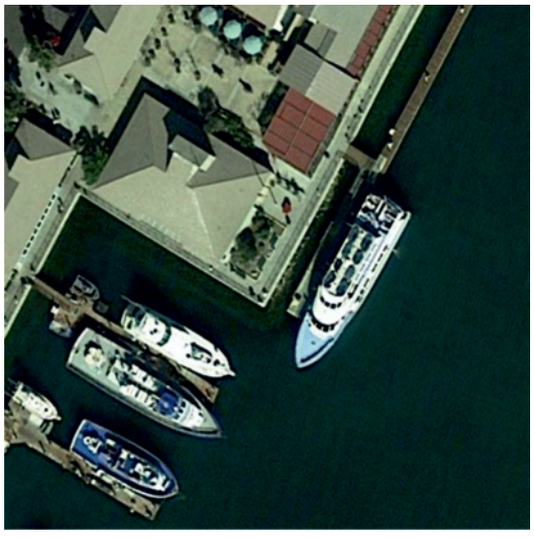

Рис. 3. Пример снимка, Kaggle Ships
Fig. 3. The sample image, Kaggle Ships

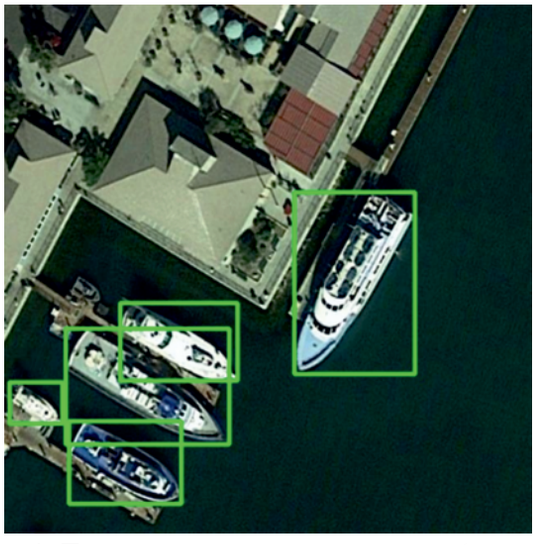

Рис. 4. Пример снимка с выделенными объектами
Fig. 4. The image example with the objects detected

Изображения представлены в формате PNG. Размерность изображений составляет от 225×225 до 930×930 пикселей. В выборке присутствуют разноракурсные фотоснимки, полученные в дневное время суток. Пример исходного снимка, а также его вид с обрамляющими окнами, полученными в результате работы алгоритма, приведены на рис. 3 и рис. 4 соответственно.

Имеющиеся наборы данных разбиваются на 3 выборки: обучающую, тестовую и валидационную. Обучающая выборка используется для формирования весовых коэффициентов нейронной сети, тестовая — для подбора оптимальных параметров в процессе обучения, а валидационная — для итоговой оценки результатов обучения.

Суммарно полученные наборы данных содержат 10671 снимок. В связи с относительно малым числом снимков в выборке произведено обогащение данных путем применения ряда трансформаций: горизонтального отражения, поворотов, различных

размытий и исключения процентной доли пикселей исходного изображения. В результате суммарно полученные наборы данных включают 100 тыс. снимков, разделенных следующим образом.

- Kaggle Ships In Google Earth — 13 747 снимков:
- обучающая — 11 685,
- валидационная — 1374,
- тестовая — 688.
- ShipRSImageNet — 30 242 снимка:
- обучающая — 25705,
- валидационная — 3025,
- тестовая — 1512.
- Huawei Ship — 56011 снимков:
- обучающая — 47 609,
- валидационная — 5602,
- тестовая — 2801.

Каждый из трех экспертов обучался на своей выборке, т. е. исходные наборы данных не смешивались.

Показатели точности

При работе с выделением объектов используется понятие IoU — отношение площади пересечения обрамляющих окон к их общей площади. В частности, данную величину используют для определения близости полученного обрамляющего окна к «истинному». Объект считается «правильно» выделенным, когда значение IoU превышает заданный порог. В качестве порога обычно используется одно из следующих значений: 0,75, 0,9, 0,95 [11].

В качестве метрики оценки результатов работы алгоритма используется точность (англ. Average Precision, AP) [11]:

$$\text{Точность} = \frac{n}{N},$$

где n — количество «правильно» выделенных объектов, N — количество объектов в выборке. Отметим, что точность вычисляется для объектов одного класса, в связи с чем в задачах выделения обычно используется понятие средней точности (англ. Mean Average Precision, mAP) — среднее значение точности.

На каждой эпохе вычисляется точность для валидационного и тренировочного наборов. Одной эпохой является полный проход по обучающей выборке.

После обучения слабых экспертов получены следующие показатели точности ансамбля:

- тренировочный набор — 100 %;
- тестовый — 91,2 %.

Необходимо учитывать, что в процессе оценки точности все слабые эксперты имели равный вес в голосовании. Это означает, что действительный уровень точности ансамбля находится несколько выше, так как в ходе валидации не учитывался контекст используемых снимков (ракурс, погодные условия).

На рис. 5 и 6 показаны графики точности слабых экспертов на тренировочном и валидационном наборах.

На рис. 7 представлено сравнение точности результатов работы предложенного метода с уже существующими решениями на тестовой выборке предложенного метода. Уже существующие решения были обучены на выборке, полученной путем

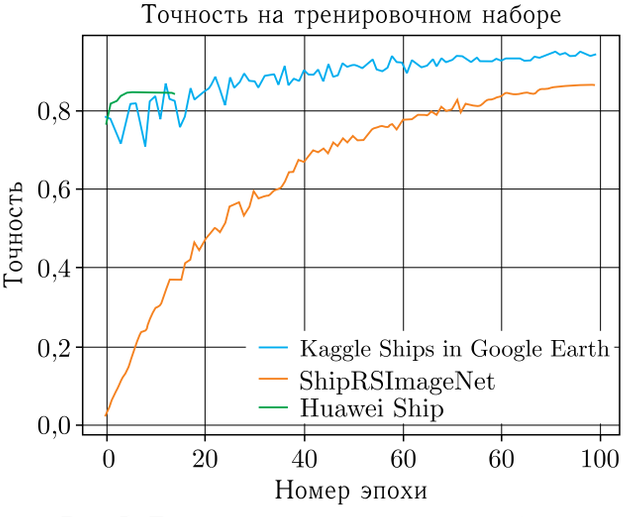

Рис. 5. Точность на тренировочном наборе
Fig. 5. Accuracy on the training set

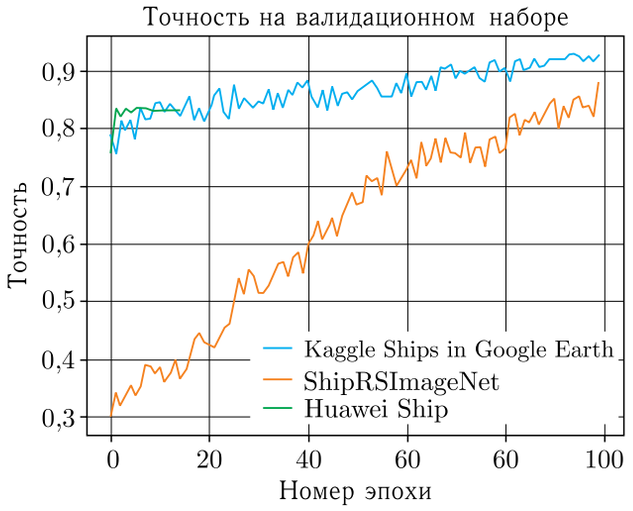

Рис. 6. Точность на валидационном наборе
Fig. 6. Accuracy on the validation set

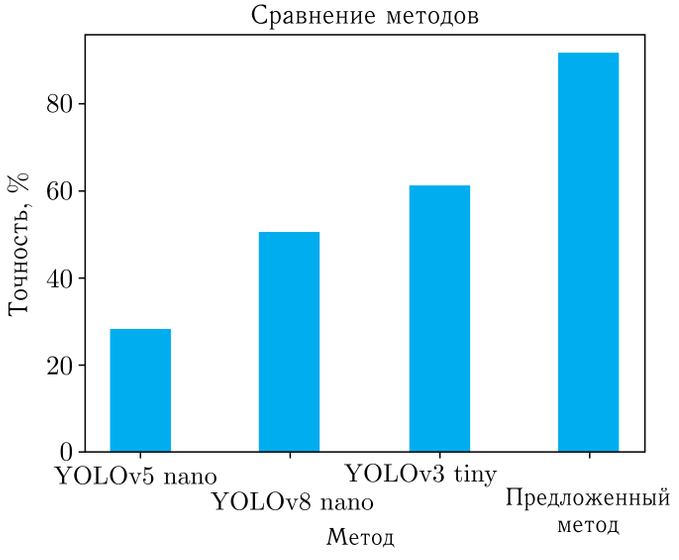

Рис. 7. Сравнение точности работы с существующими методами на основании тестовой выборки
Fig. 7. Comparison of the operation accuracy with the existing methods based on a test sample

объединения имеющихся наборов данных, в то время как предложенный метод обучался согласно приведенному выше описанию.

Заключение

Программное обеспечение, разработанное на основе предлагаемого метода, может быть использовано для избежания столкновений судов, осуществления автономного плавания, обеспечения безопасности, целостности акваторий и прочие. Разработанный метод позволяет выделить надводные объекты для дальнейшей обработки и получения информации. С помощью результатов выделения можно создать выборку для классификации надводных объектов.

К достоинствам представленного метода можно отнести: сравнительно высокую точность работы (порядка 91 %), работу с аэрофотоснимками и спутниковыми снимками, а также выделение близко расположенных к берегу и малых надводных объектов. Использование ансамбля позволяет уменьшить размер обучающей выборки. К недостаткам относится то, что работа метода проверена только на дневных снимках.

Разработанный метод имеет перспективу дальнейшего развития. Возможны улучшения, связанные с расширением выборки ночными снимками, а также добавлением классификации существующих надводный объектов.

Список литературы

1. *Lee Sung-Jun*, *Roh Myung-Il*, *Oh Min-jae*. Image-based ship detection using deep learning // Ocean Systems Engineering. № 10. 2020. P. 415–434.

2. *Chen Z.*, *Chen D.*, *Zhang Y.* Deep learning for autonomous ship-oriented small ship detection // Safety Science. № 130. 2020, 104812, 9 p.

3. *Евдокимова К.В.*, *Тассов К.Л.* Анализ алгоритмов, построенных на «слабых экспертах» // Вестник РГГУ. Сер. Информатика. Информационная безопасность. Математика. № 2. 2022. С. 33–47.

4. *Ganaie M.*, *Hu M.*, *Tanveer M.* Ensemble deep learning: A review // Engineering Applications of Artificial Intelligence. № 115. 2022, 105151, 47 p.

5. *Terven J.*, *Cordova-Esparza D. M.* A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond. arXiv:2304.00501. 2023, 34 p.

6. *Hao L.*, *Deng L.*, *Yang C.* Enhanced YOLO v3 Tiny Network for Real-time Ship Detection from Visual Image // IEEE Access. № 9. 2021. P. 16692–16706.

7. *Jiang P.*, *Ergu D.*, *Liu F.* A Review of Yolo Algorithm Developments // Procedia Computer Science. № 199. 2022. P. 1066–1073.

8. *Zheng M.*, *Ming L.*, *Yuguang W.* PAN: Path Integral Based Convolution for Deep Graph Neural Networks. arXiv:1904.10996. 2019, 7 p.
9. *Tsung-Yi L.*, *Piotr D.*, *Ross G.* A Feature Pyramid Networks for Object Detection. arXiv:1612.03144. 2017, 10 p.
10. *Zhang Z.*, *Zhang L.*, *Wang Y.* 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. № 14. 2021. P. 8458–8472.
11. *Hussain Muhammad.* YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection // Machines and Tooling. № 11. 2023, 25 p.

Дата поступления рукописи
в редакцию 19.02.2024
Дата принятия рукописи
в печать 26.08.2024