

Исследование моделей нейронных сетей для многоцелевых бортовых информационных систем радиомониторинга,

В.И. Вебер, В.Ю. Куприц, Ю.А.Светличный, А.А. Твердохлебов

Томский государственный университет систем управления и радиоэлектроники, 634050, г. Томск, пр-т. Ленина, 40

vladislav.i.veber@tusur.ru, vladimir.i.kuprits@tusur.ru

Аннотация

В статье рассматривается применение моделей нейронных сетей для обнаружения объектов в многоцелевых бортовых информационных системах радиомониторинга, включающих радиолокационные и оптические каналы. Приводятся результаты точности (precision) и полноты (recall), которые являются метриками используемыми при оценке вероятности правильного распознавания объектов различными моделями нейронной сети YOLO.

Ключевые слова

Многоцелевые бортовые информационные системы радиомониторинга, нейронная сеть, машинное обучение.

Study of neural network models for multi-purpose airborne radio monitoring information systems including optical channels

V.I. Weber, V.Y. Kuprits, Yu.A.Svetlichny, A.A. Tverdokhlebov

Tomsk State University of Control Systems and Radioelectronics, 634050, Tomsk, Lenin Ave. Lenin, 40

vladislav.i.veber@tusur.ru, vladimir.i.kuprits@tusur.ru

Annotation

The paper discusses the application of neural network models for object detection in multi-purpose airborne radio monitoring information systems that include optical channels. The results of evaluation of recognition accuracy and performance by metrics of accuracy, completeness and computational complexity are given.

Keywords

Multipurpose airborne radio monitoring information systems, neural network, machine learning.

Введение

В настоящее время одним из наиболее перспективных направлений развития автоматизированных систем обнаружения и классификации объектов являются интегрированные комплексы мониторинга (ИКМ), представляющие собой совокупность датчиков и сенсоров, систем связи и программных средств для обработки данных, которые совместно функционируют, для получения и обработки информации. Основная идея ИКМ заключается в совместной обработке данных от различных датчиков для формирования обобщенной информации об объекте и окружающем пространстве. Наиболее распространенными для применения в бортовых системах являются оптико-электронные и радиолокационные датчики.

Оптико-электронные системы (ОЭС) создают изображения объектов на основе электромагнитных волн оптического диапазона (от ультрафиолетового до инфракрасного). ОЭС включают в себя пассивные датчики, такие как телевизионные и тепловизионные, которые работают в различных диапазонах волн и обеспечивают разные возможности для мониторинга.

Инфракрасные (ИК) датчики обладают рядом преимуществ по сравнению с оптическими и радиолокационными системами. Прежде всего, они эффективны в условиях низкой освещенности или полной темноты, поскольку фиксируют тепловое излучение объектов, что делает их незаменимыми для ночного наблюдения. ИК-датчики также менее подвержены воздействию атмосферных помех, таких как дым, пыль или туман, что позволяет использовать их в сложных погодных условиях. В отличие от радиолокационных систем, ИК-датчики являются пассивными, что делает их менее уязвимыми для обнаружения и подавления.

Радиолокационные системы (РЛС) играют ключевую роль в ИКМ благодаря способности обнаруживать цели на больших дальностях, а также формировать радиолокационные изображения поверхности Земли в любых погодных и метеорологических условиях, как днем, так и ночью. Эти системы многофункциональны и используются для мониторинга воздушного, наземного и водного пространств, а также решают ряд других задач, связанных с целевым предназначением конкретных летательных аппаратов.

В настоящей статье рассматриваются результаты работы моделей нейронных сетей для распознавания наземных объектов в инфракрасном и радиолокационном диапазонах волн.

2. Система автоматического распознавания объектов на ИК и РЛ изображениях

Для практической реализации автоматической системы распознавания наземных объектов бортовыми ИКМ в инфракрасном и радиолокационном диапазонах волн необходимы размеченные базы данных соответствующих изображений, обученная на этих данных модель нейронной сети и бортовой компьютер.

Для обучения и повышения точности распознавания объектов моделью нейронной сети базы данных изображений являются одним из ключевых элементов. Без качественного и разнообразного набора данных невозможно создать эффективную автоматическую систему распознавания наземных объектов. Существует множество наборов данных с аэросъемкой для обучения моделей нейронных сетей, например, VisDrone [1] или UAV123 [2] с оптическими изображениями и FLAME [3], DroneRGBT [4], ASL-TID [5], с инфракрасными изображениями и ограниченное количество данных с радиолокационными изображениями MSTAR [6].

Большинство наборов данных не содержит информации о полете и условиях, в которых были получены изображения, таких как высота и угол визирования камеры, а также время суток (день или ночь), что не позволяет изучать влияние этих факторов на эффективность обнаружения при помощи нейронных сетей. Многие наборы данных сосредоточены на определенных аспектах, таких как конкретные категории объектов, низкая высота полета или одиночные сцены. Например, в наборе данных DroneRGBT присутствует только один класс и это человек, а все изображения в ASL-TID произведены с очень маленькой высоты.

Набор данных HIT-UAV10 [7] был создан для решения этих проблем и повышения точности обнаружения объектов на инфракрасных изображениях. Данный набор данных является общедоступным, полученный с помощью БПЛА и тепловизионной камеры DJI Zenmuse XT2. Он содержит 2989 инфракрасных изображений с пятью классами: Person, Bicycle, Car, OtherVechile, DontCare, в общей сложности 24899 аннотированных объекта. В класс DontCare вошли объекты, которые не идентифицировались во время сортировки данных по классам. Видеозапись осуществлялась в различных местах, например: у школы, дороги, парковки, спортивных площадок, высота варьировалась от 60 до 130 метров, при угле обзора от 30 до 90 градусов, полеты проводились как днем так и ночью.

Примеры изображений из этого набора данных показаны на рис. 1 для разного времени суток.

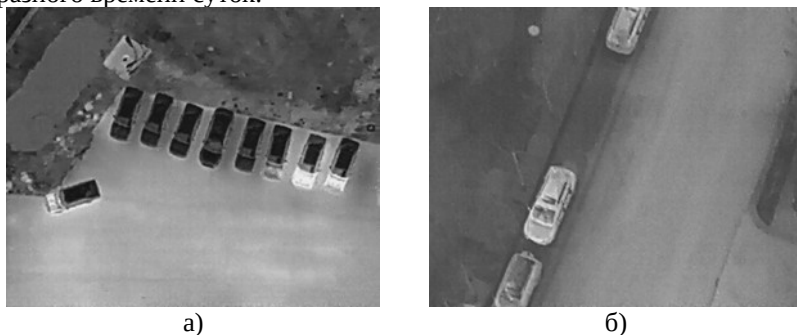


Рис. 1. Пример изображений из набора данных в ночное (а) и дневное (б) время суток

Fig. 1.Example of images from the dataset at night and day time

В последние годы из моделей нейронных сетей наиболее популярной является модель YOLO (You Only Look Once) [8]. Модель YOLO имеет архитектуру нейронной сети, предназначенной для решения задач обнаружения объектов в изображениях и видео. Основная идея YOLO заключается в том, что она рассматривает задачу обнаружения как задачу регрессии, предсказывая координаты ограничивающих рамок и классы объектов непосредственно из изображения в одном проходе. Исследования, проводимые ранее, показали превосходство модели YOLO над другими алгоритмами такими, как R-CNN [11] или SSD.

Обученные модели нейронной сети могут быть устанавливаются на бортовой микрокомпьютер с нейронным процессором. Например, на микрокомпьютер ROC-RK3588-PC на базе процессора Rockchip RK3588, оснащенного 8-ядерным процессором ARM Cortex-A76 до 2,4 ГГц, 6-ядерным графическим процессором Mali-G610 MC6 и нейронным процессором NPU с производительностью до 6 TOPS.

3. Сравнение моделей

Для оценки качества работы моделей использовались метрики таких как Recall(R) и Precision(P) [9]. Recall показывает, какую долю объектов положительного класса из всех объектов положительного класса нашел алгоритм.

$$Recall = \frac{TP}{TP + FN},$$

где TP – количество правильно идентифицированных положительных образцов; FN – количество ошибочно предсказанных отрицательных образцов, когда они на самом деле положительные.

Точность (Precision) можно интерпретировать как долю объектов, названных классификатором положительными и при этом действительно являющимися положительными.

$$Precision = \frac{TP}{TP + FP},$$

где FP – количество неправильно идентифицированных отрицательных образцов как положительных.

Recall демонстрирует способность алгоритма обнаруживать данный класс вообще, а precision — способность отличать этот класс от других классов.

При проведении исследований рассматриваются модели yolo5, yolo8, yolo8worldv2 yolo9 и их модификации. На рис. 3 и рис. 4 показаны результаты сравнения точности и полноты моделей.

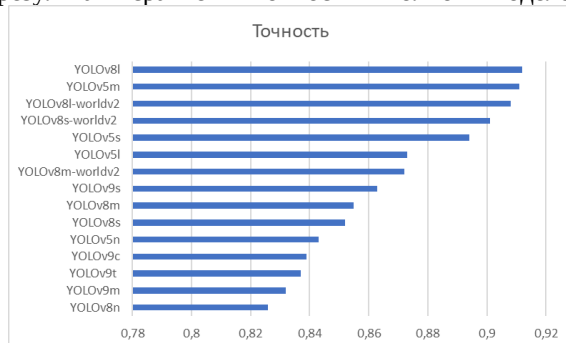


Рис. 3 Сравнение значений точности моделей

Fig. 3.Compression of model precision value

Среди рассмотренных моделей, представленных на рис.3, yolov8l и yolov5m показывают наилучший результат точности 0.912 и 0.911.

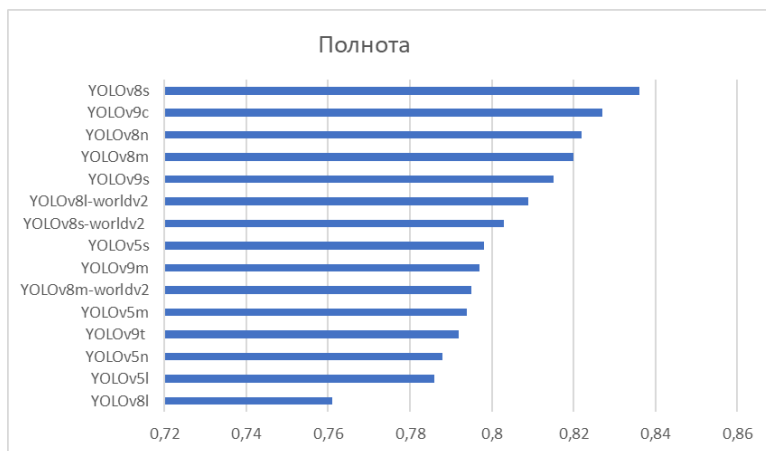


Рис. 4 Сравнение значений полноты моделей

Fig. 4.Compression of model Recall value

Как видно из рис. 4 наилучшие показатели полноты имеют модели yolov8s и yolov9c с значениями 0.836 и 0.827.

Примеры распознавания моделью YOLOv8l-worldv2 объектов на ИК изображении показаны на рис. 5.

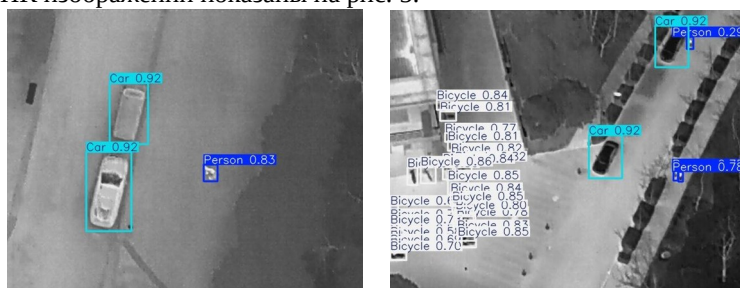


Рис. 5. Пример распознавания моделью YOLOv8l-worldv2 объектов на ИК изображении

Примеры распознавания моделью YOLO объектов на радиолокационном изображении показаны на рис. 6.

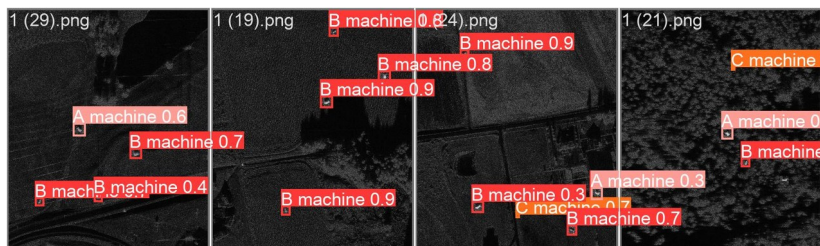


Рис. 6. Результат распознавания объектов на радиолокационных изображениях

Fig. 6. Result of object recognition a) in infrared images; b) on radar images

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ, проект FEWM-2023-0014.

Заключение

Перед проведением исследований был выполнен анализ различных наборов данных с оптическими, инфракрасными и радиолокационными изображениями, которые можно использовать для обучения моделей нейронных сетей с целью распознавания объектов. Из рассмотренных баз данных наибольший интерес для обучения моделей нейронных сетей представляет HIT-UAV10.

Результаты исследований показали, что из рассмотренных моделей нейронных сетей наилучший результат в точности классификации объектов показали yolov8l и yolov5mu с соответствующими значениями вероятности 0,912 и 0,911. Наилучшие результаты метрики полноты имеют модели yolov8s и yolov9c с значениями 0.836 и 0.827. Эти модели лучше использовать в системах где необходимо обнаруживать все объекты, даже если это приводит к увеличению числа ложных срабатываний.

Список литературы

1. Zhu, P., Wen, L., Bian, X., Ling, H. & Hu Q. Vision Meets Drones: A Challenge. Preprint at <https://arxiv.org/abs/1804.07437>, 2018.
2. Mueller, M., Smith, N. & Ghanem, B. A Benchmark and Simulator for UAV Tracking. Computer Vision—ECCV 2016: 14th European Conference, 2016, pp. 445–461.
3. Shamsoshoara, A. et al. Aerial Imagery Pile burn detection using Deep Learning: the FLAME dataset. Computer Networks 193, 1008001, 2021.

Режим доступа: <https://doi.org/10.1016/j.comnet.2021.108001>, (дата обращения: 2.08.2024).

4. Peng, T., Li, Q. & Zhu, P. RGB-T Crowd Counting from Drone: A Benchmark and MMCCN Network. Proceedings of the Asian Conference on Computer Vision (ACCV), 2020.

5. Portmann, J., Lynen, S., Chli, M. & Siegwart, R. People detection and tracking from aerial thermal views. 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014, pp. 1794–1800.

6. Dataset MSTAR. Режим доступа: <https://www.kaggle.com/datasets/atreyamajumdar/mstar-dataset-8-classes> (дата обращения: 2.08.2024).

7. Suo, J., Wang, T., Zhang, X. et al. HIT-UAV: A high-altitude infrared thermal dataset for Unmanned Aerial Vehicle-based object detection. Sci Data 10, 227, 2023. Режим доступа: <https://doi.org/10.1038/s41597-023-02066-6> (дата обращения: 2.08.2024).

8. Terven, Ju. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS / Ju. Terven, D. M. Córdova-Esparza, Ju. A. Romero-González // Machine Learning and Knowledge Extraction. – 2023. – Vol. 5, No. 4. – P. 1680-1716.

11. Timoshkin M.S. Comparison of YOLO V5 and FASTER R-CNN for detection people in the image in streaming mode / M.S. Timoshkin, A.N. Mironov, A.S. Leontev International Research Journal.- 2022. - №6 (120) - URL:

<https://research-journal.org/archive/6-120-2022-june/sravnenie-yolov5-i-fasterr-cnn-dlya-obnaruzheniya-lyudej-na-izobrazhenii-v-potokovom-rezhime> (accessed: 12.08.2024). - doi: 10.23670/IRJ.2022.120.6.020;