

АДАПТИВНЫЙ МЕТОД ДЕТЕКТИРОВАНИЯ И ОТСЛЕЖИВАНИЯ ДИНАМИЧЕСКИХ ОБЪЕКТОВ S-Y-BILSTM

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

Белорусский Государственный Университет, Минск, Беларусь

E-mail: saetchnikovivan@gmail.com

Отслеживание динамических объектов является одной из самых сложных задач компьютерного зрения, а применительно к беспилотным летательным аппаратам (БПЛА) обусловлено рядом причин. В частности, это большое расстояние до объектов слежения и разнообразие их размеров, неравномерность интенсивности, деградация изображения. В данной работе мы предлагаем адаптивный метод детектирования и слежения за динамическими объектами S-Y-BILSTM, базирующийся на использовании сети YOLOv4eff для формирования карты признаков дифференцирующего изображения, сети SSD в качестве метода обнаружения объектов, и оптимизированной двунаправленной сети на основе LSTM в качестве техники слежения. Оценка эффективности предложенного метода была проведена на самостоятельно оптимизированном наборе данных видеоряда, снятого БПЛА с высоты 15–45 метров. Доказана эффективность и надежность предложенного метода.

Ключевые слова: *динамический объект, детектирование, отслеживание, компьютерное зрение, беспилотный летательный аппарат (БПЛА), глубокое обучение, нейронная сеть.*

Введение. За последние несколько десятилетий область компьютерного зрения сильно продвинулась в таких задачах как анализ движения, захват и обнаружение объектов, слежение, в первую очередь, благодаря методам на основе глубокого обучения. Одной из наиболее сложных и перспективных областей применения являются беспилотные летательные аппараты. Несмотря на большие перспективы систем на основе компьютерного зрения, сценарии обнаружения и сопровождения объектов с беспилотных летательных аппаратов сопровождаются рядом проблем, в частности, деградацией изображения, неравномерной интенсивностью объекта, малым размером наблюдаемого объекта, проблемами в реализации в реальном времени, например, проблемами специфики перспективы, сложности фона, масштаба и разнонаправленности объектов на спутниках и пилотируемых летательных аппаратах [1].

Проблема деградации изображения: нагрузка, которую несет платформа мини-UAV, строго ограничена с точки зрения веса, объема и мощности. Быстрые изменения движения во внешней среде (такие как свет, облачность, туман, дождь и т. д.) вызывают нечеткость и зашумленность аэрокосмических изображений, что неизбежно приводит к деградации изображения [2].

Проблема неравномерной интенсивности объекта: оборудование для получения изображений с БПЛА обычно использует широкоугольные объективы с большой апертурой, фиксированным фокусным расстоянием. Более того, гибкое движение камеры приводит к неравномерной плотности захваченных объектов [3].

Проблема размера объекта: изображения, полученные с помощью БПЛА дистанционного зондирования, могут быть сняты на разных высотах, в результате чего регистрируются фотографии, содержащие наземные объекты разного размера. Это затрудняет применение классического подхода в анализе данных на основе глубокого обучения [2].

Таким образом, требуется разработать гибкий параметрический многоэтапный метод слежения за объектами, который будет обладать широкими возможностями оптимизации исходя из природы исследуемых объектов. Предложенная реализация метода S-Y-biLSTM для задачи отслеживания объектов с беспилотного летательного аппарата использует SSDeff детектор объектов, сеть YOLOv4eff в качестве метода извлечения признаков и двунаправленная LSTM сеть для отслеживания объектов по их траекториям.

R-S-BiLSTM OBJECT метод отслеживания динамических объектов. Метод R-S-biLSTM состоит из трех основных компонент: карта признаков дифференциальных изображений, созданная сетью извлечения признаков на основе YOLOv4eff [4], сеть SSD, в качестве метода обнаружения объектов для предсказания ограничивающих границ и усовершенствованная версия двунаправленной сети на основе LSTM в качестве метода отслеживания. Мы оптимизировали базовую архитектуру YOLOv4 для создания более семантически разветвленной карты признаков входного дифференцирующего изображения. Использовали 4 связи CSP с Mish активацией, экспоненциальное линейное взвешенное скользящее среднее (linearly weighted moving average - LWMA) и входную якорную сетку размером 11x11 ячеек. В качестве метода обнаружения мы предложили метод SSD_eff, который, с одной стороны, удовлетворяет условию мониторинга глубины. С другой стороны, он значительно сокращает количество параметров модели и улучшает способность магистральной сети к извлечению признаков. Мы предложили улучшенный алгоритм обнаружения объектов single shot multibox detector (SSD) - SSD_eff. Сеть извлечения признаков DenseNet-S-32-1 была разработана для замены оригинальной магистральной сети VGG-16 SSD. Нами предложен новый модуль объединения признаков для введения дополнительной контекстной информации в слои признаков пирамиды. Для каждого слоя признаков, используемого для обнаружения, добавляется модуль предсказания остатков. Двунаправленная LSTM использовалась в каче-

стве магистрального метода отслеживания благодаря возможности запоминать более долгосрочные связи, чем LSTM. В стандартной сети LSTM входные данные кодируются через полностью связанные слои и векторизуются – пространственная информация теряется, поэтому полносвязные слои были заменены на сверточные. В разработанной версии LSTM сверточные и рекуррентные операции при переходе от входа к состоянию и от состояния к состоянию полностью использовали информацию о пространственно-временной корреляции. Таким образом, удвоенная LSTM может получить доступ к дальнему контексту в обоих направлениях временной последовательности входного сигнала.

Результаты. Тесты были проведены на базе набора данных, собранного из части видеопоследовательностей с VisDrone 2019, Drone Vehicle Dataset и DTB70 [3]. Анализ проводился на базе последовательности объектов в режиме боковой съемки с БПЛА на высотах от 15 до 45 метров. Собранный набор данных включает 300 видеопоследовательностей длительностью 3,4 часа с соответствующими аннотациями объектов. Средняя площадь, занимаемая обнаруженными и отслеживаемыми объектами, составила около 3,4% пикселей от всего кадра, максимальная – 23,6%. Предобработка включала приведение к единому разрешению 1024x1024 пикселей изображения, используя предварительно обученную сеть ImresNet [5], основанную на стратегии остаточного обучения [5], нормализацию и разбивку на фрагменты (batches). В качестве функции потерь использовали функцию кросс-энтропии с параметрами α в 0,15, а значением γ в 1,5 (см. (1)).

$$FL(p_t) = -\alpha * (1 - p_t)(1 - p_t)^\gamma * \log(p_t). \quad (1)$$

Результаты отслеживания объектов представлены в таблице 1. Предложенный метод превзошел LYOLOv4eff, DeepSort и ROLO по большинству метрик MOT, представленных в таблице 1. В частности, мы значительно уменьшили количество ложной смены ID идентификаторов (IDsw: 37) по сравнению с LYOLOv4eff (IDsw: 89), что при более низком уровне удержания по метрике MT демонстрирует значительный положительный сдвиг в надежности предложенного метода по сравнению с LYOLOv4eff. Более высокое значение метрики ML достигнутое сетью LYOLOv4eff нивелируется более высокими значениями метрик FN и FP предложенным методом S-Y-biLSTM. Подводя итог, учитывая 2 основные метрики отслеживания множественных динамических объектов MOTA и MOTP, отмечаем, что предложенный метод достиг более высокой надежности и точности по сравнению с LYOLOv4eff [6] ROLO и DeepSort.

Оценка надежности предложенного подхода

Архитектура Сети	MOTA ↑ (%)	MOTP ↑ (%)	FN ↓	FP ↓
<i>ROLO</i>	39.3	59,2	28529	19423
<i>DeepSort</i>	36.9	52,8	27910	15268
<i>LYOLOv4eff</i> [6]	39.7	63,2	24523	12432
<i>S-Y-biLSTM</i>	41.7	68.6	23572	11923

Архитектура Сети	MT ↑	ML ↓	IDsw ↓
<i>ROLO</i>	134	394	57
<i>DeepSort</i>	116	415	74
<i>LYOLOv4eff</i> [6]	129	389	89
<i>S-Y-biLSTM</i>	153	375	67

В таблице представлены следующие обозначения: MOTA (multiple object tracking accuracy) – Ассурасу отслеживания объектов, MOTP (multiple object tracking precision) - Precision отслеживания объектов, FN (false negative) – количество пропущенных целей, FP (false positive) – количество “ghost” траекторий, MT (mostly tracked) – удержание отслеживания на протяжении 80% продолжительности жизни объекта, ML (mostly lost) – удержание менее 20 % продолжительности жизни объекта, IDsw (ID switches) – количество ложной смены ID.

Заключение. В данной работе была предложена сеть S-Y-biLSTM для задачи отслеживания дорожных объектов с БПЛА, состоящая из сети извлечения признаков на основе YOLOv4eff, сети SSD в качестве метода детектирования объектов, и двунаправленной сети LSTM для отслеживания динамических объектов. Анализ показал, что предложенный метод по основным метрикам отслеживания MOTA (41.7) и MOTP (68.6) превосходит ранее нами предложенный метод LYOLOv4eff [6] (39.7%; 63.2%), а также существующие подходы на основе сетей ROLO (39.3; 59.2) и DeepSort (36.9; 52.8). Предложенный нами метод S-Y-biLSTM превосходит по точности и надежности сеть LYOLOv4eff, что подтверждает эффективность использования сети на базе SSD, как детектора, и YOLOv4eff, как способа выделения информативных признаков.

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Maksai A., Fua P. Eliminating Exposure Bias and Metric Mismatch in Multiple Object Tracking // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA. 2019. P. 4634-4643, doi: 10.1109/CVPR.2019.00477.
2. Martinez-Martin E., del Pobil A. P. Object Detection and Recognition for Assistive Robots: Experimentation and Implementation // In IEEE Robotics & Automation Magazine. 2017. V. 24, N. 3. P. 123-138. doi: 10.1109/MRA.2016.2615329.
3. Kapania S., et al. Multi Object Tracking with UAVs using Deep SORT and YOLOv3 RetinaNet Detection Framework // In Proceedings of the 1st ACM Workshop on Autonomous and Intelligent Mobile Systems (AIMS '20). Association for Computing Ma-

- chinery, New York, NY, USA. 2020. Article 1, P. 1–6. DOI: <https://doi.org/10.1145/3377283.337>.
4. Saetchnikov I., Tcherniavskaia E. A., Skakun V. V. Object detection for unmanned aerial vehicle camera via convolutional neural networks // In IEEE Journal on Miniaturization for Air and Space Systems. V. 2. N. 2. P. 98-103. doi: 10.1109/JMASS.2020.3040976.
 5. Saetchnikov I., Skakun V. V., Tcherniavskaia E. A. Pattern recognition on aerospace images using deep neural networks // IEEE 7th International Workshop on Metrology for AeroSpace (MetroAeroSpace), Pisa, Italy. 2020. P. 336-340, doi: 10.1109/MetroAeroSpace48742.2020.9160198.
 6. Saetchnikov I., Skakun V. V., Tcherniavskaia E. A. Efficient objects tracking from an unmanned aerial vehicle // IEEE 8th International Workshop on Metrology for AeroSpace (MetroAeroSpace). 2021 P. 221-225. doi: 10.1109/MetroAeroSpace51421.2021.9511748.

МАШИННОЕ ОБУЧЕНИЕ ПРИ ПЕРЕМЕЩЕНИИ МОБИЛЬНОГО РОБОТА

А. В. Сидоренко, Н. А. Солодухо

Белорусский государственный университет, Минск, Беларусь

Рассмотрены вопрос моделирования при навигации с огибанием препятствий мобильного робота с использованием методов машинного обучения: Q-обучения, алгоритма SARSA, глубокого Q-обучения и двойного глубокого Q-обучения. Разработанное программное обеспечение включает средства Mobile Robotics Simulation Toolbox, Reinforcement Learning Toolbox и пакет визуализации Gazebo для моделирования среды. Результаты вычислительного эксперимента показывают, что для моделируемой среды размером 17 на 17 блоков и препятствия длиной в 12 блоков обучение при использовании алгоритма SARSA происходит с лучшей производительностью, чем для остальных

Ключевые слова: *робот, машинное обучение, Q-обучение, перемещение.*

ВВЕДЕНИЕ

При внедрении мобильных роботов в космическую, военную, производственную сферы деятельности человека одной из актуальных является проблема управления движением мобильного робота в некоторой среде при известном местоположении робота, расположении целевых точек, в которые должен переместиться робот [1]. При этом существенным является обеспечение безопасного движения робота без столкновения со встречающимися на его пути препятствиями.

При решении подобных задач, как правило, используются алгоритмы машинного обучения, включающие алгоритмы обучения с подкреплением, нейросетевые алгоритмы, алгоритмы глубокого обучения [2, 3]. Ис-