

УДК 621

**ПРОГРАММНАЯ РЕАЛИЗАЦИЯ МЕТОДА ИДЕНТИФИКАЦИИ И ОТСЛЕЖИВАНИЯ
ОБЪЕКТОВ ДЛЯ АВТОНОМНОЙ СИСТЕМЫ КОМПЬЮТЕРНОГО ЗРЕНИЯ БЕСПИЛОТНЫХ
ЛЕТАТЕЛЬНЫХ АППАРАТОВ**

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

*Белорусский государственный университет
Минск, Республика Беларусь*

Аннотация. Последние тренды в развитии микроэлектроники для малых космических и беспилотных аппаратов открывают широкие перспективы для внедрения методов идентификации и отслеживания в системах компьютерного зрения. В данной работе мы предлагаем подобный метод и его программную реализацию для автономной системы компьютерного зрения беспилотных летательных аппаратов. Метод основан на идентификации глубокой нейронной сетью YOLOv5tr и отслеживании двунаправленной сетью „долгой краткосрочная память“, дополненной сетью на базе автоэнкодера, формирующей карту признаков. Метод был обучен на отмаркированном наборе данных видеоряда с БПЛА и реализован в программной среде Python. Сравнительный анализ предложенного метода с существующими показал эффективность для практического применения на системах компьютерного зрения БПЛА.

Ключевые слова: компьютерное зрение, идентификация объекта, отслеживание объекта, нейронная сеть, беспилотный летательный аппарат (БПЛА).

**A SOFTWARE IMPLEMENTATION OF AN OBJECT IDENTIFICATION AND TRACKING
METHOD FOR AN AUTONOMOUS COMPUTER VISION SYSTEM OF AN UNMANNED
AERIAL VEHICLE**

Saetchnikov I., Skakun V., Tcherniavskaia E.

*Belarussian State University
Minsk, Republic of Belarus*

Abstract. Recent trends in the development of microelectronics for small spacecrafts and drones offer broad prospects for the implementation of identification and tracking methods in computer vision systems. In this paper, we propose a similar method and its software implementation for an autonomous computer vision system for unmanned aerial vehicles. The method is based on the identification by a deep neural network YOLOv5tr and tracking by a bi-directional network „long short-term memory, augmented by an autoencoder-based network that forms a feature map. The method was trained on a labelled UAV video dataset and implemented in a Python software environment. A comparative analysis of the proposed method with existing ones showed the effectiveness for practical application on UAV computer vision systems.

Key words: computer vision, object identification, object tracking, neural network, unmanned aerial vehicle (UAV).

*Адрес для переписки: Саечников И.В., пр. Независимости, 4, Минск 220030, Республика Беларусь
e-mail: saetchnikovivan@gmail.com*

В течение последних десятилетий область компьютерного зрения сделала заметный рыбок в задачах обнаружения объектов и их семантического анализа [1, 2]. При этом одним из наиболее перспективных областей применения являются системы компьютерного зрения [3], установленный на беспилотных летательных аппаратах. Однако, сценарии обнаружения и сопровождения объектов с беспилотных летательных аппаратов, особенно методами на основе глубокого обучения, сопровождаются требованиями устойчивости метода к спецификам перспективы съемки, сложности внешнего фона, варьировании высот съемки и т. д. [4, 5]. С другой стороны с точки зрения практического применения программная реализация должна быть аппаратно кросс-платформенна, чтобы обученные нейронные сети можно было запускать в реальном времени на борту беспилотного летательного аппарата.

Таким образом, в данной работе представлен разработанный адаптивный метод идентификации и отслеживания объектов и его программная реализация для автономной системы компьютерного зрения беспилотных летательных аппаратов.

Метод включает 3 основных этапа: формирование карты признаков кадров видеоряда путем прогона дифференциальных изображений через нейронную сеть на основе автоэнкодера, детектирование объекта техникой YOLOv5tr (дополненная 5-ю парами CSP, заменой на Mish активацию, размер входной якорной ячейки 10x10) и техника отслеживания двунаправленной нейронной сетью “долгой краткосрочная память” (LSTM) с заменой полностью связанных слоев сверточными, что улучшает понимание сети пространственно-временной связи. Двунаправленная сеть позволяет обучаться в обоих направлениях временной последовательности входного сигнала, что заметно улучшает

точность прогнозирования последующего состояния объекта на видеокдрах.

Обучение и тестирование проводилось на базе как самостоятельного набора данных, так и на предразмеченных кадрах видеорядов датасетов VisDrone 2020 и DTB70 [5]. При этом проводилось ранжирование выборок по высоте съемки от 10 до 50 м с шагом 10 для более скорого обучения сетей. Окончательный датасет включал 150 видеопоследовательностей общей длительностью 2 часа 20 минут с 40 % отмаркированными кадрами. После предобработки, включающей приведение к единому разрешению и разбивки на батчи, данные проходили процесс обучения через 3 этапа шаг за шагом.

Программная реализация метода была выполнена в среде Python с задействованием специализированных библиотек для анализа данных (например, Pandas, matplotlib, keras, и т. д.), а также библиотек для машинного обучения и построения нейронных сетей TensorFlow и Pytorch. Процесс обучения метода происходил следующим образом: после высчитывания дифференцирующего изображения он подавался на автоэнкодер, результат в виде карты признаков и соответствующих изображений шли на сеть детектирования объектов. По аннотируемым bounding boxes происходило обучение сети YOLOv5tr. Далее результаты работы сети YOLOv5tr, карты признаков построенные автоэнкодером подавалась на вход двунаправленной сети LSTM, которая в конечном счете после обучения выдавала прогноз по месторасположению объекта на следующем кадре с значением confidence score. Работа метода оценивалась по 6 метрикам [6]. Результаты работы метода представлены в таблице 1. Для сравнительного анализа эффективности метода по сравнению с конкурирующими решениями, были оптимизированы и протестированы на том же наборе выборок методы DeepSort и FairMot.

Таблица 1. Результаты работы метода идентификации и отслеживания объектов с БПЛА

	MOTA ↑ (%)	MOTP ↑ (%)	IDsw ↓
Предложенный метод	56,7	70,6	51
DeepSort	51,2	70,1	89
FairMot	44,5	67,2	153
	FN ↓ (%)	FP ↓ (%)	MT ↑
Предложенный метод	9,3	16,2	152
DeepSort	15,6	15,9	115
FairMot	11,1	25,4	124

где MOTA – Accuracy отслеживания объектов, MOTP – Precision отслеживания объектов, FN – false negative, FP – false positive, MT – mostly tracked, ML – mostly lost, IDsw – количество ложной смены ID.

Предложенный метод превзошел DeepSort и FairMot по двум основополагающим метрикам идентификации и отслеживания MOT и MOTP достиг значений 56,7 и 70,6 соответственно. Стоит отметить, что метод имеет значительное преимущество по метрике IDsw, что с нашей точки зрения связано с эффективно обученной техникой отслеживания, в частности, двунаправленной сетью, что позволяет оптимизировать долгосрочные и краткосрочные знания об объектах. Более низкое значение метрики FN нивелируется комплексной метрикой MOTA, составной частью которой она является.

Таким образом, результаты точности адаптивного метода идентификации и отслеживания показывают перспективу внедрения в рамках системы компьютерного зрения БПЛА. В частности, предварительно обученный метод, запущенный на компьютере Nvidia Jetson Nano, дополненный модулем передатчика, будет способен в режиме реального времени обрабатывать объекты на изображениях и получать данные для идентификации параметров обратных задач с последующей их передачей на землю. Совместимость программного кода с фреймворком Nano позволяет использовать DeepStream SDK на борту БПЛА.

Литература

1. Martinez-Martin, E. Object Detection and Recognition for Assistive Robots: Experimentation and Implementation / E. Martinez-Martin, A. P. del Pobil // IEEE Robotics & Automation Magazine. – 2017. – Vol. 24, №. 3. – P. 123–138.
2. The Unmanned Aerial Vehicle Benchmark: Object Detection and Tracking / Du, Dawei [et al.].
3. Online Multi-object Tracking Using CNN-Based Single Object Tracker with Spatial-Temporal Attention Mechanism / Q. Chu [et al.] // IEEE International Conference on Computer Vision (ICCV). – 2017. – P. 4846–4855.
4. Saetchnikov, I. Pattern recognition on aerospace images using deep neural networks / Saetchnikov, V. Skakun and E. Tcherniavskaia // IEEE 7th International Workshop on Metrology for AeroSpace (MetroAeroSpace). – 2020. – P. 336–340.
5. Saetchnikov, I. Object detection for unmanned aerial vehicle camera via convolutional neural networks / I. Saetchnikov, E. A. Tcherniavskaia, V. V. Skakun // IEEE Journal on Miniaturization for Air and Space Systems. – 2020.
6. Saetchnikov, I. Efficient objects tracking from an unmanned aerial vehicle / I. Saetchnikov, V. Skakun E. Tcherniavskaia // IEEE 8th International Workshop on Metrology for AeroSpace (MetroAeroSpace). – 2021. – P. 221–225.