

ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОЦЕНКИ СКОРОСТИ И УСКОРЕНИЯ БПЛА ПРИ ВЗЛЕТЕ И ПОСАДКЕ ЧЕРЕЗ МУЛЬТИКАМЕРНУЮ ДЕТЕКЦИЮ

Русаков К.Д.

Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия
rusakov@ipu.ru

Аннотация. В данной работе используются сверточные нейронные сети для мультикамерной детекции БПЛА и оценки их скорости и ускорения на этапах взлёта и посадки. Это открывает новые возможности в области навигации и контроля БПЛА.

Ключевые слова: Сверточные нейронные сети, беспилотные летательные аппараты, мультикамерная детекция, оценка скорости, оценка ускорения, взлет, посадка, навигация, контроль БПЛА.

Введение

Скорость и ускорение — это физические параметры, которые характеризуют перемещение одного объекта относительно другого и могут быть оценены количественно. В общем случае, прямое измерение скорости не является возможным, поэтому мы обычно прибегаем к оценке этих параметров с использованием косвенных методов. Оценка скорости играет критическую роль в многих прикладных сферах, например, при определении скорости беспилотного летательного аппарата. Для успешного применения метода оценки скорости в практике необходимо удовлетворить три критерия:

методика должна работать в реальном времени или близко к нему, должна быть возможность измерять скорость нескольких объектов одновременно. Благодаря прогрессу в области глубокого обучения и компьютерного зрения, некоторые сложные задачи теперь имеют новые решения. В контексте обнаружения объектов, предложены серии алгоритмов на основе сверточных нейронных сетей, такие как RCNN [1], Fast-RCNN [2], SSD [3], YOLO [4]. SSD и YOLO, в частности, могут работать в реальном времени. Эти алгоритмы способны определить ограничивающие рамки нескольких объектов на каждом кадре. Скорость объекта обычно оценивается двумя ключевыми способами: физическим (таким как радар или лазер) или через обработку изображений [5]. Каждый из этих методов имеет свои сильные и слабые стороны.

Физические методы работают по следующему принципу: радар или лазер излучают волновой или световой сигнал, который затем отражается от объекта и возвращается к приемнику. Так как скорость объекта влияет на частоту отраженного сигнала, сравнивая эту частоту, можно рассчитать скорость объекта. Этот подход активно применяется, и многие компании разрабатывают встроенные устройства для оценки скорости в реальном времени с точностью до 1 км/ч и диапазоном измерения от 18 до 300 км/ч. Однако этот метод также имеет свои недостатки:

- можно измерять только скорость транспортного объекта
- можно измерять только один объект одновременно
- расстояние измерения ограничено
- само устройство должно быть неподвижно

В контексте обработки изображений для оценки скорости, существуют различные подходы. В исследовании [6] представлен метод, который позволяет оценивать скорость нескольких объектов в режиме реального времени, используя обычную камеру, даже в движении. В исследовании [7] применяется калибровка камеры для установления связи между камерой и реальным миром, а затем через сопоставление изображений рассчитывается изменение положения объекта для оценки его скорости. В работе [8] использовалось априорное знание о размере объекта для его отслеживания и расчета скорости через разницу кадров. Авторы [9] использовали данные о фиксированной точке в камере для расчета скорости самой камеры. В исследовании [10] использовалась корреляция между двумя соседними кадрами для оценки перемещения объекта и вычисления его скорости. Аналогичным образом, в [11] применялись метод разности кадров и калибровка камеры для оценки скорости. Все эти методы применялись для измерения скорости автомобилей. В отношении других объектов, авторы [12] обнаружили, что параметр размытия в движении коррелирует со скоростью сферического объекта и использовали его для расчета скорости. Указанные методы эффективны в определенных сценариях, но имеют ряд ограничений: 1) большинство из них пригодны только для измерения скорости автомобилей или сферических объектов, 2) точность и надежность ограничены точностью и надежностью алгоритма

обработки изображения, 3) большинство методов предполагают статичность камеры, так как используют разность кадров для обнаружения объектов, что требует неподвижного фона.

1. Идентификация БПЛА на изображении

Для реализации обнаружения объектов в режиме реального времени была выбрана сверточная нейронная сеть с архитектурой YOLOv5 [13] в качестве основной модели. Схематическое представление этой сети можно увидеть на рисунке 1.

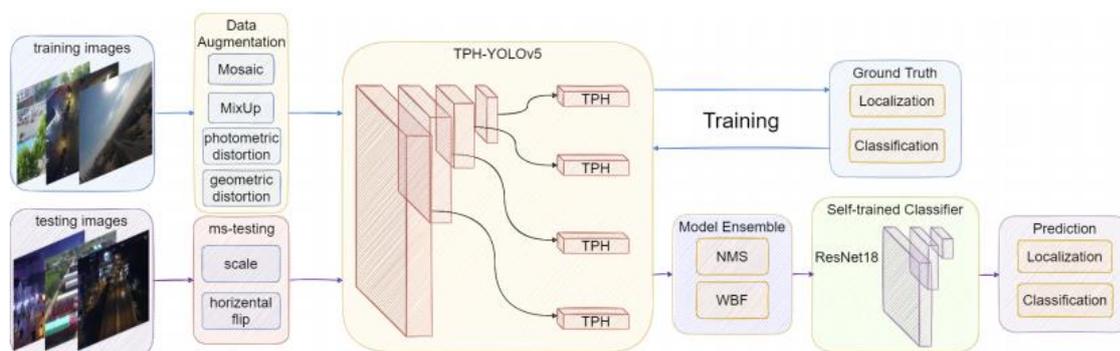


Рис. 1. Архитектура YOLOv5

Сверточные сети в основном разделяются на две составляющие: энкодер, основанный на сверточной нейронной сети, который служит для извлечения признаков из изображения, и детектор, который предназначен для классификации объектов и определения их ограничивающих рамок. К тому же, в недавно созданных детекторах объектов часто добавляются дополнительные слои между энкодером и детектором — это так называемая "шея" детектора. Энкодер в детекторе YOLOv5 может включать в себя различные архитектуры, такие как VGG, ResNet, DenseNet, MobileNet, EfficientNet, CSPDarknet53, Swin Transformer и прочие, не являясь при этом самостоятельно созданной сетью. Эти архитектуры демонстрируют высокую эффективность в извлечении признаков для классификации. "Шея" детектора разработана для оптимального использования функциональности энкодера, обрабатывая и эффективно применяя карты признаков, которые были извлечены энкодером на различных уровнях. Этот элемент является ключевым звеном в системе обнаружения объектов. Без "шеи" энкодер, действуя как классификатор, не способен выполнить задачу локализации. В данном случае детектор служит для определения положения и класса объекта на основе карт признаков, извлеченных энкодером.

Детекторы обычно подразделяются на одноступенчатые и двухступенчатые. Долгое время двухступенчатые детекторы были доминирующими в области обнаружения объектов, где RCNN является наиболее известным примером. В отличие от двухступенчатого детектора, одноступенчатый детектор одновременно прогнозирует ограничивающую рамку и класс объектов. Одноступенчатый детектор обладает преимуществом в скорости, хотя его точность может быть ниже. Среди одноступенчатых детекторов наиболее известными моделями являются серии YOLO, SSD и RetinaNet. Архитектура детектора YOLOv5 претерпела ряд улучшений по сравнению с предыдущими версиями, включая оптимизированный "головной" модуль с использованием Transformer Prediction Head (TPH) и добавленную дополнительную "голову" для обнаружения объектов различного масштаба.

Выбор в пользу данной архитектуры был обусловлен несколькими ключевыми факторами. Во-первых, YOLOv5 — это полноценная сверточная нейронная сеть, рассматривающая задачу обнаружения объектов как регрессионную задачу, что позволяет обрабатывать изображения в реальном времени, анализируя каждое изображение всего один раз. Во-вторых, YOLOv5 опирается на информацию всего изображения, вместо отдельных его частей, в отличие от таких сетей, как SSD или Faster-RCNN. Это делает YOLOv5 более эффективным при работе с движущимися объектами, что было особенно важно для данной работы.

В этом исследовании использовалась модель, обученная на наборе данных COCO, включающем 80 типов объектов, для детектирования движущихся БПЛА. Сеть успешно обнаруживала БПЛА как в статическом, так и в движущемся состоянии. Данные были разделены на обучающий и валидационный наборы, причем последний составил 10% от обучающего набора.

Аугментация данных была применена для увеличения устойчивости модели к изображениям, полученным в различных условиях. Применялись фотометрические и геометрические искажения, включая коррективку оттенка и насыщенности изображений, а также случайное масштабирование, обрезку, сдвиг и поворот. Помимо этих общепринятых методов, был использован метод MixUp [14], который основан на случайном выборе двух образцов из обучающих изображений для проведения случайного взвешенного суммирования. Метки образцов также подвергались взвешенному суммированию. Пример аугментации данных представлен на рисунке 2.

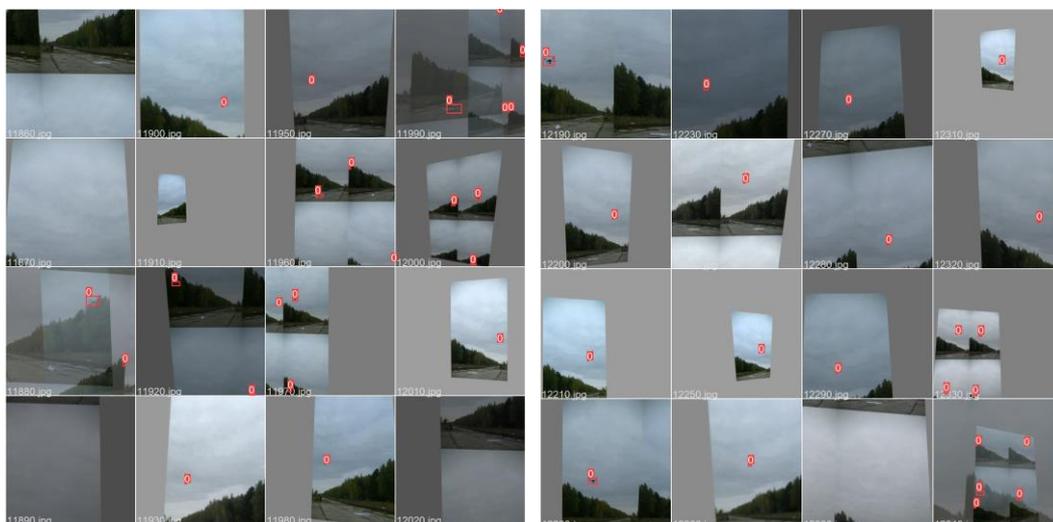


Рис. 2. Аугментированные данные

Значение скорости обучения было задано равным 0,01, обучение проводилось на протяжении 20 эпох, размер пакета данных составил 16, а размер каждого изображения - 512x512 пикселей. Для ускорения процесса обучения и обучения на небольшом объеме данных сверточная составляющая нейронной сети была заморожена.

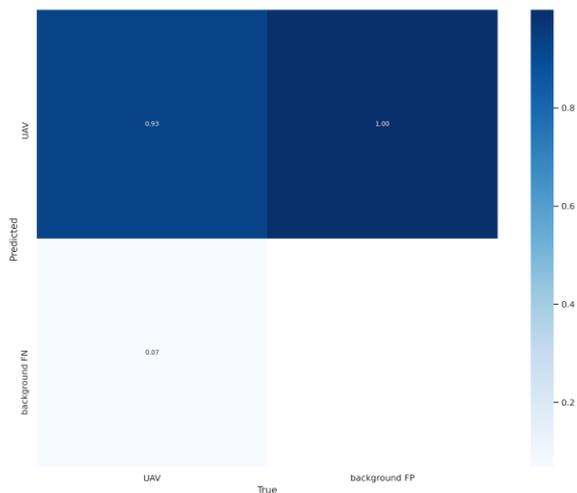


Рис. 3. Матрица ошибок

Рисунок 3 представляет собой матрицу неточностей модели. Предсказанные значения классифицируются как положительные или отрицательные, а реальные - как истинные или ложные. Матрицу неточностей обычно используют для оценки точности классификационных моделей. Тем не менее, прогнозирование и распознавание образов можно рассматривать как подмножество этой задачи, что делает матрицу неточностей полезным инструментом для измерения точности предсказаний. Особенность матрицы ошибок состоит в том, что она позволяет оценить эффективность

прогнозирования как качественно, так и количественно. Исходя из этой матрицы, можно заметить, что ошибка первого рода не превышает 7%.

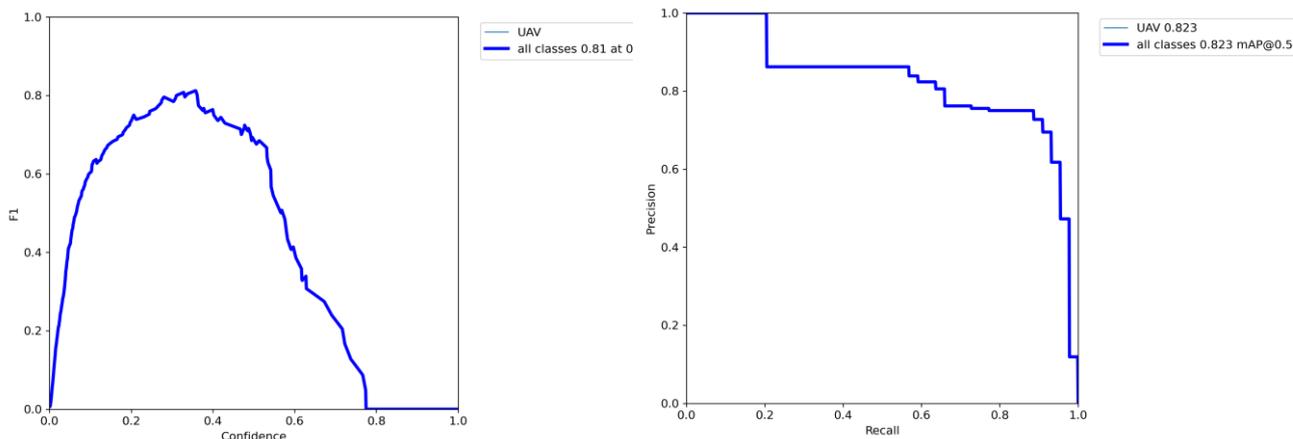


Рис. 4. F1-мера и ROC кривая

Рисунок 4 иллюстрирует графики оценки эффективности классификатора по мере F1, а также демонстрирует зависимость точности от полноты при коэффициенте Жаккара, равном 0.5. Из графика видно, что площадь под кривой составляет 0.823, что считается хорошим показателем. Рисунок 5 показывает графики обучения моделей на тренировочном и валидационном наборах данных. Эти диаграммы отражают процесс обучения и позволяют делать выводы о его адекватности. а рисунке 6 представлены результаты применения модели к валидационной выборке.

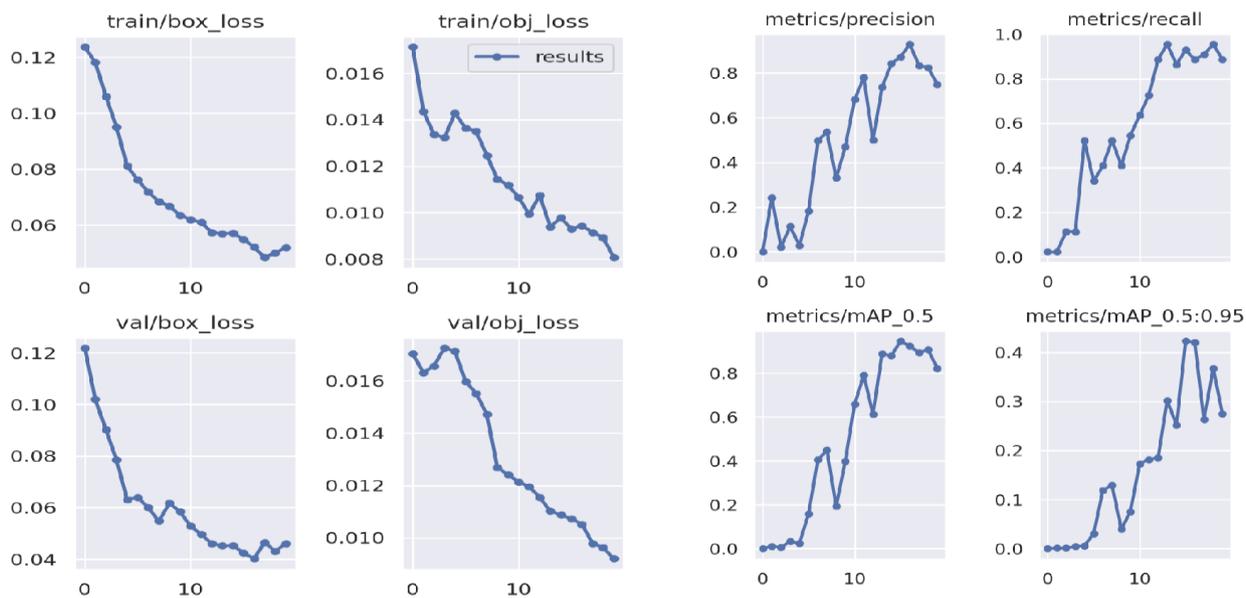


Рис. 5. Графики обучения

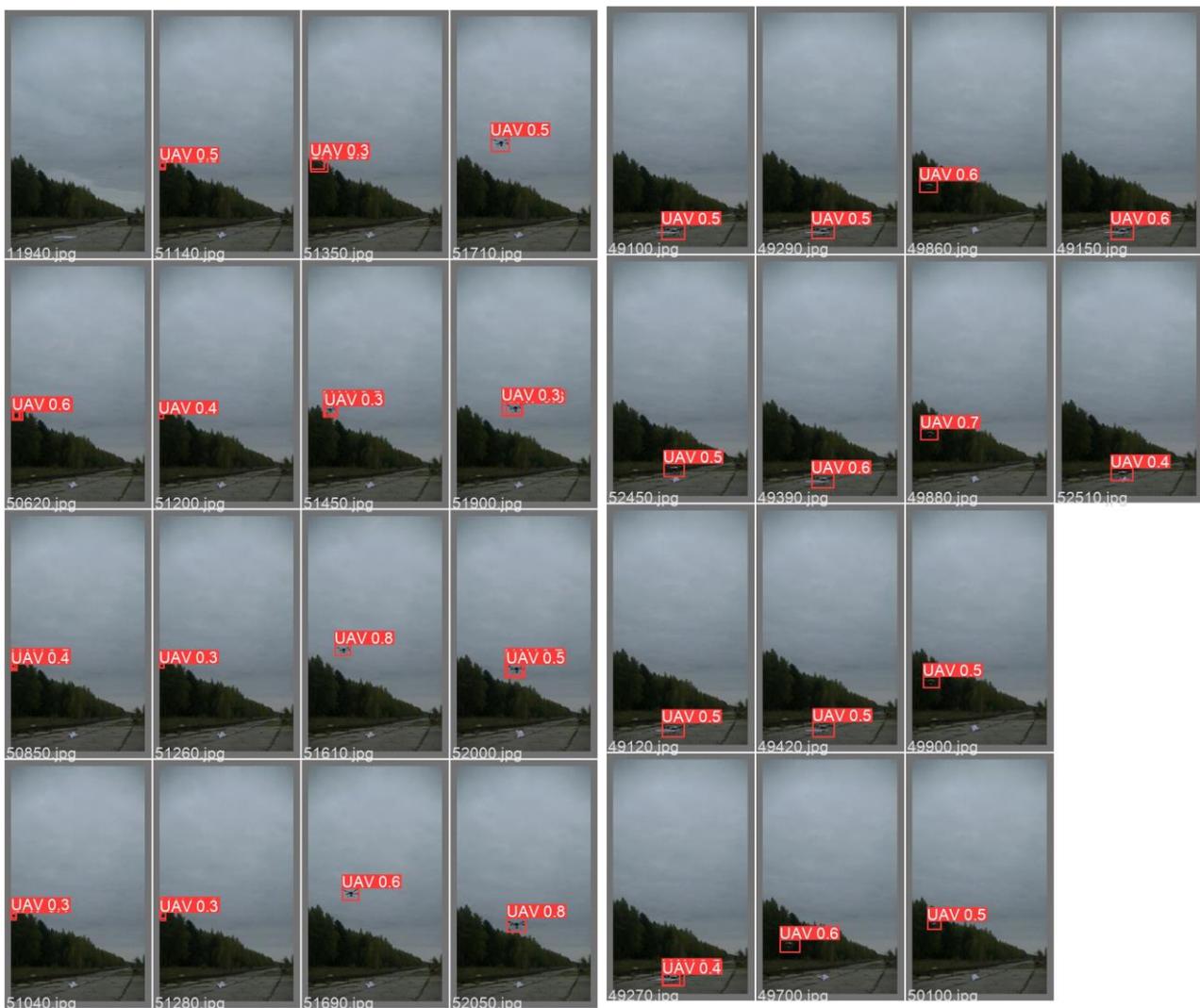


Рис. 6. Некоторые результаты применения модели

Числа рядом с меткой UAV означают уверенность модели в том, что данный объект является БПЛА.

2. Способ измерения скорости и ускорения

Из сети мы можем получить три полезные сведения: тип объекта, размер объекта и местоположение объекта, где размер и местоположение объекта содержатся в ограничивающей рамке объекта. Тип объекта используется для получения предварительных сведений о реальном размере объекта, то есть: если тип объекта – БПЛА №1 мы используем априорную информацию о размере БПЛА №1 – 30 см. Также используется соотношение пикселей объекта и пикселей всего изображения для представления.

Чтобы вычислить реальную скорость объекта по приведенным выше данным, нам нужно выяснить соответствие между параметром оптики камеры и реальным миром. Модель этого соответствия такова: предположим, что одна точка $m0$ имеет реальную скорость $v0$, тогда ее проекция на изображение $m1$ будет владеть скоростью $v1$. В единицу времени dt их путь составляет $v0dt$ и $v1dt$. Согласно теории проекции камеры, должно быть:

$$\frac{v0dt}{z} = \frac{v1dt}{f} \quad (1)$$

Где f обозначает фокусное расстояние камеры, а z обозначает расстояние между камерой и объектом. На рисунке 7 обозначен принцип фокусировки камеры.

где f – фокусное расстояние камеры, b – размер изображения объекта по вертикали, d – расстояние до объекта, B – вертикальный размер объекта. Расчет фокусного расстояния осуществляется в начале эксперимента с использованием априорных значений параметров d и b .

На втором этапе производится расчет расстояния пройденное БПЛА (рисунок 8) за единицу времени dt :

$$c = \sqrt{(d_2 - d_1)^2 + \left(\frac{c_1 B}{b_1}\right)^2} \quad (3)$$

где d – расстояния до объекта в моменты времени 1 и 2, c_1 – расстояние в пикселях от центра первого до центра второго, ограничивающего БПЛА прямоугольника, b_1 , B – размер БПЛА в момент времени 1 в пикселях и фактический соответственно.

Подставив в формулу выражение для фокусного расстояния получаем следующее:

$$c = fB \sqrt{\frac{1}{b_1^2} + \frac{2}{b_1 b_2} + \frac{1}{b_2^2} \left(\frac{c_1}{b_1 f}\right)^2} \quad (4)$$

Для вычисления скорости и ускорения необходимо учитывать характеристику количества кадров в секунду видеопотока камеры:

$$v = FPS \cdot c \quad (5)$$

$$a = \Delta v \quad (6)$$

3. Результаты эксперимента

В ходе эксперимента были получены материалы с камер установленных вдоль движения БПЛА на участках взлета и посадки (рисунок 9)

Таблица 1. Результаты измерения скорости и ускорения на участках взлета и посадки

Точка съемки	Камеры	БПЛА	Медианная скорость (м/с)	Медианное ускорение (м/с ²)
Точка 1	Камера 1, Камера 2	БПЛА № 1	11	0,01
Точка 1	Камера 1, Камера 2	БПЛА №2	15	0,01
Точка 2	Камера 3, Камера 4	БПЛА № 1	12	0,02
Точка 2	Камера 3, Камера 4	БПЛА №2	13	0,01
Точка 2	Камера 3, Камера 4	БПЛА № 3	12	0,03
Точка 2	Камера 3, Камера 4	БПЛА № 4	10	0,02
Точка 2	Камера 3, Камера 4	БПЛА № 5	20	0,01
Точка 2	Камера 3, Камера 4	БПЛА № 6	13	0,02
Точка 2	Камера 3, Камера 4	БПЛА № 7	11	0,01
Точка 3	Камера 5, Камера б	БПЛА № 1	13	0,02
Точка 3	Камера 5, Камера б	БПЛА №2	12	0,01

Как показано на рисунке 10, во время измерения скорости и ускорения БПЛА на этапе взлета, встречаются всплески значений. Эти всплески являются следствием ошибок при определении размеров ограничивающего прямоугольника, которым обозначен БПЛА. Вкратце, различия в размере ограничивающего прямоугольника могут ввести в заблуждение систему измерения скорости и ускорения, в результате чего регистрируются внезапные всплески в данных.

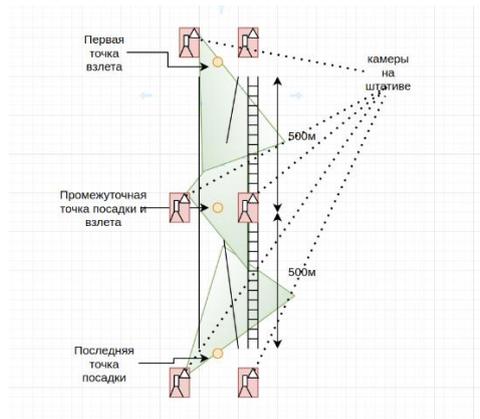


Рис. 9. Схема расположения видеокамер в ходе эксперимента

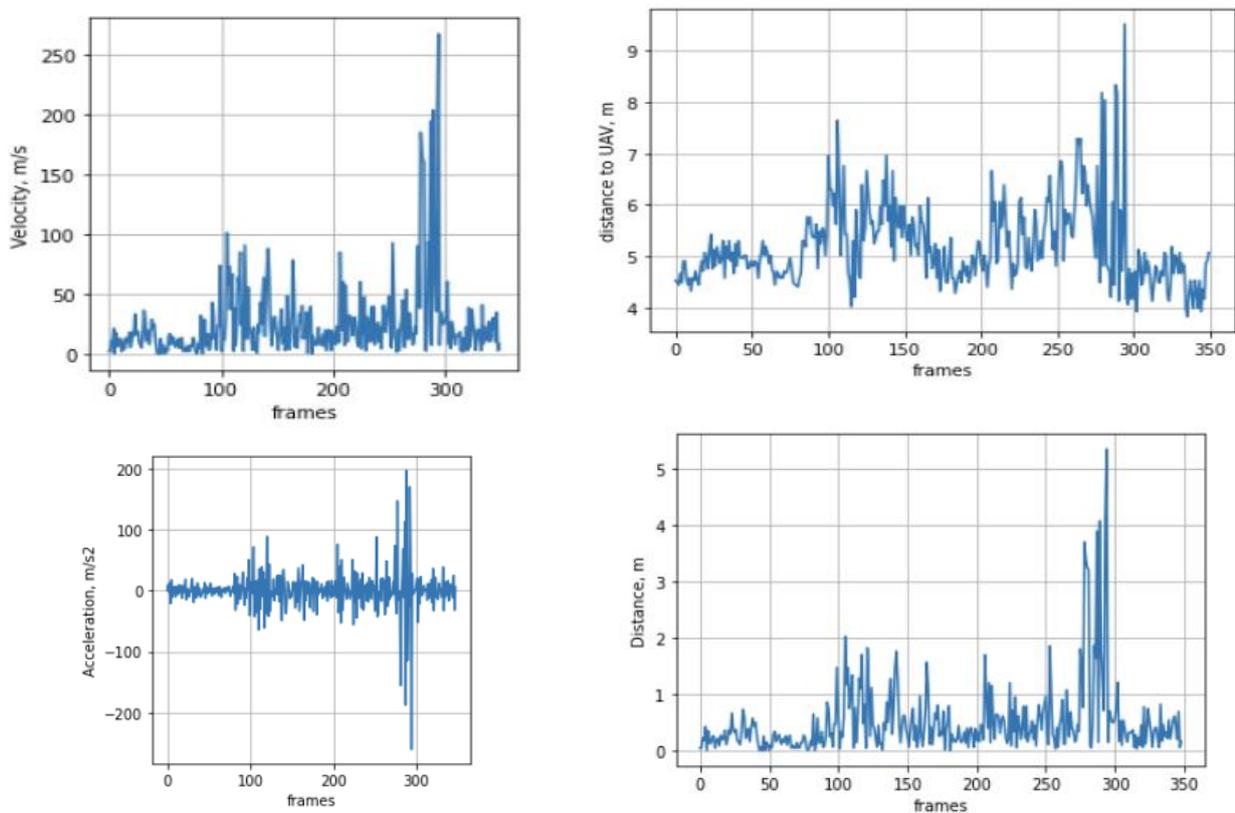


Рис. 10. Графики замеров расстояний до БПЛА, расстояний, прошедших БПЛА, скорости и ускорения на участке взлета

Для устранения этой проблемы можно применить два подхода. Во-первых, можно улучшить модель прогнозирования ограничивающих рамок. Это поможет уменьшить погрешности в определении размера прямоугольника, что приведет к более точным измерениям скорости и ускорения. Во-вторых, можно использовать дополнительные фильтры, которые будут сглаживать размер ограничивающей рамки. Это поможет уменьшить влияние непредвиденных всплесков в данных измерения скорости и ускорения.

Данные об измерении скорости и ускорения на этапах взлета и посадки в точках 1,2,3 собраны и представлены в таблице 1. Скорость и ускорение взяты как медианные значения на всем участке. Данные по каждой паре камер на участке усреднены между собой.

4. Заключение

В ходе работы были исследованы и применены современные методики детектирования и трекинга объектов, в частности БПЛА, с использованием сверточных нейронных сетей и, в частности, модели

YOLOv5. Эта модель показала эффективность в решении поставленной задачи и обеспечила хорошее качество предсказаний, что подтверждается матрицей ошибок и значением F1-меры.

Были продемонстрированы и обсуждены проблемы, возникающие при измерении скорости и ускорения БПЛА, включая всплески в данных, обусловленные погрешностями при определении размера ограничивающих рамок. Были предложены возможные решения этих проблем, такие как улучшение модели прогнозирования ограничивающих рамок и применение дополнительных фильтров для сглаживания данных.

В результате, это исследование представляет собой важный шаг в направлении повышения эффективности и надежности систем детектирования и трекинга БПЛА. Дальнейшие исследования могут включать оптимизацию модели, использование других методик обучения и оценки, а также расширение области применения модели для детектирования и трекинга других типов объектов.

Литература

1. *Girshick, Ross, et al.* Rich feature hierarchies for accurate object detection and semantic segmentation[C]// IEEE conference on computer vision and pattern recognition, Columbus, OH, USA, 2014: 580-587.
2. *Girshick R.* Fast r-cnn// IEEE international conference on computer vision, Santiago, Chile, 2015: 1440-1448.
3. *Liu W, Anguelov D, Erhan D, et al.* Ssd: Single shot multibox detector// European conference on computer vision. Springer, Cham, 2016: 21-37.
4. *Redmon J, Divvala S, Girshick R, et al.* You only look once: Unified, real-time object detection[C]// IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016: 779-788.
5. *Lin H Y and Chang C H.* Automatic speed measurements of spherical objects using an off-the-shelf digital camera[C]// IEEE International Conference on Mechatronics, Beijing, China, 2005:66-71.
6. *Liu Kaizhan, Ye Yunming, Li Xutao, Li Yan* A Real-Time Method to Estimate Speed of Object Based on Object Detection and Optical Flow Calculation// Journal of Physics: Conference Series, Volume 1004, Issue 1, article id. 012003 (2018).
7. *Yu Y Y, Li X Y and Zhao Y T.* A method to calculate the macroscopic optical flow speed of vehicle[C]// IGTA, Beijing, China, 2009:20-23.
8. *Wang J.* A method to recognize vehicle and to measure its speed based on video[D]. Northern nationalities university, 2009.
9. *Yan L P, Sun C X, Wang X L, et al.* A method to real-time measure speed based on vehiclemounted camera. CN, CN103686083A[P]. 2014.
10. *Harvey L and Cohen A.* Vehicle speed measurement using an imaging method[C]// International Conference on Industrial Electronics, Control and Instrumentation, Kobe, Japan, 1991: 1730-1733.
11. *Pumrin S and Dailey D J.* Roadside camera motion detection for automated speed measurement[C]// IEEE Conference on Intelligent Transportation Systems, Singapore, 2002: 147-151.
12. *Lin H Y and Chang C H.* Automatic speed measurements of spherical objects using an off-the-shelf digital camera[C]// IEEE International Conference on Mechatronics, IEEE, 2005:66-71.
13. *Xingkui Zhu, Shuchang Lyu, Xu Wang, Xu Wang,* TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios. ArXiv preprint ArXiv:2108.11539v1.
14. *Hongyi Zhang, Moustapha Cisse, Yann N Dauphin, and David Lopez-Paz.* mixup: Beyond empirical risk minimization. arXiv preprint arXiv:1710.09412, 2017
15. *Wei, B.; Barczyk, M.* Experimental Evaluation of Computer Vision and Machine Learning-Based UAV Detection and Ranging. Drones 2021, 5, 37