

UAV Detection Using Neural Networks

M. D. Averina¹, O. A. Levanova¹, D. V. Grushevskaya¹, K. A. Kukharev¹, D. M. Murin¹, M. A. Kalinin²

DOI: [10.18255/1818-1015-2024-2-182-193](https://doi.org/10.18255/1818-1015-2024-2-182-193)

¹P.G. Demidov Yaroslavl State University, Yaroslavl, Russia

²National Research University Higher School of Economics, Moscow, Russia

MSC2020: 68T07

Research article

Full text in Russian

Received May 6, 2024

Revised May 24, 2024

Accepted May 29, 2024

The availability of unmanned aerial vehicles (UAVs) has led to a significant increase in the number of offenses involving their use. This makes the development of UAV detection systems relevant. Solutions based on deep neural networks show the best results in detecting UAVs on video. This article presents a study of various neural network detectors and focuses on identifying objects as small as possible, up to the size of 4×4 and even 3×3 pixels. The work investigates architectures SSD (VGG16) and YOLOv3 and its modifications. Precision and recall metrics are calculated separately for different intervals of the object areas. The best result have been shown by YOLOv3 model with bbox parameters chosen as the result of object sizes clustering. Small (3×3 px) drones have been successfully identified with 76 % precision and a very small recall of 26 %. For objects between 10 and 20 pixels in area, the recall is 64 % with an accuracy of 75 %. For objects with an area more than 20px the recall is about 90 %, the precision is 89 %, and the F1 score is 90 %. These results show that it is possible to recognize even 4×4 pixel drones, which can be used in video surveillance systems.

Keywords: UAV detection

INFORMATION ABOUT THE AUTHORS

Averina, Maria D. (corresponding author)	ORCID iD: 0009-0005-3111-1488 . E-mail: maverina518@gmail.com Postgraduate student
Levanova, Olga A.	ORCID iD: 0000-0001-8078-4447 . E-mail: olaydy@gmail.com PhD, associate professor
Grushevskaya, Darya V.	ORCID iD: 0009-0006-5696-5452 . E-mail: grushevskaya.d.v@mail.ru M. Sc.
Kukharev, Kirill A.	ORCID iD: 0009-0006-8764-7860 . E-mail: kirillkukharev76@gmail.com M. Sc.
Murin, Dmitriy M.	ORCID iD: 0000-0002-8068-0784 . E-mail: nirum87@mail.ru PhD, associate professor
Kalinin, Maksim A.	ORCID iD: 0009-0007-7890-6418 . E-mail: maksim.kalinin.2110@gmail.com M. Sc.

Funding: Yaroslavl State University (project VIP-016).

For citation: M. D. Averina, O. A. Levanova, D. V. Grushevskaya, K. A. Kukharev, D. M. Murin, and M. A. Kalinin, "UAV detection using neural networks", *Modeling and Analysis of Information Systems*, vol. 31, no. 2, pp. 182–193, 2024. DOI: [10.18255/1818-1015-2024-2-182-193](https://doi.org/10.18255/1818-1015-2024-2-182-193).

Детекция БПЛА при помощи нейронных сетей

М. Д. Аверина¹, О. А. Леванова¹, Д. В. Грушевская¹, К. А. Кухарев¹, Д. М. Мурин¹,
М. А. Калинин²

DOI: [10.18255/1818-1015-2024-2-182-193](https://doi.org/10.18255/1818-1015-2024-2-182-193)

¹Ярославский государственный университет им. П.Г. Демидова, Ярославль, Россия

²Национальный исследовательский университет «Высшая школа экономики», Москва, Россия

УДК 004.93'12

Научная статья

Полный текст на русском языке

Получена 6 мая 2024 г.

После доработки 24 мая 2024 г.

Принята к публикации 29 мая 2024 г.

Доступность беспилотных летательных аппаратов (БПЛА) привела к тому, что их стали часто использовать при совершении правонарушений. Такая ситуация делает актуальной разработку систем обнаружения БПЛА. При обнаружении БПЛА применяются различные подходы на основе анализа радиочастотных и акустических сигналов, а также обработки видео данных. Лучшие результаты в обнаружении БПЛА на видео показывают решения, основанные на глубоких нейронных сетях. В этой статье мы представляем исследование различных нейросетевых детекторов и оценку возможности их практического использования в системах видеонаблюдения. Основной акцент работы направлен на выявление как можно более маленьких объектов, вплоть до размеров 4×4 пикселя. В работе представлены результаты анализа архитектур SSD(VGG16), YOLOv3 и их модификаций. В качестве метрик качества использовались полнота и точность, которые вычислялись отдельно для разных размеров объекта. Лучший результат был получен для модели YOLOv3 со значениями параметров bbox, подобранных в результате кластеризации размеров объектов. При распознавании дронов размера 3×3 удалось достичь точности 76 % при очень маленьком значении полноты 26 %. Для объектов, площадь которых составляет от 10 до 20 пикселей полнота составила 64 % при точности 75 %. Для объектов большего размера в среднем получилась полнота 90 %, точность 89 % и F1-мера 90 %. Данные результаты показывают, что распознавание дронов возможно даже при размере 4×4 , что может быть успешно использовано в системах видеонаблюдения.

Ключевые слова: детекция БПЛА

ИНФОРМАЦИЯ ОБ АВТОРАХ

Аверина, Мария Дмитриевна (автор для корреспонденции)	ORCID iD: 0009-0005-3111-1488 . E-mail: maverina518@gmail.com Аспирант
Леванова, Ольга Александровна	ORCID iD: 0000-0001-8078-4447 . E-mail: olaydy@gmail.com Кандидат физ.-мат. наук, доцент
Грушевская, Дарья Владимировна	ORCID iD: 0009-0006-5696-5452 . E-mail: grushevskaya.d.v@mail.ru Магистр
Кухарев, Кирилл Александрович	ORCID iD: 0009-0006-8764-7860 . E-mail: kirillkukharev76@gmail.com Магистр
Мурин, Дмитрий Михайлович	ORCID iD: 0000-0002-8068-0784 . E-mail: nirum87@mail.ru Кандидат физ.-мат. наук, доцент
Калинин, Максим Александрович	ORCID iD: 0009-0007-7890-6418 . E-mail: maksim.kalinin.2110@gmail.com Магистр

Финансирование: ЯрГУ (проект VIP-016).

Для цитирования: M. D. Averina, O. A. Levanova, D. V. Grushevskaya, K. A. Kukharev, D. M. Murin, and M. A. Kalinin, "UAV detection using neural networks", *Modeling and Analysis of Information Systems*, vol. 31, no. 2, pp. 182–193, 2024. DOI: [10.18255/1818-1015-2024-2-182-193](https://doi.org/10.18255/1818-1015-2024-2-182-193).

Введение

В последние годы неуклонно растет число частных лиц, имеющих беспилотные летательные аппараты (БПЛА, дроны, квадрокоптеры). По данным Федерального управления гражданской авиации США (Federal Aviation Administration, FAA) на 29 февраля 2024 года в США был зарегистрирован 781781 дрон (375226 Commercial Drones Registered, 400858 Recreational Drones Registered, 5697 Paper Registrations)¹. По данным «European ATM Master Plan» SESAR Joint Undertaking, опубликованным в декабре 2019 года, количество коммерческих дронов только в Европе вырастет с 10000 (2015 год) до почти 400000 к 2050 году [1]. В декабре 2023 года был утвержден национальный проект «Беспилотные авиационные системы»², согласно которому в России планируется к 2030 году увеличить производство беспилотников до 32000 в год.

В то же время широкое распространение квадрокоптеров привело к значительному увеличению числа нарушений безопасности в общественных местах и на объектах критически важной инфраструктуры. БПЛА могут переносить различную «полезную» нагрузку, а также могут быть оснащены различными датчиками и средствами негласного получения информации. С применением данных устройств нарушаются права на конфиденциальность персональных данных и преодолеваются установленные режимы охраны и защиты зданий и сооружений, а также охраняемых территорий и технологических объектов. В России чаще всего совершаются преступления, связанные с доставкой осужденным в учреждения уголовно-исполнительной системы запрещенных предметов с помощью БПЛА.

Обнаружение и контроль действий БПЛА становятся существенными задачами для обеспечения безопасности, тем самым стимулируя рост рынка средств противодействия им. По данным опроса компании Ernst & Young³ рост спроса на борьбу с дронами обгоняет рост спроса на них самих. Противодействие им состоит из обнаружения, идентификации и нейтрализации. На текущий момент интенсивно развиваются области обнаружения и нейтрализации. При этом системы обнаружения представляют, как самостоятельный интерес, так и включаются в системы противодействия.

Наиболее известными методами обнаружения БПЛА являются: радиолокационное, радиочастотное обнаружение, обнаружение инфракрасного излучения, электрооптическое, акустическое, и оптическое обнаружение. Оптические методы обнаружения способны эффективно справляться с их обнаружением в условиях дневного света, шума, наличия птиц. К недостаткам оптического метода можно отнести относительно небольшую дальность обнаружения.

Для раннего обнаружения данных необходима возможность распознавания объекта, имеющего малый размер. Поэтому акцент нашей работы сделан на исследовании подходов обнаружения на изображениях дронов маленького размера.

1. Системы видеонаблюдения

Для обеспечения возможности раннего обнаружения особый интерес представляет распознавание квадрокоптеров, занимающих малую площадь кадра. Пусть разрешение камеры составляет 640 пикселей, мы хотим обеспечить распознавание дрона на расстоянии 200 метров. Предположим, что для обнаружения квадрокоптера необходимо, чтобы он занимал площадь кадра 7×7 пикселей. Такая плотность пикселей может быть достигнута при угле обзора 7 градусов, тогда для кругового обзора объекта видеонаблюдения необходимо 52 камеры. В этих условиях по горизонтали и вертикали одна камера обозревает расстояние приблизительно 23 метра. Таким образом, система видеонаблюдения не сможет «увидеть» квадрокоптер, поднятый на высоту 25 метров.

¹<https://www.faa.gov/uas>

²<http://government.ru/rugovclassifier/906/events/>

³https://assets.ey.com/content/dam/ey-sites/ey-com/en_ru/topics/advisory/ey-uav-survey-eng.pdf

Теперь предположим, что для обнаружения квадрокоптера занимаемая им площадь должна составлять только 3×3 пикселя (предельный случай). Тогда угол обзора будет равен 15 градусам (для кругового обзора потребуются 24 камеры), а обзриваемое расстояние составит 53 метра. Чтобы «обойти» такую систему нужно поднять дрон уже на 55 метров. Заметим, что при размещении камер видеонаблюдения вдоль периметра объекта через 70 м высота границы обнаружения будет колебаться как раз в пределах от 35 до 53 метров. Таким образом, возможно построение системы видеонаблюдения, которая обнаруживает БПЛА на расстоянии 200 метров, при этом очень важно распознавать в видео объекты очень маленьких размеров.

Также стоит отметить, что для практического применения в системах видеонаблюдения критически важна точность распознавания алгоритма (*precision*). Если величина $1 - precision > 0.2$, то каждый пятый сигнал о наличии дрона в видео будет ошибочным, однако потребует некоторой реакции (оператора, включения сирены или средств РЭБ). Таким образом, точность ниже 80 % неприемлема для интеграции алгоритмов распознавания в системы видеонаблюдения. Если мы говорим про объекты критически важной инфраструктуры, то более важным становится показатель полноты (*recall*). Показатели точности и полноты сильно связаны, и при увеличении одного показателя обычно другой уменьшается. В данной работе не отдавалось предпочтение ни одному из этих показателей, они анализировались в совокупности.

1.1. Существующие решения

Задача создания системы видеонаблюдения с обнаружением дронов на видео не является новой. Имеющиеся подобные системы, как правило, принадлежат коммерческим организациям и являются платными, что является серьезным препятствием при оценке существующих решений.

Существующие готовые системы в открытом доступе найти достаточно трудно, отметим лишь дрон-детектор от компании «Спецлаб»⁴, однако качество обнаружения оставляет желать лучшего. Существенно проще в научных статьях найти информацию о различных подходах к задаче распознавания дронов. Благодаря соревнованию Drone vs Bird⁵ и IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS) тематика обнаружения дронов активно развивается начиная с 2017 года.

В 2024 году вышла большая обзорная статья по их обнаружению [2], как по данным с радара, так и по звуку и видео (обнаружению в видео в статье уделено мало внимания). Для распознавания БПЛА на видео было предпринято множество различных решений: использовались классические алгоритмы, например, SVM [3], нейронные сети, и комбинированный подход [4]. Есть работы, в которых используется только визуальная информация [5], в других предлагается совместное использование видео и звука [6]. Авторы еще одной статьи [4] разделили задачу на обнаружение движущихся объектов (решение основано на вычитании фона) и классификацию обнаруженного объекта на дрон, птицу и фон (за счет нейронной сети). Однако построенное решение работает только для статичной камеры. В некоторых исследованиях была важнее скорость распознавания, а в других — точность и качество обнаружения. В целом, наилучшие показатели достигнуты с помощью сверточных нейронных сетей. Также есть исследования по разработке системы обнаружения БПЛА на основе информации от двух камер (широкоугольной статической и узкоугольной поворотной с увеличением). Использование только оптической информации на основе архитектуры YOLOv3 позволило достичь хороших показателей [7].

Еще стоит отметить статью, посвященную усовершенствованию модели YOLOv5 [8] с целью уменьшения числа параметров и ускорения. В работе используется набор данных, содержащий порядка 8000 дронов среднего размера (до 32×32) и порядка 1000 штук размера 8×8 и меньше.

⁴<https://www.goal.ru/neuronet/drone-detector>

⁵<https://wosdetc2023.wordpress.com/>

Улучшенный вариант YOLOv5 показал качество mAP 90.6%, Recall 89.3% (FPS 12.6), после ускорения — mAP 87.2 % (FPS 27.6).

В статье [9] модифицировали сеть YOLOv8s и сравнили качество работы моделей с другими известными архитектурами (TIB-Net, YOLOv5-s, YOLOX-s, YOLOv7, YOLOv7-tiny, YOLOv8s). Авторам удалось достичь точности (P) и полноты (R) по 93.3 %, у оригинальной сети YOLOv8s результаты ощутимо хуже (P = 81.4 %, R = 78.1 %). Среди других моделей лучший же результат показала YOLOv5 P = 88.1 % и R = 90.9 % при тестировании уже на их наборе данных.

2. Постановка задачи

Перед авторами была поставлена задача разработки системы по распознаванию беспилотных летательных аппаратов на видео, полученному в результате статичной съемки без записи звука. Целью данной работы является исследование применимости нейросетевого детектора для систем видеонаблюдения. Поэтому для данной задачи очень важно обеспечить обнаружение дронов как можно меньшего размера, для более раннего реагирования. Также существенным является требование малого числа ложных срабатываний (*precision* > 0.8). Прежде всего необходимо решить, как оценивать качество работы построенного решения. В задачах обнаружения объектов чаще всего для определения качества модели используются классические метрики AP и mAP, которые при решении данной задачи менее информативны. Для системы видеонаблюдения привычнее оперировать такими широко известными метриками как Precision и Recall. Они дают разностороннюю и довольно полную оценку качества. Также было интересно исследовать качество обнаружения для разных размеров дронов. Поэтому отдельно для каждого интервала площади, занимаемой дроном, рассчитывалось общее число размеченных квадрокоптеров (total), число правильно распознанных дронов (TP), число ложно распознанных дронов (FP), число нераспознанных дронов (FN). И по полученным значениям уже рассчитывались основные характеристики Precision и Recall также для каждого размера. Отметим, что в различных статьях по обнаружению дронов используются разные метрики.

2.1. Набор данных

Для создания системы по распознаванию дронов на видеоизображениях необходим большой объем размеченных данных. Для качественного обучения необходим был набор, состоящий из нескольких сотен тысяч размеченных изображений / видеок кадров. В ходе работы был самостоятельно создан набор размеченных изображений, а также предприняты попытки поиска готовых наборов данных.

2.2. Готовые наборы данных

В свободном доступе было найдено несколько подходящих наборов изображений с размеченными БПЛА.

- Набор изображений дронов «Drone Dataset (UAV)» с сайта Kaggle⁶, созданный в 2019 году. Набор состоит из 1359 фотографий дронов и такого же количества файлов txt и xml с разметкой в формате, необходимом для обучения нейронных сетей. Размер набора: 725.28 Мб.
- Набор данных с беспилотными летательными аппаратами «Drone Dataset: Amateur Unmanned Air Vehicle Detection» с сайта Mendeley Data⁷. Данный набор был создан в 2019 году группой авторов для проекта «Amateur Drone Detection and Tracking». В наборе данных более 4000 снимков любительских дронов (например, квадрокоптер DJI Phantom). Кроме того, набор данных содержит «негативные» объекты, не относящиеся к дронам. Размер набора: 157 Мб.

Перечисленные наборы данных в основном содержат объекты с площадью больше 1000 пикселей и не такие многочисленные, как необходимо.

⁶<https://www.kaggle.com/datasets/dasmehdixtr/drone-dataset-uav>

⁷<https://data.mendeley.com/datasets/zcsj2g2m4c/4>

Также был проанализирован набор 2020 года с размеченными видеороликами, предоставляемый организаторами соревнования Drone vs Bird. Во втором столбце таблицы 1 представлена статистика по размерам дронов в данном наборе (после приведения картинок к размеру 608×608). Из таблицы видно, что дроны чаще всего имеют крупный размер, что не совсем подходит для решения поставленной задачи.

Поиск остальных открытых наборов данных был затруднен тем, что информация о них содержится только в научных статьях по обнаружению дронов. Так в открытом доступе был найден набор данных FL-dataset⁸, который содержит 14 черно-белых видео с дронами (38948 кадров) разрешения 752×480 , съемка с дрона, в кадре одновременно до 2-х дронов. Минимальный, средний и максимальный размеры дронов: 9×9 , 25.5×16.4 и 259×197 соответственно. Данный набор исключили, поскольку видео черно-белые и съемка с нестационарной камеры. Набор данных⁹ содержит 50 видео (70250 кадров) с несколькими дронами в кадре, типичный размер дронов от 10×8 до 65×21 . Однако съемка производилась с дрона, т. е. камера нестационарная, поэтому этот набор данных не был взят, хотя для построения детектора с покадровой обработкой он вполне подходит. В статье [10] был использован данный набор, рассмотрена модель построения как детектора, так и трекера для слежения за БПЛА. Детектор содержит много шагов, в частности используются: нахождение движения с помощью оптического потока, нахождение ключевых точек, небольшие полносвязные сети, которые обучают. Трекер строится на основе фильтра Калмана. Авторам удалось достичь следующих показателей качества: точность — 88.8 %, полнота — 89.0 %, F-мера — 88.9 %. Отметим, что в работе идет сравнение только с одной моделью из статьи [11], которую они обучили на предложенном наборе данных.

Набор данных TIB-Net¹⁰ содержит 2850 картинок, съемки на стационарную камеру с расстояния около 500 метров. Набор данных в основном содержит изображения площадью более 100 пикселей при разрешении 1920×1080 . В статье [12] предложена архитектура нейронной сети, которая показала качество $mAP = 89.2$. Набор данных состоит из отдельных изображений, но визуально можно однозначно определить, что кадры взяты из нескольких видео. Набор изначально не разделен на обучающую и тестовую выборки. Вероятно, авторы статьи владели информацией из какого видео взято изображение, и в обучающем наборе не было кадров похожих на изображения из тестовой выборки. Однако пользоваться этим набором данных можно только целиком в качестве либо тестовой выборки, либо обучающего набора. Иначе, если разделить кадры случайным образом, то очень вероятна ситуация, при которой в тестовой выборке окажутся кадры очень похожие на те, которые были в обучающей выборке. Данный набор не использовался в нашей работе, поскольку был найден уже после проведения исследований.

2.3. Авторский набор данных

Было принято решение о создании собственного набора данных. Членами команды было снято 13 видео с полетом дрона в различных локациях суммарной длительностью 3.5 часа. Съемки производились с помощью двух уличных видеокамер HIKVISION HiWatch (модель DS-I200(C) 2.8 мм; модель DS I200(B) 4 мм) с максимальным разрешением 1920×1080 пикселей и частотой смены кадров 25 к/с. В съемках участвовало два дрона: DJI Phantom 4 Pro (размер около 247.5×247.5 мм) и DJI Mini 2 Fly More Combo (размер: $203 \times 159 \times 56$ мм). В кадре всегда присутствовал один квадрокоптер.

Полученные видеофайлы были сначала предварительно размечены специально написанным для этой задачи инструментом трекинга за одним объектом. После разметка уточнялась с помощью готового инструмента Computer Vision Annotation Tool (CVAT)¹¹.

⁸<https://drive.switch.ch/index.php/s/3b3bdbd6f8fb61e05d8b0560667ea992>

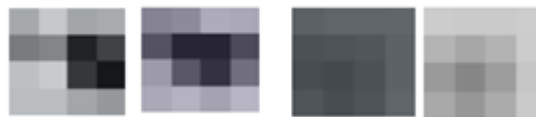
⁹https://engineering.purdue.edu/~bouman/UAV_Dataset/

¹⁰<https://drive.google.com/drive/folders/1ro-S2lwBmn83HLSppr5i-hBHLIYLAobg>

¹¹<https://github.com/openvinotoolkit/cvat>

Table 1. The distribution of the drones number by area occupied

Площадь дрона	Количество дронов Drone-vs-bird	Количество дронов авторский набор	Размер дрона, пиксели
от 0 до 10	3776	31008	$\sim 3 \times 3$
от 11 до 50	19205	67370	$\sim 5 \times 5$
от 51 до 100	18656	23790	$\sim 10 \times 10$
от 101 до 500	54414	52487	$\sim 22 \times 22$
от 501 до 1000	10775	2755	$\sim 100 \times 100$
более 1000	10994	3221	$> 100 \times 100$
Всего	117820	180619	

Таблица 1. Распределение числа дронов по занимаемой площади**Fig. 1.** Drone (left) and bird (right) are the size of 4×4 pixels**Рис. 1.** Дрон (слева) и птица (справа) размером 4×4 пикселя

Всего было размечено порядка 180 тысяч дронов. В таблице 1 в третьей колонке приведена статистика количества дронов созданного набора данных для различных размеров по площади. Как можно заметить из таблицы при создании базы упор был сделан на то, чтобы квадрокоптер в кадре был маленького размера.

Самый минимальный размер объекта в созданной базе данных 3×3 пикселя, человеку визуально невозможно определить тип такого маленького объекта. При разметке была использована дополнительная информация о траектории объекта, например, когда дрон удаляется. Начиная с размера 4×4 пикселя можно попробовать отличать объект, опираясь только на изображение. На рисунке 1 представлено несколько примеров объектов такого размера. Дроны (слева) чаще всего имеют силуэт равнобедренного треугольника с вершиной, направленной вниз, а птицы (справа) более овальные. С совсем маленькими изображениями очень легко ошибиться, наиболее уверенно можно различать изображения размером 10×10 пикселей. При таком размере уже зачастую легко распознать объект, примеры изображений можно видеть на рисунке 2. Полученная база видеоданных с разметкой беспилотных летательных аппаратов (квадрокоптеров) была зарегистрирована в 2021 году в Федеральной службе по интеллектуальной собственности (Роспатент) [13].

Финально в качестве набора данных были взяты как авторский набор, так и Drone-vs-bird, которые в совокупности содержат около 300 тысяч дронов (встречаются изображения с несколькими объектами). В тренировочную выборку вошло 80 % изображений (225277), а остальные использовались для валидации (17026) и тестирования (38201). Подчеркнем, что тестовый и валидационный наборы данных состояли из видео, которых не было в обучении.

Отметим, что при использовании авторского набора данных сравнительный анализ с другими результатами становится затруднительным. Если обучать на авторском наборе данных, а тестировать на открытом, то логично будет получить качество лучше, поскольку сеть обучалась на большем наборе данных. При обучении на общедоступных наборах результаты можно будет сравнить с другими моделями, но тогда авторский набор данных можно использовать только для тестирования. И максимум, что можно показать в этом случае, — это то, что предложенная модель работает лучше для наших данных. Авторы ставили задачу построения коммерческого решения, которое будет работать в условиях, приближенных к съемке камерой видеонаблюдения. Поэтому цель работы — это добиться наилучшего качества для нашего набора данных, в котором много небольших и да-



Fig. 2. Drone (left) and bird (right) are the size of 10×10 pixels

Рис. 2. Дрон (слева) и птица (справа) размером 10×10 пикселей

же крошечных объектов. В виду этих причин сравнение качества производилось только с работой классической базовой архитектуры, которую в дальнейшем улучшали.

3. Обнаружение дронов

Современные модели нейронных сетей, используемые для решения задачи обнаружения и классификации объектов на изображении, подразделяют на 3 основные группы.

- Двухэтапные детекторы (Two-Stage object detectors) используют алгоритм генерации регионов интересов (region-proposal). Основным представителем этого семейства является модель R-CNN (region-based convolutional networks) и ее модификации — Fast R-CNN, Faster R-CNN и Mask R-CNN.
- Одноэтапные детекторы (One-stage object detectors) — это полностью сверточные нейронные сети, что позволяет обрабатывать изображения в реальном времени. Выделяют такие модели, как SSD, YOLO, RFBNet.
- Детекторы объектов на основе точек (Points-based object detectors) основаны на обнаружении ключевых точек, по которым определяют остальные свойства объекта: центр и размер ограничивающего прямоугольника, трехмерное местоположение, ориентацию и даже позу.

Для систем видеонаблюдения немаловажным фактором является скорость работы, поскольку предполагается возможность функционирования алгоритмов в режиме реального времени. Поэтому были выбраны наиболее быстрые одноэтапные детекторы, которые основаны на концепции заранее заготовленных anchor bboxes и предсказаниях по сетке (predictions on a grid). В рамках данной работы был проведен ряд экспериментов при обучении моделей SSD и YOLOv3.

3.1. SSD

В качестве базовой модели для дообучения была взята модель SSD (backbone VGG16) с размером входного изображения 512×512 . Поэтому набор данных был преобразован к этим размерам. В таблице 2 представлена статистика, полученная для модели SSD при пороге доверия классификатору $\text{confidence} = 0.3$. Обучение проходило 20 эпох (далее наблюдалось переобучение), $\text{batch_size} = 12$, $\text{lr} = 0.0025$. Для объектов размера порядка 7×7 точность и полнота уже близки к 80 %. Для объектов меньшего размера начинает существенно падать полнота. Для размеров порядка 4×4 и 5×5 около половины дронов сеть находит, а для размера 3×3 обнаружение маловероятно.

Также были проанализированы результаты работы при сниженных порогах $\text{confidence} = 0.2$ и 0.1 (см. таблица 3). Снижением порога можно добиться увеличения полноты, но при этом ощутимо ухудшается точность.

3.2. YOLO v3

Также была протестирована модель YOLO v3, размер входного изображения 608×608 . Были рассмотрены несколько подходов. Сначала была обучена оригинальная модель, в таблице 4 представлены результаты на разных эпохах. При обучении использовались следующие значения параметров: $\text{lr} = 0.01$, $\text{batch_size} = 12$, $\text{confidence} = 0.3$. При сравнении 1-й и 5-й эпохи можно видеть, что средние показатели почти не изменились, однако произошло существенное улучшение точности для объектов 3×3 при падении точности для объектов большего размера. На 25-й эпохе видно продолжение этой тенденции. Как обсуждалось выше, дроны размера 3×3 сложно отличить

Table 2. SSD model results

Площадь	Всего	TP	FP	FN	Точность	Полнота	Размер
от 0 до 10	6490	959	1359	5531	0.41	0.15	$\sim 3 \times 3$
от 11 до 20	10005	5056	2585	4949	0.66	0.51	$\sim 4 \times 4$
от 21 до 30	6050	3724	888	2326	0.81	0.62	$\sim 5 \times 5$
от 31 до 50	4227	3360	927	867	0.87	0.80	$\sim 7 \times 7$
от 51 до 100	3253	2965	941	288	0.76	0.91	$\sim 10 \times 10$
от 101 до 150	3011	2678	326	333	0.89	0.89	$\sim 12 \times 12$
от 151 до 200	1092	1039	97	53	0.92	0.95	$\sim 14 \times 14$
от 201 до 500	3351	3247	101	104	0.97	0.97	$\sim 22 \times 22$
более 500	722	678	28	44	0.96	0.94	
Всего	38201	23706	7252	14495	0.77	0.62	

Таблица 2. Результаты модели SSD**Table 3.** SSD model test results with less confidence threshold

Confidence	0.2		0.1	
Площадь, пиксель	Полнота	Точность	Полнота	Точность
от 0 до 10	0.20	0.35	0.27	0.19
от 11 до 20	0.60	0.58	0.70	0.38
от 21 до 30	0.77	0.77	0.88	0.63
от 31 до 50	0.90	0.75	0.93	0.54
от 51 до 100	0.95	0.70	0.96	0.49
от 101 до 150	0.91	0.80	0.92	0.62
от 151 до 200	0.96	0.83	0.97	0.56
от 201 до 500	0.98	0.94	0.98	0.71
более 500	0.94	0.93	0.96	0.85
Всего	0.69	0.69	0.76	0.47

Таблица 3. Результаты модели SSD при уменьшенном confidence

от шума или птицы. А модель при обучении стала смещать внимание именно в сторону маленьких объектов (их больше в выборке). Чтобы убрать перекося, был проведен эксперимент, при котором из обучающей выборки были удалены объекты размера 3×3 пикселя. Однако в результате было получено еще большее ухудшение при обучении для больших объектов.

Именно в этот момент было принято решение о сборе подробной статистики по размерам объектов в выборке. В результате модель YOLOv3 обучили с уменьшенными размерами якорных боксов (anchor boxes): [(4, 6), (10, 10), (16, 16), (45, 20), (20, 51), (50, 79), (80, 90), (156, 198), (200, 200)]. В этот раз перекося при обучении не стало, результаты после 7-й эпохи показаны в таблице 5, левый столбец. Для дронов с площадью больше 20 удалось достичь показателей точности и полноты больше 80 %. Отметим также, что уменьшение порога IoU при обучении до 0.3 дало еще небольшое улучшение в качестве, правый столбец таблицы 5.

Поскольку уменьшение размеров базовых bbox улучшило результаты, то мы провели кластеризацию bbox на 9 классов. В качестве якорных боксов были взяты центры кластеров: [(4, 6), (11, 13), (21, 23), (39, 42), (86, 71), (225, 194), (321, 369), (558, 290), (551, 505)]. В таблице 6 указан лучший результат после 9-й эпохи обучения. Отметим, что при обучении был выставлен порог IoU = 0.5, таким образом, также удалось повысить качество работы модели благодаря подбору размеров bbox. Для объектов с размером площади более 20 пикселей точность и полнота уже превышают 87 %.

Таким образом, более точная настройка внутренних параметров модели YOLOv3 позволила улучшить качество работы для дронов маленьких размеров. Так для размеров порядка 4×4 сеть распознает 64 % дронов (улучшили на 8 % по сравнению с базовой моделью), при точности в 75 %

Table 4. Results of the YOLOv3 model for different epoch number**Таблица 4.** Результаты модели YOLOv3 при разном числе эпох

Число эпох	1		5		25	
Площадь, пиксель	Полнота	Точность	Полнота	Точность	Полнота	Точность
от 0 до 10	0.29	0.12	0.27	0.56	0.18	0.57
от 11 до 20	0.66	0.66	0.65	0.65	0.56	0.71
от 21 до 30	0.84	0.87	0.86	0.74	0.82	0.65
от 31 до 50	0.86	0.82	0.89	0.68	0.91	0.71
от 51 до 100	0.90	0.76	0.94	0.24	0.95	0.51
от 101 до 150	0.88	0.89	0.90	0.74	0.91	0.82
от 151 до 200	0.90	0.87	0.94	0.74	0.94	0.80
от 201 до 500	0.85	0.98	0.90	0.95	0.94	0.88
более 500	0.77	0.82	0.85	0.81	0.90	0.94
Всего	0.71	0.58	0.72	0.58	0.69	0.70

Table 5. Model YOLOv3 results with reduced bbox**Таблица 5.** Результаты YOLOv3 с уменьшенными bbox

IoU	0.5		0.3	
Площадь, пиксель	Полнота	Точность	Полнота	Точность
от 0 до 10	0.24	0.79	0.23	0.77
от 11 до 20	0.58	0.72	0.57	0.70
от 21 до 30	0.82	0.80	0.88	0.81
от 31 до 50	0.82	0.76	0.91	0.83
от 51 до 100	0.94	0.80	0.95	0.85
от 101 до 150	0.89	0.96	0.91	0.96
от 151 до 200	0.94	0.97	0.97	0.92
от 201 до 500	0.93	0.88	0.96	0.88
более 500	0.85	0.96	0.91	0.92
Всего	0.68	0.81	0.71	0.81

(улучшили на 4 %). Для объектов 3×3 удалось существенно поднять точность (76 % против 57 %), однако полнота остается крайне низкой (26 %, улучшили на 8 %). В целом можно отметить, что вне зависимости от используемой модели размер 3×3 распознается плохо, поэтому его возможно использовать лишь как предварительный сигнал, а принятие решения о нахождении дрона уже делать на основе трекинга такого объекта.

После обнаружения БПЛА нейронной сетью можно наложить ограничение на размер объекта для получения подходящих характеристик системы. Зададим максимальный размер БПЛА, обнаружение которых будет игнорироваться, и система не будет подавать сигнал о распознавании. В таблице 7 можно видеть сравнение качества обученных моделей в зависимости от порогового значения размера дрона. Так, например, во второй строке показаны метрики моделей, в случае когда дроны размеров 4×4 и меньше будут игнорироваться (не учитываться при оценке качества). Видно, что для таких размеров БПЛА более точная настройка внутренних параметров модели YOLOv3 позволила улучшить точность обнаружения с 70 % у базовой модели до 89 %. Для дронов размера больше 3×3 полнота была улучшена на 3 %, точность — на 15 %. А для БПЛА размера больше 5×5 лучшая модель показала качество выше 90 % (и полнота и точность). Еще раз отметим, что для всех моделей есть возможность менять соотношение метрик полноты и точности за счет изменения параметра confidence у уже обученной сети, (здесь у всех моделей задано значение 0.3).

Table 6. Model YOLOv3 results with bbox from clustering

Площадь, пиксель	Полнота	Точность	F1
от 0 до 10	0.26	0.76	0.39
от 11 до 20	0.64	0.75	0.69
от 21 до 30	0.88	0.88	0.86
от 31 до 50	0.90	0.84	0.87
от 51 до 100	0.95	0.87	0.91
от 101 до 150	0.90	0.93	0.92
от 151 до 200	0.94	0.94	0.94
от 201 до 500	0.93	0.95	0.94
более 500	0.88	0.96	0.92
Всего	0.72	0.80	0.78

Таблица 6. Результаты YOLOv3 с bbox по кластеризации**Table 7.** Comparative analysis of models

Модель	SSD 0.3		YOLOv3 base		YOLOv3 less bbox		YOLOv3 clust bbox	
Размер	П	Т	П	Т	П	Т	П	Т
~ 3 × 3	0.72	0.80	0.79	0.70	0.78	0.81	0.82	0.85
~ 4 × 4	0.82	0.84	0.90	0.70	0.87	0.84	0.90	0.89
~ 5 × 5	0.89	0.85	0.92	0.72	0.89	0.85	0.92	0.90
~ 7 × 7	0.93	0.88	0.93	0.72	0.92	0.89	0.92	0.92
~ 10 × 10	0.93	0.93	0.93	0.85	0.91	0.93	0.92	0.94

Таблица 7. Сравнительный анализ моделей

Заключение

В статье приведены результаты исследования задачи обнаружения дронов, занимающих малую площадь кадра, в видеопотоке. Были рассмотрены нейросетевые архитектуры SSD (VGG16) и YOLOv3, а также их модификации с целью улучшения качества обнаружения малых объектов.

Имеющиеся в открытом доступе наборы данных содержали недостаточное для решения поставленной задачи число размеченных дронов малых размеров, поэтому авторами был создан новый датасет, содержащий порядка 180 тысяч размеченных дронов малых размеров (до 100 × 100). Для обучения моделей он был дополнен материалами из открытых источников.

Акцент при составлении набора данных был сделан на изображениях с дронами и квадрокоптерами. При этом в наборе данных отсутствовали изображения беспилотников самолетного и вертолетного типов. Поскольку различные виды беспилотников имеют свои отличительные признаки, ожидается снижение точности распознавания для дронов самолетного и вертолетного типа. Авторами исследования не проводились тесты, связанные с другими видами беспилотных летательных аппаратов.

В ходе экспериментов было установлено, что улучшение точности обнаружения для объектов 3 × 3 может существенно влиять на точность распознавания объектов большего размера. Значения метрик качества вычислялись для разных размеров дронов в отдельности. При распознавании дронов размера 3 × 3 удалось достичь точности 76 % при очень маленьком значении полноты 26 %. Для объектов площадью из интервала (10, 20] полнота составила 64 % при точности 75 %. Для объектов большего размера в среднем получилась полнота 90 %, точность 89 % и F1-мера 90 %. Таким образом, эффективное распознавание дронов возможно даже при размере 4 × 4. Однако вне зависимости от используемой модели результаты обнаружения дрона размера 3 × 3 на текущий момент оказываются неудовлетворительными и могут использоваться лишь как предварительный сигнал.

В рамках данного проекта продолжается работа по улучшению качества обнаружения. Рассматривается возможность добавить в набор данных изображения с другими видами беспилотников. С помощью 3D моделирования в Unity уже создана модель квадрокоптера, однако реалистичность фоновой сцены требует существенного улучшения. Также планируется создать набор различных моделей квадрокоптеров, беспилотников вертолетного и самолетного типа.

Оригинальная модель YOLOv5 (без всяких настроек) была обучена, но показала результаты ощутимо хуже, поэтому результаты в статье не приведены. Сейчас производится настройка параметров модели YOLOv5, которая, по заявлению авторов, должна лучше работать с маленькими объектами, что подтверждается публикациями по распознаванию дронов (без модификаций лучше чем YOLOv8 [9]). Также планируется добавить использование трекеров с целью улучшения показателей качества распознавания.

References

- [1] *European ATM master plan: Digitalising Europe's aviation infrastructure. Executive view*, SESAR Joint Undertaking, ISBN: 978-92-9216-134-7, 2020. DOI: [10.2829/695700](https://doi.org/10.2829/695700).
- [2] N. Al-Iqubaydhi *et al.*, "Deep learning for unmanned aerial vehicles detection: A review", *Computer Science Review*, vol. 51, p. 100 614, 2024. DOI: <https://doi.org/10.1016/j.cosrev.2023.100614>.
- [3] F. Mahdavi and R. Rajabi, "Drone detection using convolutional neural networks", in *2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*, 2020, pp. 1–5. DOI: [10.1109/ICSPIS51611.2020.9349620](https://doi.org/10.1109/ICSPIS51611.2020.9349620).
- [4] U. Seidaliyeva, D. Akhmetov, L. Ilipbayeva, and E. Matson, "Real-time and accurate drone detection in a video with a static background", *Sensors*, vol. 20, no. 14, p. 3856, 2020. DOI: [10.3390/s20143856](https://doi.org/10.3390/s20143856).
- [5] A. Coluccia *et al.*, "Drone vs. bird detection: Deep learning algorithms and results from a grand challenge", *Sensors*, vol. 21, no. 8, p. 2824, 2021. DOI: [10.3390/s21082824](https://doi.org/10.3390/s21082824).
- [6] S. Jamil *et al.*, "Malicious UAV detection using integrated audio and visual features for public safety applications", *Sensors*, vol. 20, no. 14, p. 3923, 2020.
- [7] E. Unlu, E. Zenou, N. Riviere, and P.-E. Dupouy, "Deep learning-based strategies for the detection and tracking of drones using several cameras", *IPSJ T Comput Vis Appl*, vol. 11, no. 7, Unlu2019, 2019. DOI: [10.1186/s41074-019-0059-x](https://doi.org/10.1186/s41074-019-0059-x).
- [8] Y. Shang, C. Liu, D. Qiu, Z. Zhao, R. Wu, and S. Tang, "AD-YOLOv5s based UAV detection for low altitude security", *International Journal of Micro Air Vehicles*, vol. 15, p. 17 568 293 231 190 017, 2023. DOI: [10.1177/17568293231190017](https://doi.org/10.1177/17568293231190017).
- [9] X. Zhai, Z. Huang, T. Li, H. Liu, and S. Wang, "YOLO-drone: An optimized YOLOv8 network for tiny UAV object detection", *Electronics*, vol. 12, p. 3664, 2023. DOI: [10.3390/electronics12173664](https://doi.org/10.3390/electronics12173664).
- [10] J. Li, D. H. Ye, M. Kolsch, J. P. Wachs, and C. A. Bouman, "Fast and robust UAV to UAV detection and tracking from video", *IEEE Transactions on Emerging Topics in Computing*, vol. 10, no. 3, pp. 1519–1531, 2022. DOI: [10.1109/TETC.2021.3104555](https://doi.org/10.1109/TETC.2021.3104555).
- [11] A. Rozantsev, V. Lepetit, and P. Fua, "Detecting flying objects using a single moving camera", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 5, pp. 879–892, 2017. DOI: [10.1109/TPAMI.2016.2564408](https://doi.org/10.1109/TPAMI.2016.2564408).
- [12] H. Sun, J. Yang, J. Shen, D. Liang, L. Ning-Zhong, and H. Zhou, "Tib-net: Drone detection network with tiny iterative backbone", *IEEE Access*, vol. 8, pp. 130 697–130 707, 2020.
- [13] O. A. Levanova, D. V. Grushevskaya, A. A. Britnev, M. D. Averina, and D. M. Murin, *Baza videodannyh s razmetkoj bespilotnyh letatel'nyh apparatov (kvadrokopteroev). Svidetelstvo o gosudarstvennoy registratsii bazy dannyh no. 2021620329, date 2021-02-25*.