

Алгоритм автоматического детектирования объектов интереса для системы обнаружения и классификации птиц

А. Д. Безенкова

Санкт-Петербургский государственный электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)

annbeze57@gmail.com

Аннотация. В данной статье представлено нейросетевое решение задачи детектирования объектов интереса для системы обнаружения и классификации птиц. Обучение нейронной сети является основным объектом рассмотрения и может быть представлено как последовательность из трех основных этапов: выбор модели и поиск данных для обучения, процесс формирования единого набора данных и процесс обучения модели, а также тестирование и анализ результатов.

Ключевые слова: компьютерное зрение; задача обнаружения и классификации; машинное обучение

I. ВВЕДЕНИЕ

По статистике Росавиации за 2023 год произошло 1044 столкновения самолетов с птицами, из которых 90 привели к повреждению самолетов и были отнесены к авиационным инцидентам [1]. Из этих данных можно сделать вывод, что создание системы обнаружения и классификации птиц является актуальной задачей. В области компьютерного зрения обнаружение объектов является фундаментальной технологией, один из основных подходов в которой заключается в применении сверточных нейронных сетей (CNN) [2].

II. ВЫБОР МОДЕЛИ И ПОИСК ДАННЫХ ДЛЯ ОБУЧЕНИЯ

Для поставленной задачи характерна пара особенностей, накладывающая ограничения при выборе модели нейронной сети. Необходимо детектировать птиц на большом расстоянии для предотвращения столкновения, следовательно, птицы будут занимать малую площадь кадра, поэтому нужно обеспечить высокую точность предсказаний. А также, из-за быстрого перемещения объектов интереса по кадру, необходимо обеспечить высокое быстродействие алгоритма. Исходя из этих требований, была выбрана модель архитектуры YOLO, так как она отличается от других алгоритмов обнаружения объектов на основе CNN способностью детектировать область интереса в реальном времени за один проход через нейронную сеть [3]. Из-за аппаратных ограничений была выбрана архитектура YOLOv7 Tiny, представленная на рис. 1, так как на данный момент она позволяет достичь наибольшего количества обрабатываемых кадров в секунду.

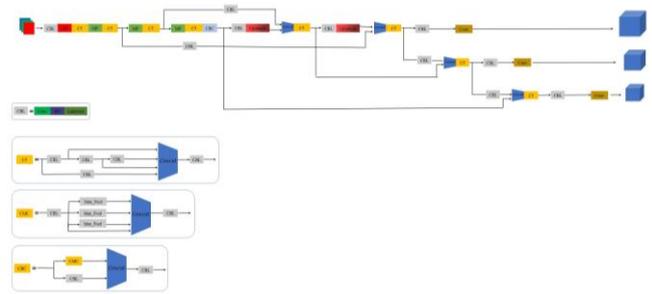


Рис. 1. Архитектура YOLOv7 Tiny

Для обеспечения высокой точности автоматической обработки кадров видеопоследовательности при обучении модели необходим набор размеченных данных, наиболее похожих на контекст решаемой задачи. Получить такой набор можно двумя способами: найти или сформировать самостоятельно. В данном случае первый подход приоритетнее ввиду высокой сложности разметки кадров. Было найдено более 170 тысяч размеченных изображений, этот объем состоит из нескольких отдельных наборов данных. Кадры полетов диких птиц на ветряной электростанции, содержащие 16400 изображений, кадры полетов птиц SOD4SD для соревнования по детектированию маленьких объектов – 39070, набор данных с камер аэропортов – 118312 [4, 5, 6]. Примеры кадров из описанных наборов данных представлены на рис. 2.



Рис. 2. Примеры изображений из найденных наборов данных

III. ФОРМИРОВАНИЕ ЕДИННОГО НАБОРА ДАННЫХ И ОБУЧЕНИЕ МОДЕЛИ

В системе обнаружения и классификации птиц использовалась камера с разрешением 3840 на 2160 пкс. При детектировании объектов интереса на дистанции не менее 1000 метров при заданных характеристиках камеры, птица будет занимать площадь больше или равную 7 на 7 пкс. Рассмотрим наборы данных: кадры у ветряной электростанции, и из набора SOD4SD имеют разрешение от 3840 на 2160 пкс до 5616 на 3744 пкс, и на них представлены птицы, занимающие площадь от 4 на 4 пкс до 570 на 244 пкс, разрешение кадров из аэропортов составляет 1920 на 1080 пкс, на которых птицы занимают меньше чем 10 на 10 пкс. Исходя из характеристик системы, было принято решение отказаться от использования набора данных с кадрами из аэропортов, таким образом, набор данных для обучения содержал 55000 изображений, разделённых на три группы: обучающая, валидационная и тестовая.

Для обучения модели были выбраны следующие параметры: в качестве начальных весов использовались веса, полученные при обучении модели на данных СОСО, изображение на входе модели масштабировалось до размера 3840 на 3840 пкс, количество эпох в одной итерации обучения составило 20, во время одной эпохи использовалось 33000 изображений из обучающей выборки, из-за высокого разрешения и технических ограничений, изображения подавались на вход модели по одному.

IV. ТЕСТИРОВАНИЕ И АНАЛИЗ РЕЗУЛЬТАТОВ

В ходе обучения внимание уделяется точности, с которой модель обнаруживает объекты, (Precision) и полноте модели (Recall), насколько стабильно происходит истинно-положительное обнаружение. Точность обнаружения показывает долю истинно-положительных решений от всех положительных решений. Полнота обнаружения отражает отношение истинно-положительных решений к сумме истинно-положительных и ложно-отрицательных решений. Метрика F1 является гармоническим средним между точностью и полнотой, ее используют для того, чтобы оценивать эти параметры совместно.

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$$

Веса, полученные в ходе 20 эпох обучения, обеспечили точность менее 0.2 при полноте обнаружения менее 0.3. На рис. 3 приведены графики изменения точности (Precision) и полноты (Recall) обнаружения на протяжении первой итерации обучения.

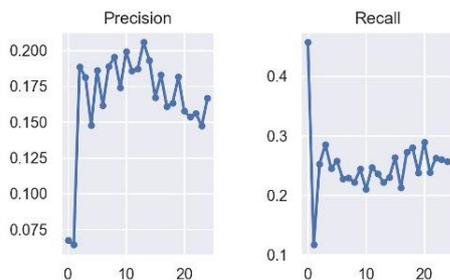


Рис. 3. Графики изменения характеристик Precision и Recall в ходе 20 эпох

При использовании данных весов для детектирования объектов на изображениях из тестовой выборки, заметно, что при обнаружении происходят ложные пропуски объектов интереса, пример приведен на рис. 4. Исходя из вышеперечисленных причин, следует, что необходимо продолжить обучение.



Рис. 4. Результат обнаружения объекта класса bird (птица) на изображении из тестовой группы после 20 эпох

Точность на тестовых данных, которую удалось достичь после 45 эпох обучения, составила 0.85, при полноте обнаружения 0.63, что показано на рис. 5. Значение F1-меры составило 0.72, это значит, что большинство птиц на кадрах из тестовой выборки детектируются верно. На рис. 6 приведен результат обнаружения птиц на изображениях из тестовой выборки.

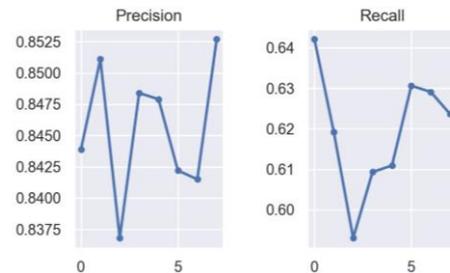


Рис. 5. Графики изменения характеристик Precision и Recall с 40 по 45 эпоху



Рис. 6. Результат обнаружения объекта класса bird (птица) на изображении из тестовой группы после 45 эпох

После получения высоких значений метрик на тестовой выборке, необходимо проверить обнаружение на кадрах, полученных в условиях предполагаемой эксплуатации. На рис. 5 приведены результаты детектирования птиц в воздушном пространстве на расстоянии менее 500 метров (а) и более 1000 метров от камеры (б).

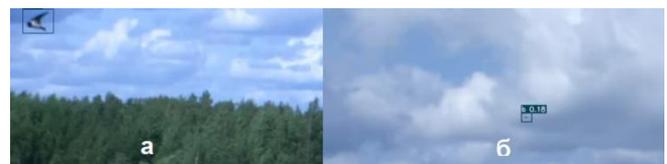


Рис. 7. Результаты обнаружения птиц на кадрах, полученных в условиях предполагаемой эксплуатации (а – дистанция менее 500 метров, б – дистанция более 1000 метров)

У. ЗАКЛЮЧЕНИЕ

В данной статье представлено решение задачи обнаружения птиц с помощью сверточной нейронной сети архитектуры YOLOv7 в рамках разработки системы обнаружения и классификации птиц для обеспечения безопасности аэропортов. В ходе решения поставленной задачи была сформирована обучающая выборка, состоящая из 55000 изображений. При тестировании результатов обучения было достигнуто высокое значение точности, составляющее 0.85. На кадрах видеопоследовательности, полученной в условиях предполагаемой эксплуатации, удалось обеспечить верное обнаружение птиц на расстоянии не более 1000 метров от камеры. Описываемая модель является частью системы обнаружения и классификации птиц и может быть улучшена для увеличения точности и дальности обнаружения.

СПИСОК ЛИТЕРАТУРЫ

- [1] Статистика столкновений с птицами и другими животными // АМРИПП Росавиации // URL: <https://favt.gov.ru/dejatelnost-bezopasnost-poletov-stolknoveniya-ptici-stat/>
- [2] Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with Deep Convolutional Neural Networks // Communications of the ACM, 2017. С. 84-90.
- [3] Bochkovskiy A., Wang C.-Y., Liao H.-Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection, 2020.
- [4] R. Yoshihashi, R. Kawakami, M. Iida, T. Naemura. "Bird detection and species classification with time-lapse images around a wind farm: Dataset construction and evaluation" // Wind Energy. 2017, вып. 12. С. 1983–1995.
- [5] Kondo, Yuki and Ukita. "MVA2023 Small Object Detection Challenge for Spotting Birds: Dataset, Methods, and Results" // 18th International Conference on Machine Vision and Applications (MVA), С. 1-4.
- [6] H. Sun, Y. Wang, X. Cai, P. Wang, Z. Huang, D. Li, Y. Shao, and S. Wang. "Airbirds: A large-scale challenging dataset for bird strike prevention in real-world airports" // ACCV, 2022. С. 1-4.