

УДК 004.855.5

ПРИМЕНЕНИЕ КАСКАДА ХААРА ДЛЯ РАЗРАБОТКИ СИСТЕМЫ РАСПОЗНАВАНИЯ ПТИЦ*А. А. Хорошева, М. Д. Федько, А. Р. Юсупов*

Донской государственной технической университет (г. Ростов-на-Дону, Российская Федерация)

В статье описано исследование применения каскада Хаара для распознавания птиц, описаны особенности формирования датасета и обучения классификатора. Отмечены достоинства и недостатки подхода.

Ключевые слова: каскад Хаара, классификатор, машинное обучение, распознавание изображений, алгоритм Виолы-Джонса.

UDC 004.855.5

APPLICATION OF HAAR CASCADE FOR BIRD RECOGNITION SYSTEM DEVELOPMENT*A. A. Khorosheva, M. D. Fedko, A. R. Yusupov*

Don State Technical University (Rostov-on-Don, Russian Federation)

This article considers the use of the Haar cascade for bird recognition, describes the features of the formation of the dataset and the training of the classifier. The advantages and disadvantages of the approach are noted.

Keywords: Haar cascade, classifier, machine learning, image recognition, Viola-Jones algorithm.

Введение. Распознавание образов — активно развивающаяся область информатики. Под понятие образа попадают предметы, явления, процессы, сигналы и ситуации.

В данной статье рассмотрено распознавание объектов на изображениях. Современные технологии позволяют производить распознавание изображения достаточно быстро, что является особо важным фактором в самых разных сферах применения, даже в области защиты природы.

Возрастающая антропогенная нагрузка на окружающую среду приводит к вымиранию или резкому сокращению числа видов животных и растений. Применяются различные природоохранные методы, в которых широко задействованы информационные технологии. Машинное обучение, в частности, распознавание объектов помогает осуществлять отслеживание и контроль за численностью тех или иных видов, например, с помощью определителей [1].

Таким образом, распознавание птиц с изображений или видеопотока является достаточно актуальной задачей, а прототипы таких интеллектуальных систем распознавания могут применяться достаточно широко — от приложений-определителей видов до умных кормушек.

Цель данной статьи — исследовать особенности построения интеллектуальной системы распознавания птиц. В качестве основного инструмента распознавания выбран каскад Хаара, который показывает хорошие результаты при распознавании объектов как с изображений, так и с видеопотока, что является важным фактором при построении такой системы [2].

Каскад Хаара. Каскад Хаара основан на алгоритме Виолы-Джонса и изначально применялся для решения задачи распознавания лиц. Однако данный алгоритм может применяться для детектирования любых объектов, как на изображениях, так и на видео. Важно отметить, что алгоритм устойчив к зашумлениям и искажениям изображения [3].

Перечислим основные принципы, на которых строится данный алгоритм [3]:

1. Применение многочисленных признаков Хаара. Для распознавания изображения в алгоритме Виолы-Джонса используются признаки Хаара. Это прямоугольные черно-белые

признаки, построенные на основе вейвлета Хаара. Они представляют из себя пространство признаков для распознавания и представлены на рис. 1.

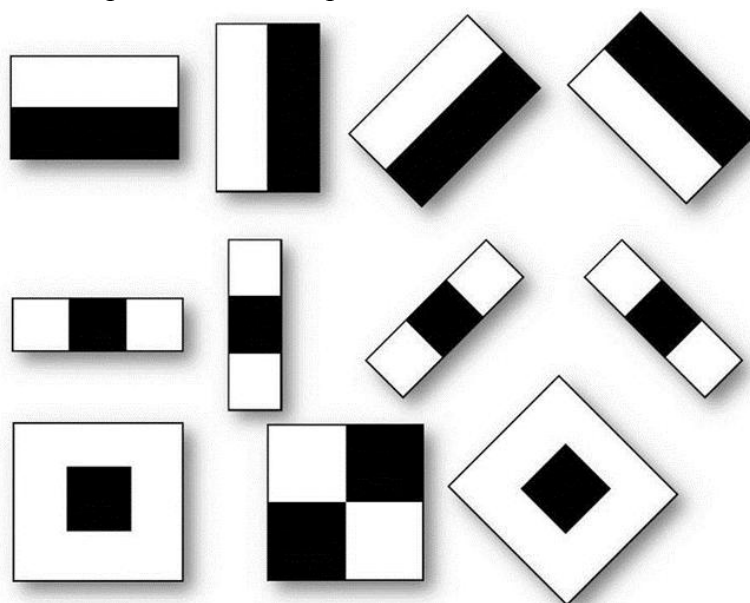


Рис. 1. Прямоугольные признаки Хаара

Значение признака вычисляется по формуле:

$$F = X - Y,$$

где X — сумма яркостей точек в светлой части признака, а Y — сумма яркостей точек в темной части признака.

2. Использование интегрального представления изображений. Для работы с большим массивом данных удобно использовать интегральное представление изображения — матрицу, в каждом элементе которой находится сумма интенсивности пикселей, размещенных выше и левее данного элемента. Рассчитать каждый элемент можно по следующей формуле:

$$L(x, y) = I(x, y) - L(x - 1, y - 1) + L(x, y - 1) + L(x - 1, y),$$

где $I(x, y)$ — яркость пикселя исходного изображения, $L(x, y)$ — полученный элемент матрицы.

Применение интегрального представления позволит резко сократить количество вычислений для расчета признаков Хаара [4];

3. Обучение с помощью AdaBoost. При обучении используются все конфигурации одного признака. Таким образом для изображения размером 24×24 пикселя число конфигураций около 40 000. Для каждой конфигурации тренируется слабый классификатор, и затем с помощью алгоритма AdaBoost из коллекции слабых классификаторов получается сильный классификатор [4];

4. Распознавание с помощью сканирующего окна. В процессе распознавания на вход алгоритма поступает тестируемое изображение, и алгоритм начинает обходить его с помощью сканирующего окна. Это окно имеет начальный размер 24×24 пикселя и вычисляет все варианты признаков. Затем окно масштабируется, и процедура обхода повторяется. Признаки попадают к обученному классификатору, который и выносит вердикт о соответствии объекта.

Особенности формирования датасета. Каскад Хаара для обучения принимает на вход два датасета — положительный и отрицательный [3]. Положительный содержит в себе изображения, содержащие искомый объект, а также файл, содержащий пути к данным изображениям и

координаты объекта на изображении. Соответственно, негативный датасет должен содержать произвольные изображения, на которых объекта нет.

В качестве искомого объекта выбрана сорока европейская (лат. *picca picca*). Изначальный датасет содержал в себе несколько сотен различных фотографий сорок. Так как положительный датасет требует координаты искомого объекта на изображении, проще всего обрезать изображение так, чтобы искомым объектом занимал большую часть. Соответственно, координаты тогда будут равны $0\ 0\ w\ h$, где w, h — ширина и высота изображения.

Таким образом, требуется произвести препроцессинг исходного датасета. Алгоритм препроцессинга можно свести к следующим этапам:

1. Перевод изображений в градации серого;
2. Создание черно-белой маски;
3. Применение детектора границ Канны к обесцвеченному изображению по маске;
4. Отбор изображений, содержащих нужный объект;
5. Уменьшение размера изображений.

Применение автоматизированной обработки изображений позволяет уменьшить количество ручной работы при формировании положительного датасета.

Соответственно, для отрицательного датасета особой подготовки не требуется.

Обучение и классификация. Для обучения классификатора воспользуемся библиотекой OpenCV. Обучение будет состоять из двух этапов [5]:

- подготовка вектора положительных изображений,
- обучение каскада.

Следует отметить, что обучение каскада на 20 стадиях занимает достаточно продолжительное время. На выходе получается подключаемый xml-файл, который можно использовать в разработке программ. Для тренировки применялось 108 положительных и 220 отрицательных изображений.

Результат работы классификатора можно увидеть на рис. 2.

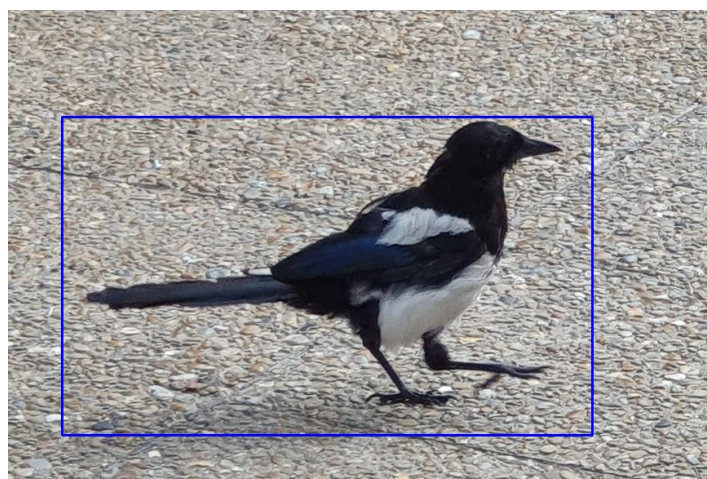


Рис. 2. Найденный на изображении объект

Малое число положительных и отрицательных примеров приводят к неточному детектированию птиц на снимках, на части изображений детектор обнаруживает птицу частями или некорректно определяет контуры (рис. 3).

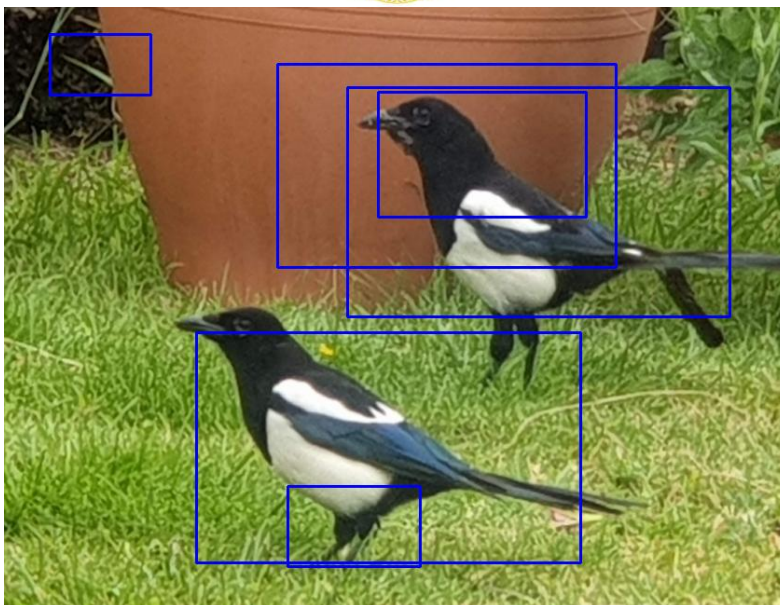


Рис. 3. Множественное ошибочное детектирование

Заключение. Применение каскада Хаара для распознавания птиц показывает достаточно хорошие результаты даже на небольших обучающих датасетах. Однако для более точной работы требуется задействовать выборки больших размеров, а также большее число уровней обучения классификатора.

Библиографический список

1. Савченко, Г. Появилось приложение, распознающее растения и животных по фотографиям [Электронный ресурс] // Bird in Flight. — Режим доступа: <https://birdinflight.com/ru/novosti/20180406-20180406-shazam-for-plants-and-animals-mobile-app.html> (дата обращения : 22.12.2019).
2. Viola P., Jones M. Robust real-time face detection. *IJCV* 57(2), 2004
3. Viola P., Jones M., Rapid Object Detection using a Boosted Cascade of Simple Features. *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001)*, 2001
4. Berger W. Deep Learning Haar Cascade Explained. <http://www.willberger.org>. 2017. Available from: <http://www.willberger.org/cascade-haar-explained/> (Accessed 22nd December 2019).
5. Docs, OpenCV. Face Detection Using Haar Cascades. *OpenCV: Face Detection Using Haar Cascades*, 2017, docs.opencv.org/3.3.0/d7/d8b/tutorial_py_face_detection.html (Accessed 22nd December 2019).

Об авторах:

Юсупов Александр Рашидович, старший преподаватель Донского государственного технического университета (344000, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), temple_black@mail.ru

Федько Максим Денисович, студент Донского государственного технического университета (344000, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), whitecoloredowl@gmail.com

Хорошева Анастасия Александровна, студент Донского государственного технического университета (344000, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), tombalabomba@yandex.ru



Authors:

Yusupov Aleksandr Rashidovich, Senior lecturer, Don State Technical University (344000, Russian Federation, Rostov-on-Don, Gagarin square 1), temple_black@mail.ru

Fedko Maksim Denisovich, Student, Don State Technical University (344000, Russian Federation, Rostov-on-Don, Gagarin square 1), whitecoloredowl@gmail.com

Khorosheva Anastasiya Aleksandrovna, Student, Don State Technical University (344000, Russian Federation, Rostov-on-Don, Gagarin square 1), tombalabomba@yandex.ru