

УДК 004.9

АВТОМАТИЗАЦИЯ КОНТРОЛЯ ЧИСЛЕННОСТИ ЦЕЛЕВЫХ ОБЪЕКТОВ ФАУНЫ С ПОМОЩЬЮ АЭРОФОТОСЪЕМКИ

Талбонен А.Н., Рогов А.А., Калинин А.В., Тимонин А.О.

Петрозаводский государственный университет, Петрозаводск, e-mail: arogov@petsru.ru

В статье описывается проблема автоматической обработки аэрофотоснимков местности и поиска на изображениях целевых объектов фауны. В данной статье предложен комбинированный метод, который состоит из двух этапов: детектирование и классификация объектов. Для целей детектирования используется алгоритм Виолы – Джонса, результаты работы которого классифицируются с помощью сверточной нейронной сети. Данный подход обладает достаточной гибкостью для применения к различным группам искомым объектов в различных условиях. Это становится возможным благодаря этапу обучения каскада Хаара на различных выборках, а также обучению сверточной нейронной сети и регулированию количества искомым классов. В статье приводится сравнение результатов детектирования и классификации, полученных в ходе исследования. Данное исследование имеет большое значение для различных мероприятий в области охраны окружающей среды.

Ключевые слова: аэрофотосъемка, обработка изображений, поиск объектов на изображениях, распознавание образов

AUTOMATIZATION OF SPECIFIC FAUNA OBJECTS COUNTING USING AERIAL PHOTOGRAPHS

Talbonen A.N., Rogov A.A., Kalinin A.V., Timonin A.O.

Petrozavodsk State University, Petrozavodsk, e-mail: arogov@petsru.ru

The paper describes the problem of automatic processing of aerial photographs and the problem of finding animals in images. The method described in this article consist of two stages: object detection and classification. Viola-Jones algorithm is used for object detection and then convolutional neural network (CNN) is used for classifying objects detected on the previous stage. The proposed approach in this paper is flexible and can be applied in different conditions. It is became possible because of cascade training with different sets of images and convolutional neural network training with different objects and different number of classes. The paper contains results of comparing different detection and classification methods. This research has a great importance for the various activities in the field of environmental protection.

Keywords: aerial photography, image processing, search for objects in images, pattern recognition

В настоящее время аэрофотосъемка с помощью беспилотных летательных аппаратов (БПЛА) становится все более популярным способом мониторинга местности. Данный способ позволяет получить информацию, которая очень важна для принятия верных управленческих решений, направленных на охрану и рациональное использование природных ресурсов, кадастровый учет, транспортный надзор, и многих других целей.

С каждым днем растет количество компаний, занимающихся поиском объектов с помощью воздушной разведки местности. Для ее проведения чаще всего используется БПЛА, оборудованный цифровой профессиональной фотокамерой, которому задается определенный маршрут. При этом в результате съемки создается огромная коллекция изображений, в которой доля искомым объектов незначительна. Большинство организаций, таких как «ZALA AERO», «Fincso AnimalCount» и «Русгеоком», предоставляющих подобные услуги, выполняют поиск

данных объектов вручную. В случае очень больших коллекций (несколько тысяч и более) обработка изображений становится достаточно трудоемкой. В связи с этим становится актуальной автоматизация поиска целевых объектов на цифровых изображениях (аэрофотоснимках) (рис. 1).

Целью исследования является разработка комбинированного метода детектирования целевых объектов на основе алгоритма Виолы – Джонса и сортировки его результатов с помощью сверточной нейронной сети.

Комбинированный метод

Процесс обнаружения объектов технически представляет собой формирование множества областей в виде набора координат прямоугольников на основе анализа содержимого исходного изображения. Участки исходного изображения, соответствующие сформированным прямоугольным областям, будут считаться найденными объектами.



Рис. 1. Пример аэрофотоснимка с выделенными объектами

Множество найденных объектов можно разделить на 2 подмножества: множество релевантных объектов, то есть тех, что соответствуют искомому классу, и множество нерелевантных объектов. На основе соотношения релевантных и нерелевантных объектов определяются количественные характеристики качества работы алгоритма обнаружения: полнота и точность.

Полнота вычисляется как отношение найденных релевантных документов к общему количеству релевантных документов. Полнота характеризует способность системы находить нужные пользователю объекты, но не учитывает количество нерелевантных объектов, выдаваемых пользователю.

Точность вычисляется как отношение найденных релевантных объектов к общему количеству найденных документов. Точность характеризует способность системы выдавать в списке результатов только релевантные объекты [1].

Зачастую, если рассматривать различные модификации одного и того же алгоритма, точность и полнота будут находиться в обратной зависимости. То есть можно добиться высокой точности, жертвуя полнотой, и наоборот. Поэтому в настоящее время для обнаружения различных объектов с высокими показателями и точности, и полноты применяется комбинированный метод.

Общая схема работы комбинированного метода выглядит следующим образом. На первом этапе изображения из коллекции подаются на вход детектора. Для достижения наилучших показателей используется алгоритм обнаружения с параметрами, соответствующими результатам с наибольшей полнотой, в результате работы которого получается набор найденных объектов. Среди

них могут быть релевантные и нерелевантные объекты. На втором этапе найденные объекты подаются на вход классификатора, задачей которого является сортировка объектов (требуется оставить релевантные объекты и отбросить нерелевантные). Таким образом удастся повысить точность поиска целевых объектов.

В данном исследовании в качестве алгоритма обнаружения рассматривается алгоритм Виолы – Джонса, который основан на использовании каскадов из слабых классификаторов, таких как детекторы Хаара [2, 9]. Данный алгоритм реализован в библиотеке OpenCV [3, 8], однако для его работы требуется обученный каскад классификаторов. Помимо детектора Хаара в качестве алгоритма поиска объектов может использоваться метод локальных бинарных шаблонов и другие методы детектирования [4].

Качество работы каждого слабого классификатора в отдельности мало отличается от случайного выбора, поэтому для построения сильного классификатора используется так называемый «бустинг» [2]. «Бустинг» (от англ. boosting – повышение, усиление, улучшение) – это процедура последовательного построения композиции алгоритмов машинного обучения, когда каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов. Основная задача такой процедуры – из множества слабых классификаторов (в данном случае – это признаки Хаара) построить эффективный сильный классификатор (каскад).

Классификатор называется сильным потому, что, в отличие от слабого, точность классификации значительно лучше чем при случайном выборе. А точность целого каскада у же стремится к 1.

Полученный в результате обучения каскад необходимо сохранить. Для этого можно использовать XML-файлы, как это сделано в библиотеке «OpenCV».

Представленный выше алгоритм обучения реализован в библиотеке «OpenCV» в виде утилиты `opencv_traincascade`. Данная утилита требует наборы позитивных и негативных образов (релевантных и нерелевантных объектов соответственно).

Далее формируется специальный конфигурационный файл с помощью утилиты `opencv_createsamples`, который затем подается на вход `opencv_traincascade`. После того, как `opencv_traincascade` завершит свою работу, обученный каскад будет сохранен в XML-файл. Это готовый обученный классификатор, который можно использовать для поиска объектов на изображении.

Сверточные нейронные сети

Сверточные нейронные сети (СНС, CNN), предложенные Яном Лекуном, в отличие от простых искусственных нейронных сетей благодаря своей архитектуре позволяют эффективно распознавать изображения [7]. В структуру такой сети входят так называемые сверточные (convolution) и субдискретизирующие (subsampling) слои. Архитектура типичной СНС представлена на рис. 2.

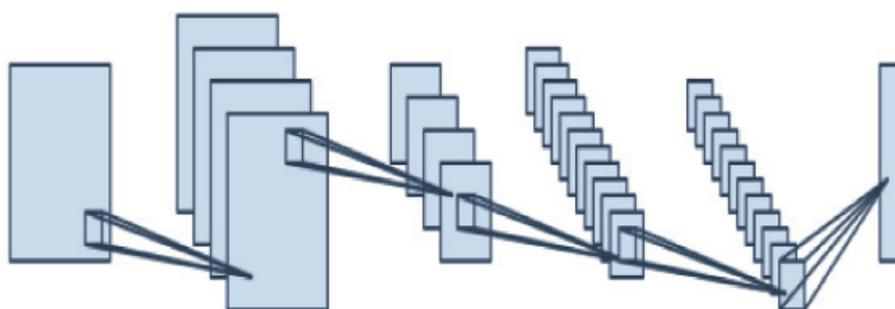


Рис. 2. Архитектура типичной сверточной нейронной сети

На данный момент существует несколько различных программных платформ и библиотек для работы с СНС. Наиболее распространенной среди них является библиотека «Caffe», разработанная в Berkeley Vision and Learning Center (BVLC) [6].

Для обучения и применения нейросетевых классификаторов библиотека «Caffe» предлагает несколько утилит.

Утилита `convert_imageset` формирует обучающую выборку в формате «LevelDB», пригодном для работы алгоритма обучения. Данная утилита принимает на вход текстовый файл, содержащий строки следующего

вида: `image.jpg < class >`, где `< class >` – числовой идентификатор класса изображения.

Для вычисления среднего значения изображения предназначена утилита `compute_image_mean`, которая на вход принимает сформированную базу данных типа LevelDB и возвращает на выходе файл, содержащий среднее значение изображения.

Для обучения сверточной нейронной сети в библиотеке «Caffe» существует программа `caffe`, которая принимает на вход специальный текстовый файл с конфигурационными параметрами. В данном файле указываются такие параметры обучения, как максимальное количество итераций (эпох обучения), периодичность и длительность тестирования промежуточных состояний и конфигурационный файл нейронной сети. Последний представляет собой текстовое описание слоев сети с указанием источника данных для обучения и тестирования и количеством выходных параметров.

Выходные значения представляют собой некоторые веса, которые вычисляются в процессе работы нейронной сети и сумма которых равна 1. Количество выходных параметров фактически определяет количество поддерживаемых классов. Оно также связано с теми значениями классов, которые указываются в файле обучающей выборки.

Практические результаты

Для обучения каскада классификаторов была подготовлена обучающая выборка размером 640 релевантных и 9800 нерелевантных объектов. Для тестирования полученного каскада была сформирована выборка из 13 цифровых аэрофотоснимков, на которых были размечены 24 релевантных объекта.

При обучении использовались параметры:
`numStages = 16` – количество уровней каскада, которые программа будет обучать;
`minihtrate = 0,99` – минимальная доля релевантных объектов, которые были корректно распознаны при обучении;

$\text{maxFalseAlarmRate} = 0,5$ – уровень ложного распознавания нерелевантных объектов, при котором обучение будет остановлено;

$\text{mode} = \text{ALL}$ – использовать полный комплект Хаар-признаков, так как нам важна ориентация объекта.

Тестирование было произведено на языке Python с помощью библиотеки OpenCV и следующего кода:

```
cascade_path = «/path/to/cascade.xml»
cascade = cv2.CascadeClassifier(cascade_path)
faces = cascade.detectMultiScale(image, scaleFactor, minNeighbors)
```

Результаты поиска объектов

№ п/п	Обозначение	Всего объектов найдено	Релевантных объектов найдено	Точность	Полнота
1	haar	340	23	0,068	0,958
2	haar_pr	50	22	0,440	0,917
3	haar_NN	83	21	0,253	0,875
4	haar_pr_NN	30	21	0,700	0,875

Результаты работы функции `detectMultiScale` зависят от входных параметров `scaleFactor` и `minNeighbors`, которые обозначают коэффициент увеличения скользящего окна и минимальное количество обнаружений одной и той же области при разных размерах скользящего окна соответственно. Для поиска оптимальных значений параметров был запущен тест с использованием перебора диапазона допустимых значений. В результате тестирования были отобраны 2 группы значений параметров, показавшие наилучшие результаты с точки зрения полноты и точности («haar» и «haar_pr»). Результаты работы алгоритма с этими параметрами представлены в таблице.

Для обучения сверточной сети, соответствующей данной предметной области, была взята за основу структура сети CIFAR-10, конфигурация которой также представлена в Caffe [5]. Единственным отличием в данном проекте является только входной слой и количество классов выходного слоя (0 – нерелевантный объект, 1 – релевантный объект).

В качестве обучающей выборки было подготовлено множество из 640 релевантных объектов и 640 нерелевантных. В качестве тестовой выборки было подготовлено множество, содержащее 23 релевантных и 31 нерелевантный объект. В результате классификации 50 объектов из 54-х были классифицированы правильно с точностью 92,6 %.

Полученный нейросетевой классификатор был использован для сортировки результатов алгоритма Виолы –Джонса и был применен к результатам, обозначенным как «haar» и «haar_pr». Полученные отсортированные выборки («haar_NN» и «haar_pr_NN» соответственно) были оценены с точки зрения полноты и точности. Результаты работы алгоритма сортировки представлены в таблице.

Заключение

В результате исследования удалось разработать комбинированный метод обнаружения целевых объектов фауны, обладающий точностью 70 % и полнотой 87,5 %. Несмотря на это, в процессе исследования обнаружились некоторые проблемы. Так, распознавание с помощью каскада Хаара с параметрами, направленными на детектирование наибольшего количества релевантных объектов («haar»), хотя и показало наибольшую полноту, точность результатов оказалась незначительной. При этом сортировка позволила повысить точность только до 25 %. Причиной такой работы каскада Хаара является недостаточное наполнение обучающей выборки изображениями различных типов из-за изначально ограниченной коллекции данных. Подобные проблемы так же проявились и при обучении нейронной сети.

В дальнейшем для решения вышеуказанных проблем предполагается использовать обучающие выборки с множеством классов, что в перспективе позволит заметно повысить точность и полноту поиска.

Работа выполнена при финансовой поддержке Программы стратегического развития ПетрГУ в рамках реализации комплекса мероприятий по развитию научно-исследовательской деятельности.

Список литературы

1. Агеев М., Кураленок И., Некрестьянов И. Официальные метрики РОМИП 2009 [Электронный ресурс]. – Режим доступа: http://romip.ru/romip2009/20_appendix_a_metrics.pdf (дата обращения: 22.05.2015).
2. Метод Виолы-Джонса (Viola-Jones) как основа для распознавания лиц [Электронный ресурс]. – Режим доступа: <http://habrahabr.ru/post/133826/> (дата обращения: 22.05.2015).
3. Распознавание образов для программистов [Электронный ресурс]. – Режим доступа: <http://recog.ru/blog/opencv/page2/> (дата обращения: 22.05.2015).
4. Рогов А.А., Талбонен А. Н., Тимонин А.О., Калинин А.В. Автоматизация контроля численности целевых объектов фауны с помощью аэрофотосъемки // Современные методы прикладной математики, теории управления и компьютерных технологий: сборник трудов VII Международной конференции. – 2014. – С. 372–375.
5. Caffe | CIFAR-10 tutorial [Электронный ресурс]. – Режим доступа: <http://caffe.berkeleyvision.org/gathered/examples/cifar10.html> (дата обращения: 22.05.2015).
6. Caffe | Deep Learning Framework [Электронный ресурс]. – Режим доступа: <http://caffe.berkeleyvision.org> (дата обращения: 22.05.2015).
7. Convolutional Neural Networks (LeNet) – DeepLearning 0.1 documentation [Электронный ресурс]. – Режим доступа: <http://deeplearning.net/tutorial/lenet.html>. (дата обращения: 22.05.2015).
8. OpenCV – библиотека алгоритмов компьютерного зрения [Электронный ресурс]. – Режим доступа: <http://develnet.ru/blog/493.html> (дата обращения: 22.05.2015).
9. Viola P., Jones M. Robust Real-time Object Detection // Second International Workshop on Statistical and Computational Theories of Vision – Modelling, Learning, Computing and Sampling // Ref. Libr. – 2001.

References

1. Ageev M., Kuralenok I., Nekrestjanov I. Oficialnye metriki ROMIP 2009 [Jelektronnyj resurs]. Rezhim dostupa: http://romip.ru/romip2009/20_appendix_a_metrics.pdf (data obrashhenija: 22.05.2015).

2. Metod Violy-Dzhonsa (Viola-Jones) kak osnova dlja raspoznavanija lic [Jelektronnyj resurs]. Rezhim dostupa: <http://habrahabr.ru/post/133826/> (data obrashhenija: 22.05.2015).

3. Raspoznavanie obrazov dlja programmistov [Jelektronnyj resurs]. Rezhim dostupa: <http://recog.ru/blog/opencv/page2/> (data obrashhenija: 22.05.2015).

4. Rogov A.A., Talbonen A.N., Timonin A.O., Kalinin A.V. Avtomatizacija kontrolja chislenosti celevykh obektov fauny s pomoshhju ajerofotos#emki // Sovremennye metody prikladnoj matematiki, teorii upravlenija i kompjuternyh tehnologij: sbornik trudov VII Mezhdunarodnoj konferencii. 2014. pp. 372–375.

5. Caffe | CIFAR-10 tutorial [Jelektronnyj resurs]. Rezhim dostupa: <http://caffe.berkeleyvision.org/gathered/examples/cifar10.html> (data obrashhenija: 22.05.2015).

6. Caffe | Deep Learning Framework [Jelektronnyj resurs]. Rezhim dostupa: <http://caffe.berkeleyvision.org> (data obrashhenija: 22.05.2015).

7. Convolutional Neural Networks (LeNet) DeepLearning 0.1 documentation [Jelektronnyj resurs]. Rezhim dostupa: <http://deeplearning.net/tutorial/lenet.html>. (data obrashhenija: 22.05.2015).

8. OpenCV biblioteka algoritmov kompjuternogo zrenija [Jelektronnyj resurs]. Rezhim dostupa: <http://develnet.ru/blog/493.html> (data obrashhenija: 22.05.2015).

9. Viola P., Jones M. Robust Real-time Object Detection // Second International Workshop on Statistical and Computational Theories of Vision Modelling, Learning, Computing and Sampling // Ref. Libr. 2001.

Рецензенты:

Питухин Е.А., д.т.н., профессор, Петрозаводский государственный университет, г. Петрозаводск;

Печников А.А., д.т.н., доцент, ведущий научный сотрудник, Институт прикладных математических исследований Карельского научного центра РАН, г. Петрозаводск.