

ПРИМЕНЕНИЕ АЛГОРИТМОВ СЕГМЕНТАЦИИ ДЛЯ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ

Дмитриев Д.В., Капранов С.Н.

ФГБОУ ВПО «Нижегородский государственный технический университет им. Р.Е. Алексеева», Нижний Новгород, Россия (603950, г. Нижний Новгород, ул. Минина, 24), e-mail: serg.kapranov@gmail.com

В работе была рассмотрена задача машинного обучения применительно к поиску объектов на изображении. Для классификации изображений использовался подход «мешок слов». В работе рассматривался ряд классификаторов: SVM, DTree, RTree, ERTree, GBTtree, Boost, Bayes, KNearest. Оригинальный алгоритм классификации был дополнен этапами сегментации и фильтрации. На этапе сегментации выполнялось отделение искомого объекта от фона и других объектов, находящихся на изображении, с помощью алгоритмов GrabCut и Watershed. Этап фильтрации служил для выделения ключевых точек, принадлежащих только объекту, который необходимо обнаружить на изображении. Предложенный алгоритм реализован с использованием библиотеки компьютерного зрения OpenCV. Внесение в оригинальный алгоритм этапов сегментации и фильтрации позволило на 20–30 % уменьшить количество ошибок классификатора.

Ключевые слова: машинное обучение, компьютерное зрение, сегментация изображений.

IMAGE CLASSIFICATION USING SEGMENTATION ALGORITHMS

Dmitriev D.V., Kapranov S.N.

“Nizhny Novgorod State Technical University n.a. R.E. Alekseev” Nizhny Novgorod, Russia (603950, Nizhny Novgorod, street Minin, 24), e-mail: serg.kapranov@gmail.com

The paper considers a machine-learning problem for image classification and detection on the scene. The “Bag of words” approach has used for image classification. The article involves the following classifiers: SVM, DTree, RTree, ERTree, GBTtree, Boost, Bayes, KNearest. The original classification algorithm was extended with segmentation and filtration steps. On the segmentation step the object key points were separated from background and another objects on the scene using GrabCut and Watershed algorithms. The filtration step used for object key points allocation. This object should be detected on the image. The algorithm was implemented using computer vision library OpenCV. These algorithm improvements allow to decrease classification errors about 20–30 %.

Keywords: machine learning, computer vision, image segmentation.

Системы компьютерного зрения все больше и больше проникают в различные области науки и техники: сельское хозяйство, метрологию, производство изделий для электронной промышленности и т.д. Основными задачами систем компьютерного зрения являются: анализ двумерного, трёхмерного изображения или видео потока, а также поиск и выделение объектов.

На протяжении последних десятилетий задаче поиска объектов на изображении уделяется большое внимание из-за того, что это позволяет производить обнаружение и классификацию объектов. На данный момент выполнено огромное количество исследований и опубликовано большое количество работ, однако, с нашей точки зрения, данная область остаётся ещё не просто не до конца изученной, а всего лишь слегка тронутый исследователями.

Методы распознавания лиц Виолы Джонса и Майкла Джонса [4], [5] до сих пор являются основными в распознавании не только лиц, но и изображений в целом. В основу их методов положен метод скользящего окна, который используется во всех бюджетных фотоаппаратах для выделения лиц.

На данный момент компьютерное зрение – это технология, позволяющая создавать информационные системы, имеющие навыки обнаружения, слежения и классификации объектов. Основная цель, стоящая перед этой технологией, – построение программного обеспечения, способного понимать, что присутствует на изображении.

Объект исследования

Для решения задачи машинного обучения бинарной классификации используется подход, называемый «мешок слов» [2]. Исходными данными для этого алгоритма выступают изображения, которые делятся на три категории:

- изображения, содержащие искомые объекты;
- изображения, не содержащие искомого объекта;
- проверочная выборка изображений.

Первым шагом алгоритма (рис.1) выступает поиск ключевых точек изображения. Затем вычисляются дескрипторы ключевых точек. В работе используются дескрипторы SIFT и SURF.

Далее выполняется обучение словаря. Признаки разбиваются на группы с помощью алгоритмов кластеризации. Результатом работы алгоритма кластеризации является словарь (центры кластеров).

После этого выполняется построение признакового описания изображений. Далее выполняется поиск ключевых точек и нахождение дескрипторов. Для каждого дескриптора выполняется поиск ближайшего центра кластера. На основе вычисленных дескрипторов формируем «мешок слов», т.е. частоту встречаемости определенных групп дескрипторов на изображении. Слово в данном случае – это часто повторяющиеся фрагменты изображения.

В качестве последнего шага выступает алгоритм машинного обучения, который использует классификаторы. В данной работе было использовано восемь классификаторов [1]: SVM, DTree, RTree, ERTree, GBTree, Boost, Bayes, KNearest.

Недостатком алгоритма является то, что для классификации используются все ключевые точки изображения. Это означает, что для обучения используются точки, которые объекту не принадлежат.

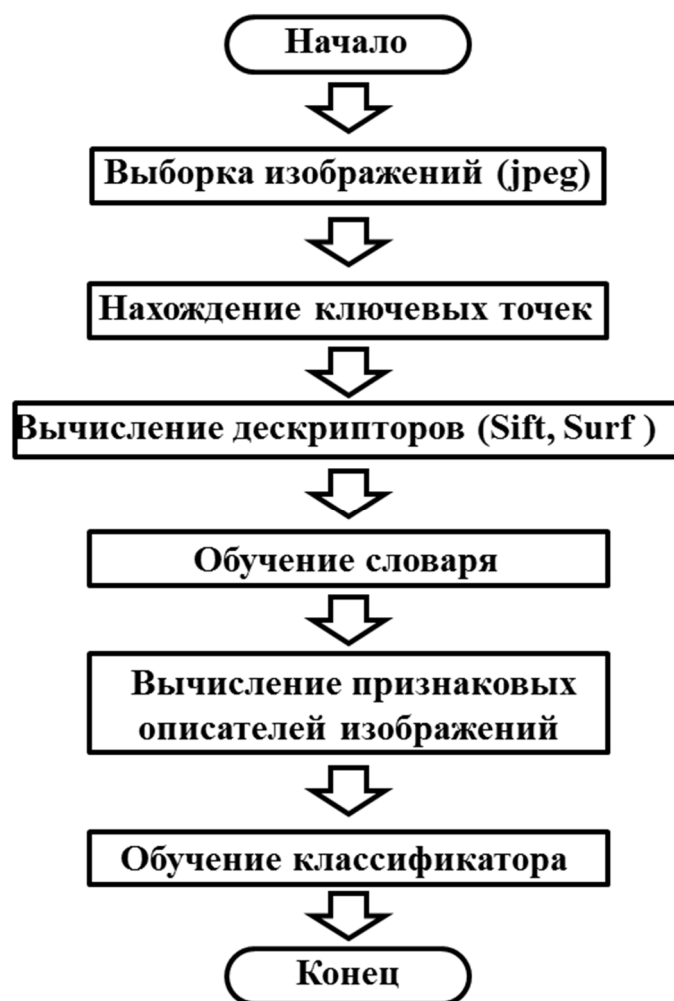


Рис. 1. Схема алгоритма машинного обучения бинарной классификации

Было предложено в алгоритм добавить шаг сегментации объекта и шаг фильтрации (рис. 2). Для этого были использованы алгоритмы сегментации [3] Watershed, GrabCut. Алгоритмы сегментации служат для выделения двух сегментов, один из них относится к объекту, другой к области, на которой ошибочно находятся ключевые точки.

Целью шага фильтрации станет фильтрация точек, которые не принадлежат объекту. Это позволит сократить процент ошибок при поиске объектов. На первой стадии алгоритма фильтрации выполняется ручная сегментация изображения, т.е. будут выделены объект и фон изображения. Вторая стадия включает в себя создание текстового файла, который будет содержать матрицу, размер которой будет равен размеру изображения, и каждый элемент будет являться точкой изображения (0 или 1). Ноль – это точка, которая принадлежит объекту. Единица – это точка, которая принадлежит фону. На основании этих данных будут фильтроваться ключевые точки.

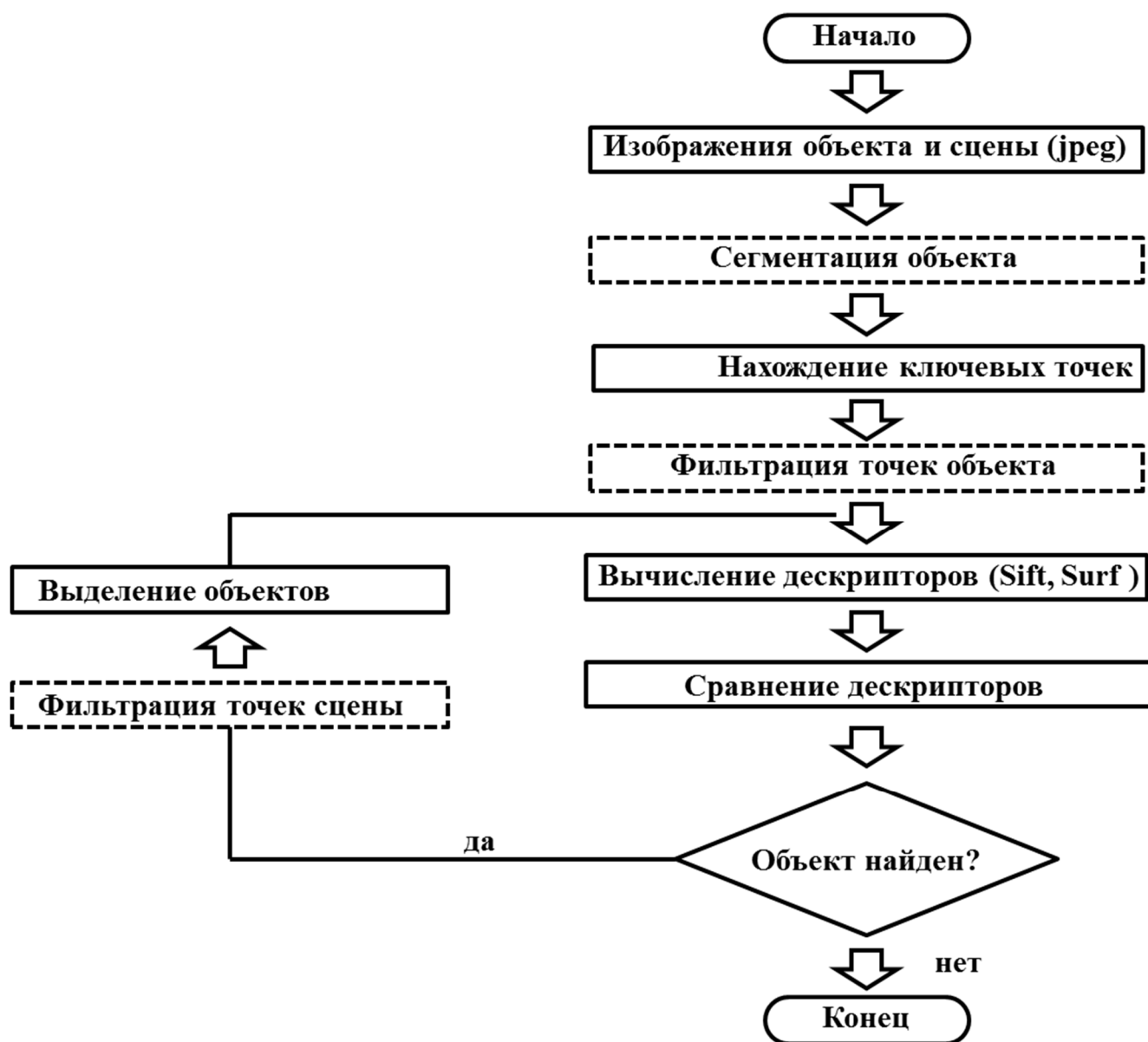


Рис. 2. Схема модифицированного алгоритма машинного обучения

На шаге фильтрации точек объекта будет выполняться считывание матрицы из текстового файла. Каждая ключевая точка имеет координату, предполагается считывать координату этой точки и сравнивать ее с значением из матрицы. Если ключевая точка принадлежит фону, то она удаляется из массива ключевых точек. Если координата ключевой точки совпадает с точкой объекта, то сохраняется в массиве ключевых точек.

На шаге фильтрации точек сцены выполняется удаление ключевых точек сцены, которые принадлежали найденному объекту. Что позволяет находить похожие объекты на сцене.

Предложенный алгоритм реализован с использованием библиотеки компьютерного зрения OpenCV. В ходе эксперимента был протестирован процент ошибочно классифицированных изображений с использованием оригинального алгоритма и алгоритма

с модификацией. На рисунке 3 представлены результаты моделирования с использованием дескрипторов SURF.

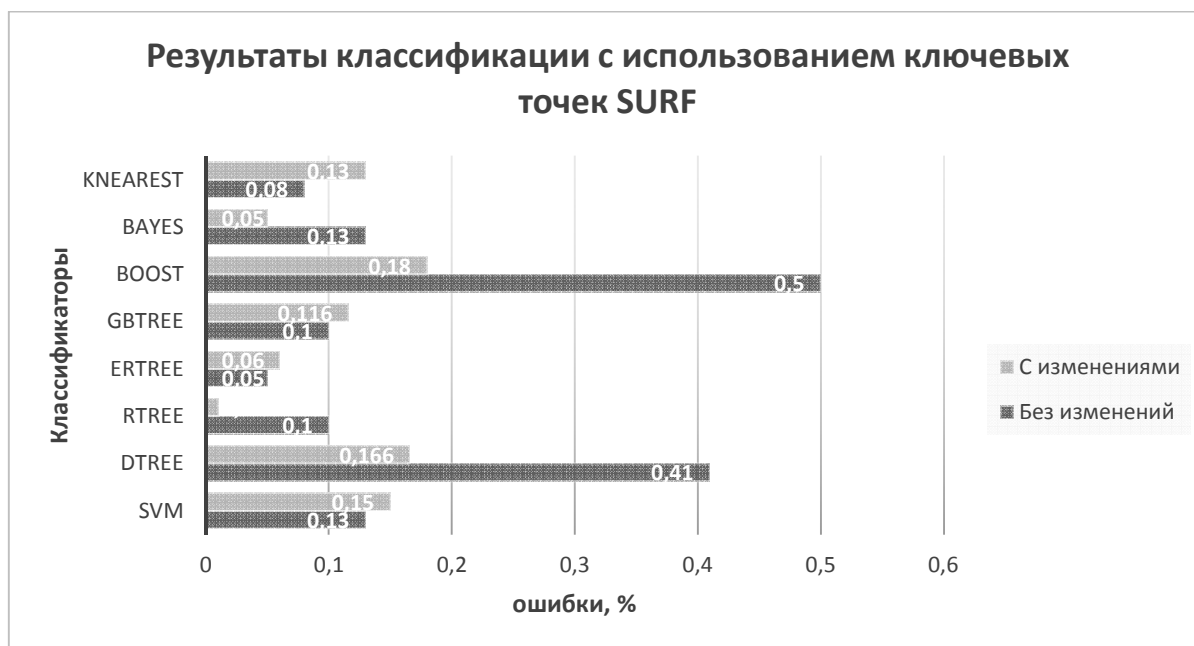


Рис. 3. Сравнительные результаты классификации с использованием ключевых точек SURF

При сравнении данных для модифицированного и оригинального алгоритмов видно, что фильтрация точек улучшила процент верно классифицированных изображений, т.е. процент ошибки уменьшился для классификаторов DTree, RTree, Boost, Bayes.

Результаты DTree улучшились с 41 % до 16.6 %. В количественном соотношении – это 250 и 100 изображений. Фильтрация, которая была предложена для модификации алгоритма, уменьшила процент ошибок на 24.4 %, т.е. 150 изображений были верно классифицированы с использованием модификации. Эти результаты показывают, что классификатор DTree чувствителен к зашумленным данным и отсеивание шума помогает улучшить результаты его работы.

Результаты RTree также улучшились с 10 % до 1.6 %. В количественном соотношении – это 60 и 10 изображений. Таким образом, фильтрация точек помогла улучшить результаты классификатора на 8.4 % или на 50 изображений. Классификатор RTree оказался менее чувствителен к данным, чем DTree.

Результаты классификатора Boost с типом бустинга REAL улучшились с 50 % до 18 %. В количественном соотношении – это 300 и 110 изображений. Это самое значительное из восьми классификаторов улучшение, полученное с использованием фильтрации. Результаты показывают, что тип бустинга REAL очень плохо работал с данными, которые не были отфильтрованы.

Результаты Bayes также улучшились с 13 % до 5 %. В количественном соотношении – это 80 и 30 изображений. Получили, что фильтрация точек помогла улучшить результаты классификатора на 8 % или на 50 изображений.

При работе алгоритма было отфильтровано 12 % ключевых точек, что позволило значительно уменьшить процент ошибок для классификаторов типа Boost и DTree.

На рисунке 4 приведены данные для классификаторов SVM и Boost.

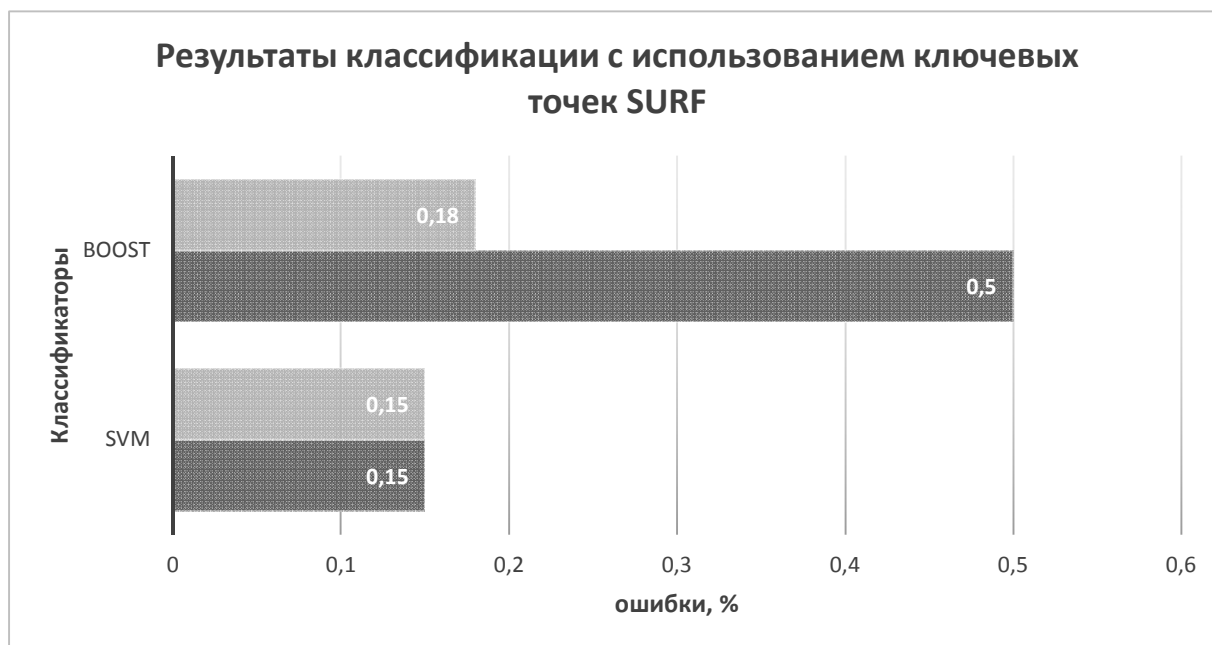


Рис. 4. Сравнительные результаты классификации с использованием ключевых точек SURF

Тип бустинга DISCRETE не приведен, т.к. процент ошибок не изменился при модификации алгоритма и составляет 3 %. Из этого можно сделать вывод, что данный тип бустинга хорошо работает с данными, которые не отфильтрованы. SVM классификатор с модификацией, как видно из рисунка 4, не изменил процент ошибок и составляет 15 %.

Заключение

В работе была выполнена реализация подхода «мешок слов». Добавлена модификация алгоритма с использованием алгоритма сегментации изображений Watershed.

В данной работе были исследованы алгоритмы классификации, а также, совместно с ними, алгоритмы сегментации изображений. При детальном изучении классификаторов выяснилось, что некоторые классификаторы: DTree (с ключевыми точками SURF) и Boost (для SIFT, SURF) с типом бустинга REAL, чувствительны к данным, которые используются при обучении. Использование фильтрации точек помогло уменьшить процент ошибок данных классификаторов на 20–30 %.

Список литературы

1. Воронцов, К.В. Лекции по машинному обучению. URL: <http://www.machinelearning.ru/> (дата обращения: 05.10.2012).
2. Конушин, А.С. Лекции по курсу Введение в компьютерное зрение. URL: <http://courses.graphics.cs.msu.ru/course/view.php?id=10> (дата обращения: 10.11.2012).
3. Конушин, А.С. Лекции по курсу Дополнительные главы компьютерного зрения. URL: <http://courses.graphics.cs.msu.ru/course/view.php?id=10> (дата обращения: 15.12.2012).
4. Viola, P., Jones, M. Rapid object detection using a boosted cascade of simple features // CVPR 2001.
5. Viola, P., Jones, M. Robust real-time face detection // IJCV 57(2), 2004.

Рецензенты:

Андреев В.В., д.т.н., профессор кафедры «Ядерные реакторы и энергетические установки», ФГБОУ ВПО «Нижегородский государственный технический университет им. Р.Е. Алексеева», г. Нижний Новгород.

Рындык А.Г., д.т.н., профессор, профессор кафедры «Информационные радиосистемы», ФГБОУ ВПО «Нижегородский государственный технический университет им. Р.Е. Алексеева», г. Нижний Новгород.