

УДК 004.931

Обучение классификатора Виолы-Джонса для локализации автомобильных номерных знаков

*Таранян А.Р., аспирант
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Информационные Системы и Телекоммуникации»*

*Научный руководитель: Девятков В.В., д.т.н., профессор
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана
bauman@bmstu.ru*

Введение

С увеличением количества автомобилей на дорогах программно-аппаратные модули их идентификации играют все большую роль в нашей жизни. Подобные модули являются ключевыми компонентами систем мониторинга трафика, контроля доступа на предприятия, автоматизации работы парковок. Наиболее часто применяемым методом идентификации автомобиля является оптическое распознавание государственного номерного знака, установленного на нем. Можно выделить следующие основные этапы оптического распознавания номерного знака на кадре видеопотока:

1. Локализация номерного знака
2. Сегментация символов на локализованном номерном знаке
3. Распознавание сегментированных символов

Одним из самых эффективных подходов для решения задачи локализации объектов на изображении является метод Виолы-Джонса [1]. С помощью данного метода во множестве работ получены очень хорошие результаты обнаружения различных сложных объектов, он имеет высокую производительность и хорошо распараллеливается.

Метод Виолы-Джонса

Рассмотрим основные идеи, применяемые в методе Виолы-Джонса:

- Использование признаков Хаара в качестве слабых классификаторов;
- Использование интегрального представления изображения;
- Использование метода бустинга для обучения сильного классификатора;

- Построение каскадного классификатора для быстрого отсека большей части областей, не являющихся искомым объектом, еще на первых уровнях каскада.

Признаки Хаара. Признак - это отображение $f: X \rightarrow D_f$, где D_f - множество допустимых значений признака. Признаки делятся на следующие типы в зависимости от множества D_f :

- бинарный признак, $D_f = \{0,1\}$;
- номинальный признак, D_f - конечное множество;
- порядковый признак, D_f - конечное упорядоченное множество;
- количественный признак, D_f - множество действительных чисел.

На рис. 1 приведено графическое представление нескольких признаков Хаара.

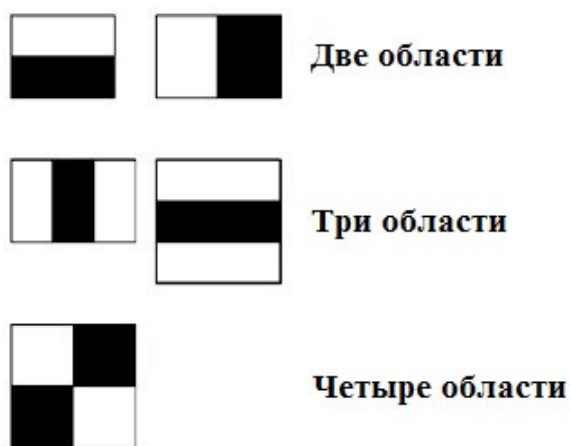


Рис. 1. Примеры признаков Хаара

Для окна фиксированного размера признак Хаара представляет собой множество прямоугольных областей белого и черного цветов. Значение признака для исследуемой области изображения вычисляется как:

$$F = W - B,$$

где W - это сумма значений пикселей, находящихся в белых областях признака, а B - сумма значений пикселей, находящихся в черных областях.

Понятно, что при вычислении значения признака Хаара, напрямую суммируя значения всех необходимых пикселей исследуемой области, будет проделываться $n * m$ операций, где n и m - это ширина и высота исследуемой области соответственно. Однако данный процесс можно значительно ускорить, если использовать так называемое интегральное представление изображения.

Интегральное представление изображения. Интегральное представление изображения является матрицей, совпадающей по размерам с исходным изображением. Значение каждой ячейки этой матрицы является суммой интенсивностей всех пикселей, находящихся левее и выше данной ячейки. Элементы матрицы вычисляются по следующей формуле:

$$I[y, x] = \sum_{i=0}^y \sum_{j=0}^x Im[i, j], \quad (1)$$

где $Im[i, j]$ - яркость пикселя исходного изображения. Каждый элемент матрицы $I[y, x]$ представляет собой сумму пикселей в прямоугольнике от $(0, 0)$ до (y, x) . Расчет интегральной матрицы производится за один проход и занимает линейное время относительно числа пикселей в изображении. Расчет возможен по следующей формуле:

$$I(y, x) = Im(y, x) - I(y-1, x-1) + I(y-1, x) + I(y, x-1). \quad (2)$$

Построив интегральную матрицу можно очень быстро вычислять сумму пикселей произвольного прямоугольника произвольной площади.

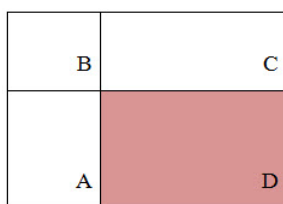


Рис. 2. Прямоугольник с областями A, B, C и D

Пусть в прямоугольнике ABCD на рис.2 есть интересующий нас объект D. Из рисунка очевидно, что сумму пикселей в прямоугольнике D можно выразить посредством суммы и разности смежных прямоугольников по следующей формуле:

$$S(D) = I(B) + I(ABCD) - I(BA) - I(BC) \quad (3)$$

Таким образом, значение любого признака Хаара можно вычислить посредством всего нескольких операций, количество которых зависит только от количества прямоугольников в признаке и не зависит от площади исследуемой области.

Обучение классификатора методом бустинга. Обучение классификатора - это процесс создания классификатора или настройки его параметров таким образом, чтобы он мог успешно классифицировать передаваемые ему объекты. В задаче локализации номерных знаков классификатор должен уметь относить объект к одному из двух классов – «номерной знак» и «не номерной знак».

Классификатор называется слабым, если результат классификации этим классификатором недостаточно хорош, но, тем не менее, он выше 50 % (т.е. классификатор показывает лучшие результаты, чем случайное угадывание).

Классификатор называется сильным, если он показывает достаточно хорошие результаты классификации.

Идея бустинга заключается в комбинировании слабых классификаторов таким образом, чтобы получить сильный классификатор. В случае метода Виолы-Джонса слабыми классификаторами являются всевозможные признаки Хаара, которые имеют размер, predetermined заранее (минимальный размер искомого объекта в кадре). Количество таких признаков Хаара очень велико, и очевидно, что при проверке значения всех признаков во время поиска объекта производительность поиска будет очень низкой. Отметим, что для области размером 24x24 пикселя количество всевозможных признаков Хаара превышает 180 тысяч [1]. С помощью алгоритма бустинга, применяемого в методе Виолы-Джонса, из всего множества признаков выбираются те, которые лучшим образом подходят для обнаружения искомого объекта.

В данной работе проводится сравнение эффективности использования следующих четырех алгоритмов бустинга при обучении классификатора для обнаружения номерных знаков: *Discrete AdaBoost* [2], *Real AdaBoost* [3], *LogitBoost* [4] и *Gentle AdaBoost* [4].

Каскадная архитектура классификатора. Во многих задачах локализации количество всевозможных объектов на анализируемом изображении значительно превышает количество искомым объектов. Для ускорения обработки изображения в подобных задачах в алгоритме Виолы-Джонса применена идея каскада классификаторов. Каждый уровень этого каскада представляет собой классификатор, называемый ансамблем, который состоит из нескольких признаков Хаара, причем требуется, чтобы ансамбли всех уровней имели очень хороший показатель обнаружения объекта. Однако допускается, чтобы ансамбли имели достаточно большой процент ложных срабатываний (обычно до 50 %). Анализируемый объект поочередно проверяется каждым уровнем каскада, и если на каком-либо уровне ансамбль определяет, что объект не является

искомым, то процедура заканчивается и объект помечается как негативный. Если же классификаторы всех уровней определяют этот объект как положительный, то объект считается положительным. Ансамбли первых уровней целесообразно составлять из малого количества слабых классификаторов, тогда так ансамбли последних уровней могут быть достаточно сложными и могут содержать десятки и сотни слабых классификаторов для отсекаания особо сложных негативных объектов. Таким образом, большая часть негативных объектов отсекается уже на первых уровнях каскада, и для них высчитываются значения лишь небольшого количества признаков Хаара.

Обучение классификатора Виолы-Джонса

Для обучения классификатора Виолы-Джонса было построено обучающее множество из изображений номерных знаков и изображений, не содержащих номерных знаков.

Для построения позитивного множества изображений (являющихся номерными знаками) было отобрано 475 изображений номерных знаков под разными углами наклона, при разных условиях освещения и разной загрязненности. Эти изображения были обрезаны таким образом, чтобы содержать минимальное количество пикселей за пределами номерного знака (рис. 3).



Рис. 3. Обрезание изображения номерного знака

Далее, для обрезанных изображений был вычислен угол наклона, и было оценено качество изображения номерного знака (освещенность, контраст, загрязнение). После этого для каждого изображения путем поворота, изменения яркости и добавления

небольших шумов было получено несколько изображений (угол наклона номерных знаков в результирующих элементах в плоскости изображения находился между -4 и 4 градусами), которые были уменьшены до размера 65x14 пикселей (рис. 4).

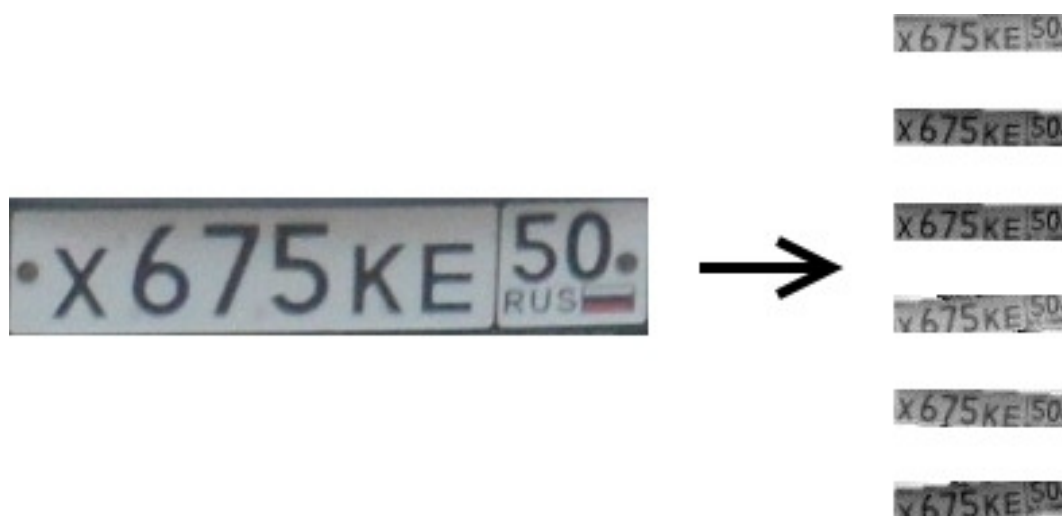


Рис. 4. Формирование нескольких позитивных изображений из одного

Таким образом, было получено 2000 позитивных элементов обучающей выборки.

Большую часть негативных элементов обучающей выборки составили объекты, которые наиболее часто встречаются при поиске номерного знака – это машины, дороги, деревья, люди, здания и т. д. (рис. 5). Крайне важно, чтобы элементы негативной выборки не содержали номерных знаков.



Рис. 5. Примеры негативных изображений

Необходимо отметить, что из каждого негативного изображения получается большое количество негативных элементов путем выбора на нем всевозможных прямоугольных областей с соотношением сторон 65:14. Всего было отобрано 3912 негативных изображений.

Обучение классификатора производилось с помощью модуля обучения классификатора Виолы-Джонса в библиотеке *OpenCV* (*opencv_traincascade.exe*) со следующими параметрами:

- *-dataplates.xml* (имя выходного обученного классификатора)
- *-vecsamples.vec* (файл с позитивными изображениями, созданный с помощью модуля *opencv_createsamples.exe*)
- *-bgnegative.txt* (файл со списком негативных изображений)
- *-numPos 1740* (количество позитивных изображений, умноженное на определенный коэффициент меньше единицы ($2000 * 0,87$))
- *-numNeg 3402* (количество негативных изображений, умноженное на определенный коэффициент меньше единицы ($3912 * 0,87$))

- -numStages 20 (максимальное количество уровней в обученном каскадном классификаторе)
- -precalcValBufSize 256 (размер буфера значений предварительно вычисленных признаков в мегабайтах)
- -precalcIdxBufSize 256 (размер буфера индексов предварительно вычисленных признаков в мегабайтах)
- -featureType HAAR (тип используемых признаков)
- -mode BASIC (тип используемых признаков Хаара)
- -w 65 -h 14 (ширина и высота изображений обучающей выборки)
- -minHitRate 0.99942 (минимальная доля правильно классифицированных позитивных объектов обучающей выборки для каждого уровня каскада (ансамбля))
- -maxFalseAlarmRate 0.5 (максимально доля ложно-позитивных срабатываний для негативных объектов обучающей выборки для каждого уровня каскада (ансамбля))
- -weightTrimRate 0.95
- -bt DAB (используемый алгоритм бустинга, где DAB - *Discrete AdaBoost*, GAB - *Gentle AdaBoost*, RAB - *Real AdaBoost*, LB - *LogitBoost*)

На рис. 6 приведено время обучения классификатора при использовании четырех упомянутых алгоритмов бустинга на компьютере с твердотельным накопителем и процессором *Intel Core i5* с тактовой частотой 3,4 ГГц.

Таким образом, каждым методом бустинга было обучено по одному классификатору.

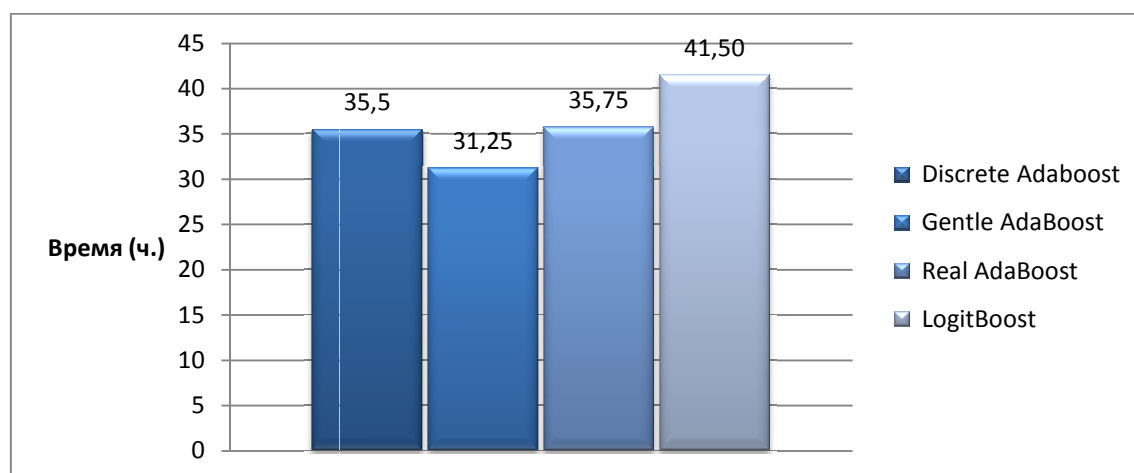


Рис. 6. Время обучения. Сравнение алгоритмов бустинга

Тестирование обученного классификатора Виолы-Джонса

Для тестирования обученных классификаторов было разработано приложение на языке C++ с использованием фреймворка QT и библиотеки OpenCV. Локализация

номерных знаков в кадре видеопотока производилась с помощью метода detectMultiScale объекта CascadeClassifier (из библиотеки OpenCV), который предварительно инициализировался одним из четырех обученных классификаторов.

Тестирование проводилось для трех разрешений видеофайлов: 360p (640x360 пикселей), 480p (480x852 пикселя) 720p (720x1280 пикселей) на компьютере с процессором *Intel Core i5* с тактовой частотой 3,4 ГГц (были задействованы два ядра процессора).

На рис. 7 представлена производительность тестируемых классификаторов.

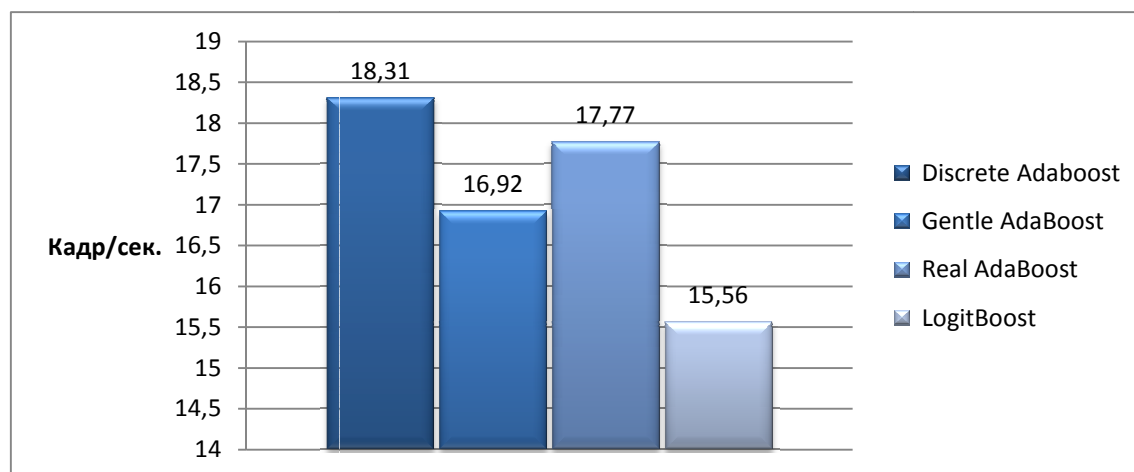


Рис. 7. Количество обработанных кадров за секунду

Для тестирования качества локализации классификаторов была подготовлена тестовая выборка из 800 изображений. Так как все классификаторы обучались для локализации номерных знаков с наклоном в плоскости изображения до 4,0 градусов, номерные знаки во всех тестовых изображениях имели угол наклона от -4,0 до 4,0 градусов в плоскости изображения, и от -20 до 20 градусов в плоскостях, перпендикулярных плоскости изображения. Тестовая выборка содержала снимки транспортных средств в различных погодных условиях, различных условиях освещения, с загрязненными и чистыми номерными знаками. На рис. 8 показан результат тестирования четырех классификаторов на подготовленной тестовой выборке.

Также, для всех четырех классификаторов было проанализировано количество ложно-позитивных срабатываний. На рис. 9 показано количество ложно-позитивных обнаружений на 4.32400×10^8 изображениях размера 65x14 пикселей.

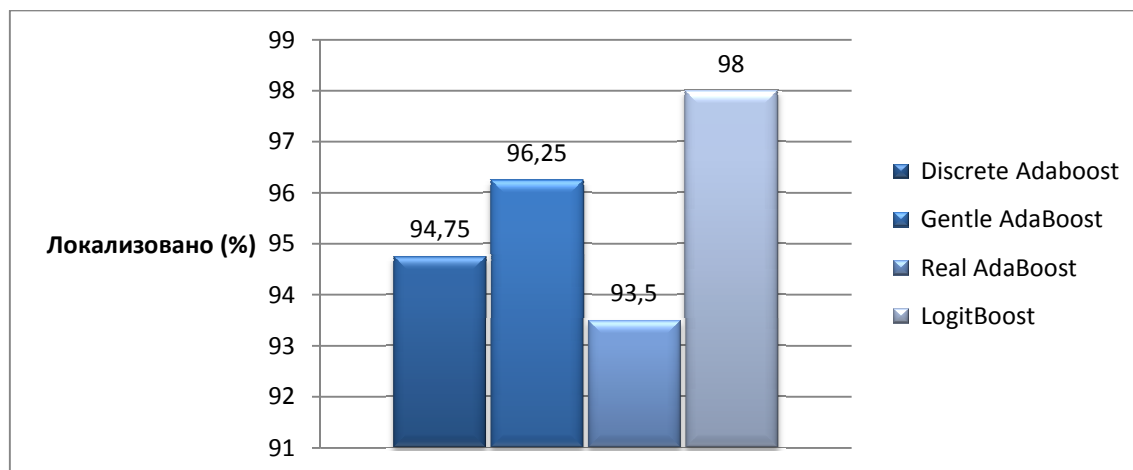


Рис. 8. Качество локализации

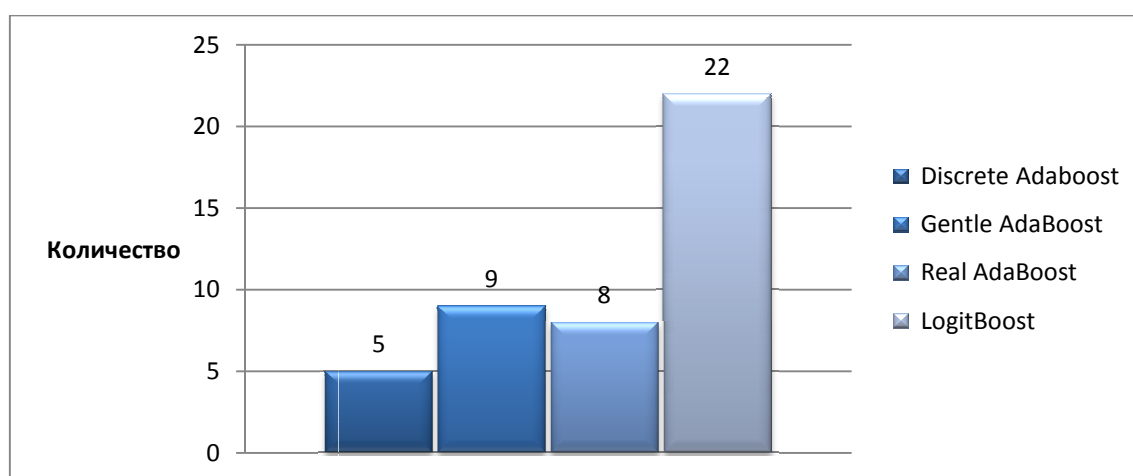


Рис. 9. Количество ложно-положительных срабатываний. Сравнение алгоритмов бустинга

Результаты тестирования показали, что классификатор, обученный с применением алгоритма бустинга *LogitBoost*, показывает лучшее качество обнаружения, однако имеет значительно большее количество ложно-положительных срабатываний и демонстрирует ощутимо меньшую производительность. На основании проведенного исследования автор считает целесообразным использование алгоритма *Gentle AdaBoost*, который позволяет получить оптимальное соотношение качества локализации, количества ложных срабатываний и производительности.

Список литературы

1. Viola P., Jones M. J. Robust real-time face detection // International Journal of Computer Vision. Vol. 57. 2004. Pp. 137-154.

2. Schapire R. E., Freund Y. A decision-theoretic generalization of on-line learning and an application to boosting // Journal of Computer and System Sciences. Vol. 55. 1997. Pp. 119–139.
3. Schapire R. E., Singer Y. Improved Boosting Algorithms Using Confidence-rated Predictions // Machine Learning. Vol. 37. 1999. Pp. 297-336.
4. Friedman J., Hastie T., Tibshirani R. Additive logistic regression: a statistical view of boosting // The Annals of Statistics. Vol. 28. 2000. Pp. 337-407.