Содержание

[Введение 2](#_Toc419043635)

[1. Постановка задачи 3](#_Toc419043636)

[1.1. Формат входных и выходных данных 3](#_Toc419043637)

[1.2. Методы оценки качества детекции 4](#_Toc419043638)

[2. Общий подход к решению задачи 5](#_Toc419043639)

[2.1. Гистограмма ориентированных градиентов (HOG) 5](#_Toc419043640)

[2.2. Модель опорных векторов (SVM) 5](#_Toc419043641)

[3. Используемые инструменты 5](#_Toc419043642)

[4. Общий подход к реализации 5](#_Toc419043643)

[5. Результаты тестирования 5](#_Toc419043644)

[Заключение 5](#_Toc419043645)

[Список используемых источников 6](#_Toc419043646)

[Приложение А. Результаты работы детектора. 6](#_Toc419043647)

[Приложение Б. Исходный код 6](#_Toc419043648)

# Введение

Задача распознавания образов, в том или ином виде, встречается в различных областях человеческой деятельности – от военного дела и систем безопасности до оцифровки аналоговых сигналов. С развитием вычислительных мощностей появляются все новые и новые задачи, в качестве подзадач включающие в себя детекцию объектов. Широкое применение такие алгоритмы получили в области дорожного движения. В частности, системы распознавания пешеходов могут применятся для анализа загруженности городских улиц, контроля за соблюдением правил дорожного движения и повышения безопасности на дорогах. Особенное значение такого рода системы имеют при разработке беспилотных транспортных средств. Основной их целью является предотвращение столкновения автомобиля с человеком.

В области распознавания пешеходов проводятся многочисленные исследования. Основными проблемами в ходе решения задачи – выбор классификаторов и способа формирования векторов признаков. На данный момент существует целое множество решений этой задачи [1][2].

Однако, задача распознавания пешеходов – не единственная задача распознавания, с которой сталкиваются разработчики беспилотных автомобилей. В рамках масштабного проекта по созданию к 2020 году беспилотного автомобиля компанией Cognitive Technologies был объявлен конкурс на разработку программного обеспечения детекции транспортных средств движущихся по дорогам, на изображениях, полученных с движущейся видео-камеры (видео регистраторa). Целью данной работы является разработка алгоритма, решающего данную задачу, основанного на уже известных подходах к детекции пешеходов.

# Постановка задачи

В рамках решения задачи предполагалась разработка алгоритма, позволяющего определить факт наличия и расположение автомобиля на фото с видео регистратора. Для заданного изображения необходимо получить список четырехугольников, описывающих изображенные на нем транспортные средства и степень уверенности в полученных координатах.

## Формат входных и выходных данных

Для решения задачи были предоставлены два вида файлов: данные для обучения детектора и тестовая выборка.

Данные для обучения детектора были представлены в виде единого архива, внутри которого были расположены следующие директории:

* image - цветные изображения размера 480x270;
* label - разметка изображений на предмет наличия.

Файлы разметки именовались по правилу <image-name>.txt и содержали следующую информацию:

* Флаг о полностью размеченном кадре (true или false)
* Список ограничивающих квадратов автомобилей в формате x y width height (четыре числа, разделенные пробелом, задающие левый верхний угол и размеры квадрата; начало отсчета – левый верхний угол)

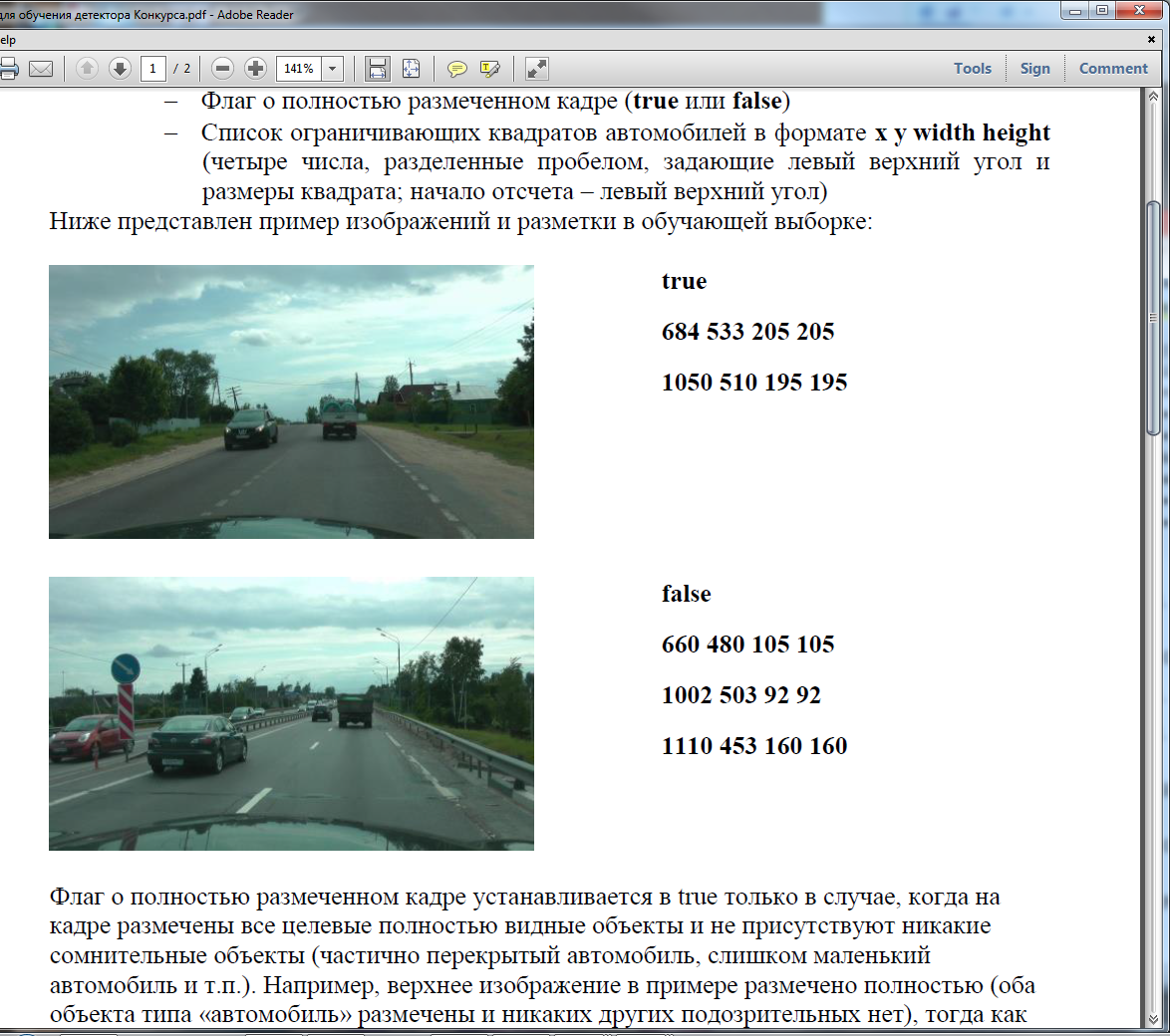


Рисунок 1.1 – Пример изображений и разметки в обучающей выборке

Флаг о полностью размеченном кадре устанавливался в true только в случае, когда на кадре были размечены все целевые полностью видные объекты и не присутствовали никакие сомнительные объекты (частично перекрытый автомобиль, слишком маленький автомобиль и т.п.). Например, верхнее изображение в примере размечено полностью (оба объекта типа «автомобиль» размечены и никаких других подозрительных нет), тогда как нижнее изображение не полностью размечено (присутствую несколько частично перекрытых автомобилей на встречной полосе).

В качестве результата работы программа должна была для каждого изображения создавать текстовый файл с результатами детектирования. Файл результатов должен был именоваться по правилам <image-name>.txt и содержать список ограничивающих квадратов найденных на изображении автомобилей в следующем формате:

x1 y1 width1 height1 score1

x2 y2 width2 height2 score2

…

xn yn widthn heightn scoren

Здесь x y width height – четыре числа, разделенные пробелом, задающие левый верхний угол и размеры квадрата, а score – число от 0 до 255, указывающее степень уверенности в найденном объекте (чем больше – тем более надежным считается найденный объект).

## Методы оценки качества детекции

Для оценки качества работы детектора была использована кривая точность-полнота (precision-recall curve), построенная на основе степени уверенности.

Для оценки того, что объект найден правильно, использовалась мера Жаккара с коэффициентом 0.5. Иными словами, пусть у нас есть идеальное расположение автомобиля на карте *Ideal*, а также результат детектирования *Real.* Тогда считаем, что результат детектирования верный, если:

Далее введем несколько понятий. Общее количество объектов в разметке будем называть «количество положительных объектов» и обозначать P (Positive). Правильно детектированные объекты будем называть «точное срабатывание» и обозначать TP (True Positive). Неправильно детектированные объекты будем называть «ложными срабатываниями» и обозначать FP (False Positive).

Таким образом, полнота (True Positive Rate, TPR) и точность (Positive Predictive Value, PPV) выражается по следующим формулам:

Для общего анализа работы детектора вычислялась F- метрика, которая представляет собой среднее гармоническое величин TPR и PPV:

# Общий подход к решению задачи

Для решения поставленной задачи было решено использовать HOG-дескрипторы. Для классификации был использован метод опорных векторов.

## Гистограмма ориентированных градиентов (HOG)

Гистограммы ориентированных градиентов - один из способа построения дескрипторов особых точек. Данный подход основан на подсчете количества направлений градиента в локальных областях изображения и отличается тем, что вычисляется на плотной сетке равномерно распределенных ячеек и использует нормализацию перекрывающегося локального контраста для увеличения точности.

Основной идеей алгоритма является допущение, что внешний вид и форма объекта на участке изображения могут быть описаны распределением градиентов интенсивности или направлением краев.

При описании фрагмента изображения оно разбивается на несколько небольших участков, называемых ячейками. В ячейках вычисляются гистограммы направленных градиентов внутренних точек. Обычно они объединяются в одну гистограмму ℎ = 𝑓(ℎ1, ..., ℎ𝑘), после чего она нормализуется по яркости. Таким образом достигается инвариантность к изменениям освещения. Для аппроксимации градиента изображения могут быть использованы различные дискретные дифференциальные операторы, такие как оператор Собеля, оператор Превитта или оператор Щарра.

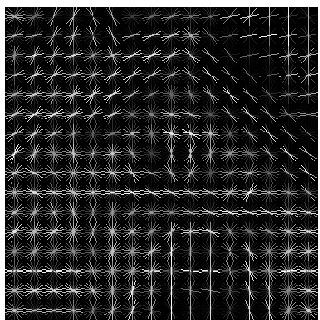


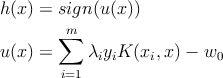
Рисунок 2.1 – Гистограмма ориентированных градиентов

* 1. Метод опорных векторов (SVM)

Данный метод является одним из решений задачи классификации, которая состоит в определении к какому классу из, как минимум, двух изначально известных относится данный объект. При таком подходе обычно предполагается, что выборка линейно разделима (задача бинарной классификации). Метод основан на построении гиперплоскости, разделяющей исходную выборку на два класса. Так как существует бесконечно много плоскостей, вводится требования, чтобы она располагалась максимально далеко от ближайших к ней точек обоих классов.

Перед началом работы, классификатор необходимо обучить на выборке, разделенной на «положительные» и «отрицательные» примеры.

Классификатор описывается следующими соотношениями



где – порог (свободный член); – коэффициент; – пара из учебного набора, (векторы , для которых > 0 называются опорными); K ( , x ) – функция ядра

Машина опорных векторов отображает входной вектор в пространство более высокой размерности и находит разделяющую гиперплоскость с максимальным зазором между точками разных классов. Отображение это осуществляется с помощью функции ядра.

Изначально SVM это линейный классификатор, т.е. он может решать только линейно разделимые задачи. Однако, применяя нелинейное ядро, можно отобразить исходные данные в пространство большей размерности, где уже может существовать оптимальная разделяющая гиперплоскость.

Обучение SVM сводится к нахождению коэффициентов λ (множителей Лангранжа) и .

Для классификации объекта достаточно определить его положение относительно гиперплоскости.

# Используемые инструменты

Существует большое количество библиотек, реализующих основные алгоритмы компьютерного зрения. Для решения данной задачи была выбрана открытая библиотека VLFeat, качественно реализующая весь необходимый инструментарий, а также предоставляющая возможность многопоточной обработки данных с использованием OpenMP. На официальном сайте [3] представлена подробная документация.

# Общий подход к реализации

# Результаты тестирования

# Заключение

# Список используемых источников

1. Dollar P. et al. Pedestrian detection: An evaluation of the state of the art //Pattern Analysis and Machine Intelligence, IEEE Transactions on. – 2012. – Т. 34. – №. 4. – С. 743-761
2. S. Walk, N. Majer, K. Schindler, and B. Schiele, “New features and insights for pedestrian detection,” in IEEE Conf. Computer Vision and Pattern Recognition, 2010. – C. 1030-1037
3. VLFeat.org. [Электронный ресурс]: - Режим доступа: http://www.vlfeat.org/

# Приложение А. Результаты работы детектора.

# Приложение Б. Исходный код