

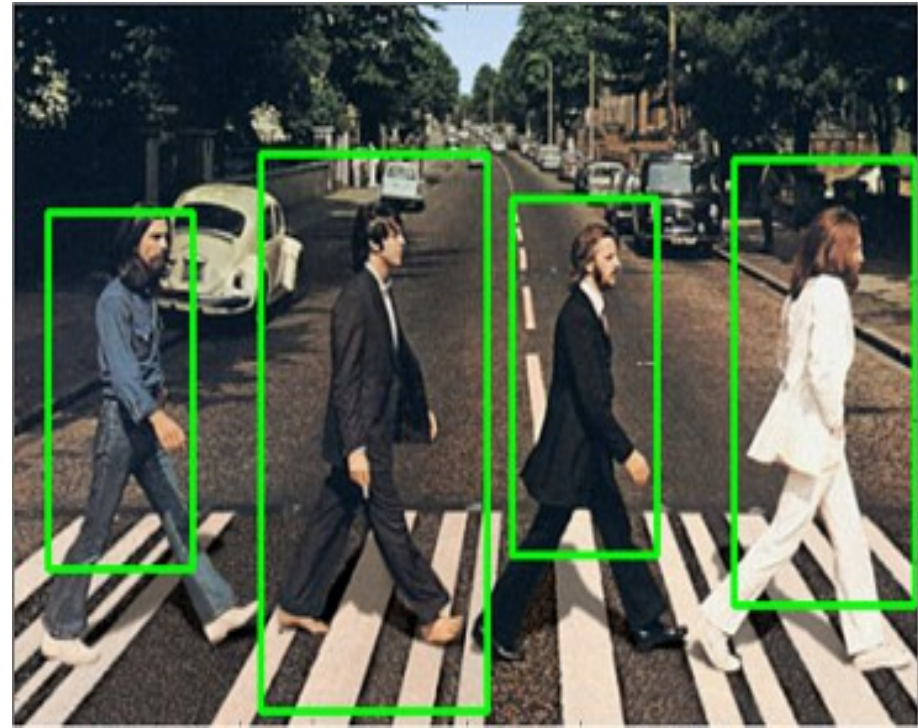
СПбПУ Петра Великого
Кафедра «Прикладная математика»

Распознавание объектов на изображении

Студент группы 43601/2 Дамир Шагаев

Введение

Имеется некоторое входное изображение.
Требуется найти области, в которых находятся интересующие объекты. Например, интересующий объект - человек.

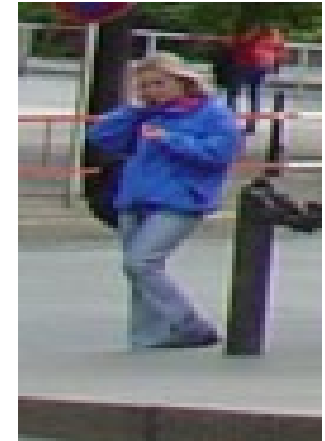


Постановка задачи

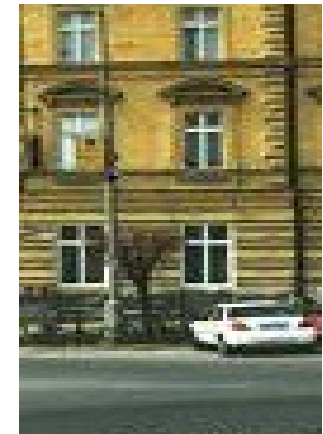
Имеется два набора тренировочных коллекций растровых размерности 64 x 128 изображений людей в полный рост:

1. Набор №1 - изображения, на которых присутствуют люди
2. Набор №2 - изображения, на которых отсутствуют люди

Имеется некоторое новое входное изображение, на котором требуется выделить человека.



Набор №1



Набор №2



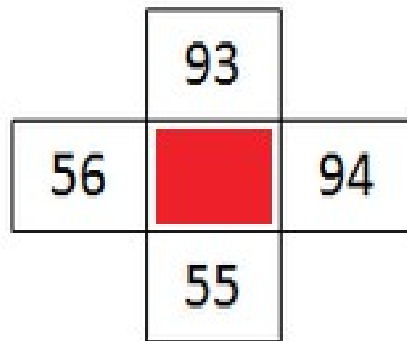
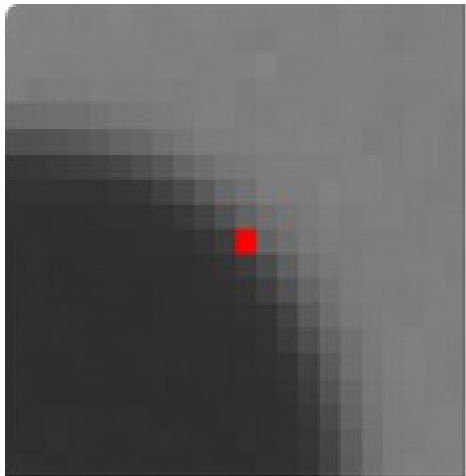
Входное изображение

HOG-дескриптор

Гистограмма направленных градиентов (HOG-дескриптор) является преобразованием изображения в многомерный вектор, таким что сопоставляет изображениям схожих объектов близкие по значению вектора.

HOG-дескриптор из исходного изображения вычисляется следующим образом:

1. Нормализация цвета и гамма коррекция. Обычно изображение приводится к полутоновому.
2. Вычисление значений градиентов по горизонтали и по вертикали. Градиент — это мера изменения значений пикселей вдоль горизонтали и вертикали для каждого пикселя. Градиент вычисляется для каждого пикселя на изображении.



$$\text{grad } f = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} 94 - 56 \\ 93 - 55 \end{pmatrix} = \begin{pmatrix} 38 \\ 38 \end{pmatrix}$$

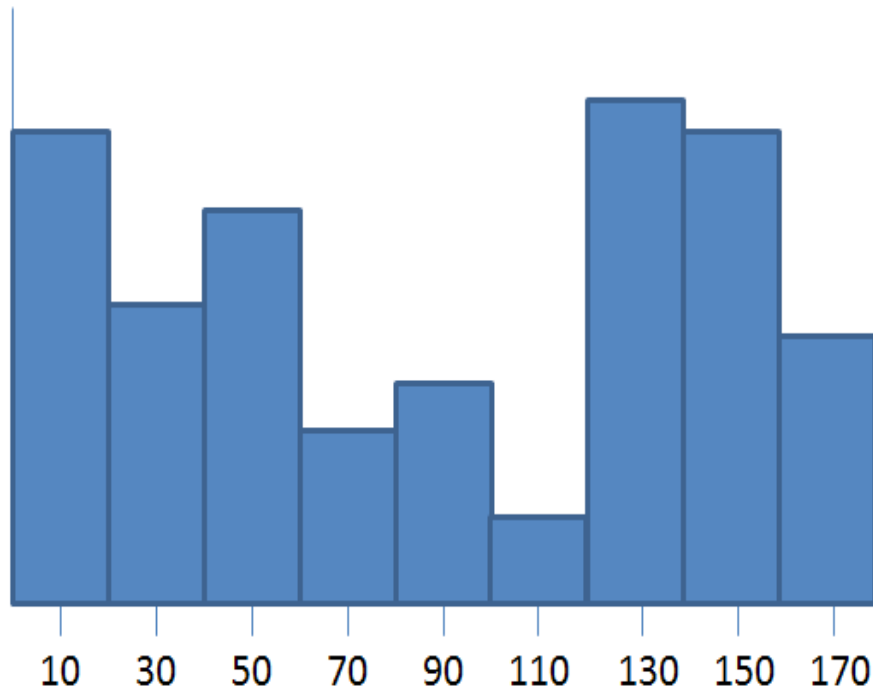
$$|\text{grad } f| = \sqrt{f_x^2 + f_y^2} = 53.74$$

$$\theta = \arctan \frac{f_y}{f_x} = 45$$

HOG-дескриптор

3. Изображение условно разбивается на ячейки(cells) размерности, например, 8 x 8 пикселей. Внутри каждой ячейки для каждого пикселя вычисляются градиент, значение градиента(норма градиента) и направление градиента по формулам из пункта 2.

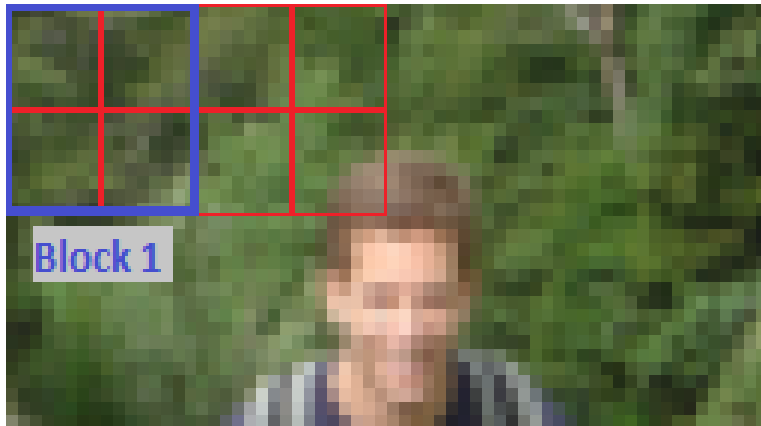
4. На предыдущем этапе получены 64 градиента. Составляем из значения градиента, направления градиента гистограмму из 9 бинов. По горизонтали шкала градусов направлений градиентов, где на один бин приходится 20 градусов, по вертикали — значения градиентов.



При построении гистограммы происходит объединение между ближайшими бинами. Например, если градиент имеет направление 85 градусов, то $1/4$ его значения добавляется к бину с центром в точке 70 и $3/4$ его значения к бину с центром в точке 90.

HOG-дескриптор

5. Объединение ячеек в блоки(blocks). Каждые четыре ячейки изображения объединяются в блоки, которые могут накладываться друг на друга(overlap of blocks). Например, блок размерности 16 x 16 пикселей состоит из 4 ячеек размерности 8 x 8 пикселей.



6. Блок 16 x 16 пикселей состоит из 4 ячеек, которые представимы в виде гистограммы, характеризуемой набором 9 значений бинов.

Итог: каждый блок представим в виде набора 36 значений гистограмм, т.е. получаем вектор из 36 значений.

Вычислим евклидову норму этого вектора и поделим каждое значение этого вектора на значение нормы. Получим нормализованный вектор, а также нормализованные гистограммы. Нормализация проводится с целью получения инвариантного к изменениям освещения дескриптора.

НОG-дескриптор

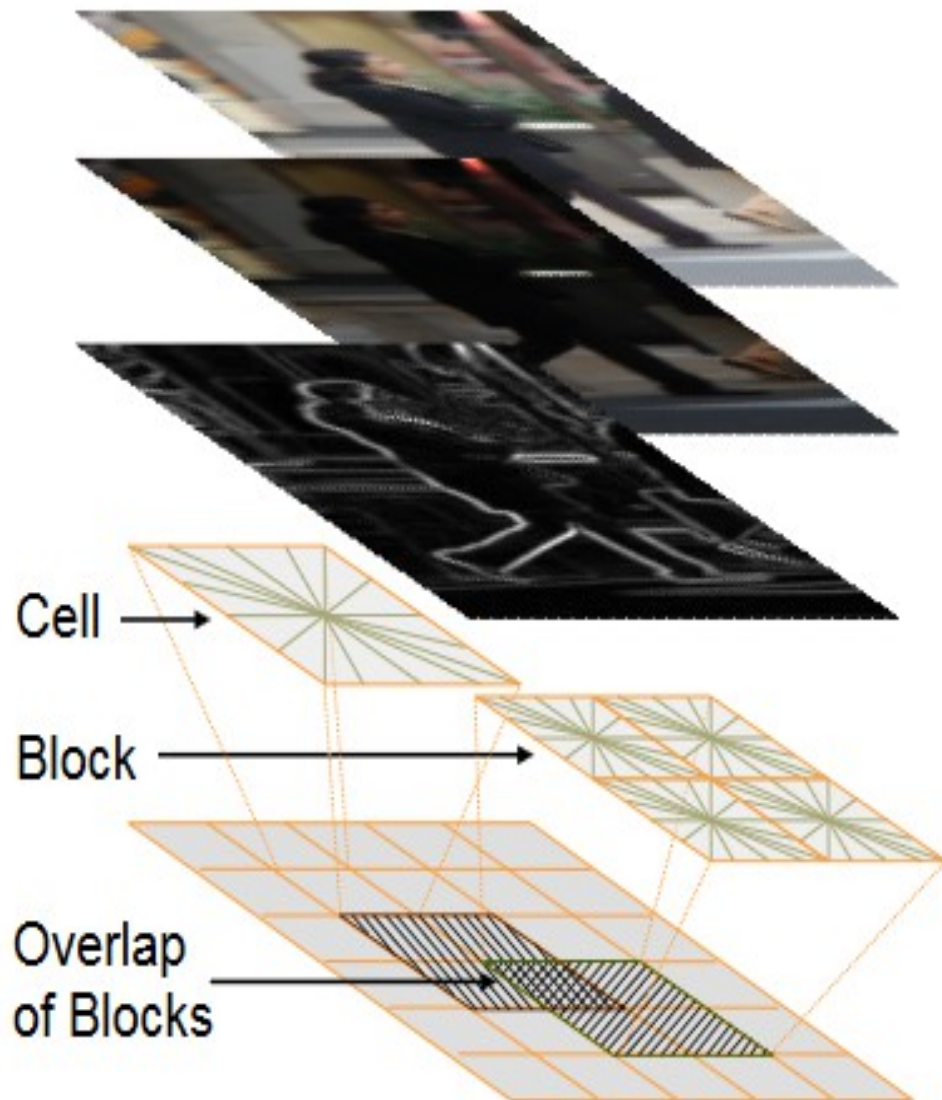
7. Изображение размерности 64×128 пикселя состоит из 7 блоков по горизонтали и 15 блоков по вертикали.

Всего 105 блоков размерности 16×16 пикселей.

В каждом блоке 4 ячейки, каждая ячейка состоит из 9 бинов гистограммы.

Итоговый вектор состоит из 3,780 значений.

НОG-дескриптор получен.



Этап построения классификатора

Задача классификации: по входному объекту, который представлен n -мерным вектором, определить, к какому классу он принадлежит.

Вход: два набора тренировочных коллекций

1. Вычисление HOG-дескрипторов для нормализованных изображений каждой тренировочной коллекции.
2. Расстановка меток для каждой тренировочной коллекции.
3. В случае SVM-классификатора, при обучении по тренировочной коллекции строится гиперплоскость, разделяющая пространство объектов на классы. Она задается следующим уравнением: $\langle w, x \rangle = b$, где w - перпендикуляр к разделяющей гиперплоскости, b - скалярный порог;
4. Построение классификатора.

Выход: приближенная функция, которая сопоставляет произвольному объекту класс, которому этот объект принадлежит.

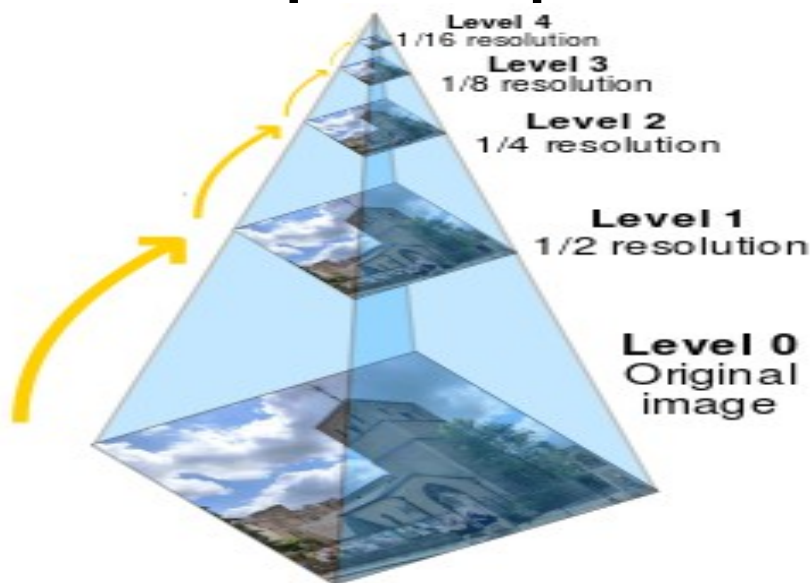
Input image



Histogram of Oriented Gradients

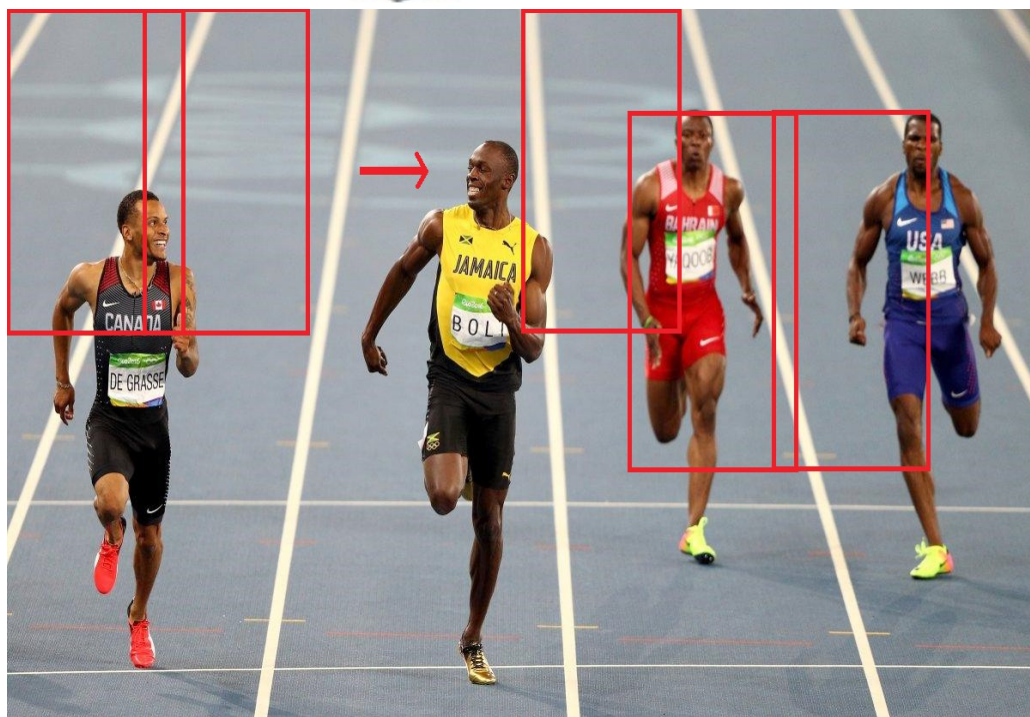


Алгоритм распознавания объектов



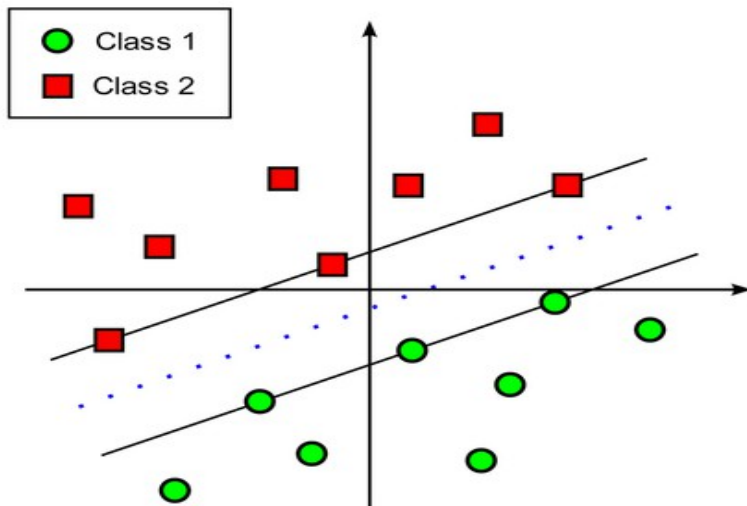
Вход: некоторое новое изображение

1. Масштабирование входного изображения. Результатом является набор изображений разной размерности. Данный шаг обусловлен тем, что на изображении объекты могут иметь разный размер.



2. Метод скользящего окна: по масштабированному изображению происходит перемещение окна, имеющего размерность изображений из тренировочной коллекции. В каждом новом положении окна вычисляется значение HOG-дескриптора и запоминаются его координаты. Каждое изображение, имеющее размерность скользящего окна, представляется в виде n — мерного вектора.

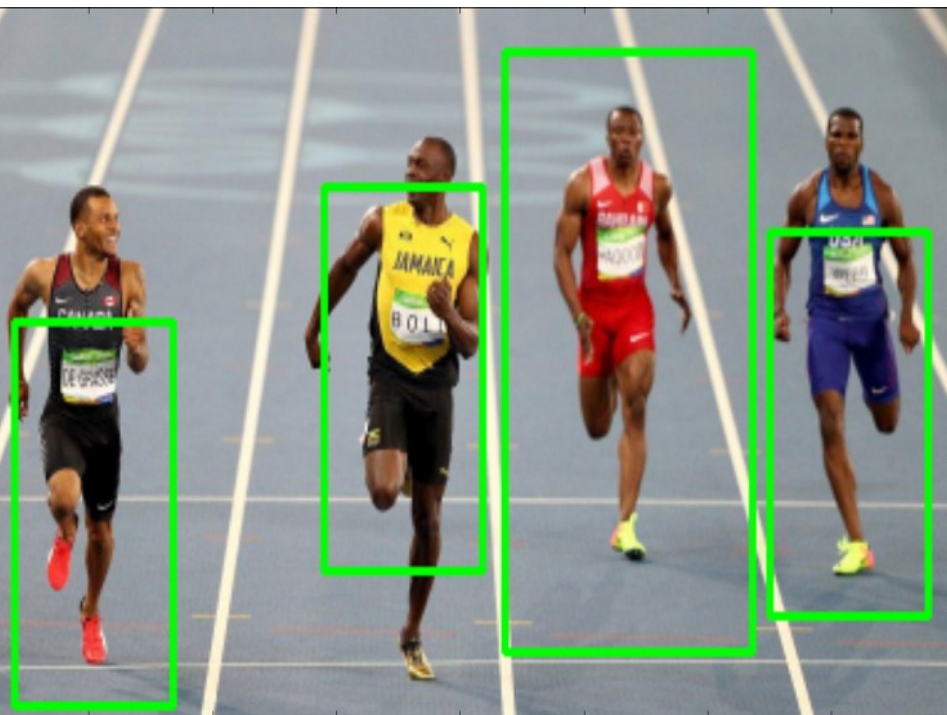
Алгоритм распознавания объектов



3. Полученные на предыдущем этапе значения HOG-дескриптора подаются на вход классификатору.

4. Классификатор сопоставляет каждому HOG-дескриптору некоторое число — метку класса, которому соответствует изображение, полученное на втором шаге алгоритма.

Выход: координаты скользящего окна, где были обнаружены интересующие объекты.



Список литературы

- [1] N. Dalal, B. Triggs. Histograms of oriented gradients for human detection, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.
- [2] Kishor B.Bhangale, R.U.Shekokar. Human Body Detection Using Histogram of Oriented Gradients and SVM. International Journal of Electronics Communication and Computer Engineering, Volume5, Issue (4), 2014.
- [3] Joe Minichino, Joseph Howse. Learning OpenCV 3 Computer Vision with Python - Second Edition, Packt Publishing Ltd, 2015, p. 356