

Базовые принципы в Computer Vision

Problems outlook. Image descriptors, HOG Filters.

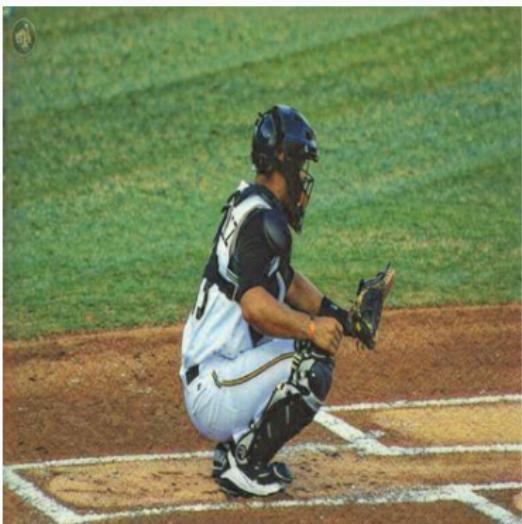
Дмитрий Котовенко

Center for modern beamer themes

Проблемы компьютерного зрения

Проблема представления.

Как представлены изображения.



| | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|
| 71 | 91 | 63 | 79 | 31 | 01 | 92 | 21 | 43 | 69 |
| 12 | 80 | 65 | 74 | 36 | 84 | 54 | 89 | 52 | 11 |
| 73 | 93 | 90 | 81 | 33 | 03 | 44 | 23 | 45 | 68 |
| 04 | 34 | 67 | 70 | 22 | 48 | 56 | 02 | 32 | 13 |
| 75 | 95 | 46 | 83 | 35 | 05 | 82 | 25 | 47 | 57 |
| 94 | 06 | 98 | 24 | 14 | 64 | 58 | 53 | 72 | 15 |
| 77 | 97 | 26 | 85 | 37 | 07 | 16 | 27 | 49 | 61 |
| 28 | 50 | 00 | 30 | 20 | 62 | 60 | 55 | 66 | 17 |
| 10 | 40 | 59 | 87 | 39 | 09 | 96 | 29 | 51 | 41 |
| 08 | 18 | 42 | 38 | 76 | 99 | 88 | 78 | 86 | 19 |

Задачи. Классификация.

- MNIST/SVHN/FashionMNIST
- CIFAR-10/CIFAR-100
- ImageNET ILSVRC/Places



Задачи. Детекция.

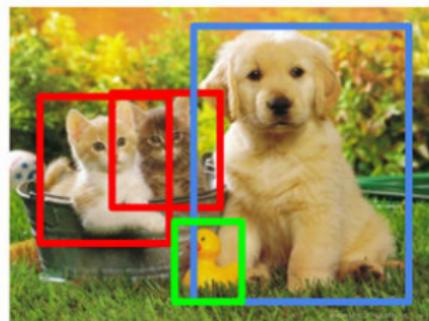
- COCO
- ImageNet Object Detection Challenge

Classification



CAT

Object Detection



CAT, DOG, DUCK

Задачи. Сегментация.

- Various Medical Images
- Aerial images
- ImageNet Object Detection Challenge
- COCO
- Cityscapes

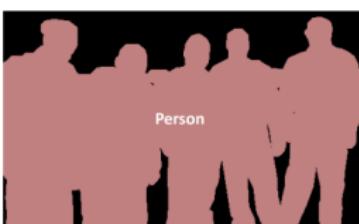


Задачи. Instance Segmentation.

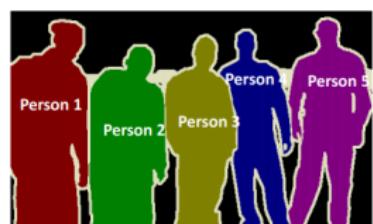
- Various Medical Images
- COCO



Object Detection



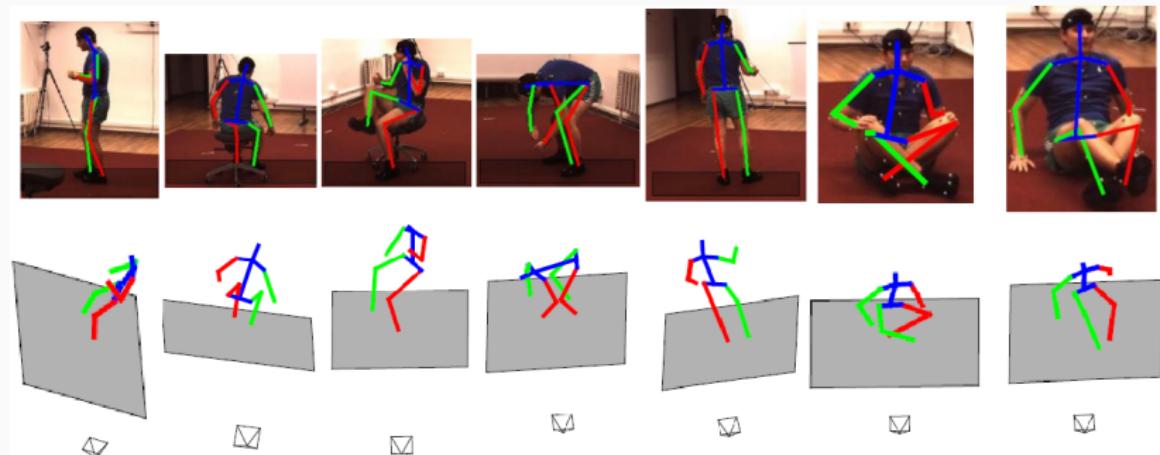
Semantic Segmentation



Instance Segmentation

Задачи. Pose estimation.

- Сложности с оценкой качества
- 2D vs 3D
- Нет абсолютно точного GT
- Как оценивать невидимые части



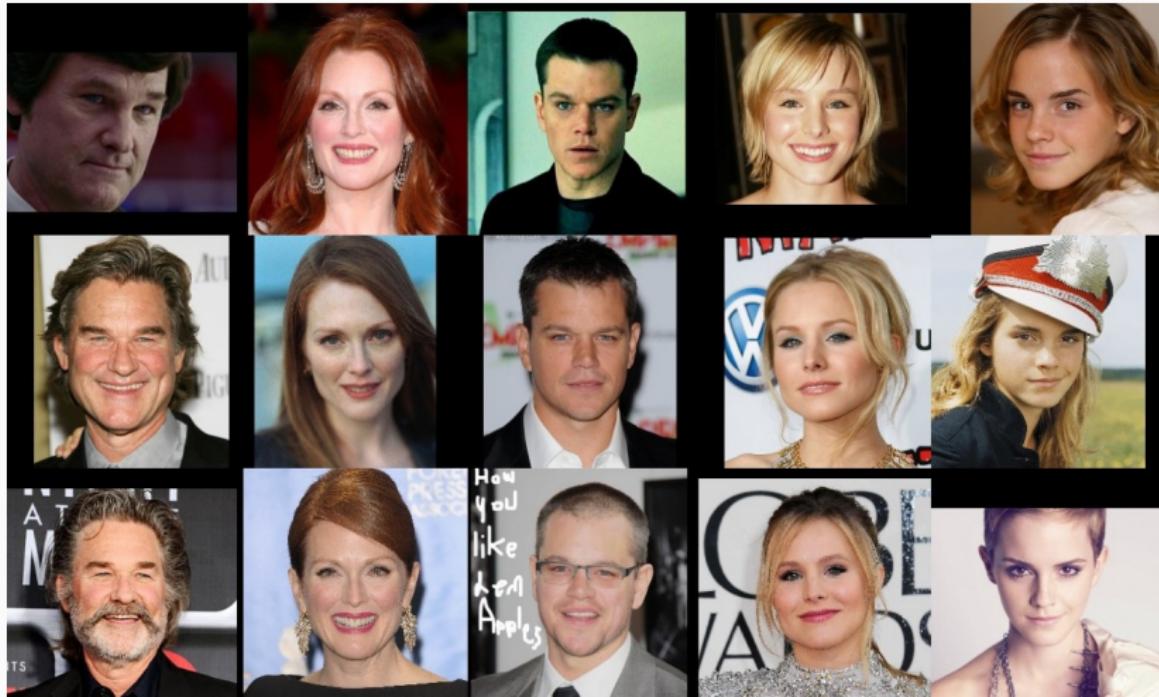
Задачи. Face recognition.

- Сложности с оценкой качества
- 2D vs 3D
- Нет абсолютно точного GT
- Решённая задача. Никакого DL



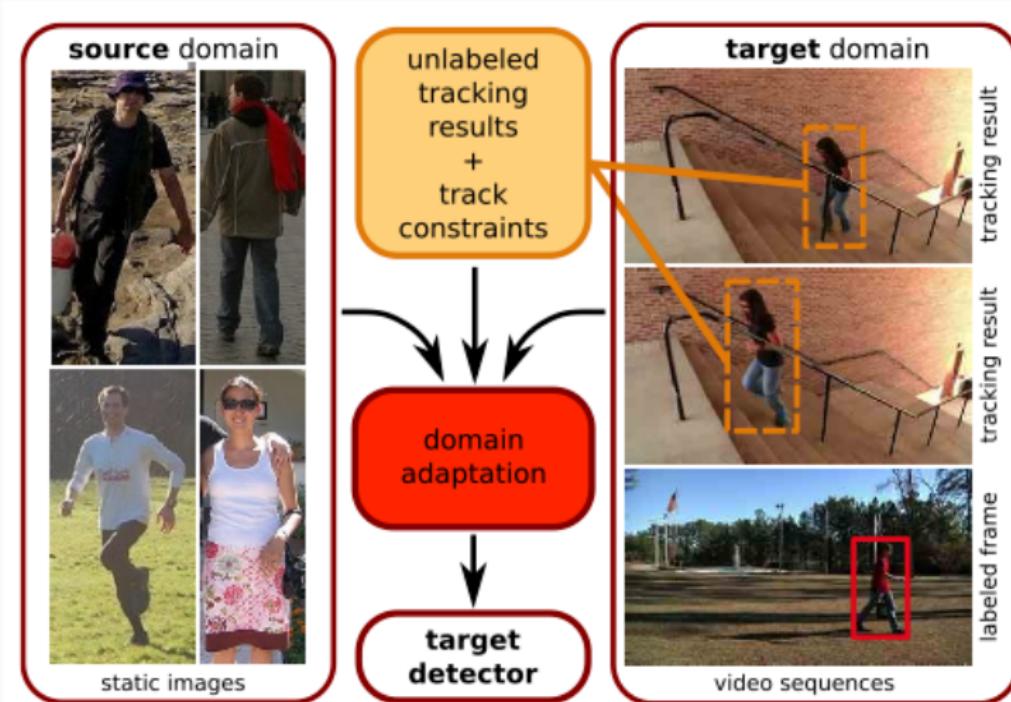
Задачи. Face identification.

- Решённая задача
- Проблемы с большими масштабами
- Генерализация



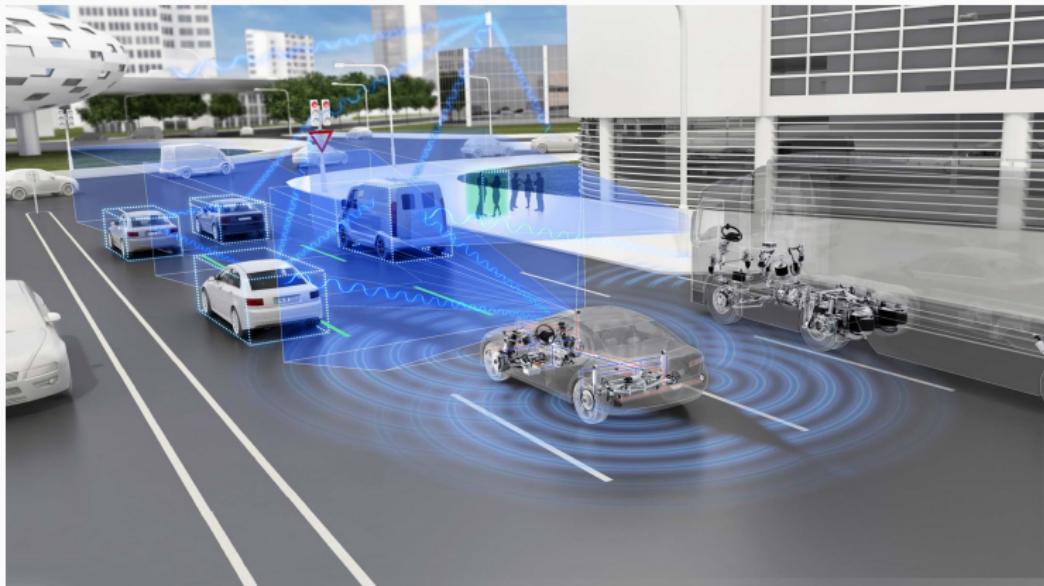
Задачи. Domain Shift.

- Не нужно размечать данные
- Решение проблемы обобщения признаков на новые классы



Задачи. Автономное вождение.

- Комбинация предыдущих проблем
- Нестандартная метрика качества
- Этические проблемы



Задачи. GANs.

- Модно, стильно, молодёжно
- Локально, но не глобально
- Проблемы с трэйном
- Ничего лучше пока нет



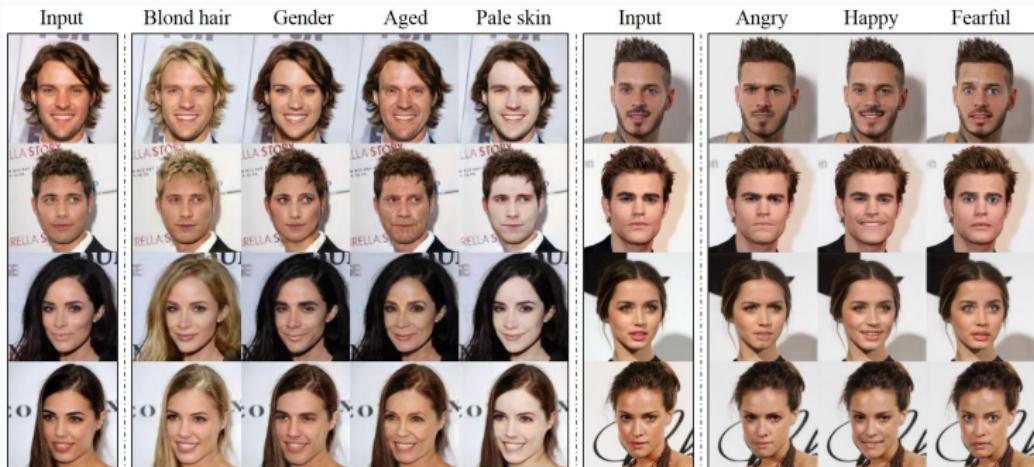
Задачи. GANs.

- Хорошо работает на датасэтах с малой вариативностью
- Вариативность растровая, но не векторная
- Копирует или генерирует?



Задачи. GANs.

- Разнообразие приложений
- Semi-supervised



Задачи. GANs.

- Междисциплинарная область
- Может помочь решить обе задачи



Задачи. Style Transfer.

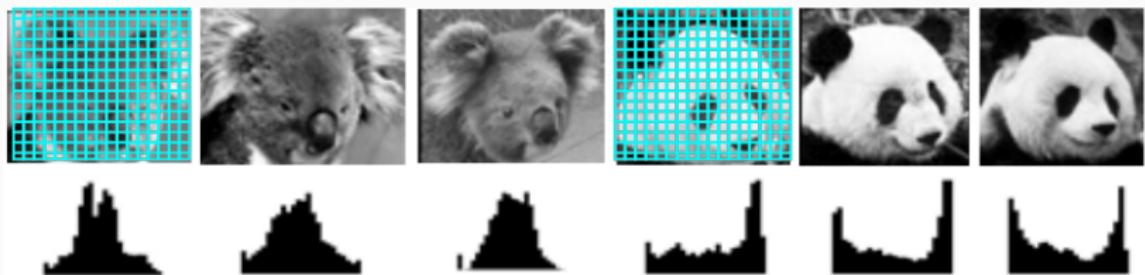
- Междисциплинарная область
- Исследование особенностей сетей
- Простые приложения



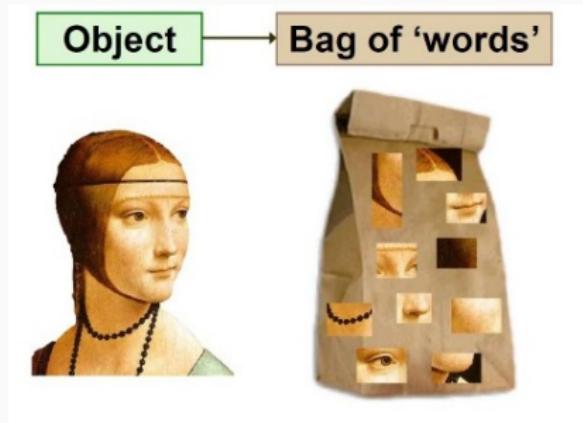
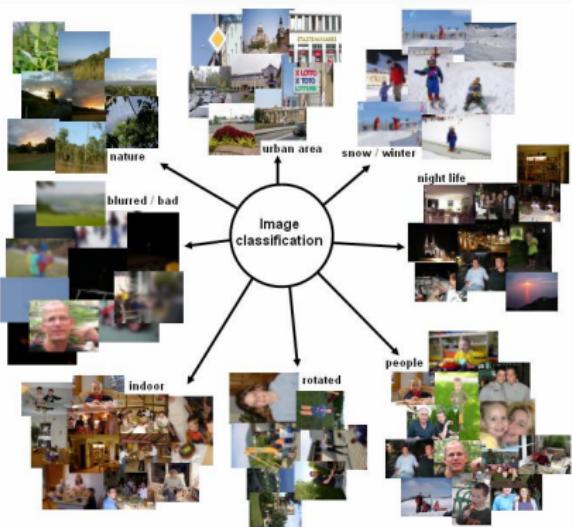
Классификация.

Классификация изображений. Глобальный подход.

- Гистограмма интенсивностей как признак
- На цветном или ч/б изображении
- Чувствительно к освещению и вариации внутри класса

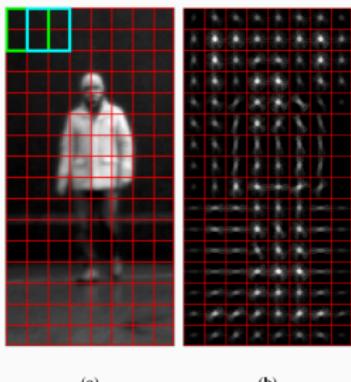


Классификация изображений. Локальный подход.



Классификация изображений. До глубинного обучения.

- Посчитать локальные дескрипторы

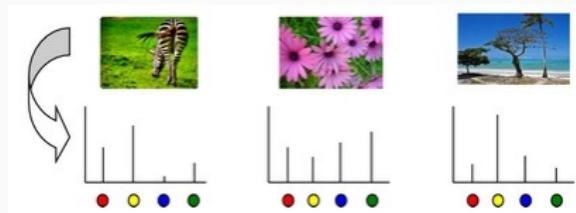


(a)

(b)

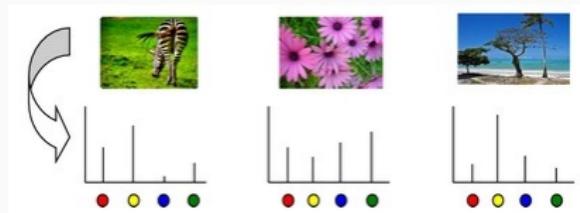
Классификация изображений. До глубинного обучения.

- Посчитать локальные дескрипторы
- Компрессия дескрипторов в словарь



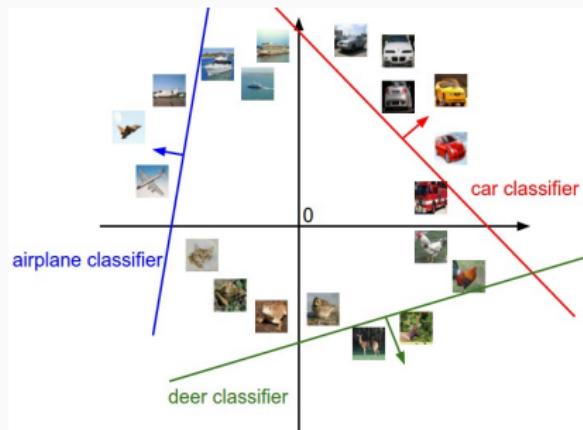
Классификация изображений. До глубинного обучения.

- Посчитать локальные дескрипторы
- Компрессия дескрипторов в словарь
- Собрать массив изображений и соответствующих слов



Классификация изображений. До глубинного обучения.

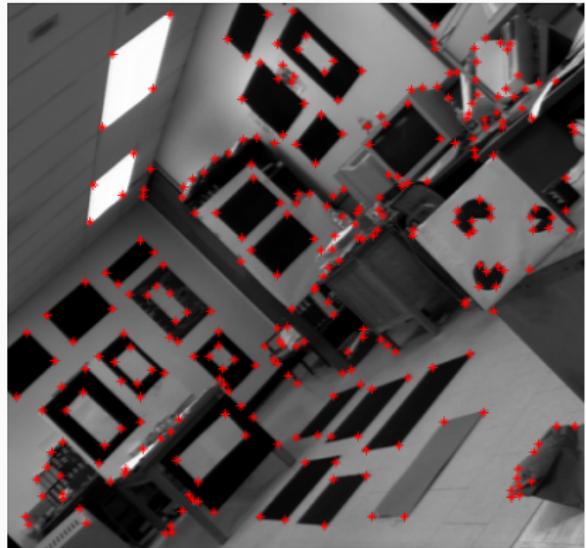
- Посчитать локальные дескрипторы
- Компрессия дескрипторов в словарь
- Собрать массив изображений и соответствующих слов
- Натренировать ML алгоритм на словах



Дескрипторы

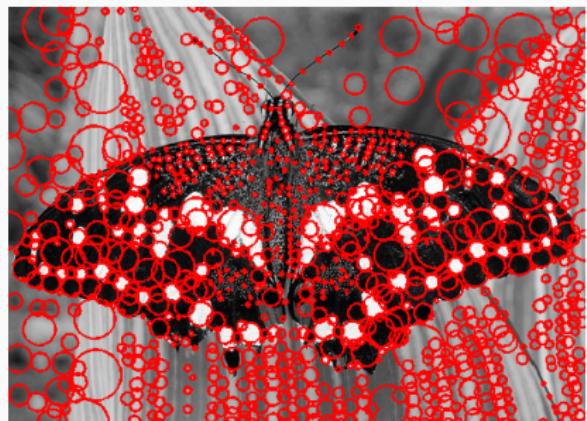
Локальные дескрипторы.

- Corner detectors.



Локальные дескрипторы.

- Corner detectors.
- Blob detectors.



Локальные дескрипторы.

- Corner detectors.
- Blob detectors.
- Edge detectors: Sobel, Prewitt, Canny



Edge detector. Конволюция.

Для $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ и $t \in \mathbb{R}^n$

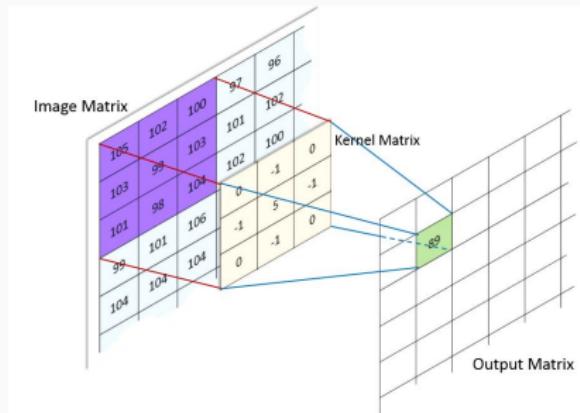
конволюция определяется:

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} f(\tau)g(t - \tau) d\tau = \\ = \int_{\mathbb{R}^n} f(t - \tau)g(\tau) d\tau.$$

В дискретном случае для g определённой на

$\{-M, -M + 1, \dots, M - 1, M\}$:

$$(f * g)[n] = \sum_{m=-M}^M f[n - m]g[m]$$



Edge detector. Фильтр Собеля

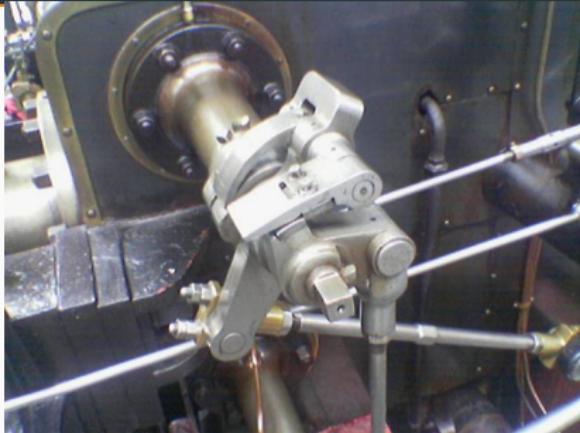
$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix}$$

$$G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$G = \sqrt{G_x^2 + G_y^2}$$

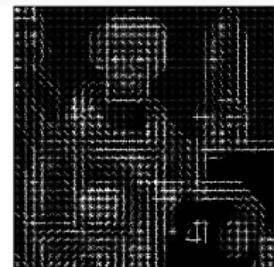
Вычислительная
эффективность:

$$\begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} +1 & 0 & -1 \end{bmatrix}$$



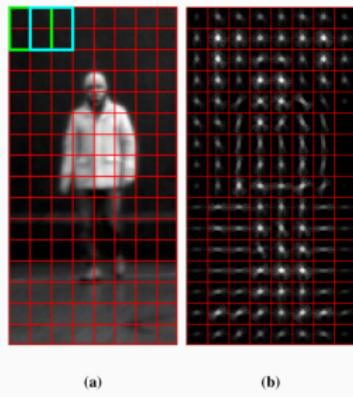
HOG. Мотивация.

- Подход: объект характеризуется контуром а контур считается детектором эджей
- Проблема: слишком много данных
- Решение: агрегировать градиенты в небольшом окне



HOG. Алгоритм.

- Нормализация входной картинки
- Для каждого подокна посчитать градиенты. Угол градиента: $\theta = \text{atan} \left(\frac{G_y}{G_x} \right)$

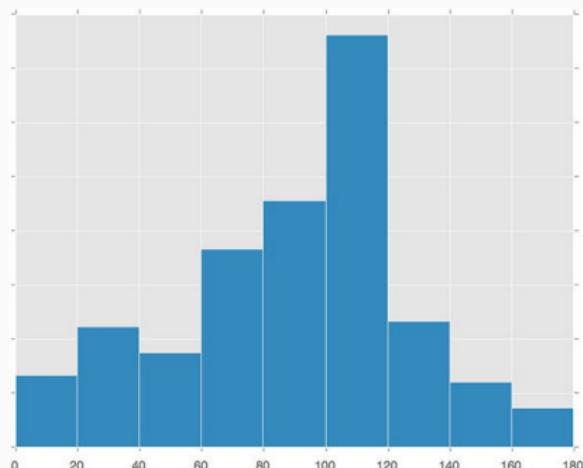


(a)

(b)

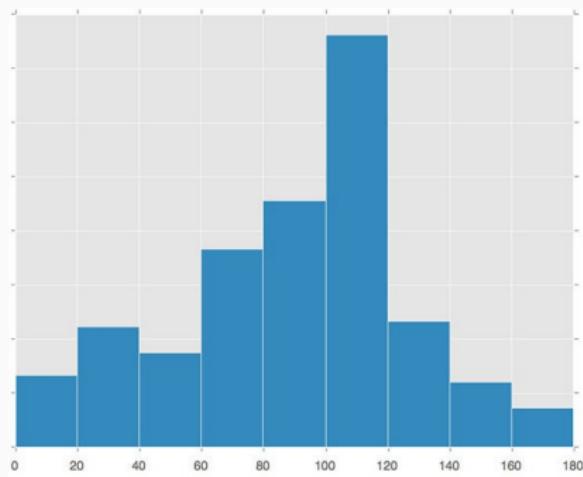
HOG. Алгоритм.

- Нормализация входной картинки
- Для каждого подокна посчитать градиенты. Угол градиента: $\theta = \text{atan} \left(\frac{G_y}{G_x} \right)$
- Для каждого окна гистограмма направлений



HOG. Алгоритм.

- Нормализация входной картинки
- Для каждого подокна посчитать градиенты. Угол градиента: $\theta = \text{atan} \left(\frac{G_y}{G_x} \right)$
- Для каждого окна гистограмма направлений
- Нормализация гистограмм по L_2 норме $f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}}$



HOG. Достоинства и недостатки.

- Нечувствительно к освещению
- Инвариантность к смещению, повороту
- Нечувствительность к цвету
- Плохо работает на текстурах
- Проблемы с фоном
- Небогатый локальный арсенал