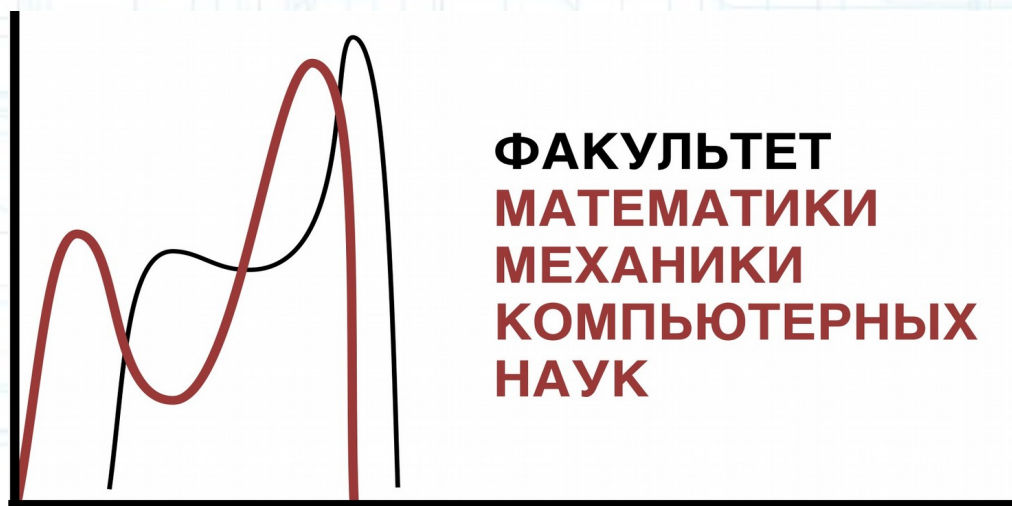


Машинное обучение в компьютерном зрении



**ФАКУЛЬТЕТ
МАТЕМАТИКИ
МЕХАНИКИ
КОМПЬЮТЕРНЫХ
НАУК**

Содержание

- Постановка задачи МО
- Подходы к решению задачи
- МО в компьютерном зрении

1974–1980 – зима для Искусственного интеллекта

- Наивный оптимизм:
 - 1954 через 3 — 5 лет задача машинного перевода будет полностью решена
 - 1958 через 10 лет компьютер станет чемпионом мира по шахматам; докажет важную математическую теорему
 - 1970 через 3 — 8 лет будет создан искусственный интеллект общего назначения, сравнимый с интеллектом среднестатистического человека
- парадокс Моравека:
 - относительно легко достичь уровня взрослого человека в таких задачах как тест на интеллект или игре в шашки, однако сложно или невозможно достичь навыков годовалого ребёнка в задачах восприятия или мобильности

Задача обучения

X — множество объектов

Y — множество ответов

$y : X \rightarrow Y$ — неизвестная зависимость
(target function)

Дано:

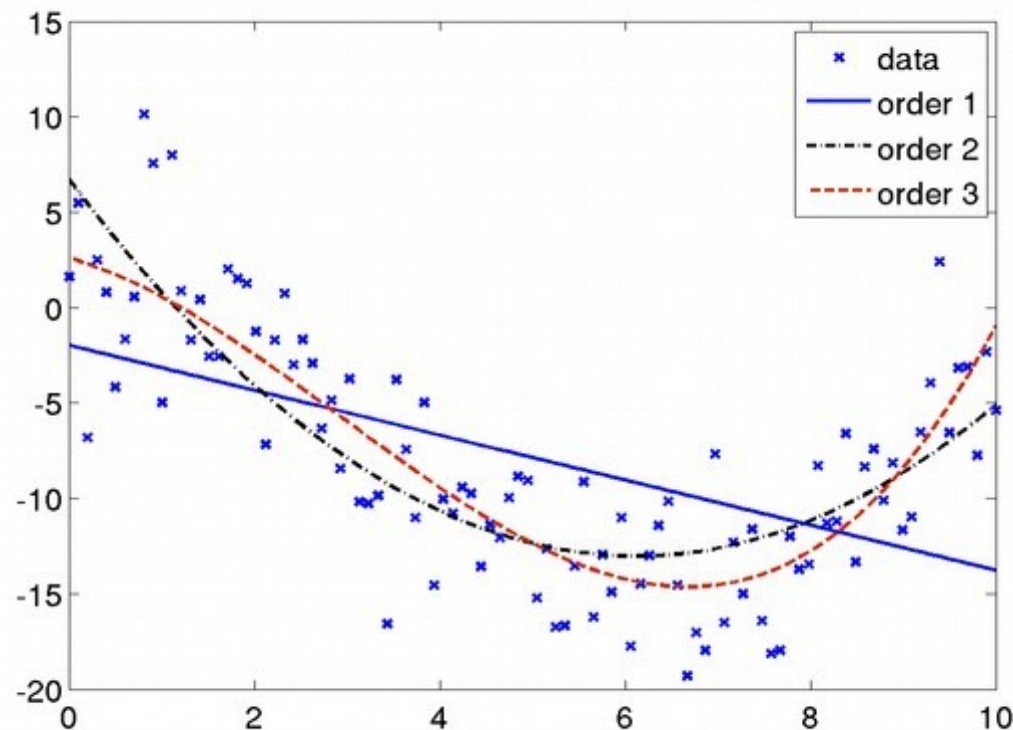
$\{x_1, \dots, x_\ell\} \subset X$ — обучающая выборка
(training sample)

$y_i = y(x_i), i = 1, \dots, \ell$ — известные ответы

Задача обучения

Найти:

$a : X \rightarrow Y$ — алгоритм, решающую функцию (decision function), приближающую y на всём множестве X



Типы задач

Задачи классификации (classification):

$Y = \{-1, +1\}$ — классификация на 2 класса

$Y = \{1, \dots, M\}$ — на M непересекающихся классов

$Y = \{0, 1\}^M$ — на M классов, которые могут пересекаться.

Задачи восстановления регрессии (regression):

$Y = \mathbb{R}$ или $Y = \mathbb{R}^m$

Задачи ранжирования (ranking):

Y — конечное упорядоченное множество

Признаки

- Компьютер всегда имеет дело с признаковым описанием объектов. Например: пациента можно описать признаками: имя, возраст, номер полиса, жалобы, давление, температура, результаты анализов

- $f : X \rightarrow D_f$

- Типы признаков:

- бинарный
- номинальный
- порядковый
- количественный

Матрица объектов-признаков:

$$\begin{pmatrix} f_1(x_1) & \dots & f_n(x_1) \\ \dots & \dots & \dots \\ f_1(x_\ell) & \dots & f_n(x_\ell) \end{pmatrix}$$

Модель и алгоритм обучения

- **Модель** – это семейство “гипотез”

$$A = \{g(x, \theta) \mid \theta \in \Theta\}$$

одна из которых (как мы надеемся)
приближает целевую функцию

- **Алгоритм обучения**

$$\mu: (X \times Y)^\ell \rightarrow A$$

находит гипотезу в модели, которая
наилучшим образом приближает
целевую функцию, используя известные
значения (обучающую выборку)

Пример - регрессия

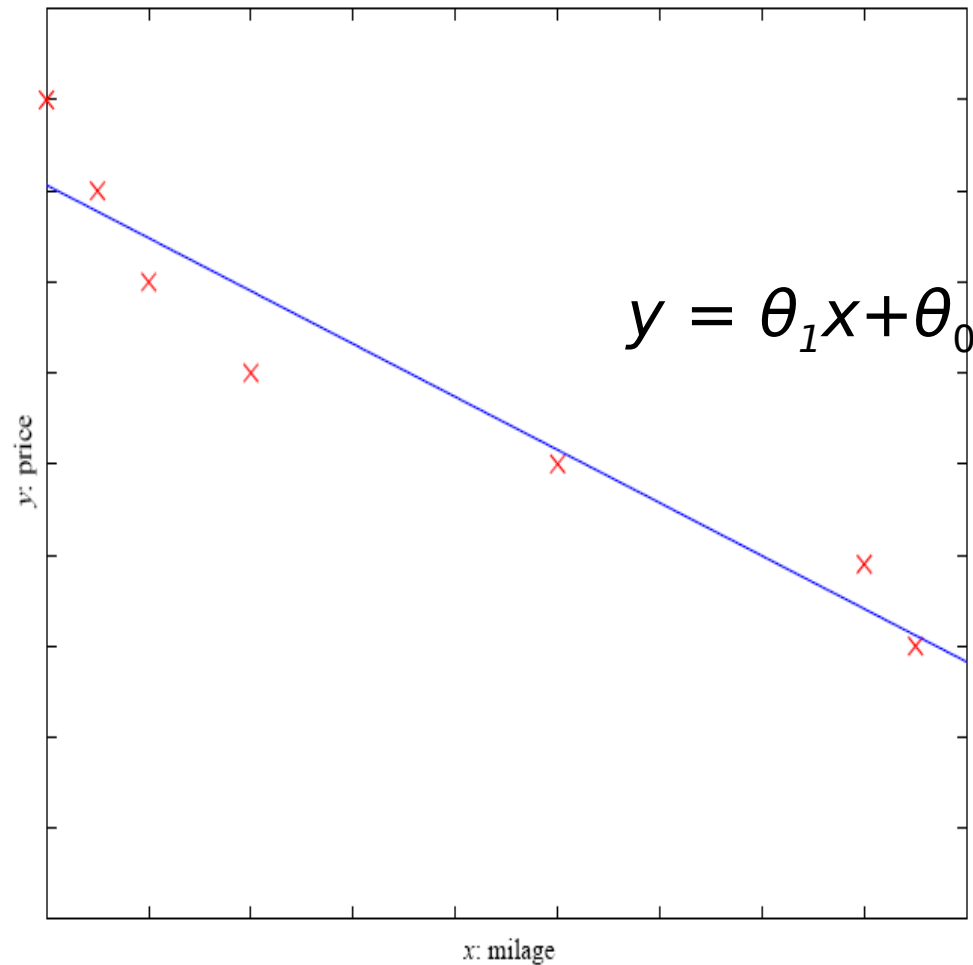
- Цена автомобиля
- x : признаки (пробег)

y : цена

$$y = g(x, \theta)$$

$g(\cdot)$ - модель,

θ_0, θ_1 - параметры



Обучение на основе минимизации эмпирического риска

- Функция потерь $\mathcal{L}(a, x)$ - величина ошибки гипотезы a на объекте x .

Примеры:

- бинарная

- $\mathcal{L}(a, x) = |a(x) - y^*(x)|$

- $\mathcal{L}(a, x) = (a(x) - y^*(x))^2$

- Эмпирический риск: $Q(a, X^\ell) = \frac{1}{\ell} \sum_{i=1}^{\ell} \mathcal{L}(a, x_i)$

- Самый популярный алгоритм обучения – минимизация эмпирического риска:

$$\mu(X^\ell) = \arg \min_{a \in A} Q(a, X^\ell)$$

Метрические алгоритмы

- Гипотеза: близким объектам соответствуют близкие ответы
- Нужна метрика, определяющая понятие близости
- Метод ближайшего соседа: объект x принадлежит тому же классу, что и ближайший в обучающей выборке

Линейные алгоритмы

- Для двух классов +1 и -1:

$$a(x) = \text{sign}(\langle w, x \rangle + w_0)$$

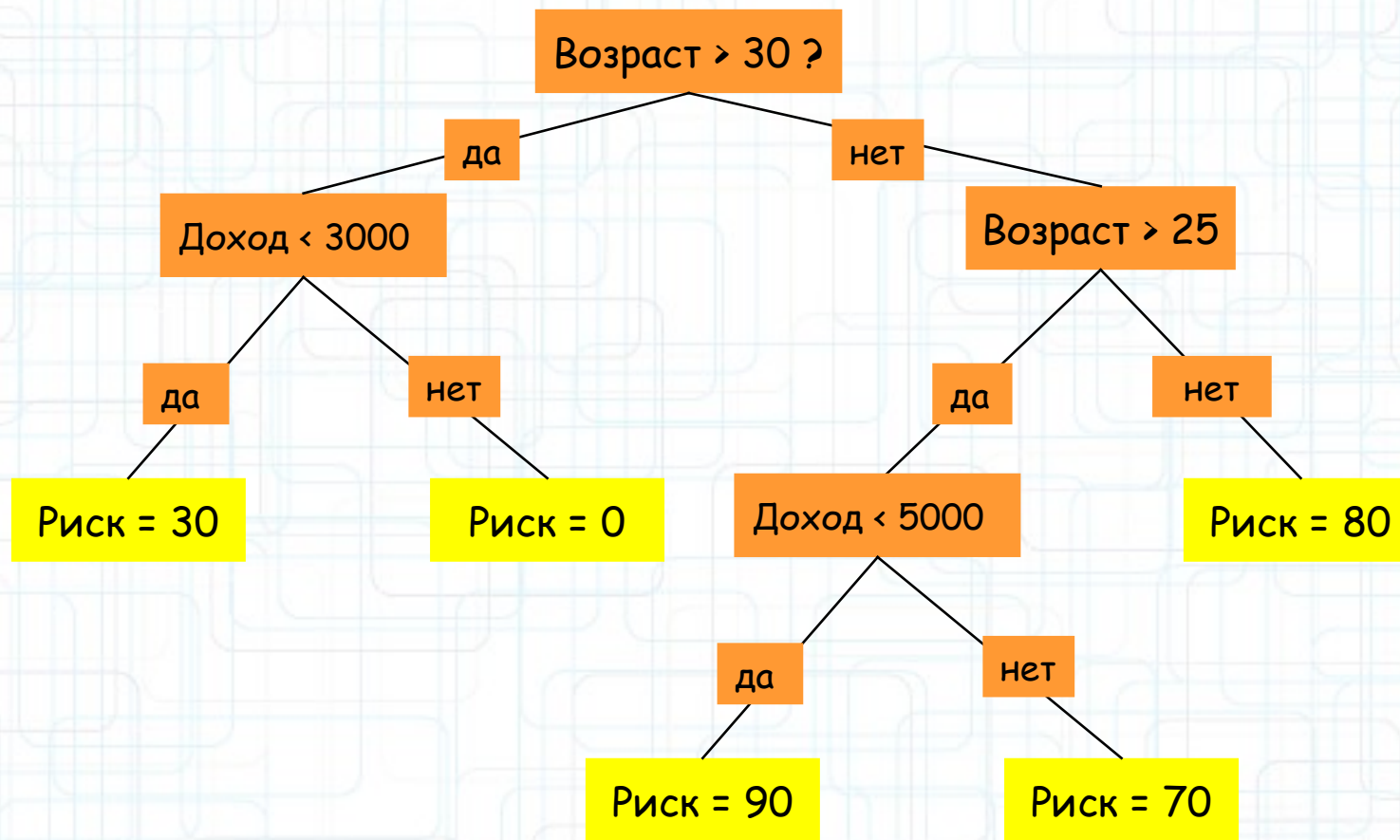
- Вектор w и параметр w_0 находится по обучающей выборке
- Примеры:
 - однослойный перцептрон
 - логистическая регрессия
 - SVM – метод опорных векторов
 - линейная регрессия

Байесовские алгоритмы

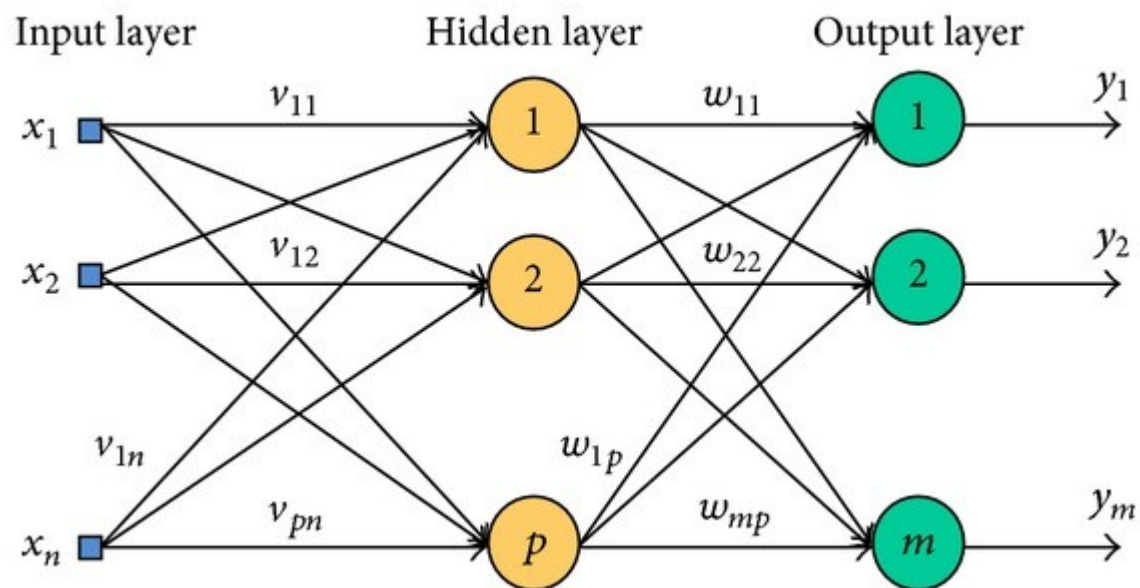
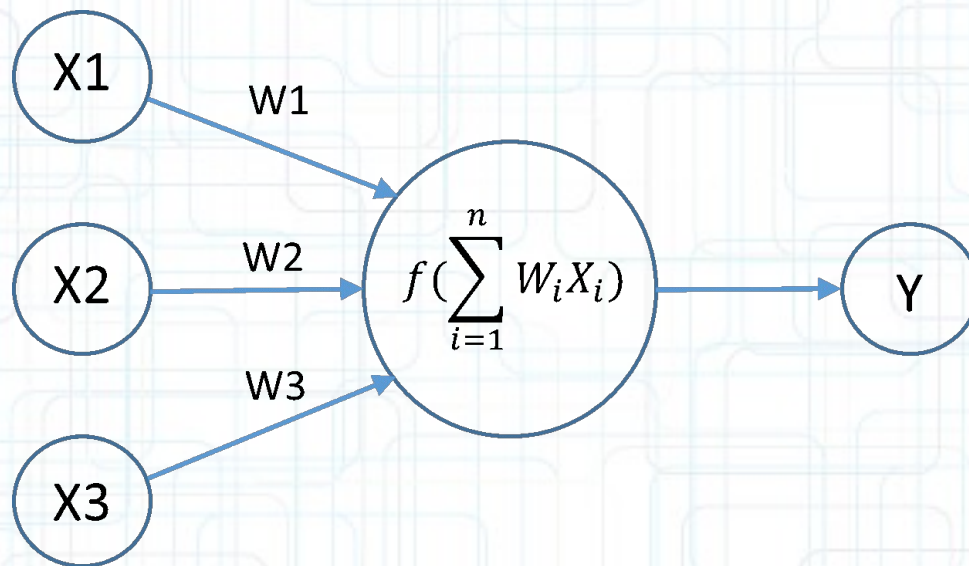
- Восстанавливают плотность распределения $p(x,y)$
- Пример:
наивный байесовский классификатор

Решающие деревья

- Пример — кредитный скоринг



Нейронные сети



Выступ объекта

- Характеризует соответствие предсказания реальному значению
- Отрицательные значения обычно означают ошибку классификации
- Типы объектов:
 - эталоны – объекты с большим положительным выступом (характерные объекты своего класса)
 - случайные выбросы – объекты с большим по модулю отрицательным выступом (аномалии)
- Выступы зависят от алгоритма МО

Примеры эталонов и аномалий

- Предсказание судьбы пассажиров Титаника
 - эталоны:
 - Mr. Peju, class=3, 36 лет, погиб
 - Ms. Emma, class=1, 24 года, выжила
 - Mr. George Floyd, class=2, 23 года, погиб
 - аномалии:
 - Mr. Ernst Ulrik, class=3, 25 лет, выжил
 - Miss. Annie Clemmer, class=2, 38 лет, погибла
 - Edvin Rojj Felix, class=3, 2 года, погиб

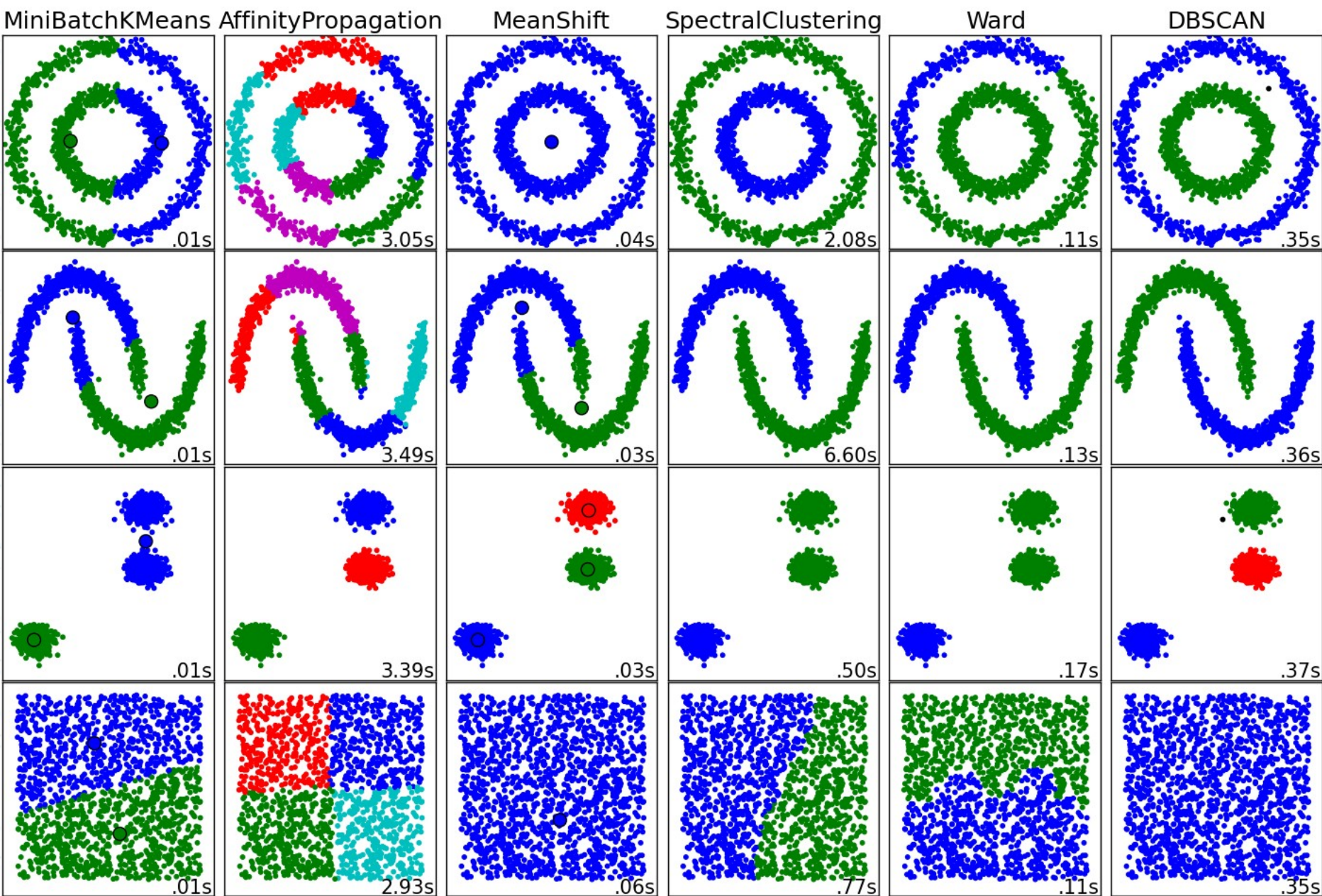
Примеры эталонов и аномалий

- Сентимент-анализ отзывов по методу “bag of words”
 - эталон: (предсказание = +1, реально = +1)
Действительно хороший фильм. Очень жестокий и реалистичный
 - аномалия: (предсказание = -1, реально = +1)
«Придурки» продолжают свое победное шествие по большим экранам. Их игрища конечно на любителя, но их второе пришествие меня разочаровала по сравнению с первой частью. Иногда смешно, иногда очень смешно, то чаще противно, тупо и ужасно и.... Вывод - поржать можно, но от употребления пищи при этом следует отказаться, а то постигнет разочарование...

Кластеризация

- Дано:
 - пространство объектов X
 - обучающая выборка X^{ℓ}
 - метрика между объектами
- Найти:
 - множество кластеров Y
 - алгоритм кластеризации $a : X \rightarrow Y$
- Каждый кластер должен состоять из близких объектов
- Объекты разных кластеров должны быть существенно различны

Пример: результаты работы алгоритмов кластеризации



Использовать МО или нет?

Доводы в пользу МО:

- не нужно думать над алгоритмом/функцией решения задачи, она получается автоматически
- алгоритмы МО апробированы и отлажены, написать самому алгоритм аналогичного качества (изобрести велосипед) – очень сложно
- измеряется результативность алгоритма

Причины использования самописных алгоритмов:

- понятны принципы, по которым он работает
- не нужно строить **хорошую** обучающую выборку
- сложно разобраться в алгоритмах МО
- сложно увидеть, как применить МО
- кажется, что тот метод, который я придумал, гораздо лучше всего, что было когда-либо придумано

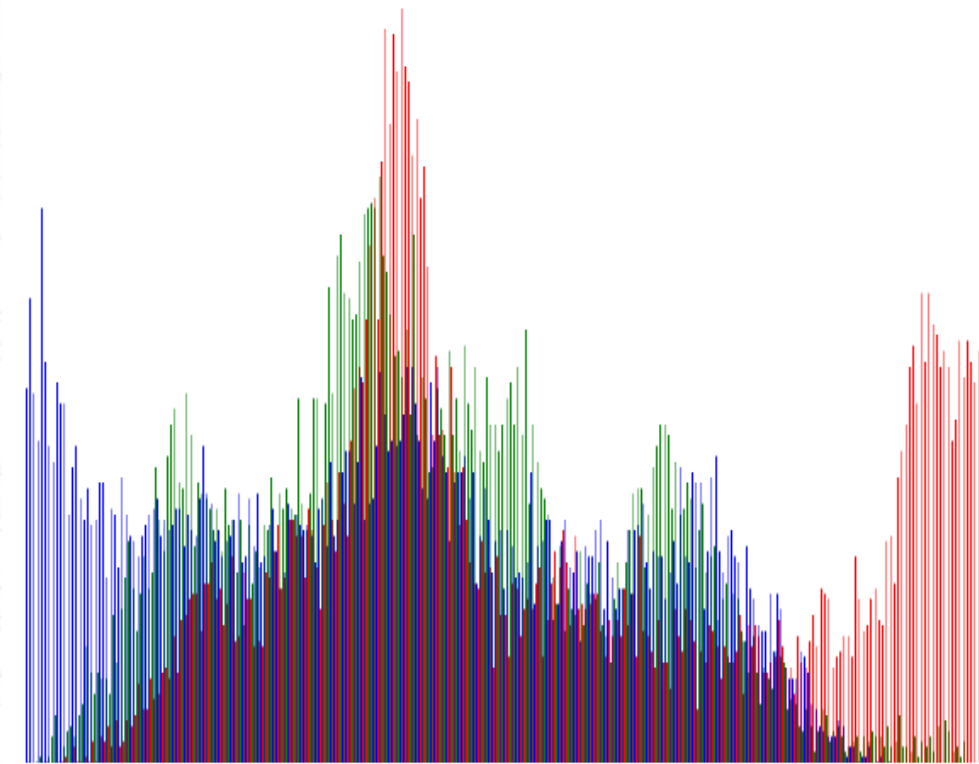
Задачи компьютерного зрения

- **Классификация:**
 - распознавание лиц, объектов, жестов
 - распознавание рукописного текста
 - поиск по изображению
- **Регрессия:**
 - определение положения и ориентации объекта в пространстве
 - стерео-реконструкция
 - восстановление зашумленных изображений
- **Выбросы:**
 - выявление аномалий
 - реферирование видео
- **Кластеризация:**
 - сегментация изображений
 - поиск дубликатов (фотоподделки)

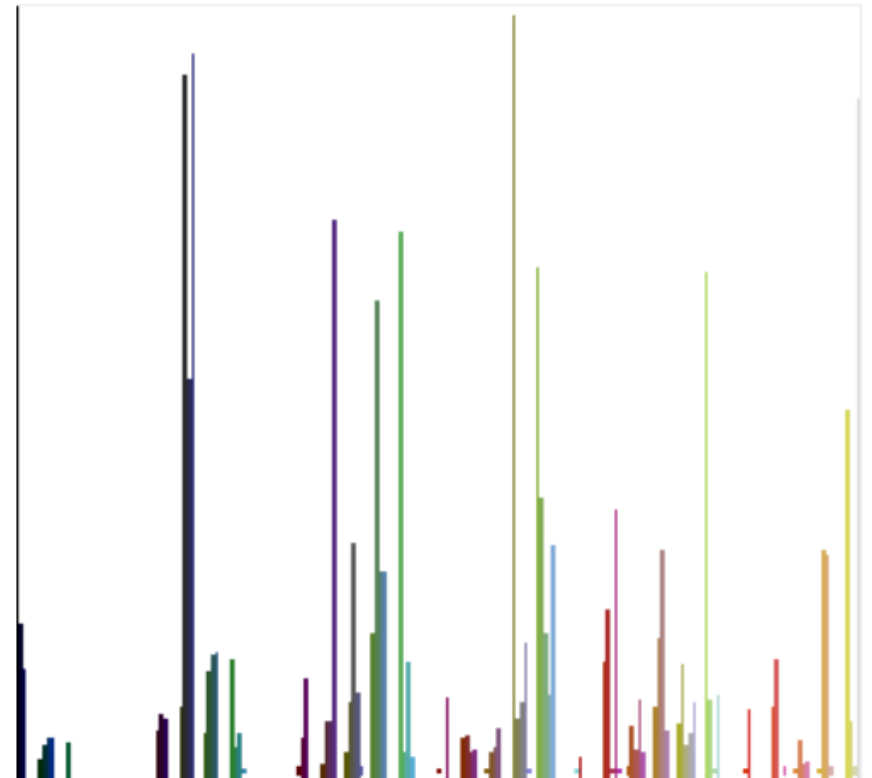
Признаки изображений

- Глобальные:
 - полноцветные гистограммы
 - контекст формы
 - GIST
- Локальные:
 - Детекторы: DoG, MSER, Hessian Affine, KAZE, FAST
 - Дескрипторы: SIFT, GLOH, SURF, LIOP, BRIEF, ORB, FREAK, BRISK, CARD
- Свертки с ядрами

Полноцветная гистограмма

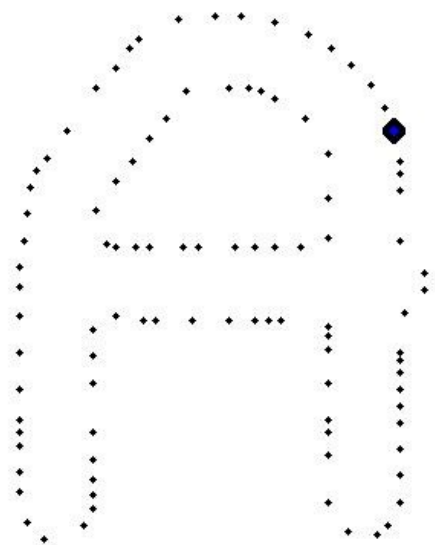


Обычная RGB-гистограмма

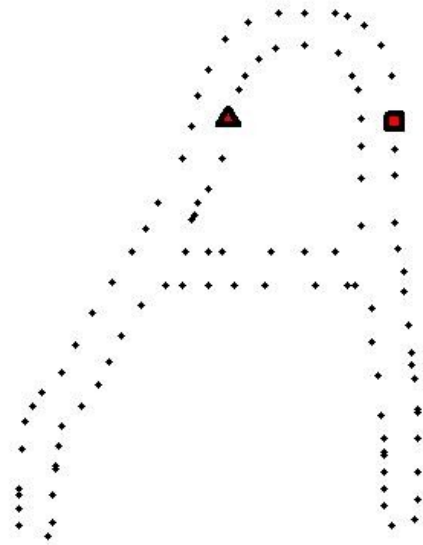


Полноцветная гистограмма

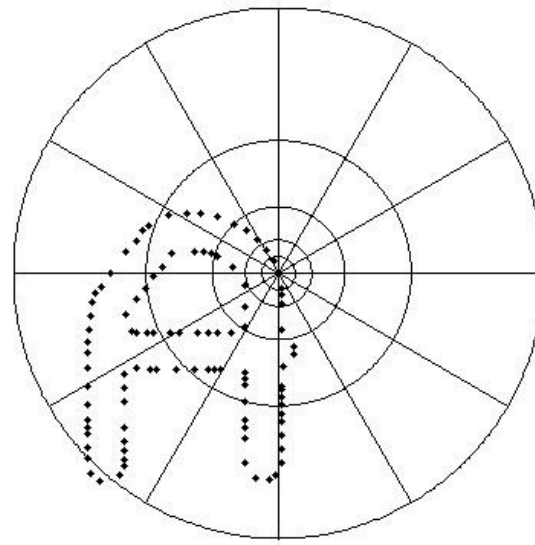
Контекст формы



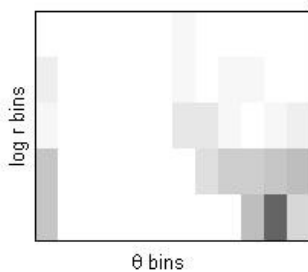
(a)



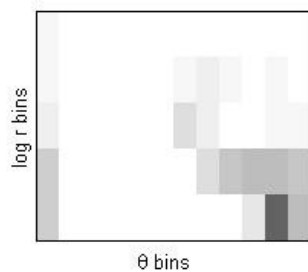
(b)



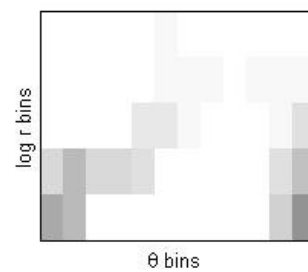
(c)



(d)

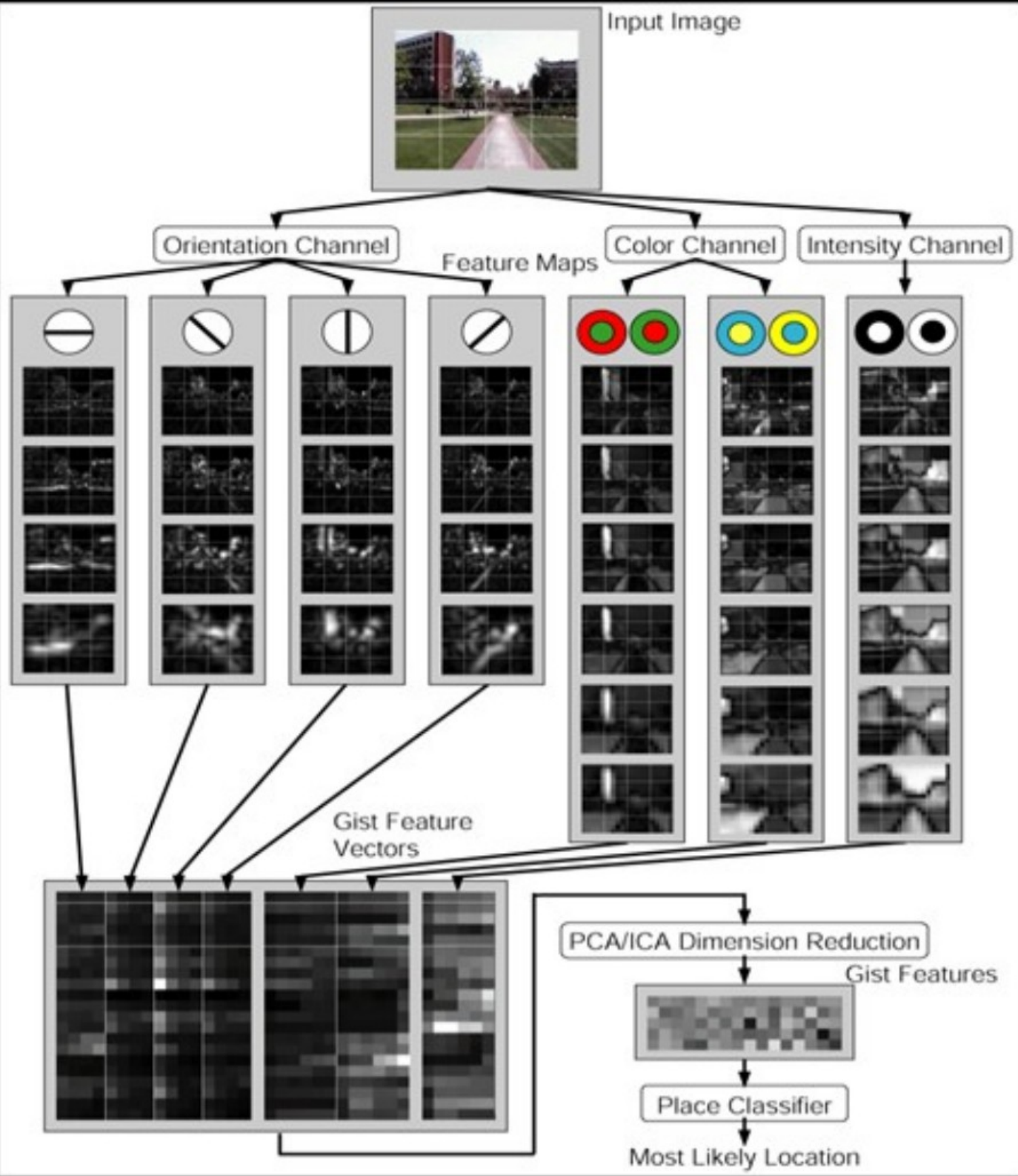


(e)

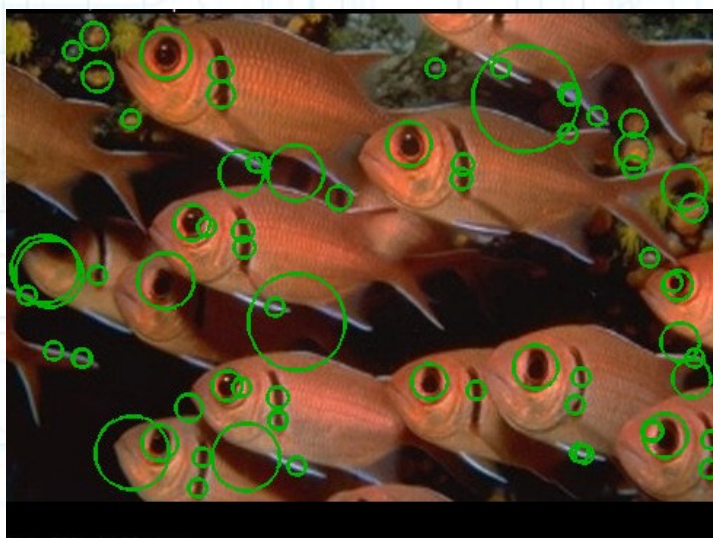
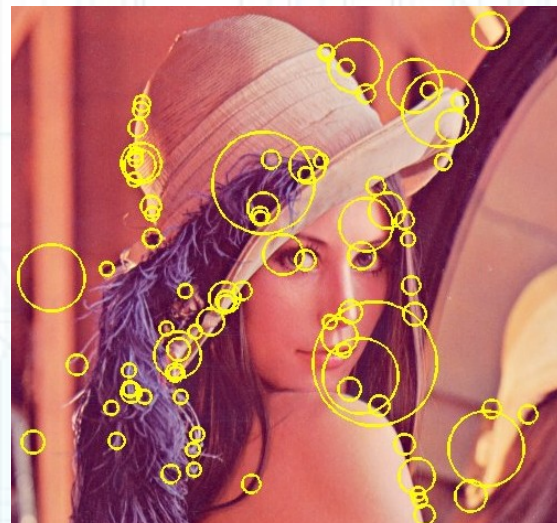
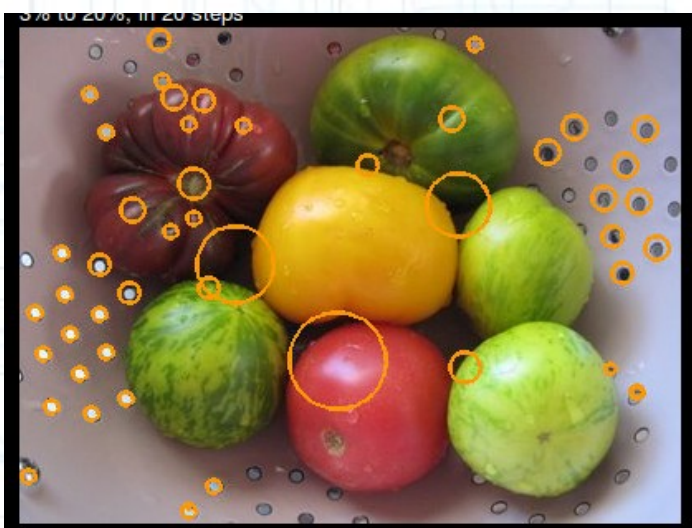


(f)

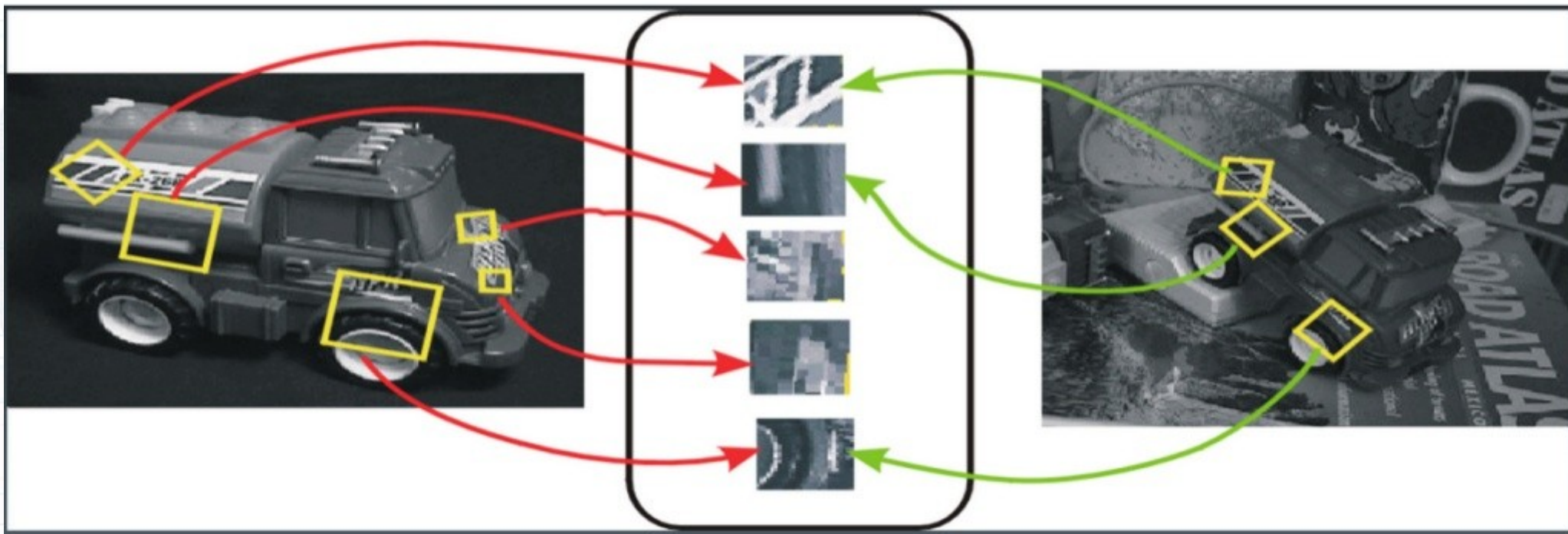
GIST



Детектирование особых точек (Blob detection)



SIFT – Scale Invariant Feature Transform



Свертки с ядрами



Свертки с ядрами

Сверточные нейронные сети

