Object class recognition

Выполнили:

Веслогузова Александра
Лазаревич Андрей
Сергеев Павел
Юргин Павел

Руководитель:

Тузова Екатерина

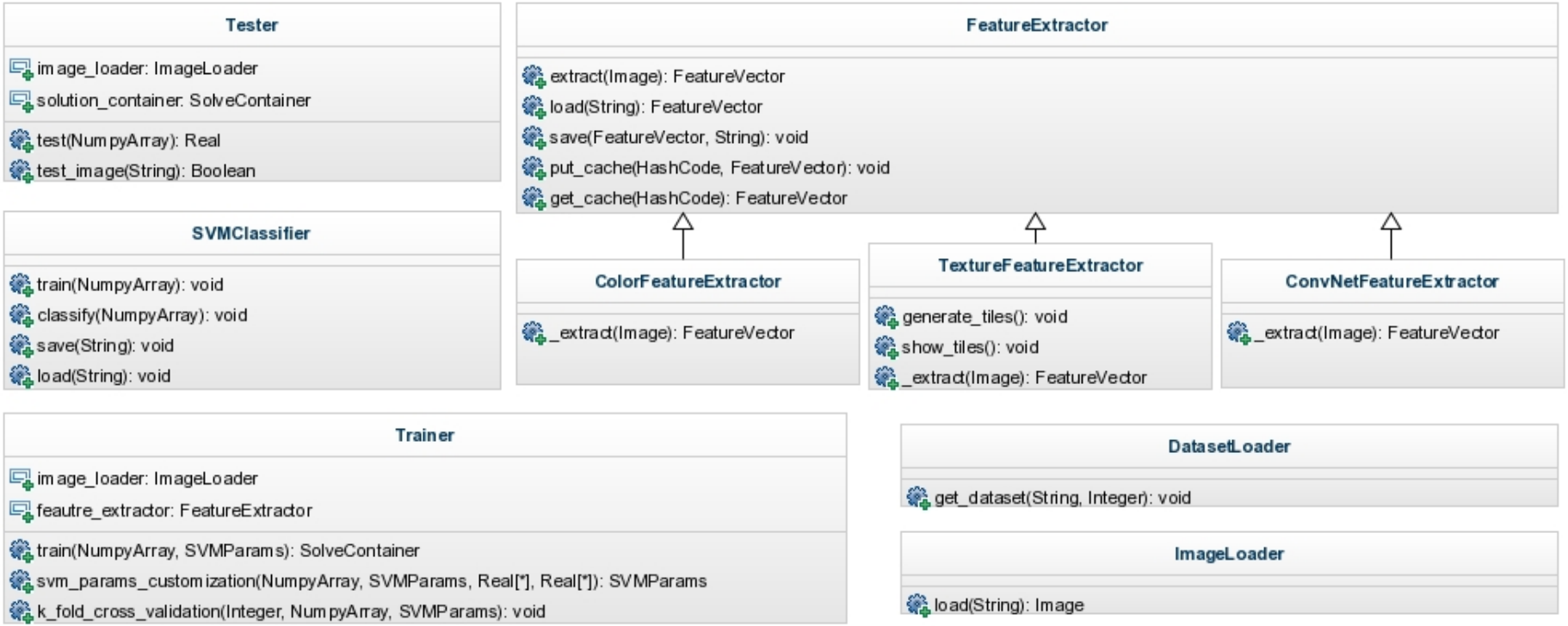
Задача

- Определение класса объекта изображенного на фотографии
- Частный случай данной задачи - определение принадлежности объекта на фотографии к одному из двух классов

Этапы построения алгоритма

1. Выбор алгоритмов обработки изображения
2. Подбор методов выделения вектора признаков
3. Выбор классификатора
4. Определение методов тестирования

Архитектура приложения
Основные модули

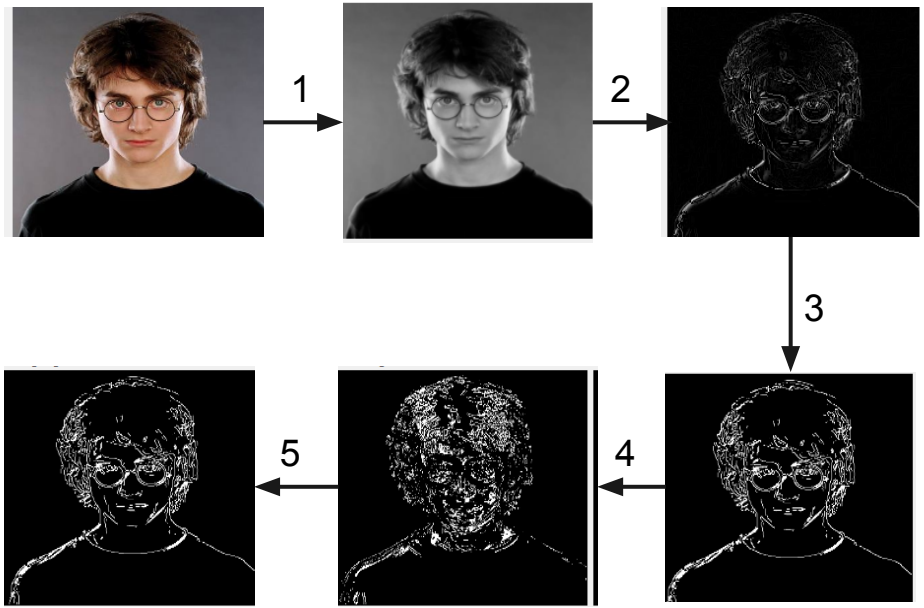


ImageLoader

- Масштабирование изображения без искажения пропорций
- Поиск вероятного нахождения объектов на изображении
- Сохранение уже обработанных изображений в кэш, для ускорения дальнейшей работы.

Оператор Кэнни

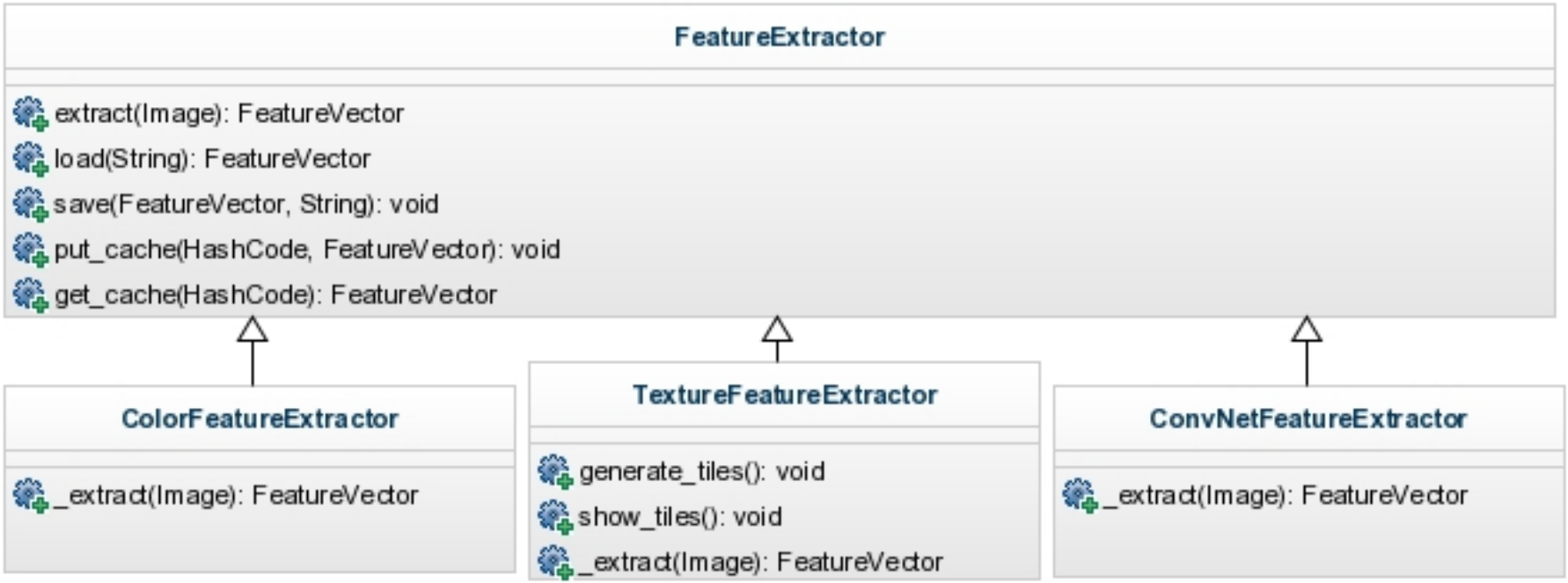
1. Размытие изображения для уменьшения шумности
2. Поиск градиентов
3. Подавление немаксимумов
4. Двойная пороговая фильтрация
5. Трассировка неоднозначностей



Feature Extractor

$$\longrightarrow \quad x \in \mathbb{R}^n$$

FeatureExtractor



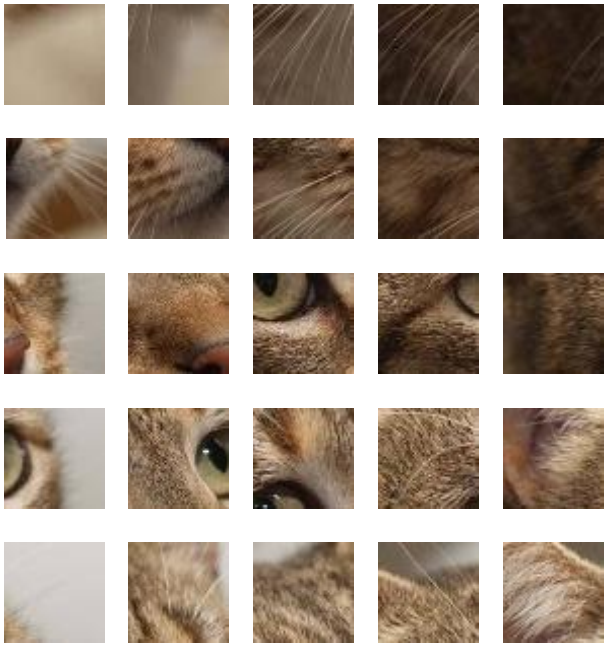
TextureFeatureExtractor

Генерация вектора признаков на основе сравнения изображения со словарем текстур (тайлов).

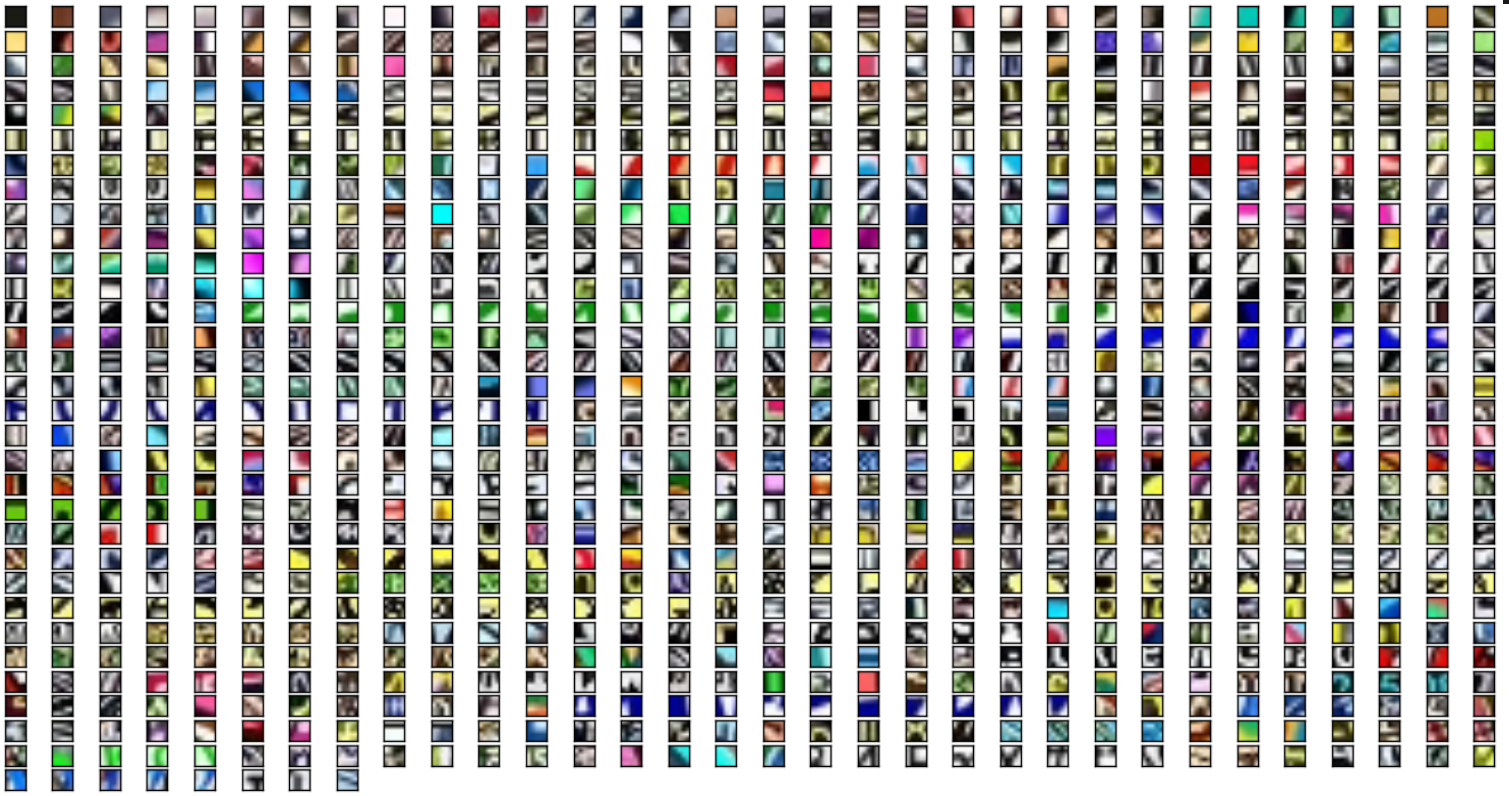
Функции:

- Генерация словаря тайлов по обучающей выборке.
- Сравнения изображения с конкретным тайлом.

Генерация тайлов



Пример словаря



Похожесть изображений

Были использованы следующие метрики похожести изображений:

- Минимум максимумов разностей евклидова расстояния между пикселями.

$$\rho_1(T, I) = \min_{x,y} \max_{x',y'} \sqrt{< T(x + x', y + y'), I(x', y') >}$$

$T(x, y),\ I(x, y)$ - RGB тайл и изображение

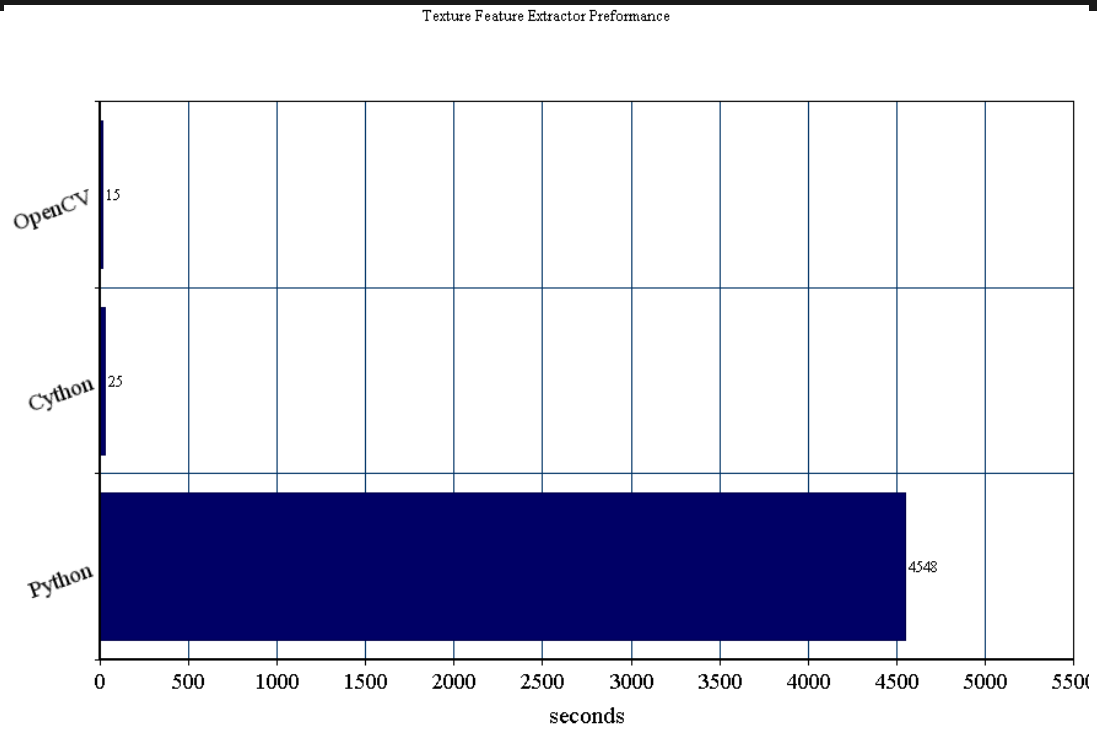
- Максимум нормированной корреляции:

$$\rho_2(T, I) = \min_{x,y} \max_{x',y'} \frac{\sum_{x'y'}(T'(x', y') \cdot I'(x + x', y + y'))}{\sqrt{\sum_{x',y'} T'(x', y')^2 \cdot \sum_{x',y'} I'(x + x', y + y')^2}}$$, где

$I'(x, y)$ - нормированное по яркости черно-белое изображение

$T'(x, y)$ - нормированный по яркости черно-белый тайл

Performance

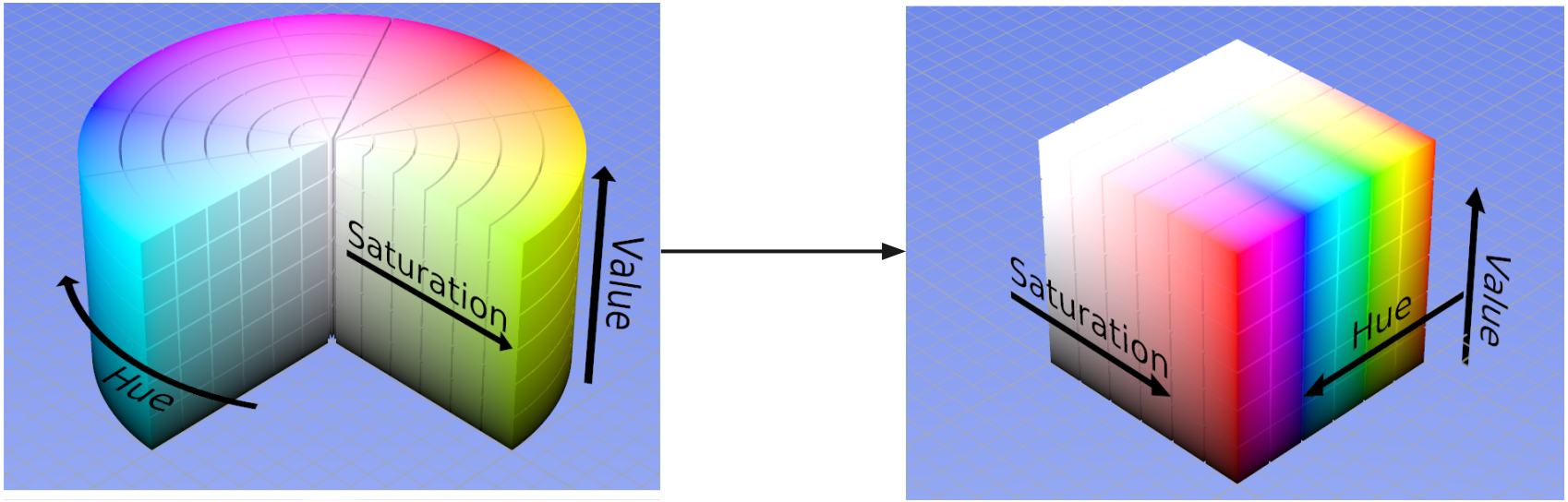


ColorFeatureExtractor

Генерация векторов признаков, основанный на цветовых характеристиках изображения.

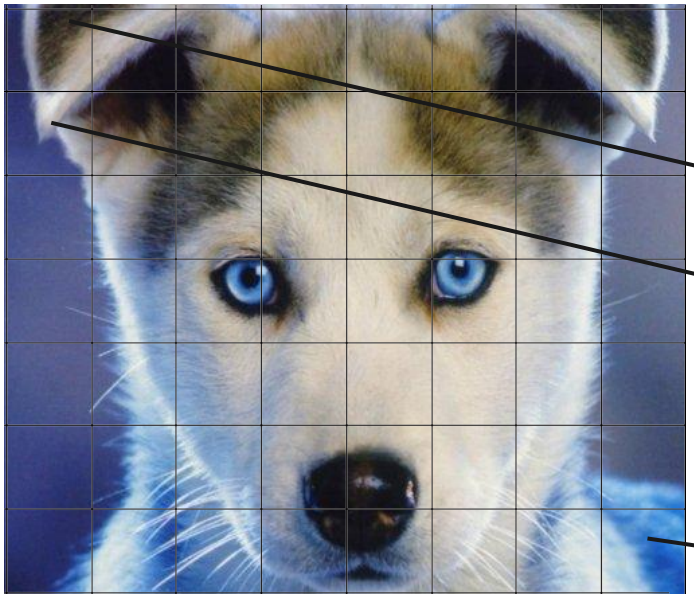
ColorFeatureExtractor

1. Разбиваем цветовое пространство HSV на $n \cdot m \cdot l$ частей.



ColorFeatureExtractor

1. Разбиваем изображение на части.
2. Извлекаем вектор признаков.

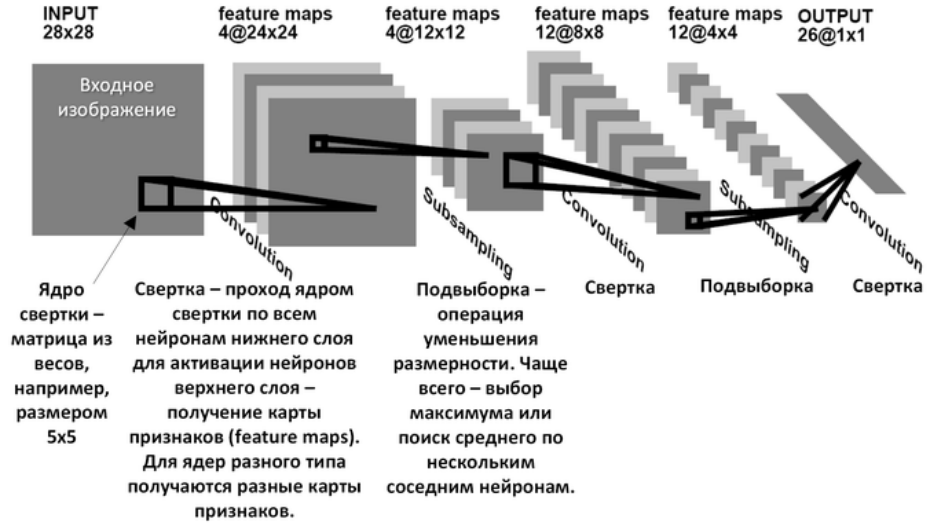


1	1	0	1	0	1	0	1	...	1	1	1	1	0	1	0	0	0
1	0	1	0	1	0	1	1	...	1	0	0	0	0	0	0	0	0
.																	
.																	
.																	
0	1	1	1	0	1	1	0	...	0	0	0	0	0	0	1	1	0

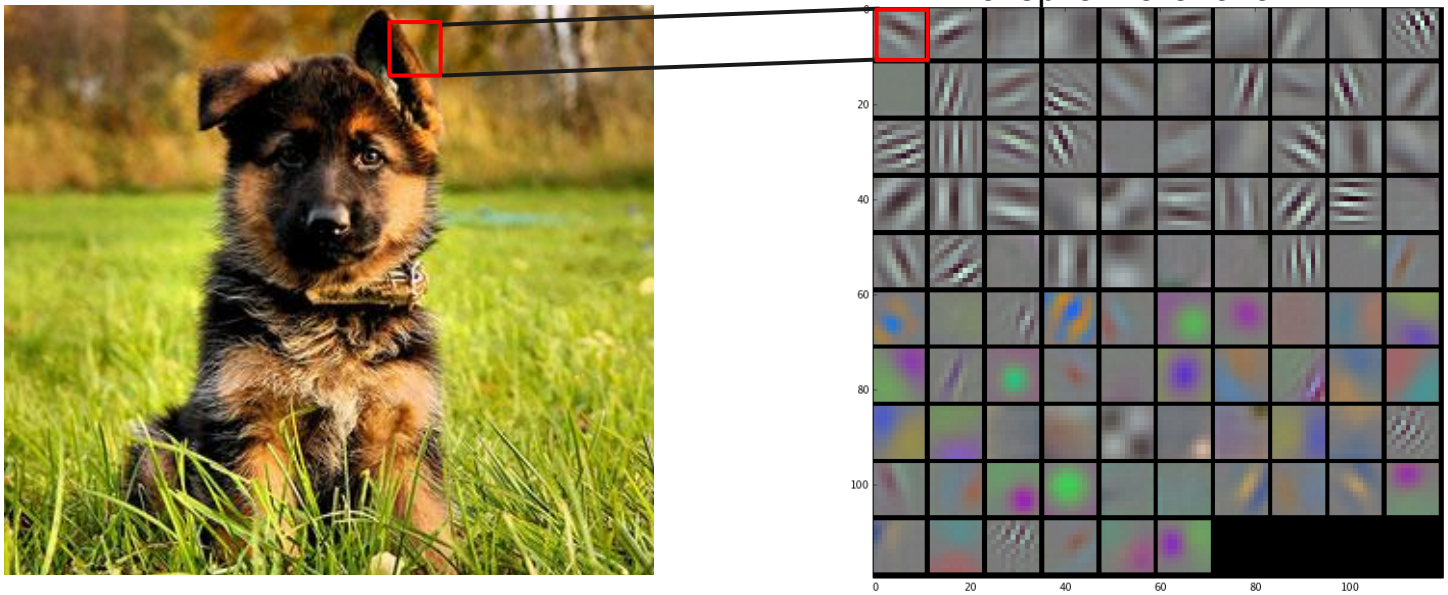
k

ConvNetFeatureExtractor

Для извлечения вектора признаков используется обученная сверточная нейросеть ILSVR 2012.



ConvNetFeatureExtractor



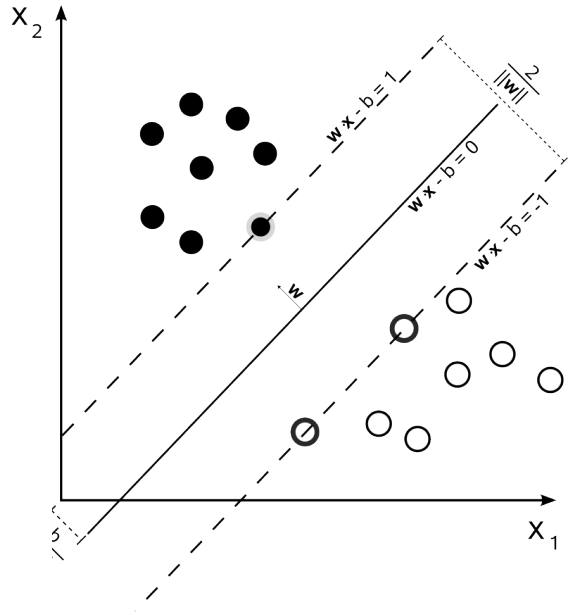
SVMClassfier

$$a(x) = sign(\sum_{j=1}^{n} w_j x_j - w_0)$$

где $x = (x_1, ..., x_n)$ вектор признаков,

$w = (w_1, ..., w_n)$ параметры алгоритма

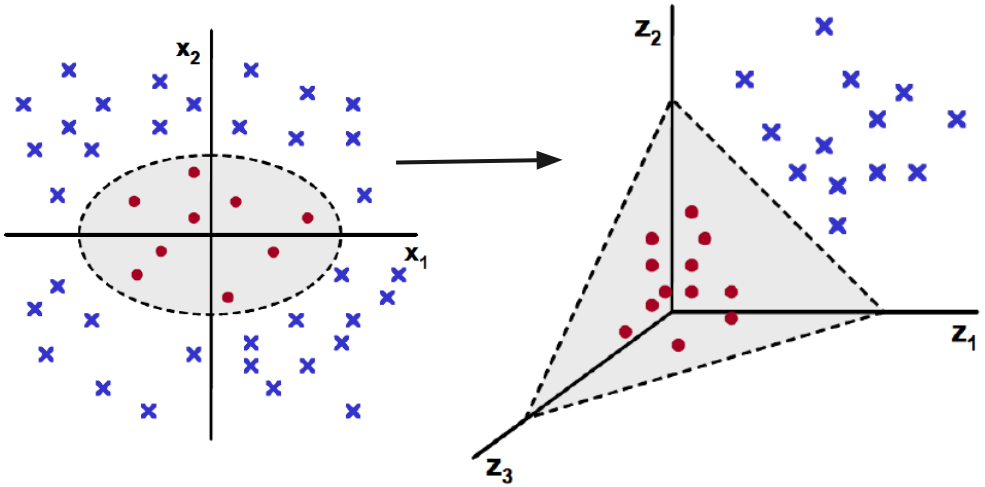
$$\begin{cases} \frac{1}{2}\langle w, w\rangle + C\sum_{i=1}^{\ell}\xi_i \to \min\limits_{w,w_0,\xi}; \\ y_i\big(\langle w, x_i\rangle - w_0\big) \geqslant 1 - \xi_i, \quad i = 1, \ldots, \ell; \\ \xi_i \geqslant 0, \quad i = 1, \ldots, \ell. \end{cases}$$



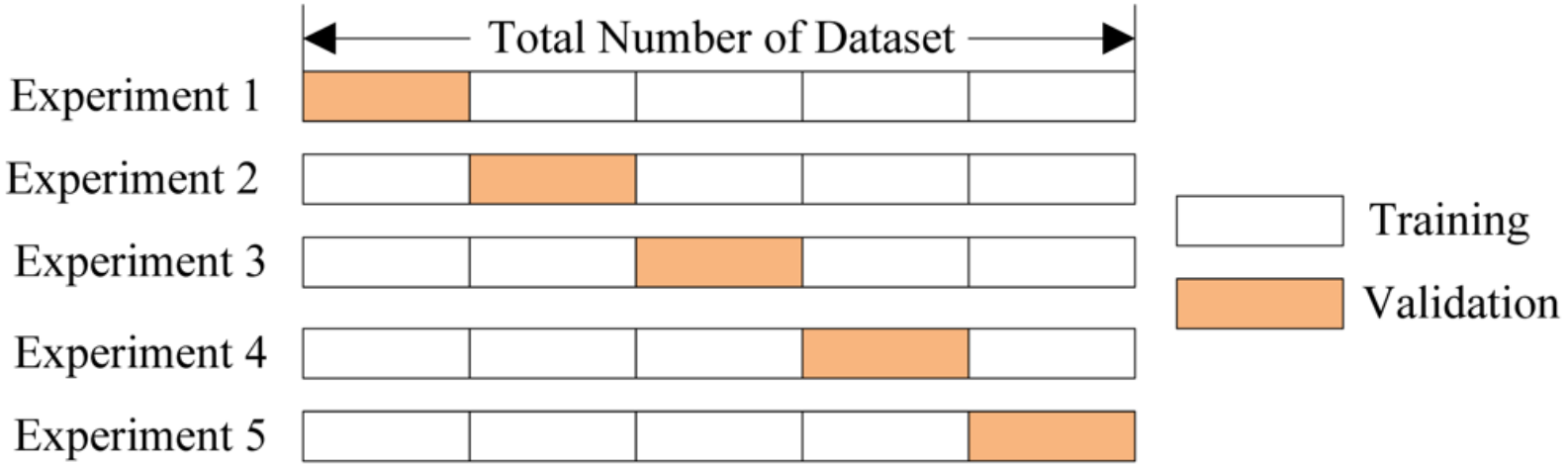
Параметры SVM

- Управляющий параметр алгоритма
- Ядро:
  - Линейное
  - RBF
  - Poly
- Параметры ядра

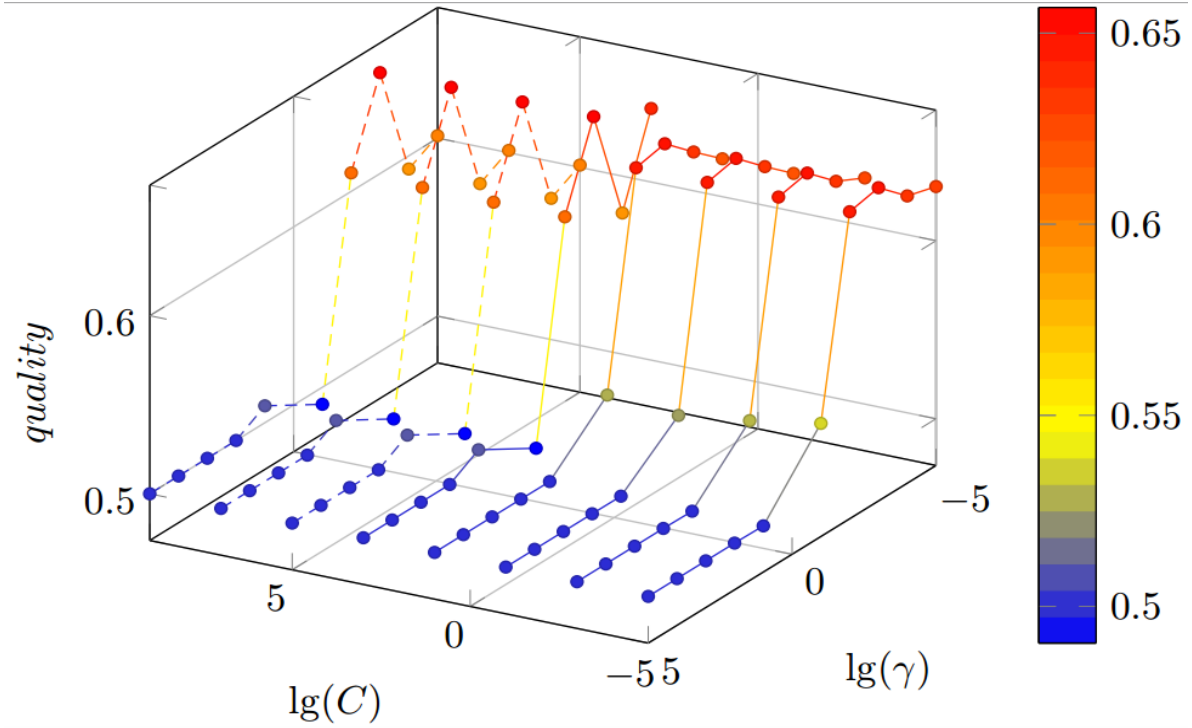
$$K(x, x') = \exp(-\gamma \|x - x'\|^2)$$



5-Fold Cross Validation



Подбор настроек SVM и результаты. ColorFeatureExtractor

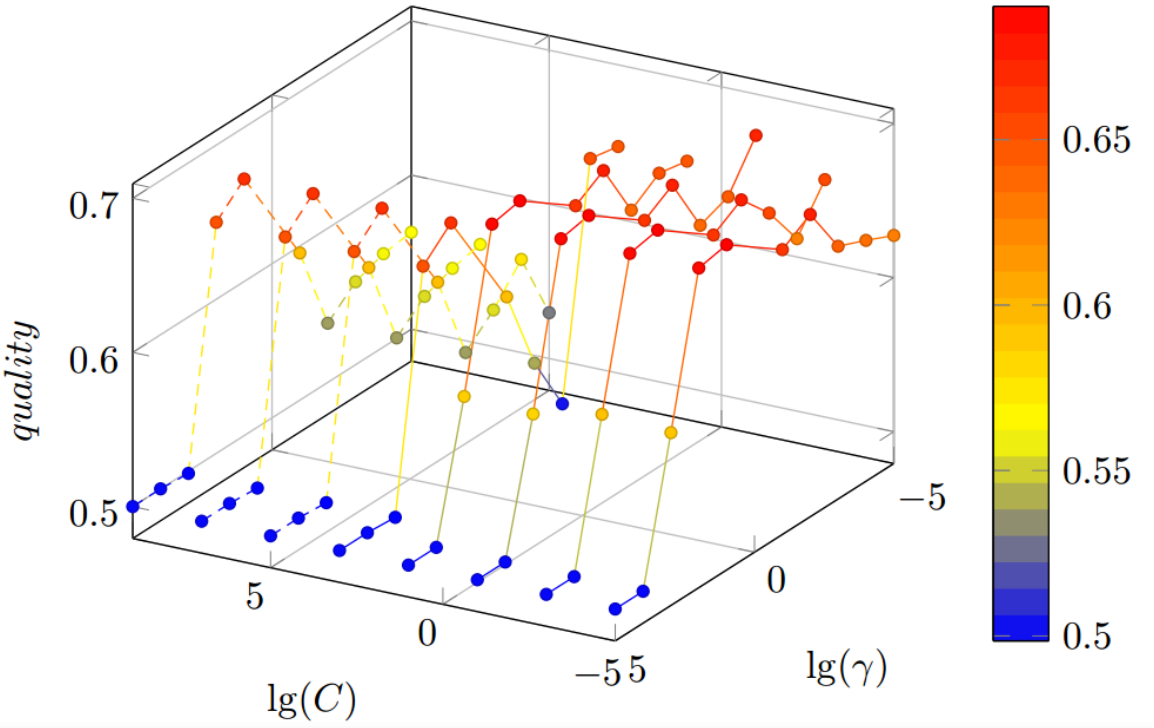


c = 10 … 10^{10}

gamma = 0.001

quality = 0.6567

Подбор настроек SVM и результаты. TextureFeatureExtractor



C = 0.00001 … 0.001

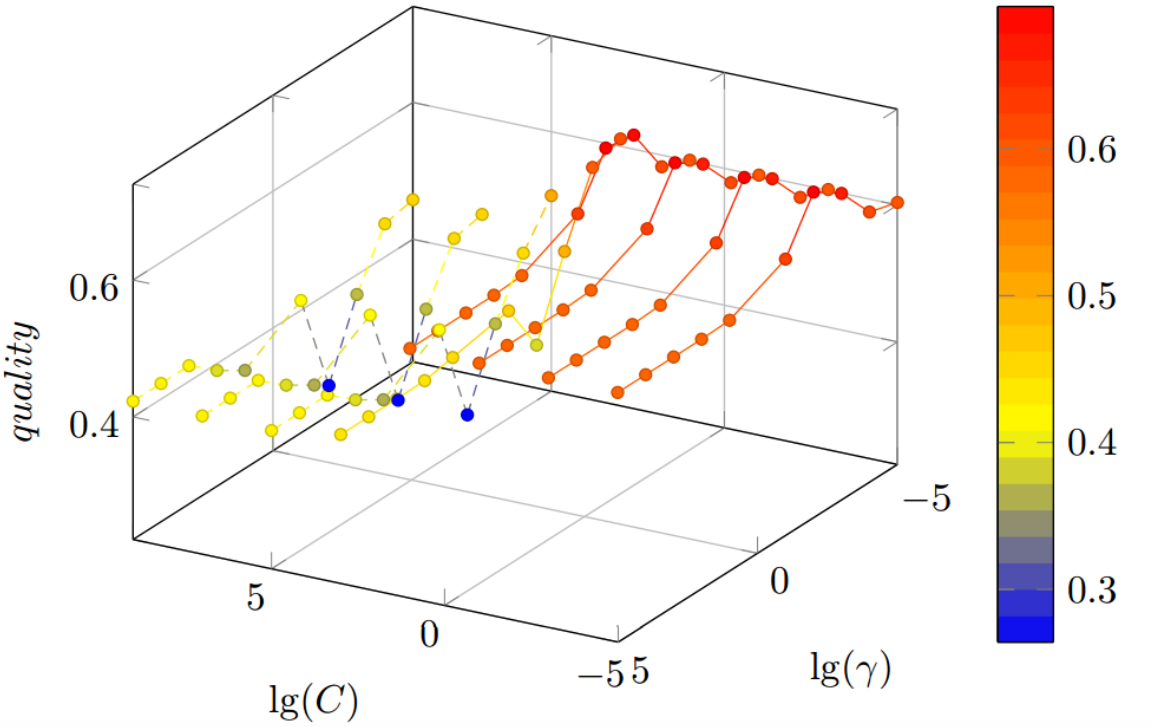
gamma = 0.01

quality = 0,6754

Результаты

Feature Extractor	Kernel	Params	Result
ConvNetFeatureExtractor	Poly	c = 1, degree = 3, gamma = 0	94%
ColorFeatureExtractor	RBF	c = 1000 gamma =0.001	**65,67%** **F_1(cat)=0.54** **F_1(dog)=0.55**
TextureFeatureExtractor	RBF	c = 0.001 gamma = 1	69% **F_1(cat)=** **F_1(dog)=**

Результаты обучения для нескольких классов



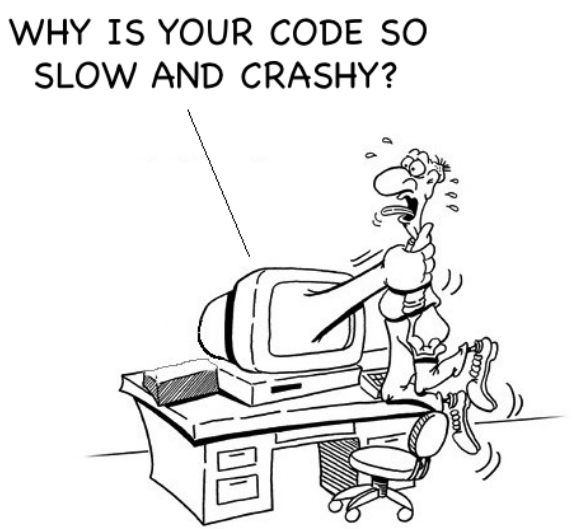
Классы:

- собаки
- лошади
- медведи
- слоны

quality = 0.6967

Lessons learned

1. Работа в команде
2. Python Perfomance
3. Чтение научной литературы



Репозиторий

https://github.com/ktisha/object_class_recognition

Литература

- Computer Vision: Algorithms and Applications, Richard Szeliski
- Learning OpenCV: Computer Vision with the OpenCV Library, Gary Bradski (Author), Adrian Kaehler
- OpenCV-Python Tutorials

Литература

- Алгоритмы обработка изображений:
  - Lindeberg, Tony (2001), "Edge detection", Encyclopedia of Mathematics, Springer, ISBN 978-1-55608-010-4
  - Rafael C. Gonzalez, Digital Image Processing 2001, ISBN-13: 978-0131687288, стр. 335

Литература

- Сверточные нейросети
  - arXiv:1310.1531v1 [cs.CV] 6 Oct 2013

    DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition
  - http://deeplearning.net
  - http://caffe.berkeleyvision.org/

Литература

- Классификация и верификация:
  - К.В. Воронцов “Лекции по линейным алгоритмам классификации”
  - А. Я. Червоненкис “Компьютерный анализ данных”