Отчет по контесту «Классификация изображений»

МФТИ, ФИВТ | группа 396а

по дисциплине «Анализ изображений»

Иван Чиликин

2018

Оглавление

[1. Задание 2](#_Toc503865138)

[2. Ход выполнения 2](#_Toc503865139)

[3. Заключение 4](#_Toc503865140)

1. Задание

Произвести распределение изображений по 6 классам:

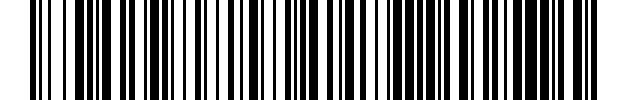
1. Aztec Code:



1. штрихкод вида:



1. штрихкод вида:



1. PDF417:



1. QR Code:



1. прочее.

Побить baseline 0.95 в соревновании miptqrcodesclf по классификации на платформе kaggle.

2. Ход выполнения

Для каждого класса изображений реализована соответствующая функция распознавания.

Aztec Code классифицируется по центральной «мишени»: подсчитывается сумма значений одной из компонент цвета в трех областях, и на основе этого определяется их цвет. На обучающем наборе классификатор выдал достоверный результат, однако на тестовом – не все изображения были корректно классифицированы, поскольку на некоторых из них имелись черные полосы по краям. После этого, была создана специальная функция проверки краев изображения. С ее применением распознавание тестовых картинок 0 класса стало выполняться без ошибок.

Далее, ввиду своей простоты, было реализовано распознавание изображений 2 класса. Соответствующая функция подсчитывает количество вертикальных черных линий шириной в 1 пиксель. Если их количество больше заданного (изменялось в ходе разработки), картинка относилась к данному классу. С его точки зрения, обучающий и тестовый наборы отличаются тем, что второй содержит перевернутые изображения. После обнаружения этого нюанса, стали подсчитываться также и горизонтальные линии, а их количество сравниваться с конкретным числом. В результате вышеописанной доработки все изображения тестового набора, относящиеся ко 2 классу, стали успешно распознаваться.

Класс 1 отличается от 2 тем, что на изображениях первого имеется текст. Это было принято во внимание при создании новой функции распознавания. Было замечено, что ширина текста приблизительно равна трети ширины кода/картинки, а высота кода составляет около 70% от общей высоты. В ходе первичной реализации была допущена неточность: подсчитывалось количество сплошных черных линий шириной в 1 пиксель, высотой не менее 0.6 и не более 0.8 от высоты картинки, ранее описанным способом и сравнивалось с третью ширины изображения, что вызвало коллизию с классификатором 0 класса. В итоговом варианте функции учитывается именно ширина текста и кода, что обеспечивает безошибочное распознавание картинок соответствующего класса.

Штрихкоды PDF417 имеют отличительную особенность: два прямоугольника по обоим краям кода, имеющие строго заданные размеры – ширина/высота в 21 и 24 пикселя. Этот критерий стал ключевым при распознавании изображений 3 класса.

Распознавание картинок, содержащих QR Code, аналогично применяемому для 0 класса за тем исключением, что на изображении должно находится 3 строго заданных квадрата («мишени»).

Тестовый набор данных состоит из картинок всех классов, которые могут быть перевернуты, а также содержать черные линии по краям, что было принято во внимание при реализации вышеупомянутых функций-классификаторов.

Важной особенностью программы является порядок вызовов функций классификации, который составлен в соответствии со строгостью их критериев, а именно: 0, 4, 3, 2, 1 и 5. Стоит отметить, что для удобства проверки ее работы производилась группировка распознанных картинок по папкам.

Все изображения, не удовлетворяющие критериям ни одного из первых пяти классов, относятся к 6.

3. Заключение

В результате была разработана программа, осуществляющая распознавание 5 классов изображений со штрихкодами и прочих (итого 6) со 100% точностью. В ходе ее реализации были закреплены навыки владения языком программирования Python и использования различных библиотек, расширяющих его возможности.