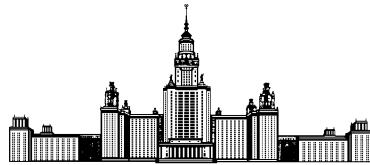


Московский государственный университет имени М. В. Ломоносова



Факультет Вычислительной Математики и Кибернетики
Кафедра Математических Методов Прогнозирования

ОТЧЁТ

«Генерация признаков формы объектов на изображении»

Лабораторная работа №2 по курсу
«Обработка и распознавание изображений»

Выполнил:

студент 3 курса 317 группы
Зиннуров Артём Ринатович

Научный руководитель:

к.ф-м.н.
Китов Виктор Владимирович

Москва, 2022

Содержание

1 Постановка задачи	2
2 Описание данных	2
3 Описание метода решения	3
4 Описание программной реализации	6
5 Эксперименты	8
6 Выводы	10

1 Постановка задачи

Требуется разработать и реализовать программу для классификации изображений ладоней.

Программа должна обеспечивать:

- ввод и отображение на экране изображений ладоней в формате **TIF**;
- сегментацию изображений на основе точечных и пространственных преобразований;
- определение «позы» ладоней;
- определение «линий пальцев» ладоней.

2 Описание данных

Исходные данные — набор из 67 цветных изображений левых ладоней разных людей, полученных с помощью сканера.

Вход программы — изображения ладоней в формате **TIF**.

Выход программы — исходное изображение ладони с нанесённым на нём кодом позы ладони, нарисованной линией пальцев и файл *Results.txt* следующего формата.

- Первая строка — код позы ладони: **1*2*3*4*5**
- Вторая строка — координаты точек ломаной линии пальцев:

```
!,0000.tif,T Xt1 Yt1,T Xt2 Yt2,T Xt3 Yt3,T Xt4 Yt4,T Xt5 Yt5,V Xv1 Yv1,V  
Xv2 Yv2,V Xv3 Yv3,V Xv4 Yv4,?
```

где (X_{ti}, Y_{ti}) и (X_{vi}, Y_{vi}) — координаты пикселей i -й найденной точки, определяющих кончики пальцев (T) и основания пальцев (V). Отсчёт идёт от большого пальца.

Код позы ладони — код вида **1*2*3*4*5**, где значок * обозначает -, если пальцы разомкнуты, или +, если они прижаты друг к другу. Пальцы нумеруются от 1 до 5, начиная с большого пальца против часовой стрелки.

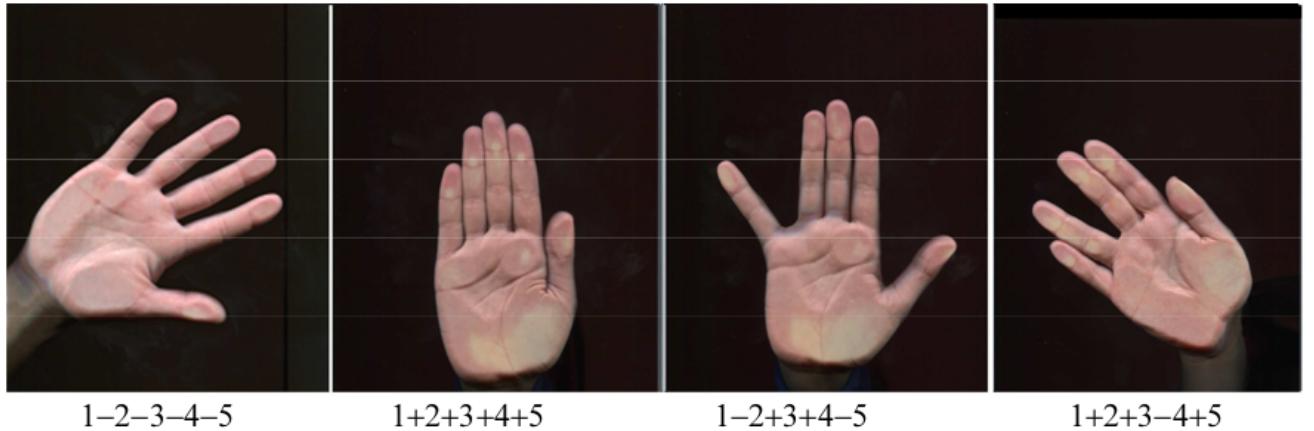


Рис. 1: Пример кода позы ладони.

Линия пальцев — ломаная линия, соединяющая точки на кончиках пальцев (tips) с точками в основаниях пальцев (valleys).

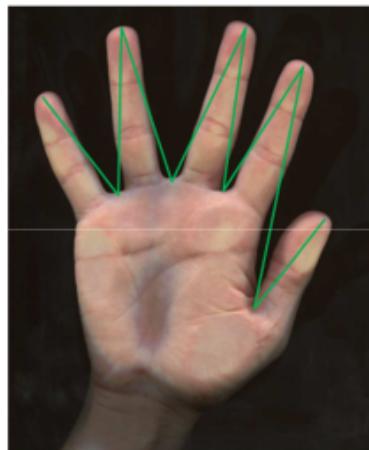


Рис. 2: Пример линии пальцев.

3 Описание метода решения

Решение данной задачи разбито на две функции: функция обучения и функция предсказания.

Функция обучения:

Функции обучения на вход передаётся имя директории, в которой расположены 67 исходных изображений левых ладоней разных людей. Строим скелет ладони специального вида, используя нейросетевую библиотеку **mediapipe**.

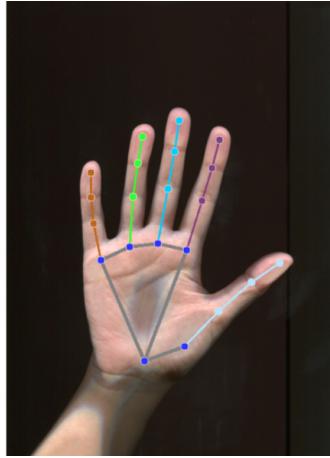


Рис. 3: Пример построенного скелета.

Данный скелет представляет из себя набор из 21 точки, которые заданы своими координатами

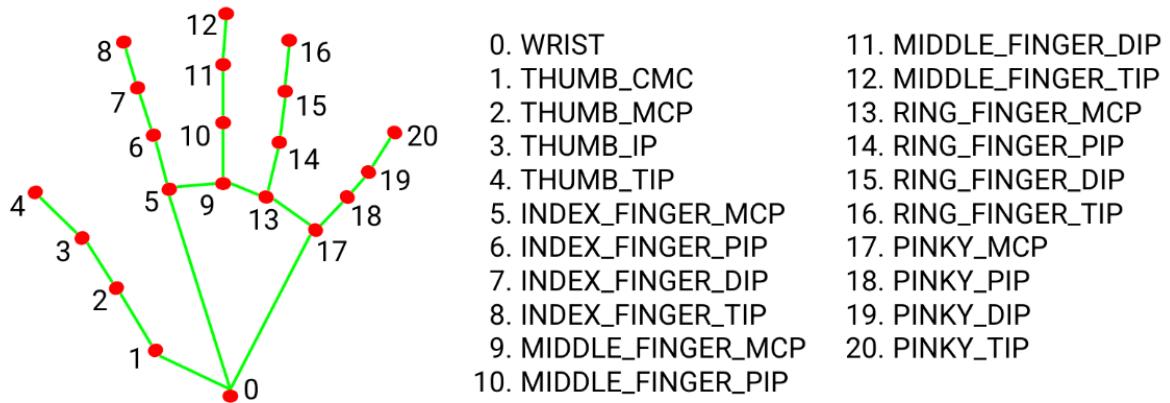


Рис. 4: Расположение, наименование и нумерация точек скелета.

Для определения взаимного расположения пальцев (прижаты или разомкнуты) произведём подсчёт расстояний между пальцами с помощью евклидова расстояния между следующими точками: 4 – 5, 8 – 12, 12 – 16, 15 – 20. Далее будем использовать полученные данные в качестве признаков для обучения классификаторов на одном признаке. Для этого необходимо иметь правильные метки, на которых будут

учиться алгоритмы и по которым в дальнейшем будут производиться замеры качества. Автор статьи лично разметил данные и приложил их в архив задания. Мы обучаем 4 бинарных классификатора для каждой последовательной пары пальцев. В процессе разметки встречались откровенно спорные случаи, но автор их почему-то не выкинул...



Рис. 5: Считаем ли мы большой палец прижатым?

Были опробованы несколько алгоритмов машинного обучения, реализованные в библиотеке `sklearn`: линейный классификатор, логистическая регрессия и SVM с RBF ядром. Все методы показывают примерно одинаково хорошие результаты точности на тестовой выборке. В качестве решающего алгоритма была выбрана логистическая регрессия. В процессе работы данной функции печатаются значения точностей для соответствующих пар пальцев. Результатом работы функции обучения являются обученные бинарные классификаторы. При выполнении данной функции происходит также удаление файла `Results.txt`, если он существует.

Функция предсказания:

Функции предсказания передаётся либо имя файла с изображением ладони, либо имя директории, в которой лежат файлы с изображением ладоней, а также обученные бинарные классификаторы. Для каждого изображения ладони происходит его считывание, построение скелета и получение координат всех точек скелета. Далее вычисляются необходимые характеристики. Идёт подсчёт расстояний между соответствующими пальцами для определения взаимного расположения пальцев (прижаты или разомкнуты), к результатам применяются обученные бинарные классификаторы

и формируется код позы ладони, который рисуется на выходном изображении. Далее происходит подсчёт координат оснований пальцев как среднее соответствующих координат следующих пар точек скелета: 5 – 9, 9 – 13, 13 – 17, которое сдвинуто в направлении $\frac{1}{8}$ вектора-суммы ближайших костяшек: 6 – 5 и 10 – 9, 10 – 9 и 14 – 13, 14 – 13 и 18 – 17 соответственно. Отдельно обрабатывается впадина у большого пальца, которая должна быть сильно ближе к большому пальцу. Для вычисления её координат производится взвешенное суммирование точек 2 и 5 с весами $\frac{3}{4}$ и $\frac{1}{4}$. В качестве кончиков пальцев берутся следующие точки скелета: 4, 8, 12, 16, 20, которые сдвинуты к краю на $\frac{1}{4}$ от длины последней костяшки: 4 – 3, 8 – 7, 12 – 11, 16 – 15, 20 – 19. На выходном изображении рисуется линия пальцев, после чего оно выводится. В конце происходит запись необходимых данных в файл *Results.txt*. При применении данной функции к целой директории все записи производятся в один файл, а также перед записью данных удаляется старый файл *Results.txt*, если он существовал.

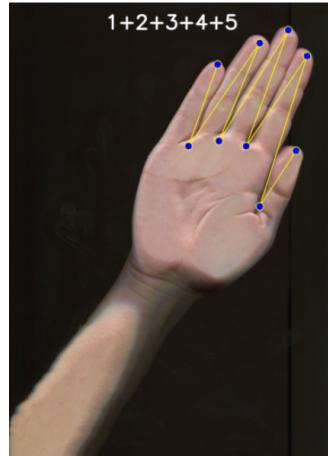


Рис. 6: Пример результата.

4 Описание программной реализации

Весь код написан на языке **Python** в интерактивной вычислительной среде **Jupyter Notebook**.

При решении данного задания используются следующие библиотеки:

- библиотека машинного обучения **scikit-learn**;
- библиотека обработки изображений **scikit-image**;

- библиотека компьютерного зрения **Open CV**;
- библиотека детекции рук **mediapipe**;
- библиотека визуализаций **matplotlib**;
- и другие библиотеки: **os**, **numpy**, **pathlib**.

Функции и классы из сторонних библиотек:

- `skimage.io.imread()` - считывает изображение;
- `plt.imshow()` - выводит изображение;
- `cv.putText()` - рисует текст на изображении;
- `cv.line()` - рисует линию на изображении;
- `cv.circle()` - рисует круг на изображении;
- `numpy.linalg.norm` - считает векторную норму;
- CLASS `mp.solutions.hands.Hands` - класс, описывающий ладонь;
- CLASS `mp.solutions.drawing_styles` - класс стилей отображения;
- CLASS `mp.solutions.drawing_utils` - класс для рисования;
- `sklearn.metrics.accuracy_score()` - считает точность;
- `sklearn.model_selection.train_test_split()` - разделяет выборку на обучающую и тестовую;
- CLASS `sklearn.linear_model.RidgeClassifier` - класс линейного классификатора;
- CLASS `sklearn.linear_model.LogisticRegression` - класс логистической регрессии;
- CLASS `sklearn.linear_model.SVC` - класс SVM классификатора.

Собственно реализованные функции:

- `print_img(img, figsize=(10, 10))` - печатает изображение;
- `OiR_fit(dir_name, test_size=0.33, random_state=42)` - функция обучения;
- `OiR_predict_solo(file_name, clfs)` - функция обработки одного изображения ладони;
- `OiR_predict_all(dir_name, clfs)` - функция обработки всех изображений ладони в данной директории.

5 Эксперименты

Проведены эксперименты на исходных данных.

```
Test accuracy 12: 1.0000
Test accuracy 23: 0.9130
Test accuracy 34: 0.9130
Test accuracy 45: 0.9565
```

Рис. 7: Значения точности классификации пальцев на тестовой выборке, используя разметку автора.

Подсчёт точности построения линии пальцев не представляется возможным.

Ниже представлены несколько примеров работы программы:

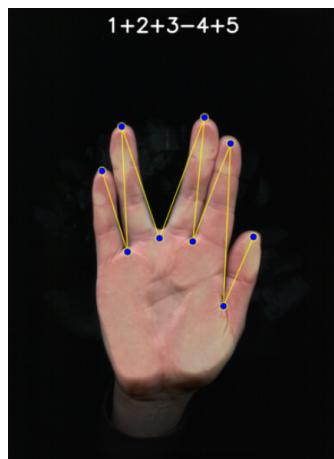


Рис. 8: Пример №1.

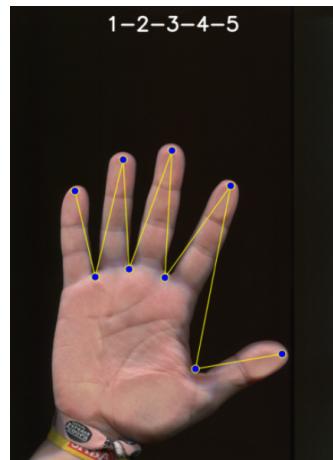


Рис. 9: Пример №2.

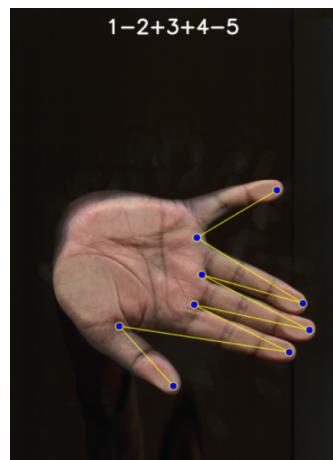


Рис. 10: Пример №3.

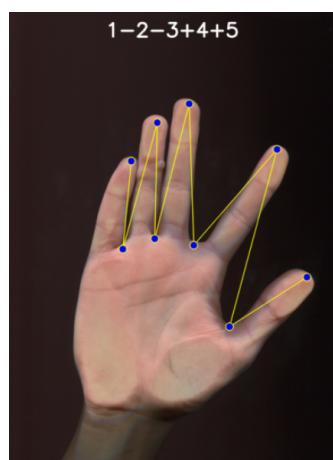


Рис. 11: Пример №4.

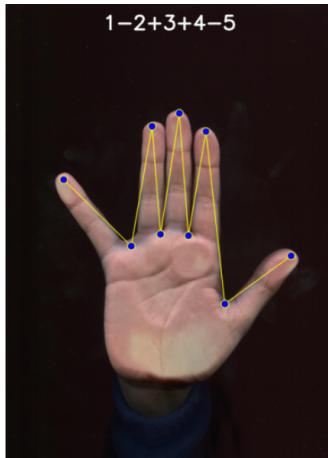


Рис. 12: Пример №5.

6 Выводы

Трудно сделать какие-либо выводы о качестве данного решения: тестовых данных было совсем не много, к тому же большинство неправильных ответов в задаче определения кода позы ладони были даны в спорных ситуациях, в которых даже сам автор не знает точного ответа. Для оценки качества решения задачи построения линии пальцев необходимо знать критерии оценки и правильные координаты. А этого мы не знаем.

Таким образом, в ходе выполнения данного задания произведено знакомство с популярными библиотеками компьютерного зрения **OpenCV** и **scikit-image**. Изучена новая нейросетевая библиотека для детекции рук **mediapipe**. Применены алгоритмы машинного обучения для решения поставленной задачи. Получен практический опыт по написанию программ на языке **Python**, составлению отчётов в системе компьютерной вёрстки **LaTeX**.