

Санкт-Петербургский Государственный Университет
Факультет Прикладной Математики и Процессов Управления

Отчет по практическому занятию №3
“СОЗДАНИЕ ПРОГРАММЫ ДЛЯ МОДЕЛИРОВАНИЯ
СИСТЕМ РАСПОЗНАВАНИЯ ЛЮДЕЙ ПО ЛИЦАМ”
в рамках курса “Прикладные задачи построения
современных вычислительных систем”

Выполнила студентка
группы 17.Б12-пу
Барташук Анастасия Евгеньевна

Содержание

Содержание	2
Цель работы	3
Описание программы	4
Анализ признаков	6
Гистограмма яркости	6
DFT	8
DCT	10
Scale	11
Градиент	13
Кросс-валидация. Поиск оптимальных значений.	16
Гистограмма яркости	16
DFT	17
DCT	18
Scale	19
Градиент	20
Исследование изменения размера тестовой выборки	21
Гистограмма яркости	21
DFT	21
DCT	22
Scale	22
Градиент	23
Заключение	24

Цель работы

Написать программу для классификации лиц, основывающуюся на признаках, задающих признаковое пространство для изображений лиц: гистограмма яркости, DFT, DCT, Scale, Градиент; так чтобы классификатор реализовывался по критерию минимума расстояний программы. Провести эксперименты (для каждой категории признаков). Запустить написанную программу и провести ее тестирование, подобрать значения параметров для близкого к 100% распознавания. Дальнейшие исследования провести в режиме кросс-валидации исходных данных. Определить параметры системы, обеспечивающие результат, близкий к 100%. Сформировать график искомых зависимостей (результат распознавания в зависимости от числа изображений). Сделать выводы по проделанной работе.

Описание программы

Программа, описанная в целях данной работы, была реализована на языке python 3 в Visual Studio Code при использовании следующих библиотек: numpy, matplotlib.pyplot, random, sklearn, os, scipy, cv2, PIL, tkinter (для пользовательского интерфейса), io и еще несколько других. В качестве базы данных лиц была взята база данных Olivetti, в ней 400 изображений (40 человек по 10 фотографий для каждого).

Пользовательский интерфейс (Рисунки 1 и 2) состоит из меню, расположенного слева, области вывода результатов справа и ползунков снизу и справа, предназначенных для скроллинга рабочей области, когда результаты выходят за размеры окна. В меню можно выбрать интересующий нас метод, задать ему некий параметр, чтобы посмотреть, как он работает на примере из 5 изображений (2 лица одного человека, 2 лица другого и 1 третьего) (кнопка “Далее”) и на всей базе данных (кнопка “FULL”), а также запустить автоматическое определение оптимального параметра для метода и наилучшего числа эталонов в каждом классе (кнопка “Вычислить”).

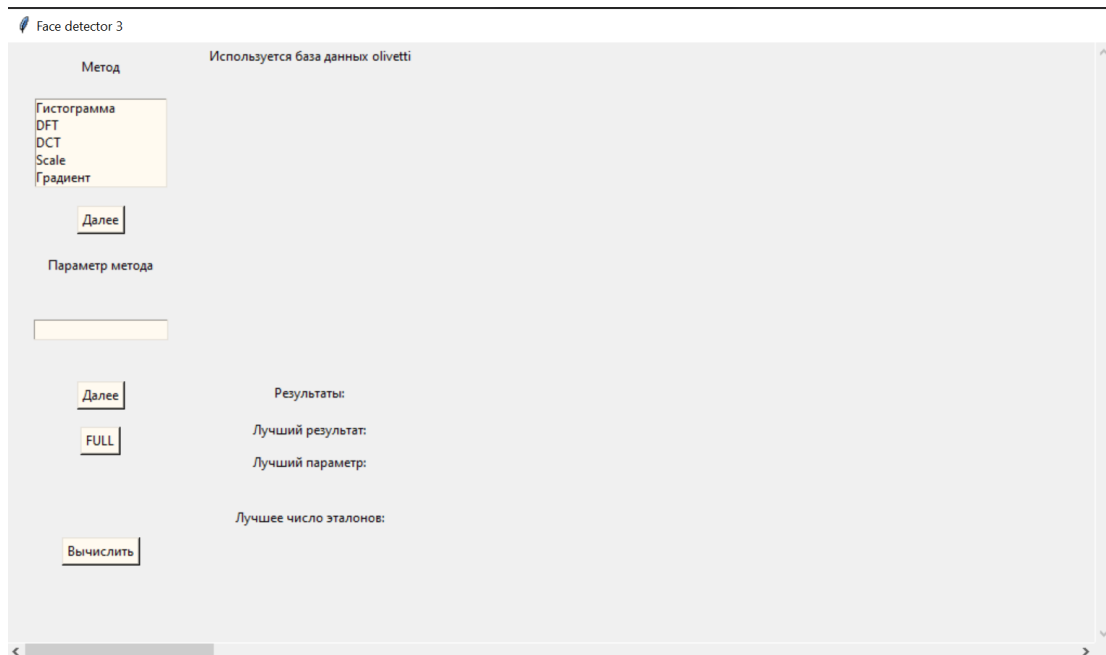


Рисунок 1. Интерфейс программы до выбора метода

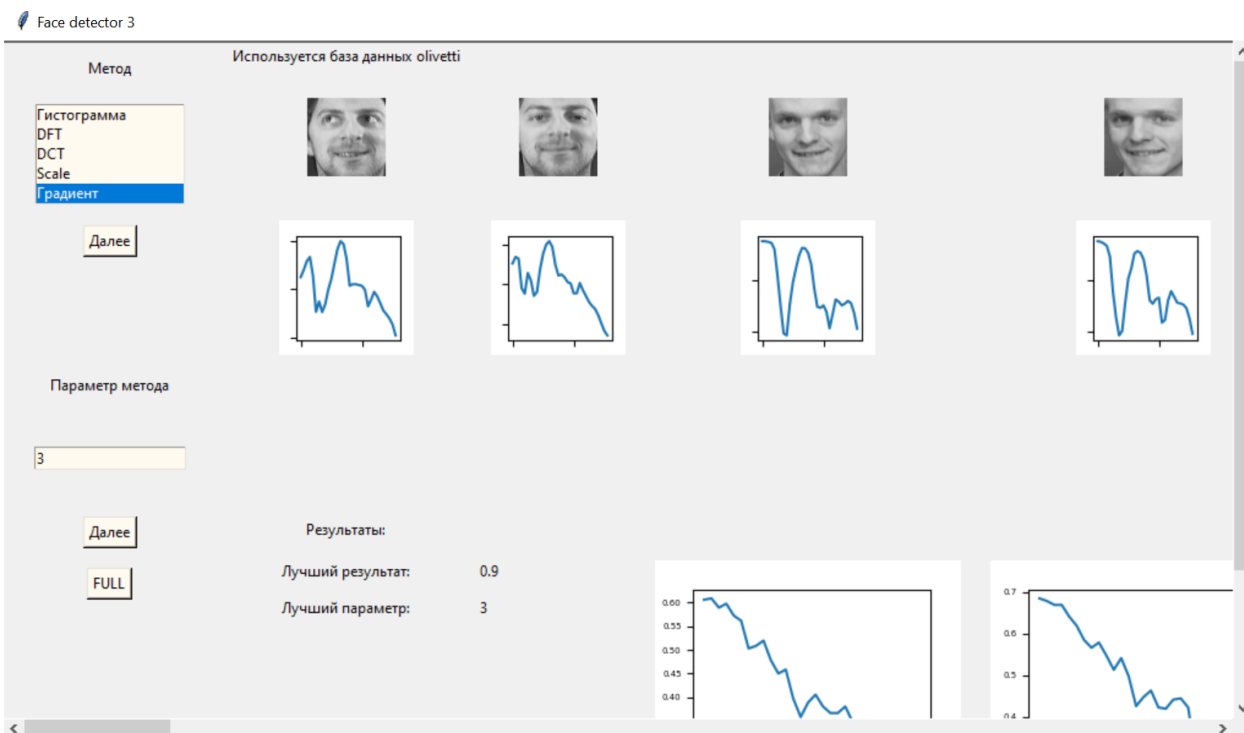


Рисунок 2. Интерфейс программы после нажатия кнопок “Далее” и “Вычислить”

Анализ признаков

Признаковое пространство для изображений лиц задается у нас следующими признаками: гистограммой яркости, DFT, DCT, Scale и градиентом. Результаты их работы для удобства мы рассмотрим на 5 изображениях из базы данных: две пары лиц, принадлежащих двум людям, и одно, принадлежащее третьему человеку. Результаты работы над остальными изображениями можно посмотреть, нажав на кнопку “FULL” в нашем приложении.

Гистограмма яркости

Это график статистического распределения элементов цифрового изображения с различной яркостью, в котором по горизонтальной оси представлена яркость, а по вертикали - относительное число пикселей с конкретным значением яркости. Параметром тут будет являться количество “столбцов” на оси X.

Рассмотрим различные значения параметров. Логично, что значения меньше 8-9 брать не стоит, так как различия между гистограммами будут минимальными, без каких-то отличительных особенностей (Рисунок 3). Возьмем в качестве примера нижней границы параметра значение равное 9 (Рисунок 4).

Теперь возьмем какое-нибудь очень большое значение параметра, например, 350 (хотя, на самом деле, где-то после 290 гистограммы меняются не сильно) (Рисунок 5).

Далее возьмем два значения где-то между 9 и 350: 50 и 70 (Рисунки 6 и 7). Как мы видим, в промежутке между этими двумя значениями у нас пропадает некая гладкость графика. Поэтому предположим, что оптимальный параметр находится где-то между 9 и 50.

Также интересным наблюдением будет то, что в обоих случаях, когда у нас лица принадлежат одним людям, графики гистограмм достаточно похожи внешне. Между лицами же разных людей наблюдаются различные тенденции, например, у первого человека три глобальных пика: в начале, конце и посередине, а у последнего только два: в начале и конце, гистограмма же изображения лица второго человека похожа на пень.

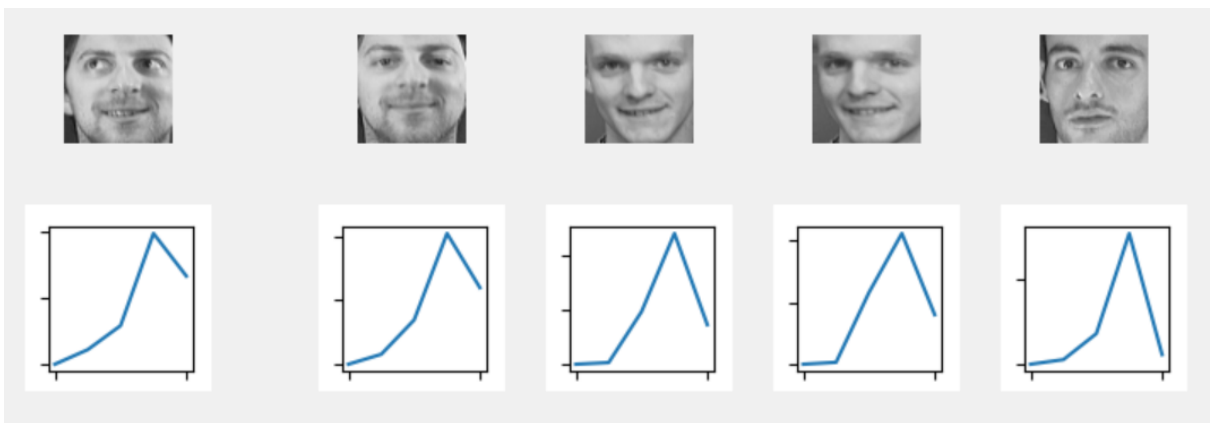


Рисунок 3. Результат работы гистограммы для параметра = 5

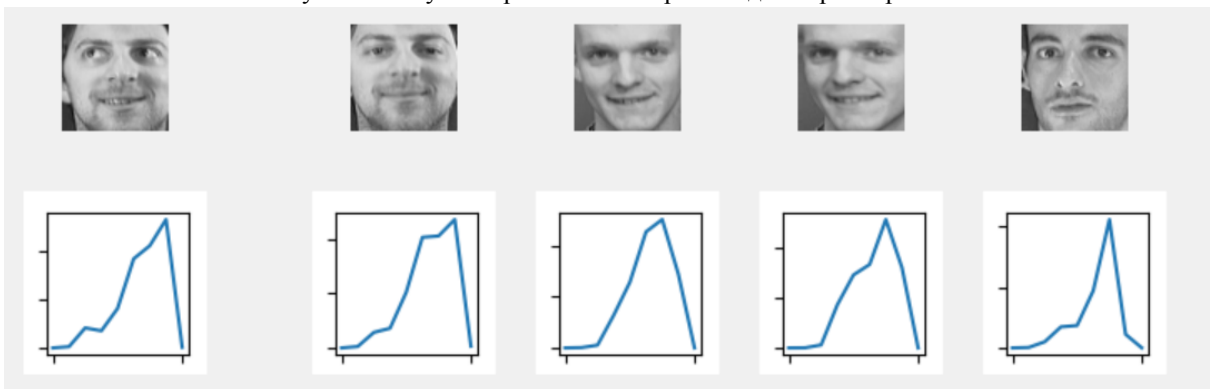


Рисунок 4. Результат работы гистограммы для параметра = 9

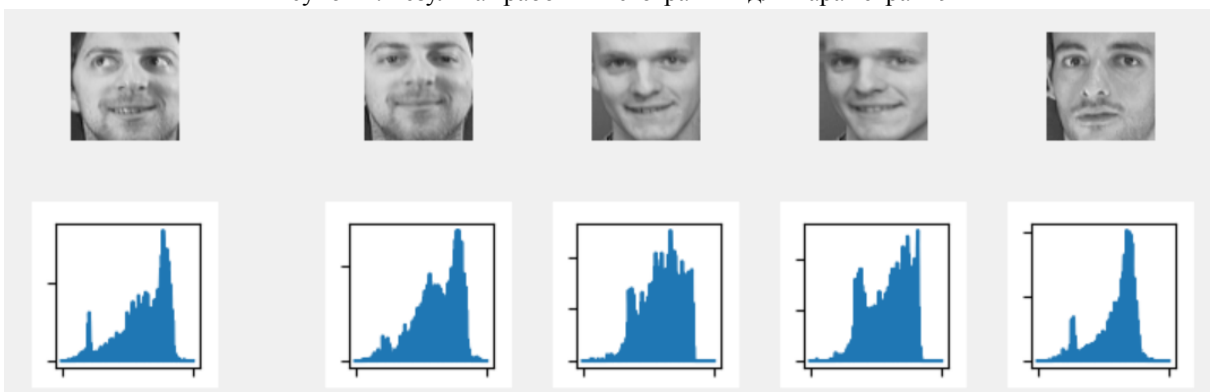


Рисунок 5. Результат работы гистограммы для параметра = 350

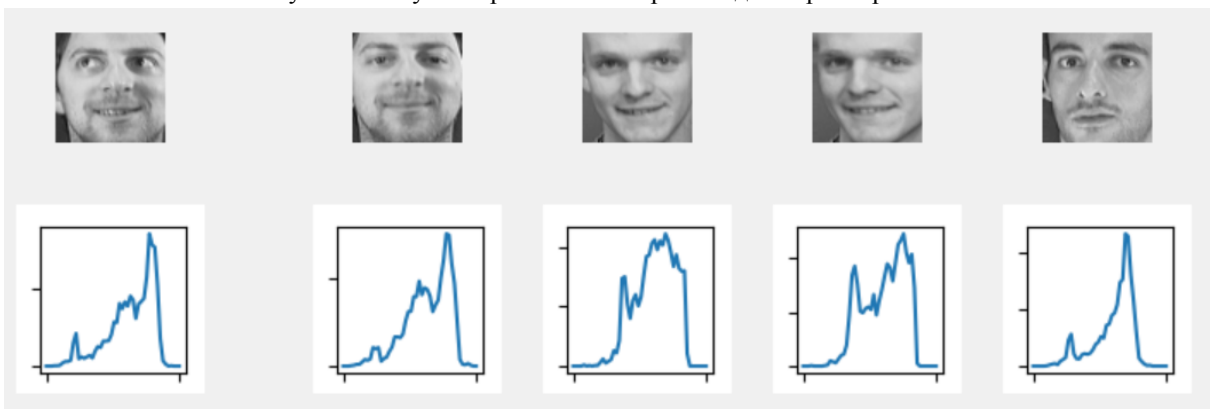


Рисунок 6. Результат работы гистограммы для параметра = 50

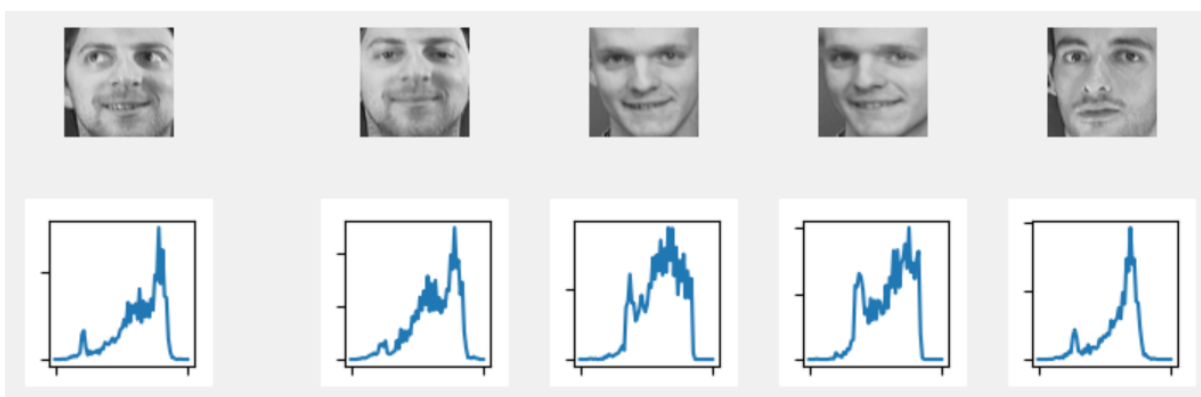


Рисунок 7. Результат работы гистограммы для параметра = 70

DFT

Дискретное преобразование Фурье на практике осуществляется с помощью быстрого преобразования Фурье (FFT, Fast Fourier Transform). Для прямоугольного дискретного изображения f (матрицы, массива) алгоритм FFT реализован на питоне функцией `fft2()`. Она возвращает преобразование Фурье той же размерности. Наиболее выраженным участком полученной матрицы является ее левая верхняя подматрица, нам остается только определить ее оптимальную размерность (что и будет нашим искомым параметром).

Теперь посмотри как работает данный метод на различных параметрах. Минимальным параметром, который нам необходимо взять, чтобы что-то увидеть, будет 3 (Рисунок 8), однако, как мы видим, в нашей задаче он нам мало чем поможет. Увеличим параметр в два раза 4 раза и посмотрим на результаты (Рисунки 9-12). Как мы видим, уже при значениях больше 24 мы имеем достаточно большое белое поле в правом нижнем углу, что выглядит достаточно неэффективным, следовательно оптимальный параметр будет где-то около 12 и 24.



Рисунок 8. Результат работы DFT для параметра = 3

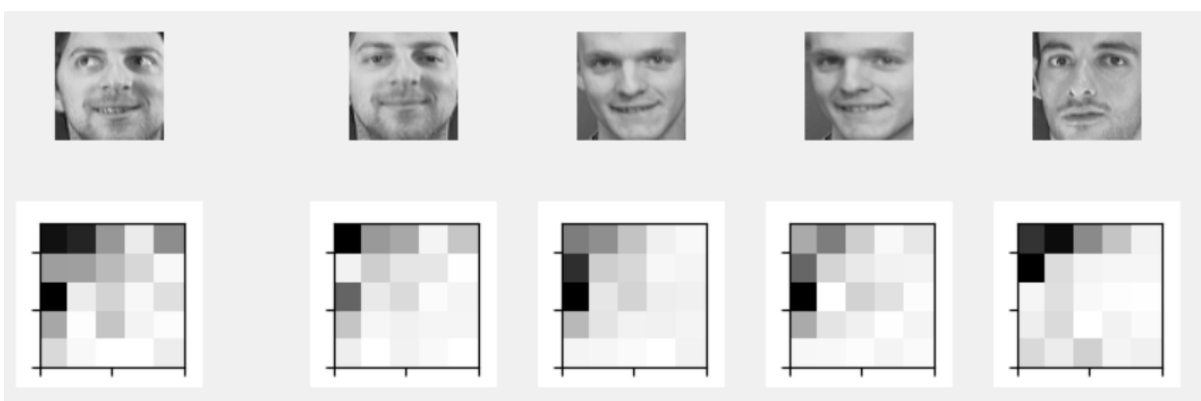


Рисунок 9. Результат работы DFT для параметра = 6

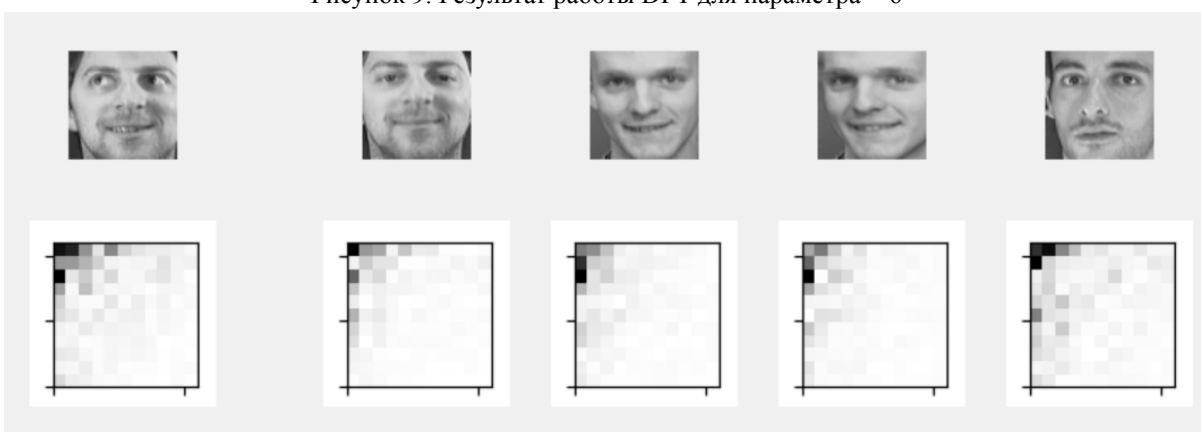


Рисунок 10. Результат работы DFT для параметра = 12

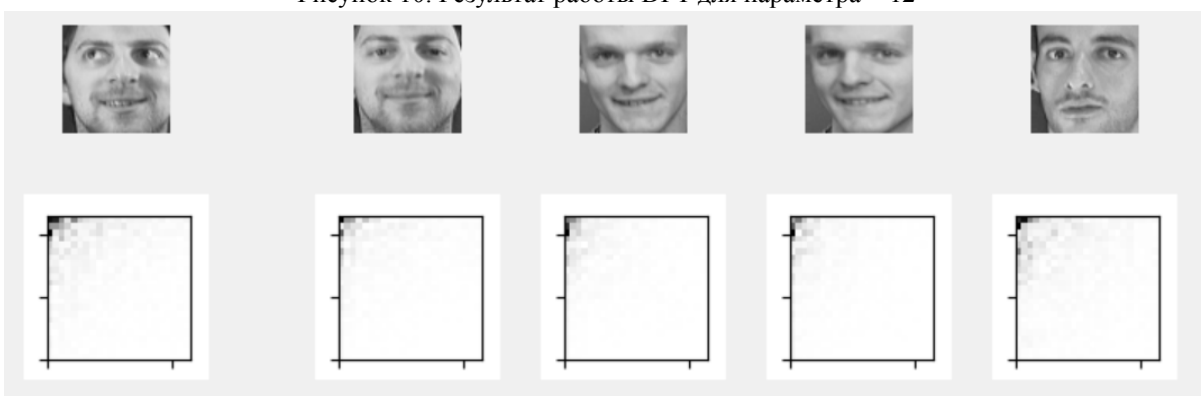


Рисунок 11. Результат работы DFT для параметра = 24

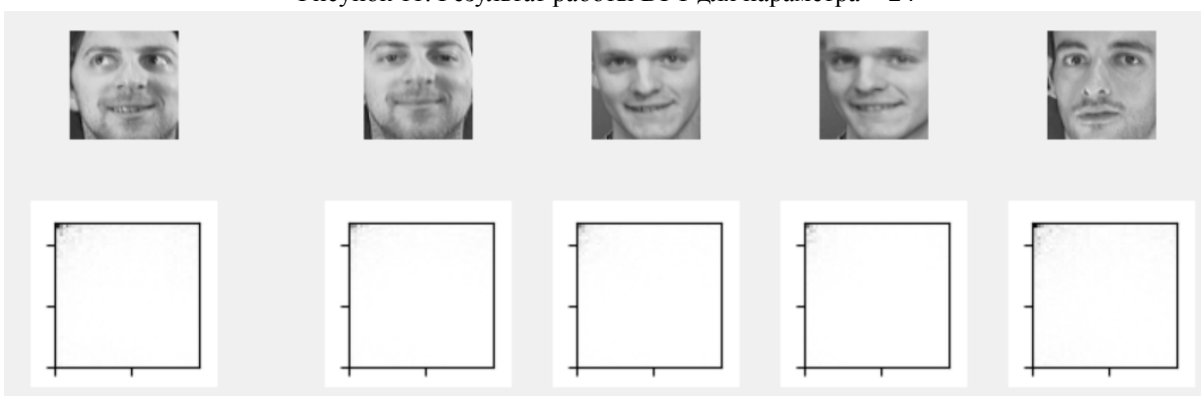


Рисунок 12. Результат работы DFT для параметра = 48

DCT

Дискретное косинусное преобразование (англ. Discrete Cosine Transform, DCT) — одно из ортогональных преобразований. Вариант косинусного преобразования для вектора действительных чисел. Это преобразование тесно связано с дискретным преобразованием Фурье и является гомоморфизмом его векторного пространства. Также как и в предыдущем признаке мы будем рассматривать левую верхнюю подматрицу и менять ее размерность.

Исследовать параметр мы будем с теми же значениями: 3, 6, 12, 24 и 48.

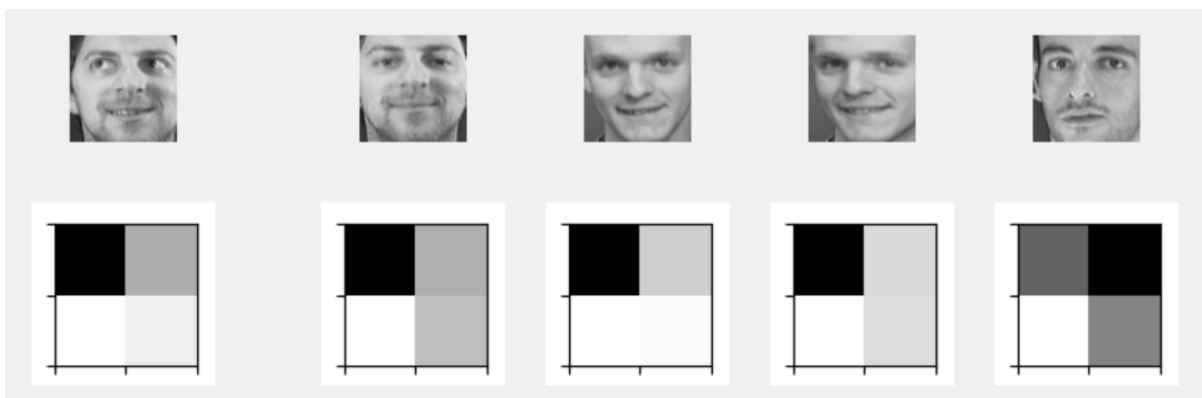


Рисунок 13. Результат работы DCT для параметра = 3

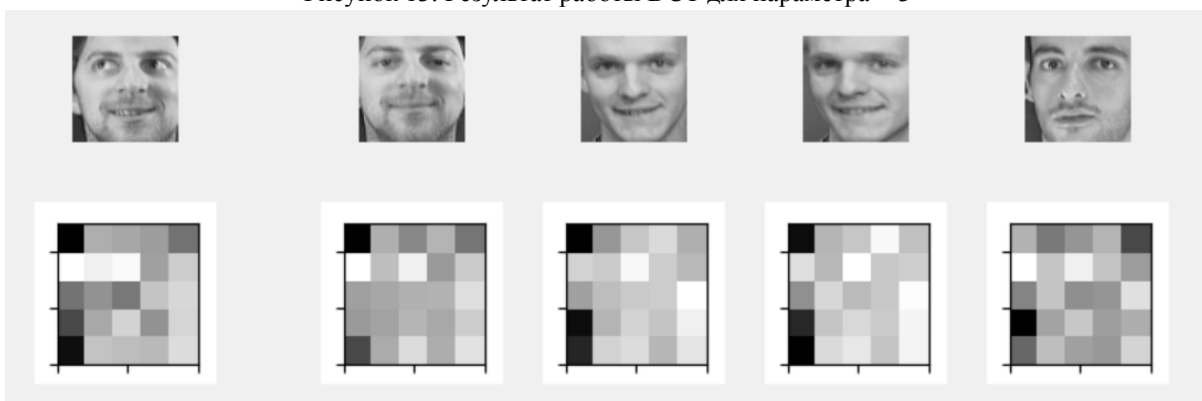


Рисунок 14. Результат работы DCT для параметра = 6

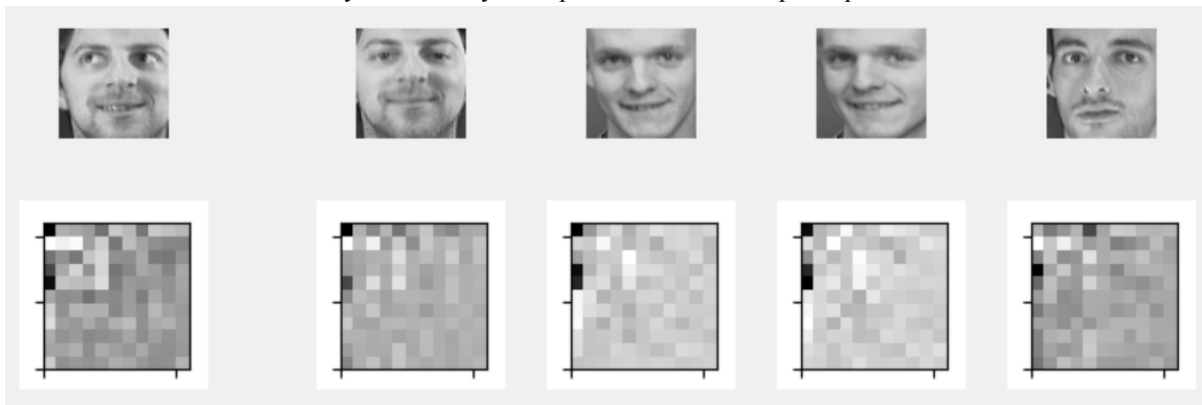


Рисунок 14. Результат работы DCT для параметра = 12

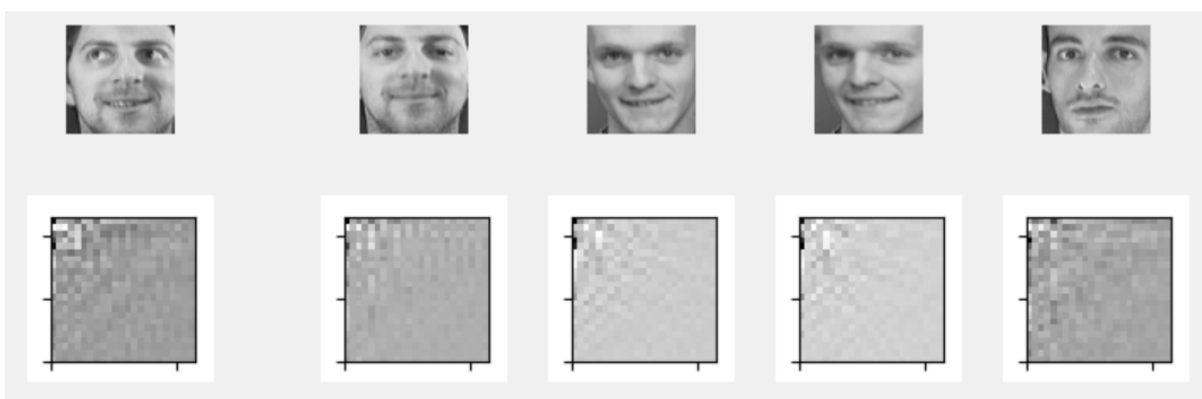


Рисунок 15. Результат работы DCT для параметра = 24

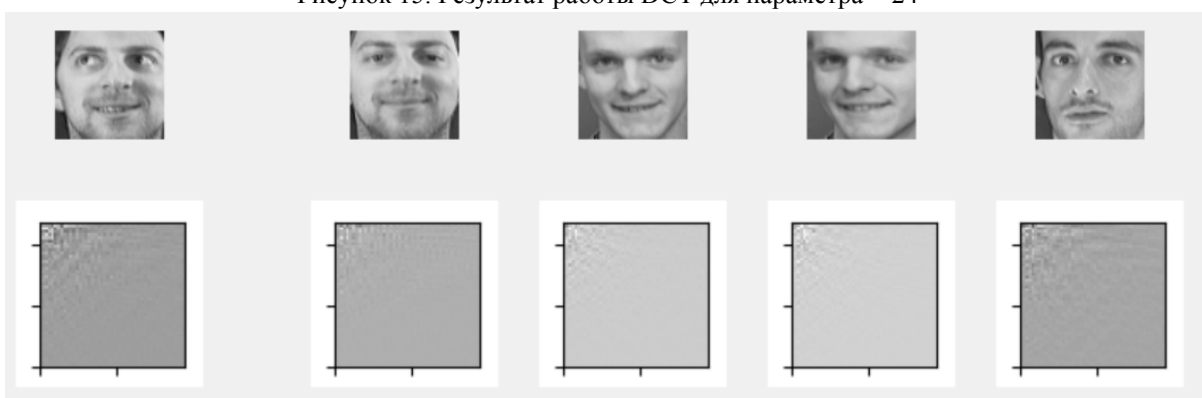


Рисунок 16. Результат работы DCT для параметра = 48

Получаем результаты аналогичные предыдущему признаку, однако вместо белого цвета в этом случае мы имеем монотонные оттенки серого в правом нижнем углу.

Scale

Метод масштабирования изображения: превращает его в картинку с меньшей размерностью. Чаще всего масштабирование данных происходит в диапазоне от 0 до 1. Мы будем исследовать оптимальное значение из этого диапазона.

Параметр равный 1 просто возвращает нам исходное изображение (Рисунок 17). При параметре 0.05 изображение сильно сжимается, часть информации теряется, поэтому он также не подойдет (Рисунок 18). Рассмотрим еще три варианта: 0.1, 0.5 и 0.75 (Рисунки 19-21). Как мы видим, при 0.75 изображение, как будто, немного блюрится, при 0.5 - уже видна пиксельность, но лицо все еще угадывается, а при 0.1 - изображение уже сильно разложено на пиксели, лицо трудно, но угадывается. Отсюда можно предположить, что оптимальный параметр будет лежать где-то между 0.1 и 0.5, наверное, ближе к 0.1.

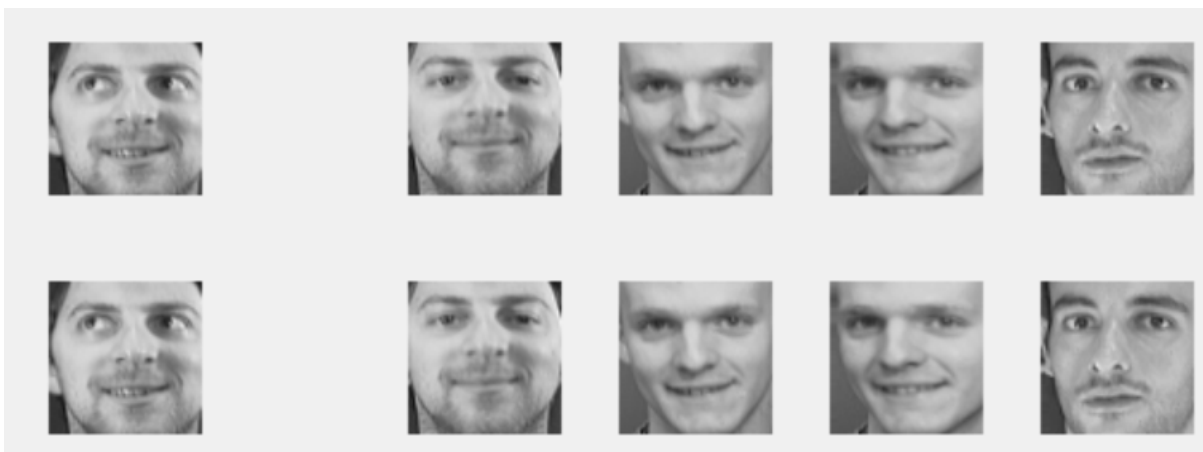


Рисунок 17. Результат работы Scale для параметра = 1



Рисунок 18. Результат работы Scale для параметра = 0.05

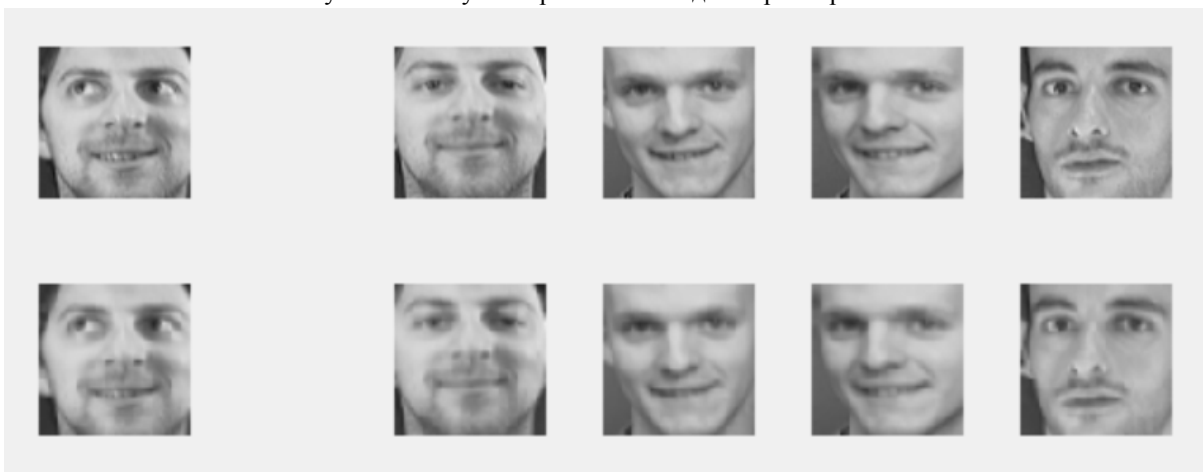


Рисунок 19. Результат работы Scale для параметра = 0.75

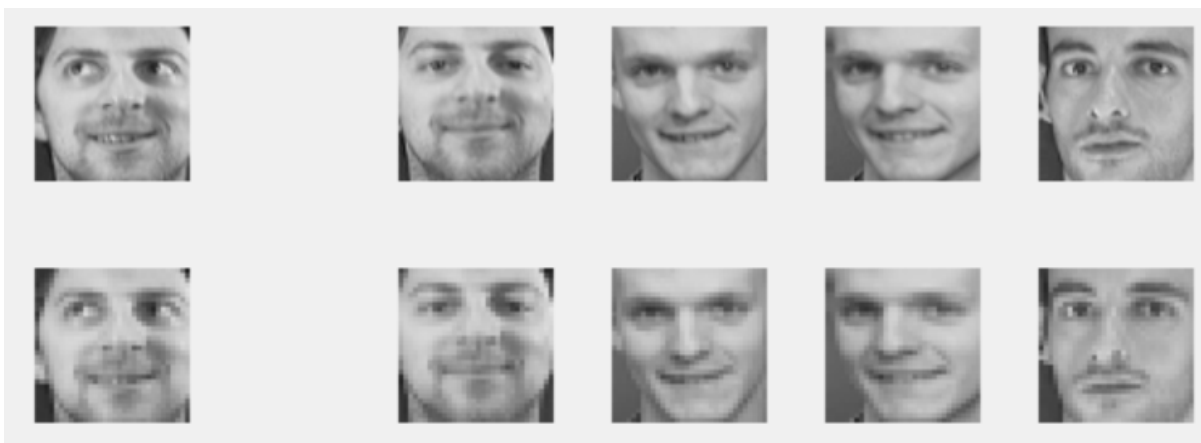


Рисунок 20. Результат работы Scale для параметра = 0.5



Рисунок 21. Результат работы Scale для параметра = 0.1

Градиент

Данный подход основан на подсчете направлений градиента яркости (интенсивности) в локальных областях изображения и базируется на том факте, что распределение градиентов яркости на каком-либо участке изображения дает представление о внешнем виде и форме объекта, расположенного на этом участке. Суть метода состоит в том, что изображение разбивается на области некоторой высоты, для каждой из которых вычисляется сумма яркостей. Совокупность построенных сумм будет являться градиентом в виде гистограммы. Наша цель: найти оптимальную высоту такой сетки.

Значение данного параметра ограничено снизу 2, так как меньше не имеет смысла, а сверху - размером нашего изображения. В нашем случае размерность равна 64. В качестве рассматриваемых значений возьмем: 32, 16, 8, 4, 2 (Рисунки 22-26).

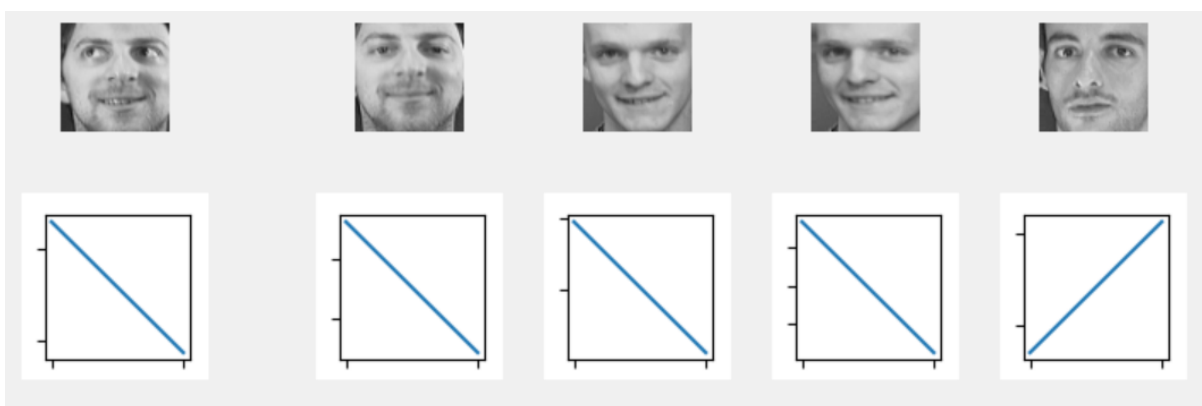


Рисунок 22. Результат работы градиента для параметра = 32

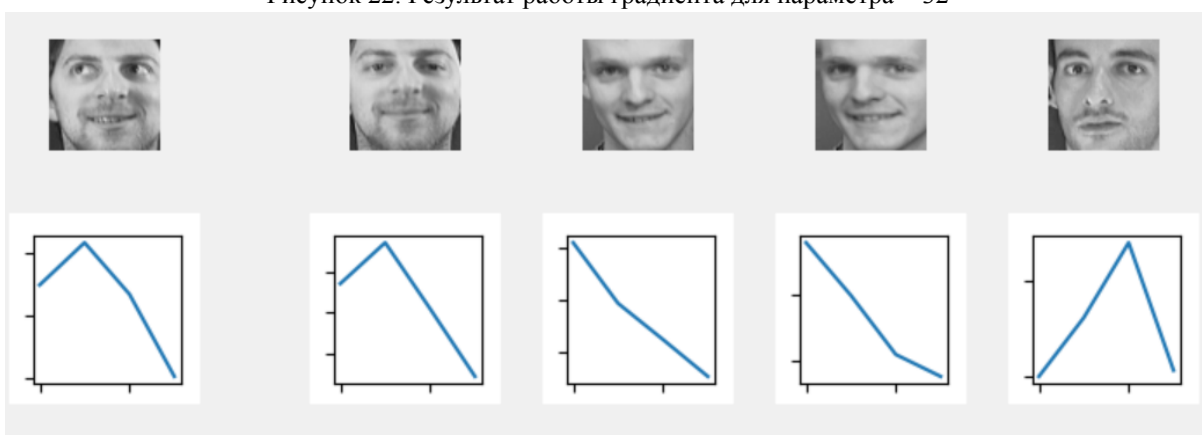


Рисунок 23. Результат работы градиента для параметра = 16

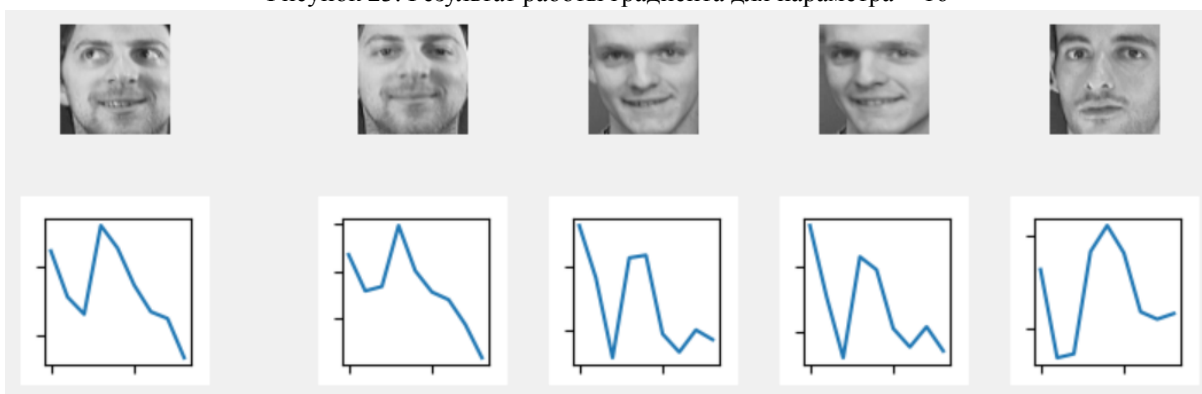


Рисунок 24. Результат работы градиента для параметра = 8

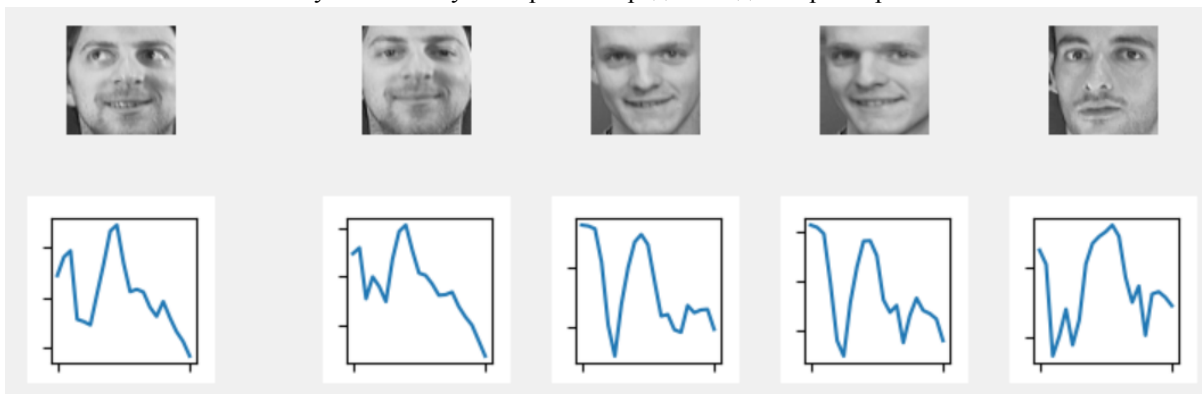


Рисунок 25. Результат работы градиента для параметра = 4

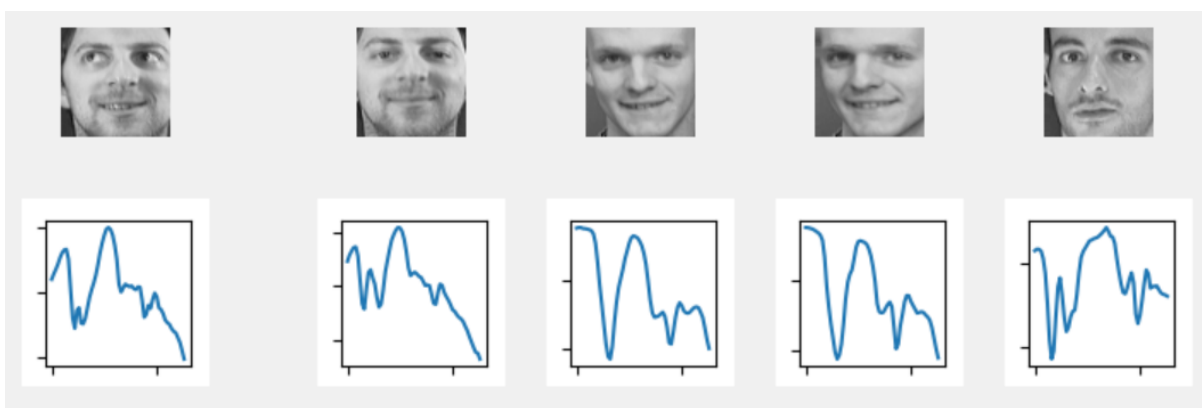


Рисунок 26. Результат работы градиента для параметра = 2

Судя по результатам, представленным на рисунках 22-26, мы имеем ситуацию схожую с признаком “гистограмма яркости”: у одного человека при разных положениях головы схожие графики. Однако оптимальный параметр в нашем случае будет меньшим (близким к 2), так как при меньших значениях мы имеем более подробные графики.

Кросс-валидация. Поиск оптимальных значений.

У нас есть 40 классов (людей), в каждом по 10 изображений лиц. Берем по одному лицу из каждого класса в эталон, остальные - в тестовую выборку. Получаем параметр и результат классификации. Затем проводим аналогичные расчеты для двух лиц в эталоне и так далее до 9 включительно. Находим при каком количестве лиц из класса получаем лучший результат. Выводим графики зависимостей точности распознавания от значения параметра при различных количествах эталонов.

Гистограмма яркости

По результатам вычислений имеем нижепредставленные графики (Рисунки 27-29), а также следующие результаты: лучшая точность - 0.95, оптимальный параметр - 43, лучшее количество эталонов в классе - 9. Наше предположение в “Анализе признаков” насчет данного параметра подтвердилось.

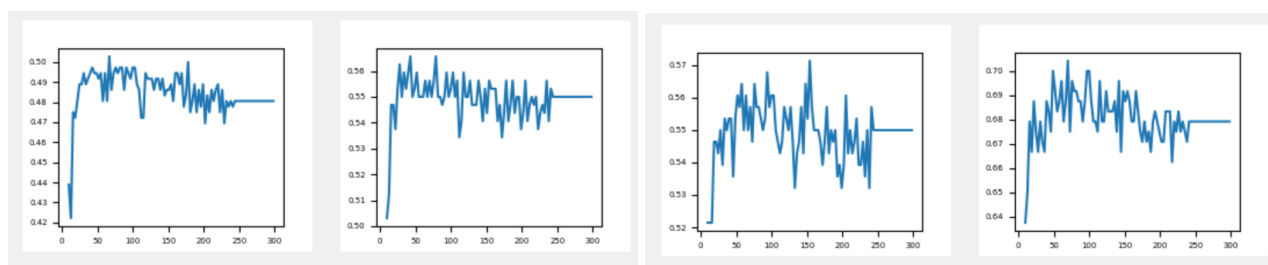


Рисунок 27. Результат работы для количества эталонов = 1 - 4

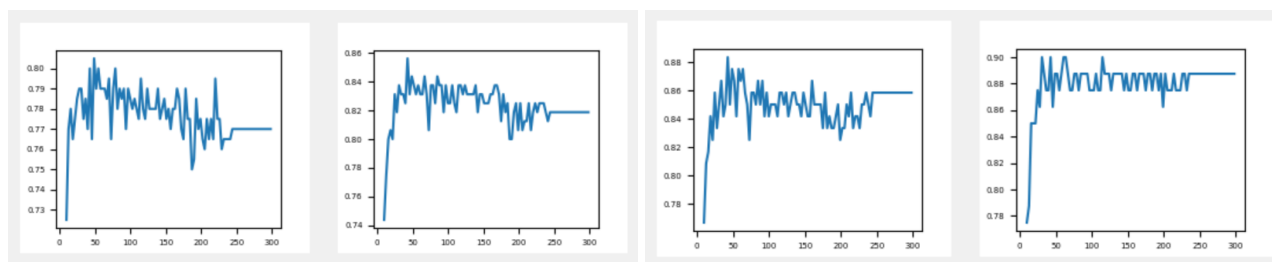


Рисунок 28. Результат работы для количества эталонов = 5 - 8

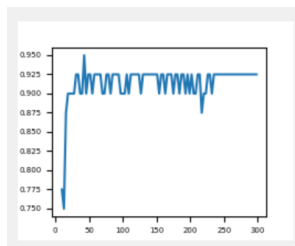


Рисунок 29. Результат работы для количества эталонов = 9

DFT

Для данного метода мы имеем следующие оптимальные значения: точность - 0.95, параметр - 13, количество эталонов - 8. Графики исследований для каждого числа изображений обучающей выборки представлены на рисунках 30-33. Предположение о значении параметра подтвердилось.

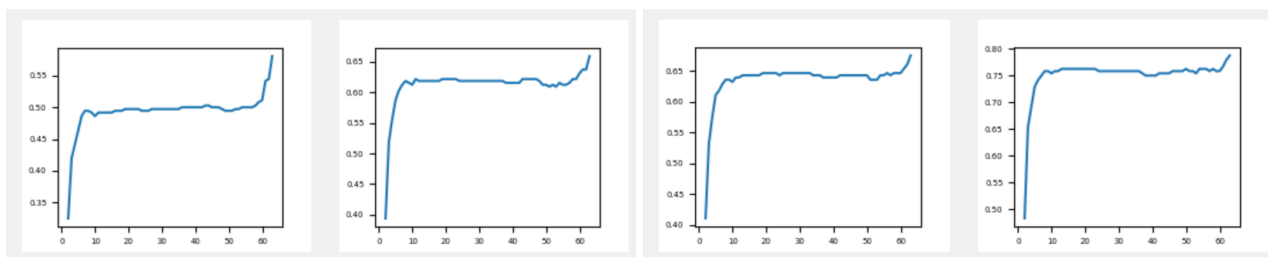


Рисунок 30. Результат работы для количества эталонов = 1 - 4

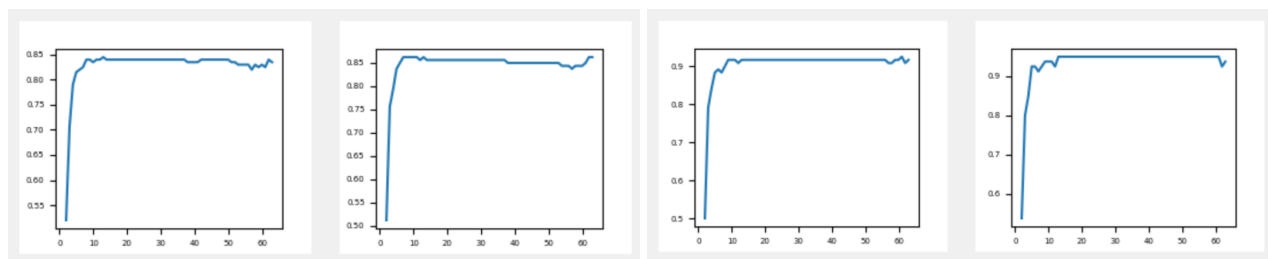


Рисунок 31. Результат работы для количества эталонов = 5 - 8

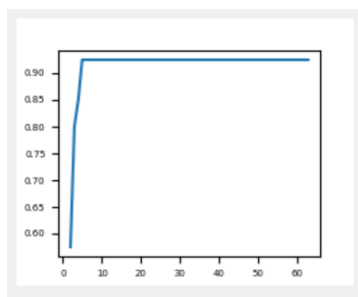


Рисунок 33. Результат работы для количества эталонов = 9

DCT

При исследовании DCT получили, что лучшая точность - 0.925, оптимальный параметр - 8, а лучшее число эталонов - 9. Графики с результатами работы представлены на рисунках 34-36. Интересно, что при увеличении параметра качество распознавания сначала возрастает, а затем, после определенного порога (пика), значительно падает. Предположение о значении параметра не подтвердилось.

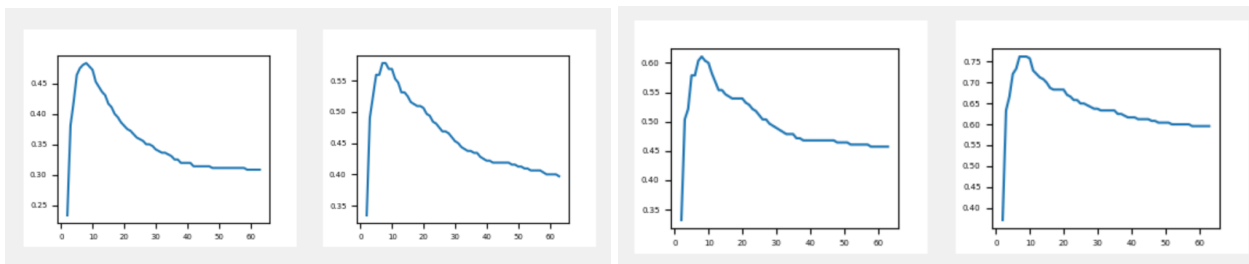


Рисунок 34. Результат работы для количества эталонов = 1 - 4

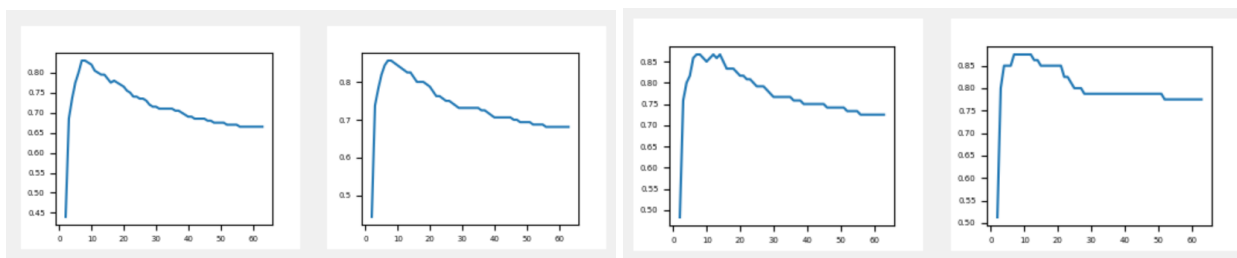


Рисунок 35. Результат работы для количества эталонов = 5 - 8

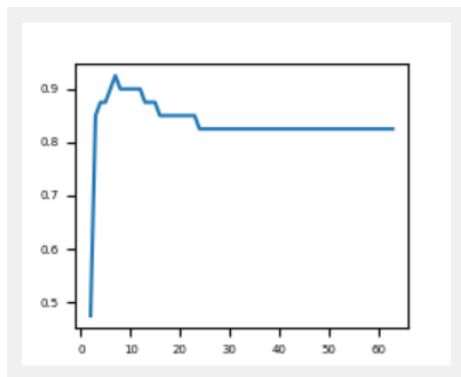


Рисунок 36. Результат работы для количества эталонов = 9

Scale

Для данного метода мы получили: точность - 0.925, параметр - 0.3 и число эталонов - 9. Предположение о значении параметра частично подтвердилось.

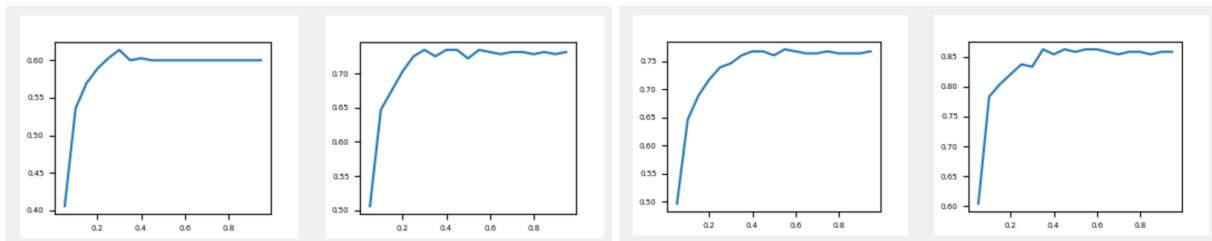


Рисунок 37. Результат работы для количества эталонов = 1 - 4

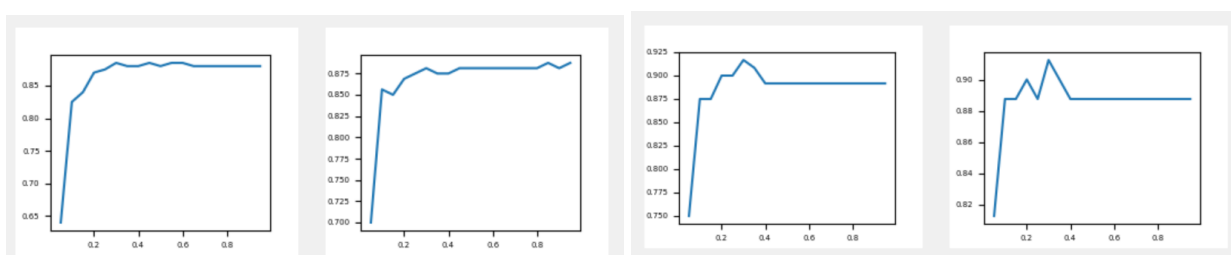


Рисунок 38. Результат работы для количества эталонов = 5 - 8

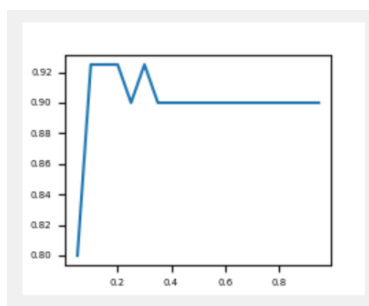


Рисунок 39. Результат работы для количества эталонов = 9

Градиент

Для данного метода мы получили: лучшую точность - 0.9, оптимальный параметр - 3, лучшее число эталонов - 7. Предположение о значении параметра подтвердилось.

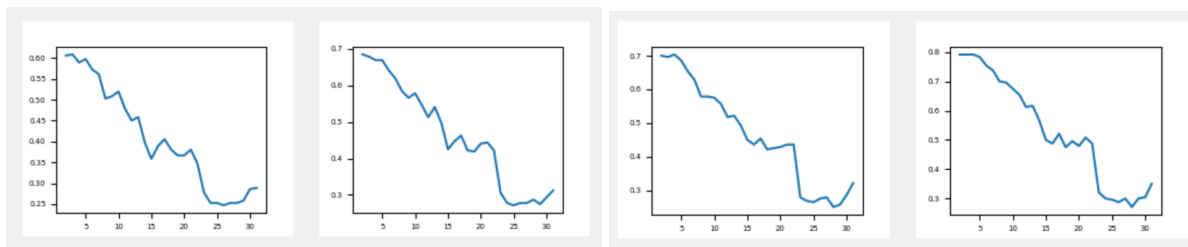


Рисунок 40. Результат работы для количества эталонов = 1 - 4

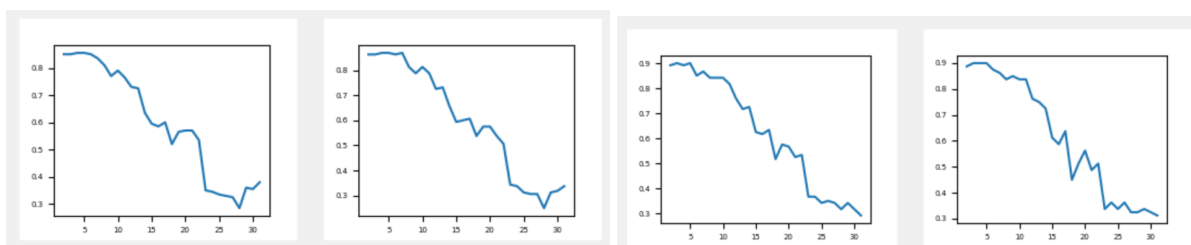


Рисунок 41. Результат работы для количества эталонов = 5 - 8

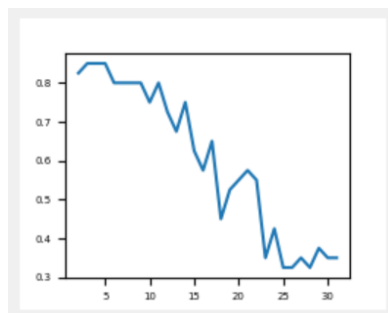


Рисунок 42. Результат работы для количества эталонов = 9

Исследование изменения размера тестовой выборки

Гистограмма яркости

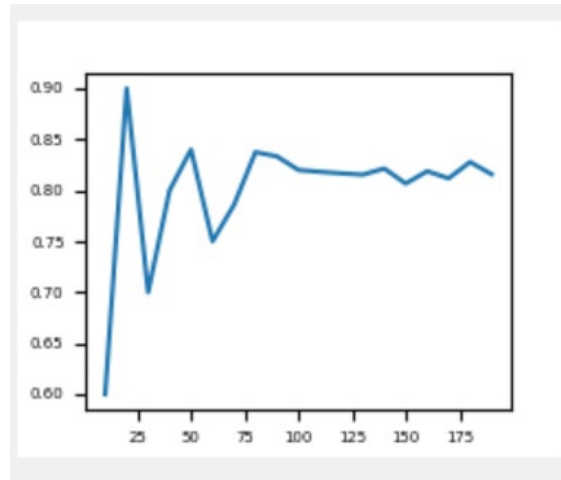


Рисунок 43. Изменение размера тестовой выборки для гистограммы яркости

Как мы видим из графика, где на оси X расположен размер тестовой выборки, а на оси Y - точность распознавания, при размере выборки больше 80 точность принадлежит области от 0.8 до 0.85, что является очень хорошим результатом.

DFT

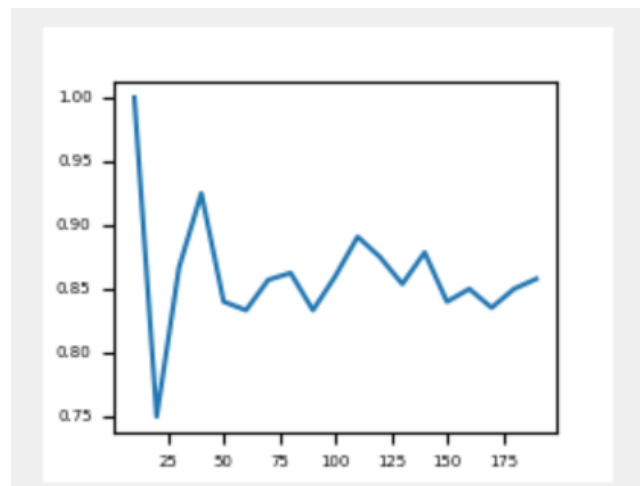


Рисунок 44. Изменение размера тестовой выборки для DFT

Данный график сообщает нам, что при размере выборки больше 50 разброс точности принадлежит промежутку $[0.85, 0.9]$, что является более хорошим результатом.

DCT

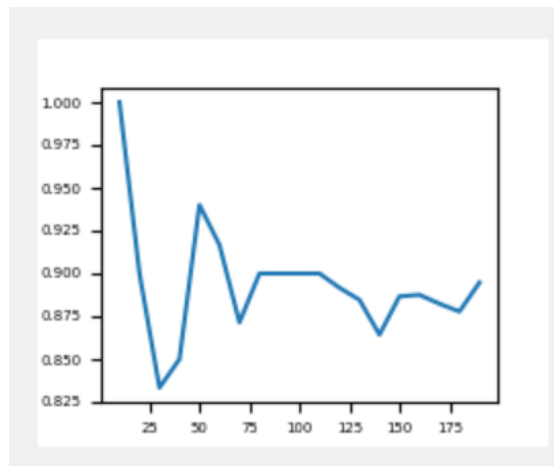


Рисунок 45. Изменение размера тестовой выборки для DCT

Как мы можем наблюдать, для данного метода при размере выборки больше 70 точность распознавания будет принадлежать промежутку от 0.875 до 0.9.

Scale

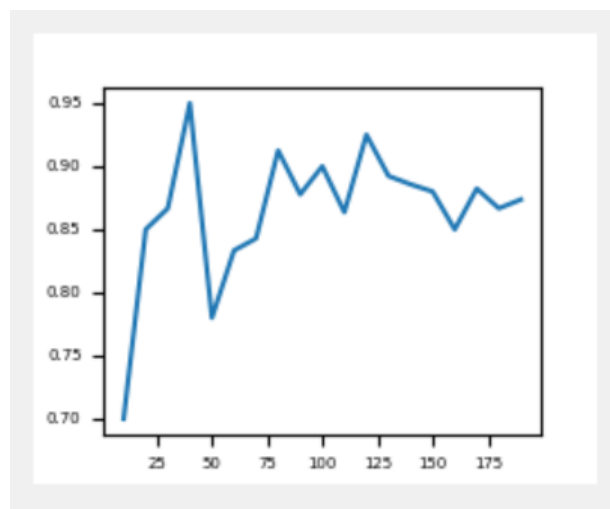


Рисунок 46. Изменение размера тестовой выборки для Scale

Данный график показывает, что при размере больше 75 мы имеем точность в диапазоне $[0.85, 0.95]$. Не смотря на то, что это пока что самый высокий результат, стоит помнить, что на самом деле мы имеем достаточно большие скачки точности внутри этого промежутка.

Градиент

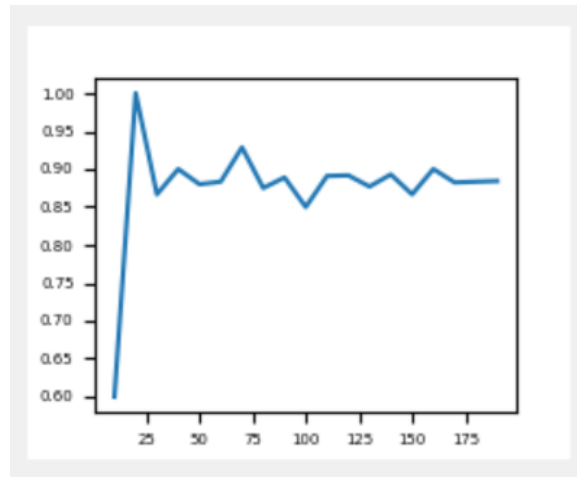


Рисунок 47. Изменение размера тестовой выборки для градиента

Проанализировав данный график, мы можем заметить, что уже при размере выборки 25 мы имеем диапазон точности $[0.85, 0.95]$. Как мне кажется, из всех результатов этот самый лучший, так как классификатор, основанный на методе градиента не только обладает качеством, близким к 1, но и быстро к нему сходится, а также в отличие от большинства других у него скачки внутри диапазона не такие значительные.

Заключение

В данной практической работе была создана программа для моделирования систем распознавания людей по лицам, изучены пять методов, на которых основывался наш классификатор, получены оптимальные значения параметров для каждого из признаков, исследована оптимальная размерность тестовой выборки, а также было получено оптимальное число эталонов изображений лиц для каждого человека для каждого из методов.

По результатам исследования, были получены наилучшие результаты с помощью метода градиента: были выделены все тенденции, принадлежащие каждому отдельному человеку, показаны отличные результаты при кросс-валидации, а также при исследовании размерности тестовой выборки точность распознавания метода была близка к 1. Наилучший для градиента параметр - 3, число эталонов для каждого класса - 7.