

SAINT-PETERSBURG STATE UNIVERSITY

Mathematics&Mechanics Faculty

Department of Software Engineering

Kavokin Aleksandr

Gender Recognition by photo

Bachelor's Thesis

Admitted for defence.

Head of the chair:

Professor Andrey Terekhov

Scientific supervisor:

Assistant professor Stanislav Sartasov

Reviewer:

Dmitry Mordvinov

Saint-Petersburg

2014

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Математико-механический факультет
Кафедра системного программирования

Кавокин Александр Сергеевич

Распознавание пола по фотографии

Бакалаврская выпускная работа

Допущен к защите

Зав.кафедрой:

д.ф.-м.н., профессор А.Н.Терехов

Научный руководитель:

ст. преподаватель Саргасов С.Ю.

Рецензент:

Мордвинов Д.А.

Санкт-Петербург

2014

Оглавление

1. Введение.....	4
2. Описание предметной области	5
2.1. Метод опорных векторов.....	5
2.2. Локальные бинарные шаблоны.....	6
2.3. Стандартный способ распознавания пола.....	7
3. Постановка задачи	9
4. Реализация	9
4.1. Opendata: база данных с лицами	9
4.2. Предложенный алгоритм распознавания пола	10
4.3. Модуль, считающий значения ЛБШ и гистограммы.....	11
4.4. Обучение классификаторов SVM	11
4.4.1. Поиск ключевых областей.....	12
4.4.2. Определение пола по ключевым областям.....	14
4.5. Тестирование классификатора SVM	16
4.6. Сравнение со стандартным алгоритмом	16
4.7. Особенности и трудности	17
5. Результаты	17
Список литературы	19
Приложение 1. Глоссарий	20
Приложение 2. Пример гистограммы для обучения SVM.....	21

1. Введение

Автоматизированная идентификация человека является одной из современных задач современного машинного обучения и компьютерного зрения. С каждым годом к этой задаче обращаются все чаще и чаще. Сферами применения результатов в данной области является обеспечение безопасности, поиск преступников, находящихся в розыске, сбор статистической информации, спортивные тесты и многие другие.

Распознавание пола является одной из подзадач идентификации человека. В современном мире даже человек не всегда способен определить, кто перед ним находится, поскольку такие привычные для нас границы между мужчиной и женщиной как длина волос, наличие сережек, косметика на лице, борода и пр. стерлись. Известны случаи, когда молодые женщины, вдохновившись подвигами Жанны Д'Арк или Шурочки из «Гусарской баллады», служили в армии, выдав себя за мужчин [2]. Аналогичные ситуации происходят и в спорте. Например, в 2009 году южноафриканскую спортсменку заставили пройти гендерный тест, засомневавшись, не является ли она мужчиной. Есть еще более необычный пример: в феврале 2014 года произошел скандал, связанный с четырьмя мужчинами, игравшими за женскую сборную Ирана. Для их идентификации пришлось проводить анализы крови, с целью проверки наличия соответствующих гормонов. Такая процедура достаточно трудна и длительна по времени. Хотелось бы найти более дешевый и быстрый способ для решения подобных проблем. Многие могут спросить, как же машина сможет определить пол, если с этим не может справиться даже человек. Но у компьютера есть преимущество, поскольку он не использует человеческие стереотипы. Его можно обучить распознаванию, анализируя некоторые ключевые области на лице, которые всегда анатомически различаются у мужчин и женщин.

Для идентификации пола требуется обучить программу на некоторой базе данных с фотографиями мужчин и женщин, и только после этого её

протестировать. Для полноты результатов тестирование проводится на множестве данных, которое использовалось для обучения, а также множестве, состоящем из фотографий, не входящих в обучающую выборку.

Современные способы распознавания пола позволяют достичь точности более чем 90%, что сравнимо с человеческими возможностями. Но добиться цифры выше 96% практически никакая программа пока что не смогла. Для улучшения этих показателей требуется модифицировать стандартные алгоритмы.

2. Описание предметной области

Существует много различных алгоритмов для распознавания пола человека. Программы могут анализировать видео, звук, изображение. Но, поскольку видео является модифицированным случаем использования обычных изображений, то чаще всего используют фотографии для идентификации человека. Самым распространенным способом является использование метода опорных векторов и локальных бинарных шаблонов для анализа фотографии лица.

2.1. Метод опорных векторов

Метод опорных векторов (англ. SVM, support vector machine) — набор схожих алгоритмов обучения с учителем, использующихся для задач классификации и регрессионного анализа. Принадлежит к семейству линейных классификаторов, может также рассматриваться как специальный случай регуляризации по Тихонову. Особым свойством метода опорных векторов является непрерывное уменьшение эмпирической ошибки классификации и увеличение зазора, поэтому метод также известен как метод классификатора с максимальным зазором [1].

Основной идеей метода опорных векторов (Support Vector Machine, SVM) является перенос исходных векторов, содержащих данные, в пространство

более высокой размерности. Затем осуществляется поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве.

На рисунке 1 изображены 3 прямые, которые разделяют двухмерное пространство с двумя видами множеств. Прямая L2 является разделяющей гиперплоскостью, поскольку разделяет два множества с максимальным зазором.

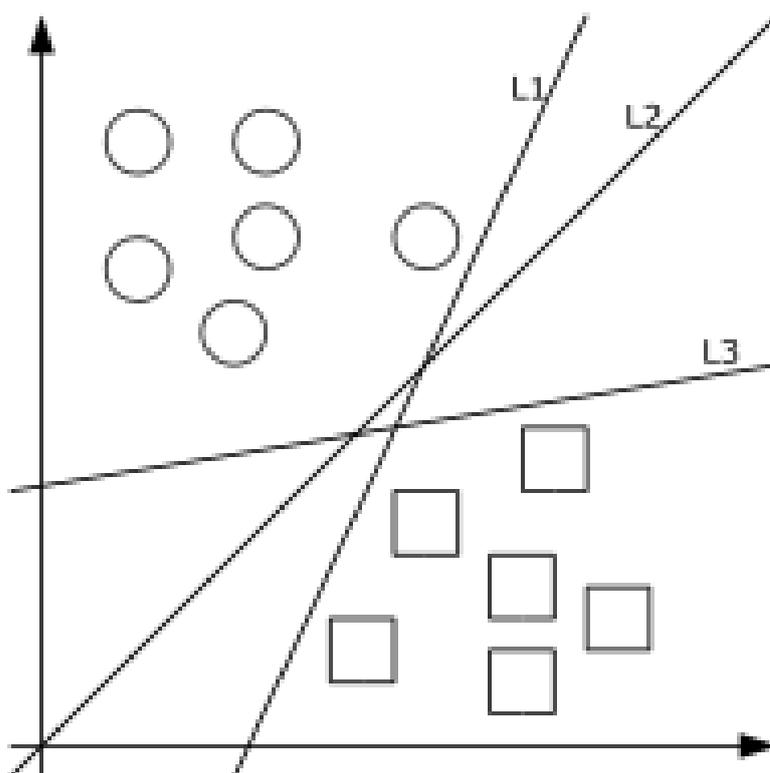


Рисунок 1: Пример использования метода опорных векторов.

2.2. Локальные бинарные шаблоны

С помощью локальных бинарных шаблонов (Local Binary Pattern, LBP) каждый пиксель изображения может быть представлен в виде восьмиразрядного бинарного кода. В качестве порогового значения принимается значение интенсивности пикселя. Затем рассматриваются 8 пикселей в окрестности и их значения интенсивности. Если значение

интенсивности пикселя больше или равно центральному, в качестве результата записывается «1», иначе – «0».

На рисунке 2 у центрального пикселя с интенсивностью «33» значение локального бинарного шаблона равно «00011110».

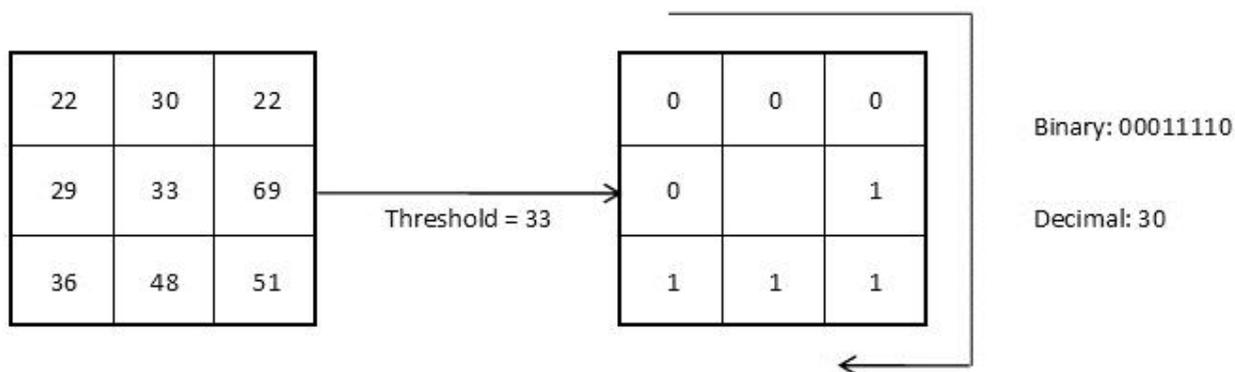


Рисунок 2: Пример локального бинарного шаблона.

Некоторые ЛБШ несут в себе больше информации, чем другие. Локальный бинарный шаблон называется равномерным, если не содержит трех подряд нулей или единиц. Использование таких бинарных кодов позволяет детектировать только наиболее существенные элементы изображения, такие как углы, линии, резкие смены цветов, и значительно экономят память.

2.3. Стандартный способ распознавания пола

Стандартным для обучения распознавания пола человека по фотографии является следующий алгоритм [3]:

- вычислить значение локального бинарного шаблона для каждого пикселя изображения;
- разбить фотографию на непересекающиеся квадратные ячейки;
- построить гистограмму по полученным значениям равномерных локальных бинарных шаблонов для каждой ячейки. Все неравномерные ЛБШ записываются в отдельный столбец гистограммы;
- преобразовать полученные гистограммы в вектора для SVM;

- обучить программу с помощью метода опорных векторов определять пол человека на основе некоторого множества, где известно, на какой фотографии изображен мужчина, а на каком женщина.

На рисунке 3 изображено разбиение изображения лица на непересекающиеся клетки, используемое в стандартных способах распознавания пола.

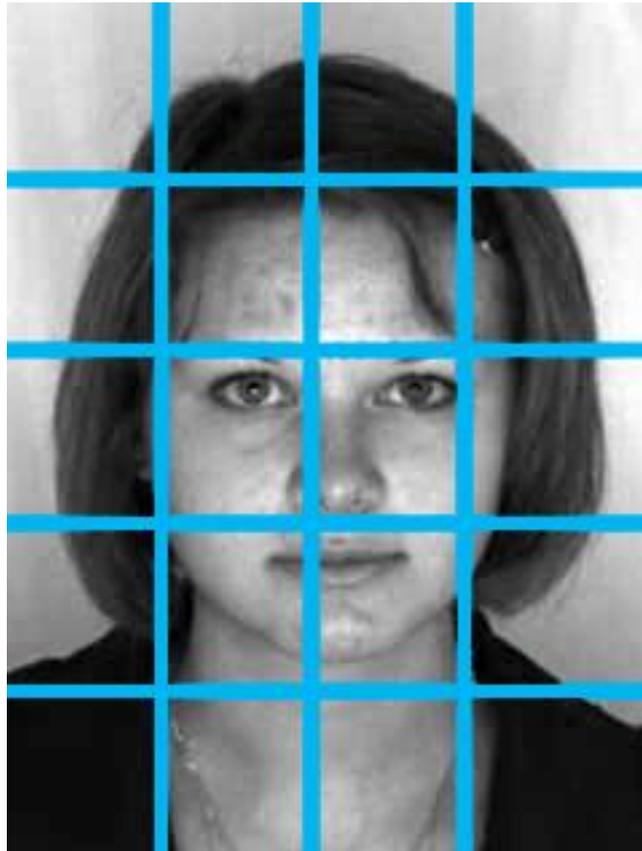


Рисунок 3: Разбиение фотографии лица на клетки.

Существуют различные работы, основанные на данном методе. В работе Boosting Sex Identification Performance [8], проведенной под патронажем Google Research Inc. был достигнут результат в 93.1% точности распознавания пола. В 2008 году в работе Gender Recognition from body [4] **Ошибка! Источник ссылки не найден.**, проведенной в Университете штата Иллинойс в Урбана-Шампейн была создана программа, определяющая возраст человека. При определении одним из факторов является распознавание пола. Неплохого

результата с ошибкой всего в 3.4% добились авторы в работе Gender Classification with Support Vector Machine [3].

3. Постановка задачи

Целью данной работы является разработка алгоритма распознавания пола человека по фотографии на основе SVM и в дальнейшем его реализация.

Основываясь на изложенной выше цели, были поставлены следующие задачи.

1. Выбрать открытую базу данных с лицами людей.
2. Придумать и реализовать алгоритм распознавания ключевых областей на фотографии.
3. Написать модуль, считающий значения ЛБШ и создающий гистограммы
4. Реализовать обучение классификаторов SVM.
 - a. Обучить один классификатор SVM поиску ключевых областей.
 - b. Обучить другой классификатор SVM распознаванию пола на основе ключевых областей.
5. Протестировать обученный классификатор SVM.

4. Реализация

4.1. Opendata: база данных с лицами

В качестве базы данных была выбрана БД color FERET [9]. Данная БД содержит 2,5 тысячи различных фотографий 856 людей, собранных в период между декабрем 1993 и августом 1996. На каждого человека есть от 5 до 33 фотографий с разными гримасами, углами поворота и освещением. К каждой фотографии прилагаются метаданные, содержащие такую информацию как расположение некоторых частей лица, пол, возраст, внешний вид (в очках/без очков) и нации. Дополнительным преимуществом данной БД является то, что она находится в свободном доступе.

4.2. Предложенный алгоритм распознавания пола

В стандартном алгоритме на результат определение пола возможно влияние случайно попавших в кадр артефактов. Чтобы исключить такую ситуацию, было предложено использовать не всю фотографию, а только некоторые области. Для каждой фотографии в базе данных были предоставлены метаданные. Анализируя их, можно получить информацию о поле и расположении глаз, носа и рта. Исходя из этого, было решено сперва находить ключевые области на фотографии человека, а затем, используя только обнаруженные части лица, определять пол. В качестве ключевых областей берутся глаза, рот и нос, как наиболее выделяющиеся на лице человека. Есть и другие характерные отличающиеся черты, например кадык, уши, но они довольно часто закрыты одеждой или волосами.

Для поиска ключевых точек изображение следует разбить на пересекающиеся клетки. Если клетки не пересекаются, то возможна ситуация, когда какой-нибудь значимый элемент лица, например рот, не попадает целиком ни в одну клетку. Размер клеток 80x80 пикселей. Области меньшего размера не включают в себя полностью весь ключевой элемент лица, а в области большего размера могут попасть какие-то лишние элементы, которые могут негативно повлиять на распознавание. Клетки пересекаются друг с другом с нахлестом в 15 пикселей.

Затем для каждого пикселя считается значение локального бинарного шаблона. Для дальнейшей работы берутся только равномерные ЛБШ. Для каждой клетки считается гистограмма, в которой каждому значению равномерного ЛБШ соответствует отдельный столбец. Если строить гистограмму по всему изображению, то будет известна лишь информация о наличии каких-то особенностей, но нельзя будет ничего сказать об их расположении. Для этого и требуется разбиение фотографии на клетки.

Для обучения поиску ключевых областей данный алгоритм выполняется для некоторого тренировочного множества, где заранее известно расположение

частей лица. Полученные гистограммы записываются в файл и используются в качестве векторов, содержащих данные для SVM. Обученный на таком множестве SVM способен определять ключевые области на фотографии лица.

После того, как ключевые области будут найдены, необходимо определить по ним пол человека. Для этого в отдельный файл записываются гистограммы ключевых областей мужчин и женщин. Таким образом, получается 8 видов векторов: мужской нос, женский нос и аналогично для рта и глаз. По этим данным обучается второй SVM, с помощью которого затем можно будет определять пол владельца каждой ключевой области.

После того, как алгоритм найдет ключевые области на лице и распознает пол владельца каждой из них, требуется окончательно определить пол изображенного человека. Поскольку возможна ситуация, когда SVM распознает две области из четырех как мужские, а другие две как женские, вводится система весов. Каждой области соответствует свой вес. И просуммировав данные веса можно указать пол человека.

4.3. Модуль, считающий значения ЛБШ и гистограммы

Для расчета локальных бинарных шаблонов был написан модуль на языке С#. Если окрестность пикселя выходит за рамки картинка, значение интенсивности в этих окрестных пикселях принимается за 0. После того, как в текущей клетке посчитаны все значения ЛБШ, из них отбираются равномерные шаблоны и записываются в гистограмму. В качестве гистограммы используется Dictionary<Key, Value>, где ключом является десятичное представление оператора ЛБШ, а значением – количество пикселей в области, имеющих такой же параметр ЛБШ.

4.4. Обучение классификаторов SVM

Для распознавания пола было обучено два классификатора SVM. Первый должен находить на фотографии лица ключевые области. С помощью второго SVM определяется гендерная принадлежность каждой из ключевых областей.

4.4.1. Поиск ключевых областей

Для обучения первого классификатора было выбрано 200 человек из базы данных colorFERET, по 5 фотографий на каждого. В выборке было 112 мужчин и 88 женщин. Для каждого изображения был выполнен алгоритм, описанный выше. Полученные в результате гистограммы ЛБШ записывались в файл для обучения SVM. Поскольку было 4 типа ключевых областей, было выбрано 5 классов (еще 1 класс как не ключевая область), на которые разделяется все множество гистограмм. Для работы с SVM была использована библиотека libsvm [5] для языка C#. Гистограммы записывались в строки по следующему шаблону:

<label> <index1>:<value1> <index2>:<value2>...# <xPos> <yPos><gender>,

где

<label> - натуральное число, характеризующее класс. 1 – рот, 2 – нос, 3 – левый глаз, 4 – правый глаз, 5 – не ключевая область

<index> - целое число, характеризующее столбец гистограммы

<value> - целое число, характеризующее количество элементов в соответствующем столбце гистограммы

<xPos> - позиция центра клетки по оси OX

<yPos> - позиция центра клетки по оси OY

<gender> - целое число, характеризующее пол. 1 – мужчина, 2 - женщина

В Приложении 2 можно увидеть пример записанной таким образом в файл гистограммы.

На рисунках 4 и 5 изображены примеры удачного и неудачного поиска ключевых областей.

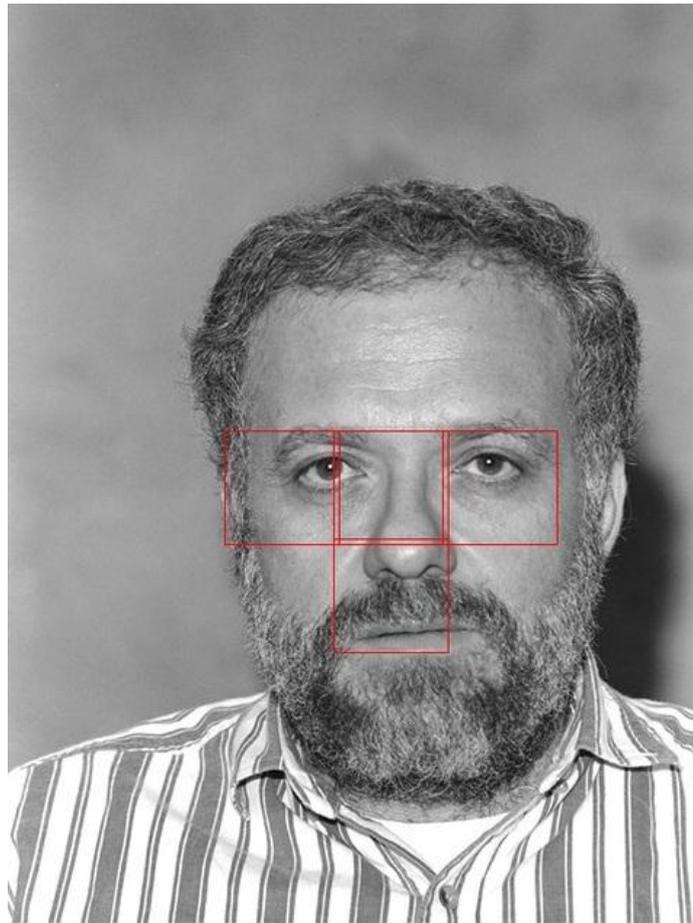


Рисунок 4: Пример удачного поиска ключевых областей

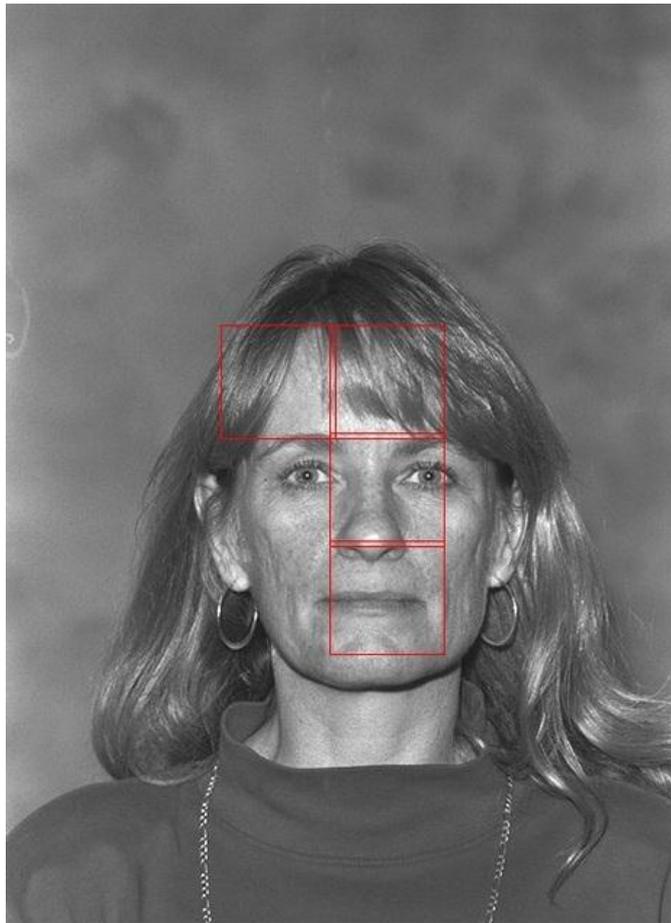


Рисунок 5: Пример неудачного поиска ключевых областей

4.4.2. Определение пола по ключевым областям

Для обучения второго классификатора использовалось 8 классов областей: мужской нос, женский нос, и аналогично для рта и глаз. Данные для обучения создавались аналогичным образом, как и для первого классификатора.

После того, как программа определяла гендерную принадлежность каждой из найденных ключевых областей, считались суммарные веса этих областей. Эмпирическим путем было выявлено, что наиболее весомым при определении пола человека является рот, затем идут глаза и нос. Таким образом, область со ртом имела вес в 40 единиц, с каждым глазом – по 20, а с носом – 10. Посчитав эти веса, можно окончательно определить пол человека.

Примеры работы программы можно видеть на рисунках 6 и 7.

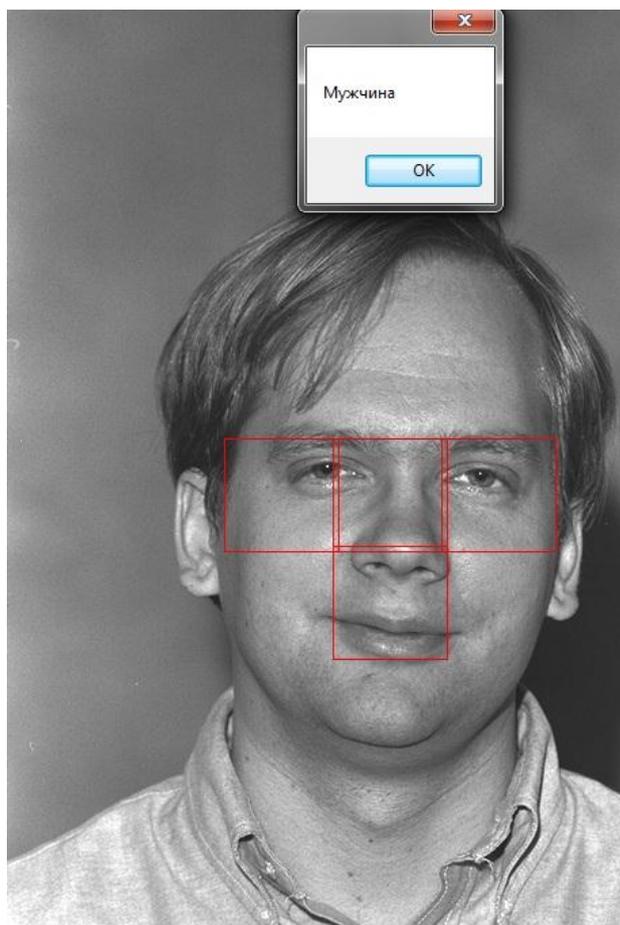


Рисунок 6. Пример распознавания мужчины.

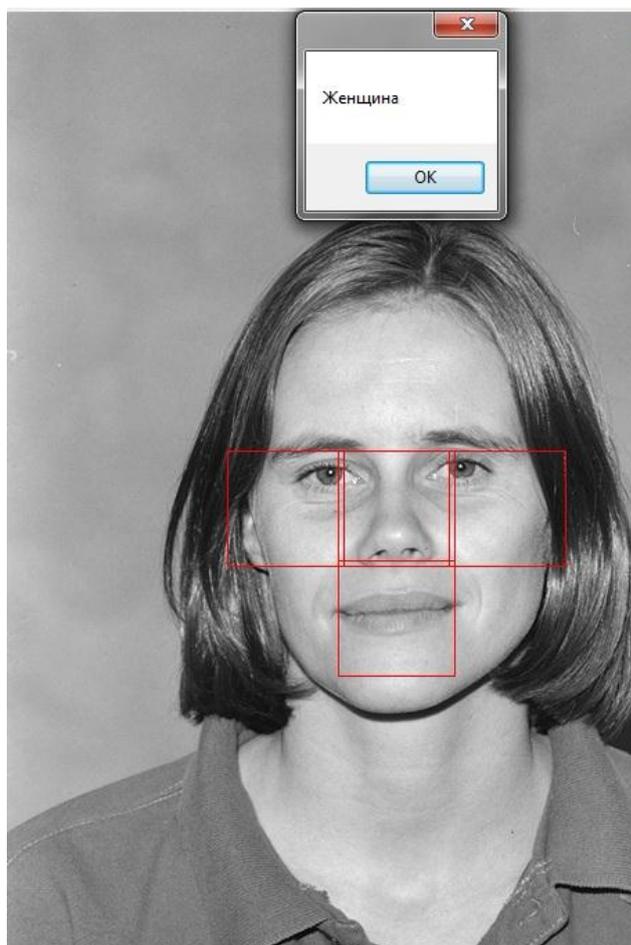


Рисунок 7. Пример распознавания женщины.

4.5. Тестирование классификатора SVM

Для тестирования программы были использованы 2 выборки из базы данных colorFERET:

1. Закрытое множество, на котором проводилось обучение.
2. Тестовая выборка: 300 человек, по 5 фотографий на каждого.

На закрытом множестве ключевые области были найдены в 75.8%. В тестовой выборке ключевые области были найдены в 51.4%. Пол человека правильно был определен в 98.9% случаях на тех фотографиях, где удалось найти ключевые области на закрытом множестве. На тестовой выборке пол корректно был определен в 97.4% случаях.

4.6. Сравнение со стандартным алгоритмом

Был написан модуль, реализующий стандартный алгоритм. Для сравнения результатов предложенного алгоритма и стандартного, было

проведено обучение стандартного алгоритма на такой же выборке и с такими же размерами клеток. Пол человека правильно был определен в 84.6% на обучающем множестве и в 69% на тестируемом множестве.

4.7. Особенности и трудности

В процессе работы были выявлены следующие особенности и трудности:

- 1) .NET Framework не поддерживает без дополнительных библиотек формат изображений .ppm, в котором хранятся фотографии в базе данных. Для решения этой проблемы была использована библиотека ShaniSoft.Drawing.PNM.dll [6]
- 2) Для обучения использовались большие объемы данных и значительные вычисления, которые трудно проводить на домашнем компьютере. Пришлось использовать дополнительные мощности компьютеров на стороне.

5. Результаты

В ходе работы были получены следующие результаты.

1. Выбрана и использована база данных лиц colorFERET, создан парсер для метаданных базы данных.
2. Предложен и реализован алгоритм распознавания ключевых областей.
3. Реализован модуль, оперирующий с операторами ЛБШ
4. Обучены классификаторы SVM для распознавания ключевых областей и распознавания пола человека.
5. Было выполнено тестирование на тренировочном множестве и обычной выборке из базы данных colorFERET

Как видно из результатов, в случаях, когда удается найти ключевые точки на фотографии, точность распознавания пола выше, чем в существовавших

ранее реализациях. Для улучшения результатов предлагается в дальнейшем использовать другой алгоритм для поиска ключевых областей на изображении. Например, алгоритм Виолы-Джонса[7].

Предложенный способ поиска ключевых точек имеет такие низкие результаты, поскольку абсолютным большинством в тренировочном множестве являются клетки пятого типа, не имеющие ключевых точек. Данная проблема так и не была решена с помощью штрафных коэффициентов для классификации пятого типа областей. Поэтому в дальнейшем предлагается использовать более надежный способ для обнаружения ключевых точек на лице, а уже затем использовать предложенный в работе алгоритм определения пола человека.

Список литературы

[1] Метод опорных векторов,

http://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D1%8B%D1%85_%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BE%D0%B2

[2] Женщины французских легионов,

<http://annales.info/evrope/small/zhfl.htm>

[3] Baback Moghaddam and Ming-Hsuan Yang. Gender Classification with Support Vector Machines,

<http://www.cs.utexas.edu/~grauman/courses/378/handouts/moghaddam2000.pdf>

[4] Liangliang Cao, Mert Dikmen, Yun Fu and Thomas S. Huang. Gender recognition from body,

<http://dl.acm.org/citation.cfm?id=1459470>

[5] LIBSVM -- A Library for Support Vector Machine,

<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

[6] PixelMap Class and PNM Image Viewer,

<http://www.codeproject.com/Articles/18968/PixelMap-Class-and-PNM-Image-Viewer>

[7] Rapid Object Detection using a Boosted Cascade of Simple Features,

<https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf>

[8] Shumeet Baluja and Henry A. Rowley. Boosting Sex Identification Performance,

<http://www.cs.cmu.edu/afs/cs/usr/har/www/ijcv2007-sex.pdf>

[9]The Color FERET Database,

<http://www.nist.gov/itl/iad/ig/colorferet.cfm>

Приложение 1. Глоссарий

- Local Binary Pattern (LPB) – локальный бинарный шаблон. простой оператор, используемый для классификации текстур с компьютерном зрении
- Support Vector Machine (SVM) – метод опорных векторов. Набор схожих алгоритмов обучения с учителем, использующихся для задач классификации и регрессионного анализа.
- Равномерный ЛБШ – локальный бинарный шаблон, у которого нет подряд более трех нулей или единиц.

Приложение 2. Пример гистограммы для обучения SVM

3 36:5 38:3 41:1 44:7 45:1 50:3 51:1 54:2 74:1 76:2 90:1 91:1 100:3 101:1 106:1
107:1 108:4 109:1 146:1 147:6 149:1 153:4 154:1 155:2 178:1 179:2 201:1 203:4
204:1 205:2 210:1 211:9 213:1 214:2 217:5 219:3 # 317 317 1

3 – номер ключевой области (левый глаз).

36:5 – значение ЛБШ. Всего в клетке 5 пикселей со значением ЛБШ равным 36.

317 317 1 – координаты центра клетки (317, 317) и номер пола человека (1 - мужчина).