|  |  |  |  |
| --- | --- | --- | --- |
| Министерство образования Республики Беларусь  Учреждение образования  Белорусский Государственный Университет Информатики и Радиоэлектроники | | | |
| Факультет | Компьютерных сетей и систем | | |
| Кафедра | Информатики  Дисциплина: Конструирование те технологии электронных вычислительных средств | | |
|  |  | | |
| **РЕФЕРАТ**  по курсу «Машинное обучение»  **Задача распознавания лиц** | | | |
| Студент:  гр. 758641  Ярош Г.И. | |  | Проверил:  Заливако С. С. |
| Минск, 2019 | | | |

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 3

Сферы применения 3

КЛАССИЧЕСКИЕ АЛГОРИТМЫ РАСПОЗНАВАНИЯ ЛИЦ 4

1. Eigenfaces 4

Описание алгоритма 4

2. Fisherfaces 5

Описание алгоритма 6

РАСПОЗНАВАНИЕ ЛИЦ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ 7

1. FaceNet 7

Сиамские сети 7

Триплет потерь 8

Архитектура сети 9

2. VGG-Face 10

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 12

# ВВЕДЕНИЕ

Задача распознавания лиц является одной из наиболее приоритетных и популярных задач в сфере машинного обучения и компьютерного зрения. Каждый день мы встречаемся с системами уличного видеонаблюдения, системы распознавания лиц применяются в офисах вместо ключей и магнитных карт, мобильные телефоны используют системы распознавания лиц для разблокировки.

Современные системы распознавания лиц могут даже определять является ли рассматриваемое лицо реальным, что предотвращает попытки взлома, основанные на использовании фотографий и других изображений человека.

К задаче распознавания лиц иногда относят следующие смежные задачи:

* Обнаружение лица на изображении. Обычно данная задача предшествует самому распознаванию лица. Результатом ее обычно является область или области изображения, где находятся лица людей, которые можно впоследствии использовать для распознавания.
* Верификация лица. Является более узкой проблемой, и ставит задачу определить сходство входного изображения лица и некоторого проверочного. Распознавание лиц в общем случае ставит задачу по входному изображению определить входит ли оно в некоторую группу известных людей.

### Сферы применения

Распознавание лиц находит применения во многих сферах жизни. Вот только некоторые из них:

* Распознавание лиц используется при выдаче документов, удостоверяющих личность, в сочетании с другими биометрическими технологиями.
* Идентификация лиц используется при проверке на таможне для сопоставления лица человека с изображением в биометрическом паспорте.
* Распознавание лиц широко применяется в сфере обеспечения правопорядка для поиска правонарушителей, используя записи уличных камер наблюдения.
* Распознавание лиц используется в различных системах авторизации, разблокировке мобильного телефона, получение доступа в офис.

# КЛАССИЧЕСКИЕ АЛГОРИТМЫ РАСПОЗНАВАНИЯ ЛИЦ

## Eigenfaces

Eigenfaces относится к подходам к решению задачи распознавания лиц, который основан на подсчете дисперсии коллекции изображений лиц и использует эту информацию для кодирования и сравнения изображений отдельных лиц в целом (в отличие от основанного на деталях или основанного на особенностях метода). В основе Eigenfaces лежит метод главных компонент, где главные компоненты распределения пикселей изображений лиц называются собственными лицами или, другими словами, собственными векторами [ковариационной](http://www.scholarpedia.org/article/Covariance) матрицы множества изображений лиц, где изображение *N* пикселей считается точкой (или вектором) в *N* мерном пространстве. Идея использования метода главных компонент для представления человеческих лиц была разработана Сировичем и Кирби (Sirovich and Kirby 1987) и использовалась Турком и Пентлендом (Turk and Pentland 1991) для обнаружения и распознавания лиц. Подход Eigenface, как полагают многие, является первой рабочей технологией распознавания лиц, и он послужил основой для одного из лучших коммерческих продуктов технологии распознавания лиц. Со времени его первоначальной разработки и публикации было создано много расширений оригинального метода и много новых разработок в системах автоматического распознавания лиц. Eigenfaces все еще часто рассматриваются как метод для сравнения по базовой линии, чтобы продемонстрировать минимальную ожидаемую эффективность такого рода систем.

### Описание алгоритма

Для корректной работы алгоритма изображения лица нормализуются для выравнивания глаз и рта, а затем все повторно дискретизируются с одинаковым разрешением в пикселях. Собственные лица затем извлекаются из данных изображения с помощью анализа главных компонентов (PCA) следующим образом:

1. Пусть дано *N* изображений лиц, каждое из которых представлено вектором размером *d*:
2. Для всех векторов подсчитываем среднее значение :
3. Вычисляем ковариационную матрицу :
4. Вычисляем собственные значения и собственные векторы ковариационной матрицы :
5. Выбираем собственных векторов, которые соответствуют наиболльшим собственным значениям.
6. Вычисление главных компонент вектора происходит по формуле:

Распознавание изображения с помощью Eigenface происходит следующим образом:

1. Для всех тренировочных изображений посчитываются собственные лица :
2. Для распознаваемого изображения лица также подсчитываются подсчитывается собственные лица :
3. В качестве критерия похожести распознаваемого изображения на изображение в тренировочной выборке можно использовать евклидово расстояние:

Если расстояние меньше изначально заданного порога, то можно рассматривать, что входное изображение распознано, как .

Подход Eigenface был важным шагом на пути к распознаванию лиц в компьютерном [зрении](http://www.scholarpedia.org/article/Vision) . Тем не менее, метод чувствителен к изменению освещения, масштаба, позы, выражения лица и окклюзии. Чтобы алгоритм работал хорошо, лицо должно быть представлено в виде спереди, в соответствующем масштабе, в аналогичном освещении, в определенном (обычно нейтральном) выражении и не включенным. Для решения множества проблем, с которыми сталкиваются реальные системы распознавания лиц, было предложено множество модификаций и расширений с момента первоначальной реализации метода.

## Fisherfaces

Метод главных компонент (PCA), который является основой метода Eigenfaces, находит линейную комбинацию функций, которая максимизирует общую дисперсию данных. Хотя это, несомненно, мощный способ представления данных, он не учитывает какие-либо классы, и поэтому при отбрасывании компонентов может быть потеряно много дискриминационной информации. Алгоритм Fisherface использует метод, называемый линейным дискриминантным Анализом. Метод выполняет уменьшение размерности с учетом класса объекта и был изобретен знаменитым математиком Р. А. Фишером. Он успешно использовал его для классификации цветов в своей работе 1936 года «Использование нескольких измерений в таксономических задачах». Чтобы найти комбинацию признаков, которая лучше всего разделяет классы, линейный дискриминантный анализ максимизирует соотношение между классами и разброс внутри классов, а не максимизирует общий разброс. Идея проста: одни и те же классы должны близко располагаться друг к другу, в то время как разные классы находятся как можно дальше друг от друга в пространстве более низкой размерности.

### Описание алгоритма

Пусть дано классов изображений . Алгоритм Fisherface заключается в следующем:

1. Различия внутри класса могут быть оценены с помощью матрицы рассеяния, вычисляемой следующим способом:
2. Похожим образом вычисляются межклассовые различия используя межклассовую матрицу рассеяния:
3. Далее находится такие базисные векторы , что значение минимизируется, а значение максимизируется. Решение может быть найдено с помощью разложения по собственным значениям:

*,*

где - матрица собственных значений, - матрица собственных векторов.

1. Найденные собственные векторы, соответствующие ненулевым значениям называются фишер-лица.

Распознавание изображения с помощью Fisherfaces происходит следующим образом:

1. Для всех тренировочных изображений посчитываются фишер-лица :
2. Для распознаваемого изображения лица также подсчитываются подсчитывается фишер-лица :
3. В качестве критерия похожести распознаваемого изображения на изображение в тренировочной выборке можно использовать евклидово расстояние:

Если расстояние меньше изначально заданного порога, то можно рассматривать, что входное изображение распознано, как .

# РАСПОЗНАВАНИЕ ЛИЦ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

В настоящее время активным образом развиваются технологии нейросетевого моделирования. Многие задачи, которые раньше представляли сложность для классических алгоритмов, теперь решаются с высокой эффективностью.

Задача распознавания лиц получила новые эффективные решения благодаря появлению и развитию глубоких сверточных нейронный сетей, а также постоянно увеличивающееся вычислительной эффективности современных компьютеров.

В данной части предлагается к рассмотрению подходы, применяемые при построении и обучении нейронных сетей для решения задач распознавания лиц.

## FaceNet

Одной из фундаментальных архитектур нейронных сетей, для распознавания лиц стала FaceNet. Она была опубликована группой исследователей компании Google в 2015 году. В ней были применены новые приемы обучения, такие как триплет потерь, сиамские сети. Данная архитектура проложила путь для создания следующего поколения систем распознавания лиц с использованием машинного обучения.

### Сиамские сети

FaceNet использует специфический тип архитектуры, получившей название сиамской сети (рис. 3).

Сиамские сети состоят из двух идентичных нейронных сетей, каждая из которых имеет одинаковые точные веса. Во-первых, каждая сеть принимает одно из двух входных изображений в качестве входных данных. Затем выходы последних слоев каждой сети отправляются в функцию, которая определяет, степень подобия изображений

В процессе обучения, алгоритм оптимизации применяется не к каждому выходу сети, а к общей функции потерь, принцип работы которой описан в следующем пункте.

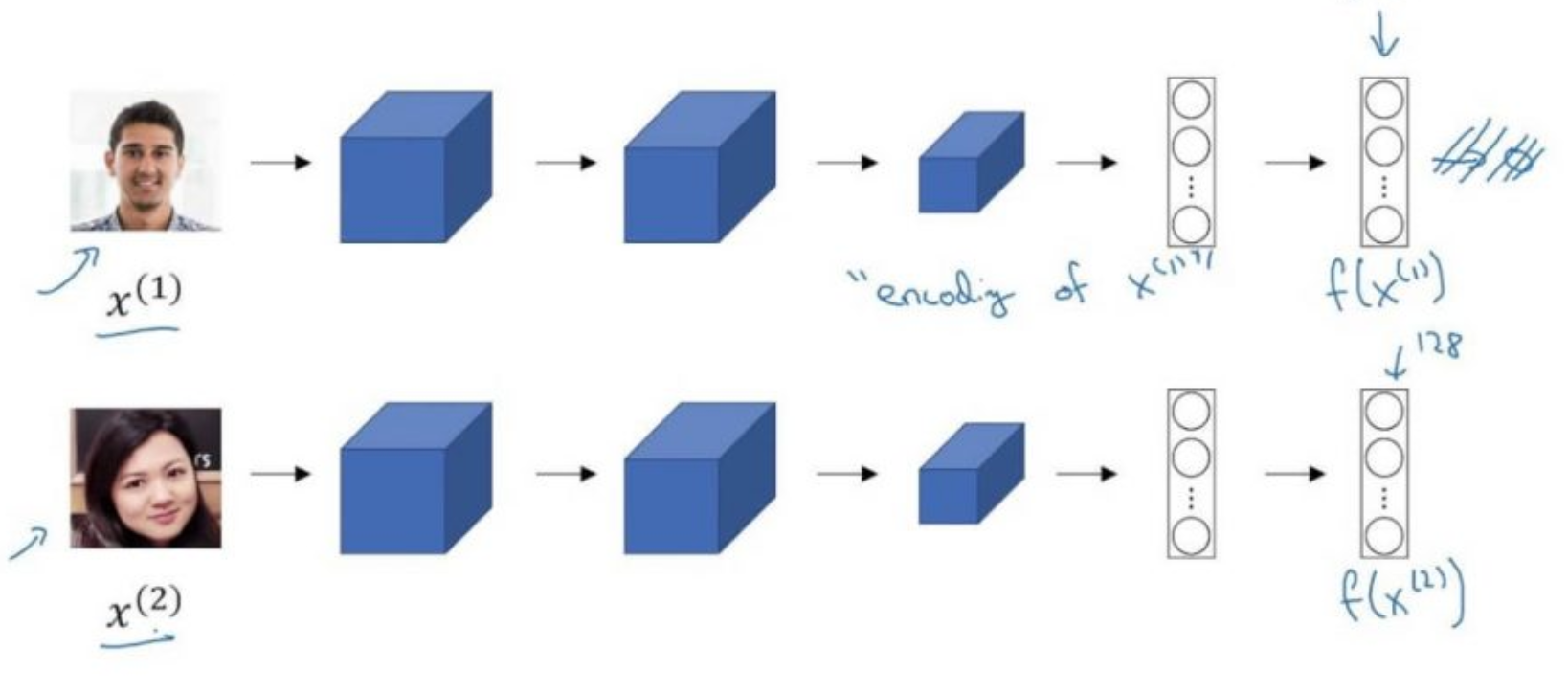


Рис. 3. Принцип работы сиамских сетей.

### Триплет потерь

В качестве функции потерь, FaceNet использует специфическую функцию, которой поставлена цель минимизировать степень различия между лицами, принаджещими к одному классу и максимизировать различия между лицами одного класса.

Данная функция использует следующий принцип формирования обучающей выборки. В выборке выделяются тройки изображений, одно из которых является тестовым (якорь), другое относится к тому же классу, а третье относится к другому (рис. 4). При чем два последних изображения подаются на вход нейронной сети, и соответственно функция потерь вычисляет степерь похожести «верного» и неверного изображения на якорь.

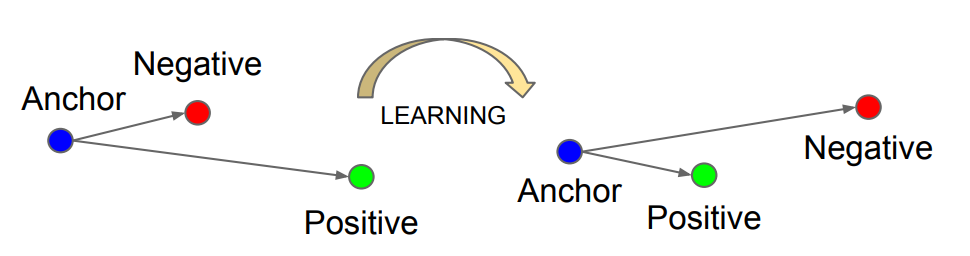


Рис. 4. Принцип формирования троек изображений.

Также имеет смысл выбирать тройки таким образом, чтобы верное и неверное изображения были максимально похожи.

Сама функция потерь для такого типа обучения сети выглядит следующим образом:

,

где – значение энкодинга якоря,

– значение энкодинга верного изображения,

– значение энкодинга неверного изображения.

Коэфициент – константа обеспечивающая, что сеть не будет оптимизировать функцию потерь следующим способом .

При выборе троек изображений следует стремиться к следующему соотношению: при фиксированном , необходимо найти такие и , что:

Выбор троек изображений в процессе обучения сети может происходить следующим образом:

* Выбирать тройки изображений каждые n шагов обучения. Для подсчета сходства можно использовать последнее состояние обучаемой сети.
* Выбирать тройки в процессе итерации из текущего минибатча изображений. При использовании данного подхода следует формировать батчи большого размера.

Правильный выбор троек изображений для обучения критичен для быстрой сходимости.

### Архитектура сети

В описанной выше работе рассматривалось несколько архитектур глубоких сверточных сетей для распознавания лиц. Пример базовой архитектуры приведен на рисунке 5.

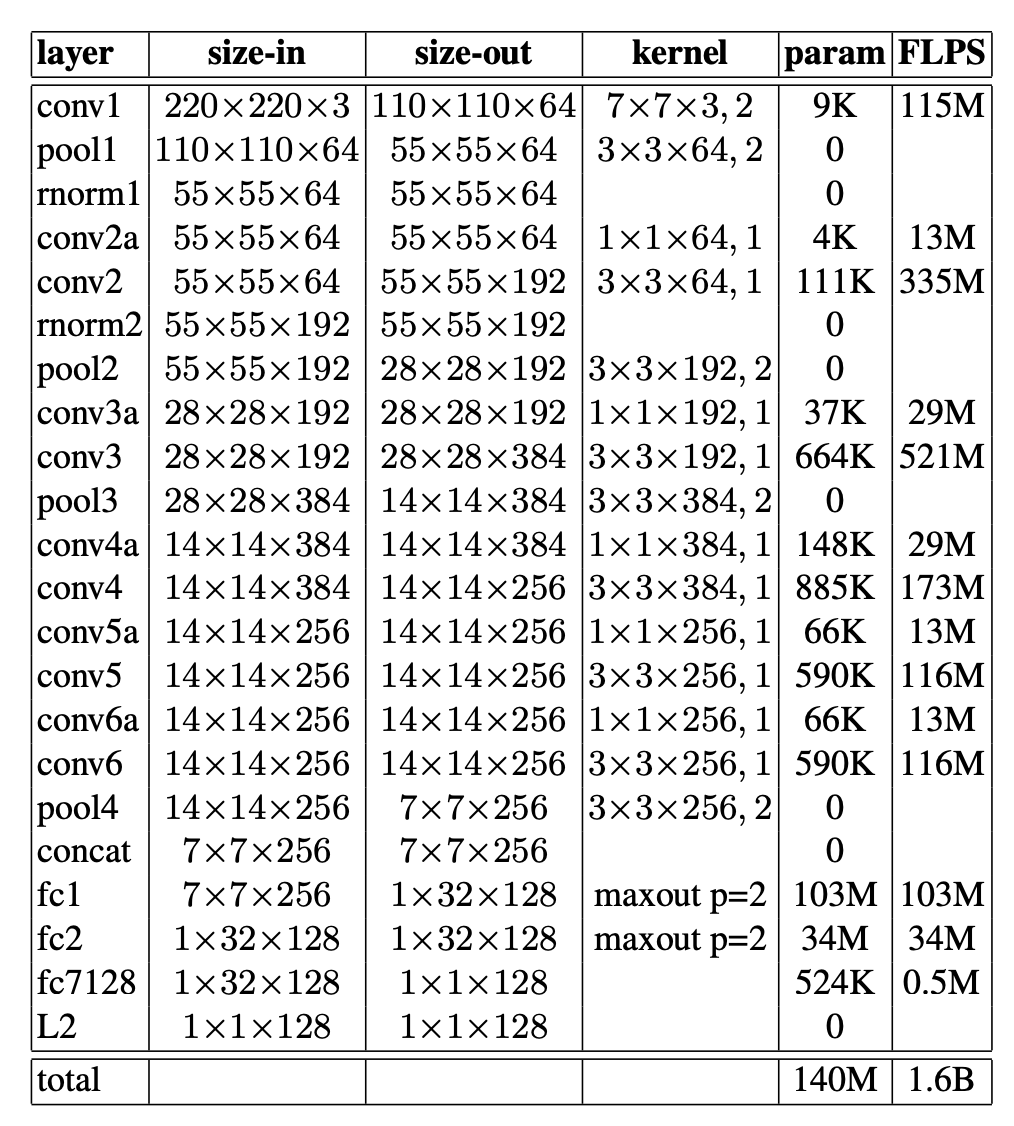


Рис. 5. Пример арихтектуры сети, рассматривающейся в работе.

Другие архитектуры сетей, приведенные в работе состоят из Inception слоев и имеют меньшее число параметров. Сравнение результатов тестирования приведено на рисунке 6.

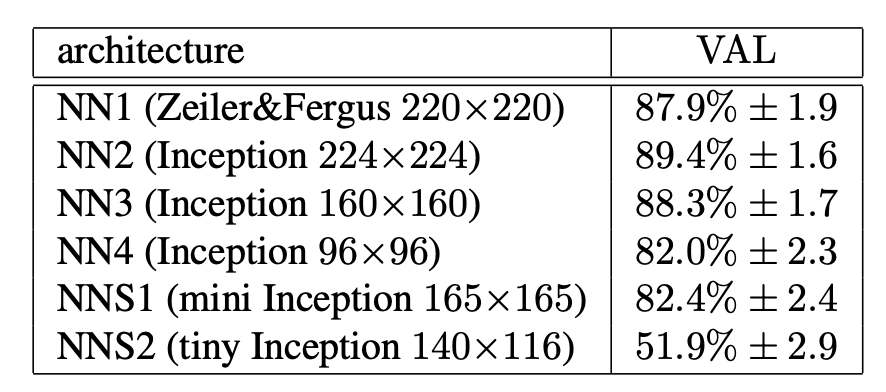
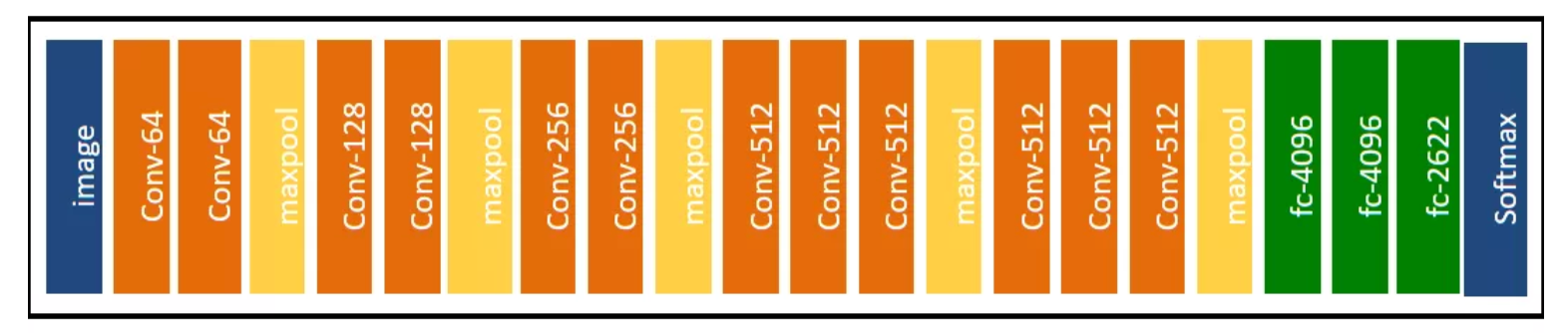


Рис. 6. Сравнение точностей моделей, рассматриваемых в работе.

## VGG-Face

Другой архитектурой глубокой сверточной нейронной сети, использующейся для распознавания лиц является VGG-Face. Она была разработана группой исследователей в университете Оксфорда.

VGG-Face использует более глубокую архитектуру, чем остальные популярные сети для распознавания лиц (рис. 7).



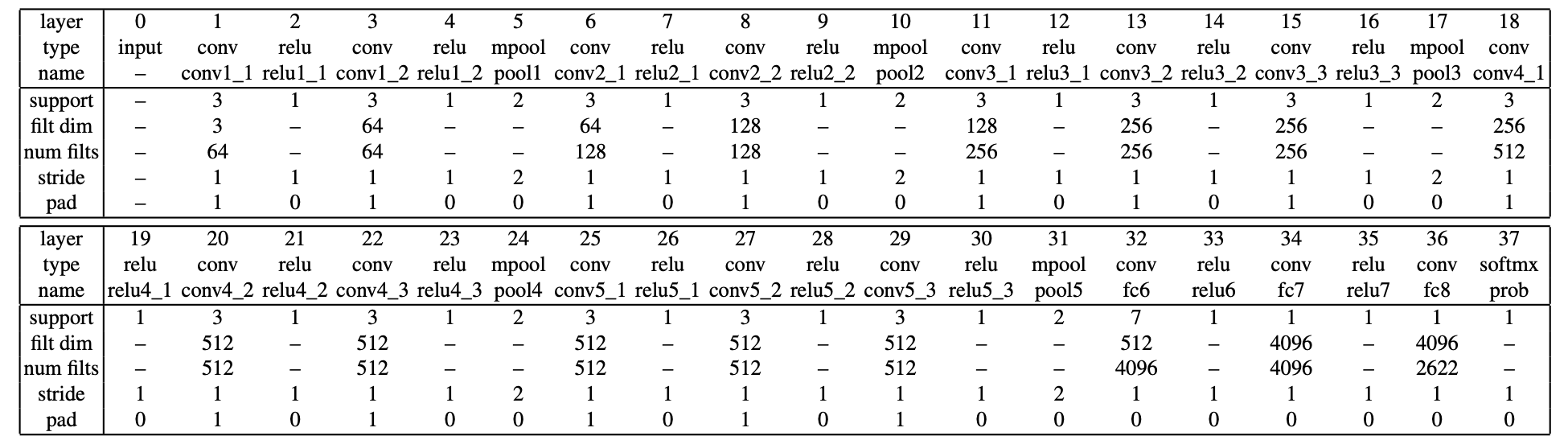


Рис. 7. Архитектура VGG-Face

VGG-Face также использует триплет потерь при обучении. Выбор троек изображений осуществляется на каждой итерации из батча. Сравнение VGG-Face с другими популярными архитектурами приведены на рисунках 8, 9. В сравнении использовались наборы данных YouTube Faces In the Wild и Labeled Faces In the Wild.

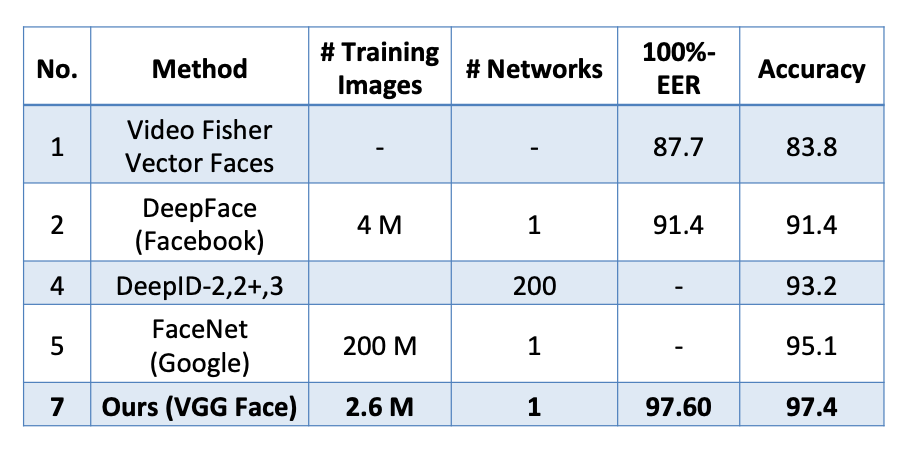


Рис. 8. Сравнение эффективности на наборе данных YouTube Faces In the Wild.

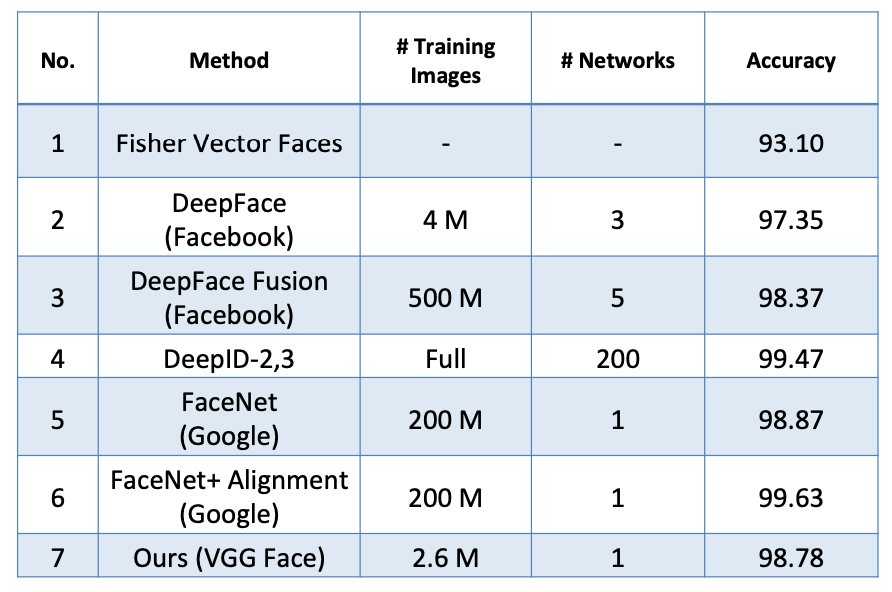


Рис. 9. Сравнение эффективности на наборе данных Labeled Faces In the Wild.

# СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

[1] Eigenfaces [Электронный ресурс]. – Электронные данные. – ­Режим доступа: http://www.scholarpedia.org/article/Eigenfaces.

[2] Fisherfaces [Электронный ресурс]. – Электронные данные. – ­Режим доступа: http://www.scholarpedia.org/article/Fisherfaces

[3] FaceNet: A Unified Embedding for Face Recognition and Clustering [Электронный ресурс]. – Электронные данные. – ­Режим доступа: https://arxiv.org/pdf/1503.03832.pdf.

[4] Deep Face Recognition [Электронный ресурс]. – Электронные данные. – ­Режим доступа: http://www.robots.ox.ac.uk/~vgg/publications/2015/Parkhi15/

parkhi15.pdf.

[5] Deep Face Recognition Poster [Электронный ресурс]. – Электронные данные. – ­Режим доступа: http://www.robots.ox.ac.uk/~vgg/publications/2015/Parkhi15/

poster.pdf.