# Введение

Реконструкция трехмерных объектов в настоящее время находит применение в обширном списке сфер деятельности. Она активно используется в программных продуктах для архитекторов, медицинских работников, в сфере обучения. 3D-реконструкция имеет много применений в поиске объектов, понимании сцены, в распознавании и слежении за объектом, навигации, проектировании, взаимодействии человека и компьютера и виртуальное обслуживание. Получение облаков точек возможно благодаря технологиям 3D-сканирования и фотограмметрии.

 Разница между двухмерными и трехмерными системами – в количестве получаемой информации о лице. 2D-системы работают с «плоской картинкой», 3D –системы – с объемной моделью. За счет обработки большей информации о лице – не только о его поверхности, но и о форме -  можно добиться большей точности распознавания.

Идентификация человека по биометрическим признакам представляет собой большую проблему в области компьютерного зрения. В последние годы большую популярность взыскали системы по распознаванию людей по лицу. Появляется все больше персональной техники с модулями защиты на основе данных технологий. Наиболее перспективными кажутся продукты для распознания человека по трехмерному скану лица.

Целью курсовой работы является проверка некоторых методов оптимизации нейронных сетей для уменьшения количества неправильных ответов нейронной сетью.

Для достижения данных целей, были поставлены следующие задачи

1. Тестирование нейронной сети на базе данных BosphorusDB;
2. Примерение Pruning для данной нейронной сети и тестирование результата;
3. Реализация поворотов 3D облаков;
4. Тестирование нейронной сети с фреймворком Keras.

VGG-Face – «сверхглубокая» сверточная нейронная сеть из 16 слоев. Данная сеть была обучена на базе порядка 10 миллионов изображений. Рассмотрим архитектуру сети VGG-Face представленную на рисунке 4. Сеть состоит из 11 слоев, каждый из которых содержит линейный оператор, за которым следует один или несколько нелинейных, таких как *ReLU*, и слой пуллинга (*maxpool*) – нелинейное уплотнение карты признаков. Первые восемь таких блоков называются сверточными, поскольку линейный оператор представляет собой группу линейных фильтров (линейная свертка). Последние три слоя называются полно связными (*Fully Connected, FC*). Они аналогичны сверточному слою, но размерность фильтров соответствует размерности входных данных, т.о. каждый фильтр «воспринимает» данные из всего изображения. За всеми слоями свертки следует слой выпрямления (*ReLU*). Первые два выходных слоя *FC* имеют размерность 4096, а последний слой FC имеет размерность 2622 либо 1024, в зависимости от функции потерь, используемых для оптимизации. В первом случае результирующий вектор передается в слой *softmax* для вычисления апостериорных вероятностей класса. На вход сети подается изображение лица *224*×*224* из которого вычли значение среднестатистического лица (вычислено на основе обучающей выборки).

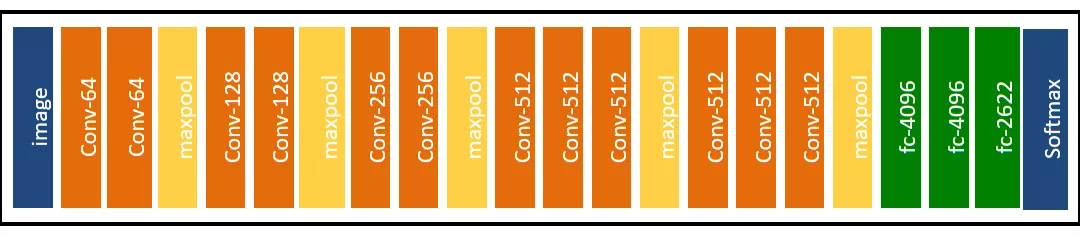


Рис. 4. Архитектура сети VGG-Face.

Чтобы научить вышеописанную сеть распознавать 3D изображения, была сгенерирована база карт глубины – ортогональные проекции трехмерных изображений лиц на двумерную плоскость. Для того, чтобы сделать систему более устойчивой к ошибкам, к каждой трехмерной фигуре были применены следующие модификации

1. Повороты трехмерной модели;
2. Добавление на случайные места модели дополнительных объектов, для имитации перекрытий (например, волосы на лице, покрытие руками или артефакты);
3. Деформация частей трехмерной модели;
4. Добавление «черных дыр» на полученные проекции – участки с отсутствующим значением глубины.

Примеры искажений, примененных к моделям, показаны на рисунке 5.

# Методы оптимизации, которые могут быть применены к нейронным сетям в данной задаче.

## Pruning.

Данный метод оптимизации используется для уменьшения объёма памяти занимаемым весами нейронной сети. Данный способ состоит из двух ключевых моментов:

1. Определение нейронов, не влияющих на результат действия всей сети.
2. Обрезка данных нейронов, чтобы они не вносили вклад в результат.

В эксперименте проводился pruning слоёв fc6, fc7, fc8 с адаптивным уровнем.

Результаты эксперимента оказались следующими: удалось снизить объём занимаемого весами диского пространства с 538 942 КБ до 533 388 КБ. При этом ошибка распознавания нейронной сетью снизилась с 16.4% до 16.2%.

## Поворот изображений.

Выравнивание изображений может способствовать качеству распознавания их нейронной сетью. Для выполнения выравнивания используется стандартная реализация алгоритма IPC в пакете MATLAB.

ICP алгоритм используется для сведения к минимуму разницы между двумя облаками точек.

ICP алгоритм состоит из следующих шагов:

1. Точки объединяются по критерию ближайшего соседа.
2. Оцениваются параметры преобразования с помощью среднеквадратичной функции стоимости.
3. Объединённые точки преобразуются с помощью оценочных параметров.
4. Многократные итерации (заново связывая точки и так далее).

Для тех облаков точек, на которых при стандартных параметрах, нейронная сеть ошиблась применялось по 10 итерации ICP. Всего из 338 элементов, на которых нейронная сеть ошибалась ранее, после поворота смогла распознать 144 правильно и вновь ошиблась на 194.

## Применение фреймворка машинного убучения Keras.

На сегодняшний день задачи машинного обучения в подавляющем большинстве случаев осуществляются с помощью нейросетевых библиотек и фреймворков. Данные программные пакеты реализуют все необходимые функции и методы оптимизации нейронных сетей, а также позволяют вести собственные разработки.

Исходная нейронная сеть обучалась и применялась с использованием Caffe.

Заключение

В данной работе рассмотрены некоторые методы модификации слоёв нейронной сети а также данных в задаче распознавания 3D-лиц человека. Применение методов, рассматриваемых в задаче позволило незначительно снизить количество ошибок нейронной сети на базе данных THE BOSPHORUS DATABASE.

Ссылки:

[1] <http://bosphorus.ee.boun.edu.tr/default.aspx>

[2] <https://arxiv.org/abs/1703.10714>