**Филиал Московского Государственного Университета**

**имени М. В. Ломоносова в городе Ташкенте**

**Факультет прикладной математики и информатики**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_



Бакалаврская программа “Прикладная математика и информатика”

Курсовая работа

**“ Распознавание эмоций по изображению человека с помощью нейронной сети при не большом объеме обучающих данных ”**

Работу выполнил

студент **Нишанов Эмиль Рустамович**

Научный руководитель:

к.ф.-м.н. **Алексеев Дмитрий Владимирович**

Ташкент

2019­­­­

**Содержание**

Содержание

Введение

Приоритетные задачи

1. Формализация задачи

1.1 Постановка задачи

1.2 Обзор существующих алгоритмов

2. Теоретические обоснования

2.1 Сверточные нейронные сети

3. Реализация алгоритма

3.1 База изображений

3.2 Сверточная нейронная сеть

3.3 Использование алгоритма

4. Заключение

4.1 Перспективы

Список литературы

**Введение**

Что такое эмоции? Сможем ли мы верно ответить на этот вопрос? А какова вероятность, что мы точно сможем распознать лицевую экспрессию у своего собеседника? Данный вопрос давно интересует не только психологов современного мира, так же этой проблемой давно занимаются и исследуют ученые всего мира. Проведенные исследования на эту тему были затронуты еще в книге «Выражение эмоций у человека и животного» Чарльза Дарвина, где он утверждает, что эмоции скорее варьируются от вида к виду, то есть это видовой признак, чем культурный. В начале 20 века существовала достаточно фундаментальная теория о том, что не существует «универсальных» эмоций. Данная теория опиралась на зависимость представителя от типа культуры. На эту же тему в 1969 году была издана одна из самых важных литературных работ, проделанная психологами и учеными. Это исследование Пола Экмана и Уоллеса Фризена, которые были явными оппонентами существующих канонов эмоционального происхождения. Они доказали, что существует 6 универсальных эмоций, которые не зависят от культурной и национальной принадлежности: удовольствие, страдание, страх, гнев, удивление и отвращение. Конечно же, к данным эмоциям всегда относят нейтральное состояние лица человека.

В последние годы распознавание эмоций стало принимать популярный характер не только в методах и специфике психологии и социологии. С учетом всемирной и всепоглощающей идеи об искусственном интеллекте ученые и программисты всего мира пришли к консенсусу необходимости достаточно точного определения эмоций с помощью программных средств. Распознавание любого вида проявления человеческого отклика и реакции это одно из самых важных препятствий для создания искусственного интеллекта. На данный момент такой вектор исследования очень важен и имеет огромное значение для всего человечества, особенно в отраслях бизнеса и военных структурах.

Принимая во внимание нынешнее пристрастие к автоматизации процессов, ученые, и не только, стали заниматься проблемами распознавания и детектирования большого количества объектов, окружающих нас, но даже, в какой-то мере, это приносит хорошие результаты. Распознавание номеров у машин уже вошло в обиход и активно используется правоохранительными органами, детектирование лиц с помощью программ и компьютера уже давно не новинка, так как во многих крупных компаниях стоит система безопасного входа по определению сотрудника посредством фотографий. Но один из самых важных прорывов во взаимодействии безопасности и детектирования был совершен при введении на особо важные объекты камер с распознаванием отрицательных, враждебно настроенных эмоций, которые помогут вычислить преступника, грабителя или террориста. Нельзя не упомянуть про прорывы в распознавании голоса, тональности текста, речи и так далее. Исходя из этого, видно, что данное направление не стоит на месте и успешно развивается. На данный момент значительную нишу в этом научном направлении занимает машинное обучение. Уже давно признали, что машинное обучение одно из самых успешных, конструктивных, универсальных и оптимальных методов для распознавания объектов на фотографии, видео фрагментах, фильмах и электронных записях. Машинное обучение показало и продолжает показывать огромное увеличение точности алгоритмов по детектированию. Но главенствующее положение на данный момент в области машинного обучения занимают нейронные сети, которые безупречно доказали свою эффективность, и не перестают улучшать результат, поэтому мы будем использовать именно этот метод. Данный метод в основном требует либо огромного

3

количества данных, которые послужат тренировочными и тестовыми образцами, либо идеально выстроенной нейронной сети, которая учитывает все нюансы и специфические моменты требуемой задачи. Важной проблемой в области распознавания лицевых экспрессий – это отсутствие согласованной и полностью учитывающей всех особенностей эталонной базы данных, на которой бы строились все алгоритмы. Данный эталон служил бы некоторым главенствующим инструментом, который обеспечивал бы возможность сравнительной характеристики ныне существующих методов и подходов. Было сделано множество попыток создать данную базу фотографий, но ученые и программисты так и не пришли к общему паттерну. Возможно, проблема кроется в глобализации маркетинга и поискам выгоды в каждой научной отрасли.

Идея нейронных сетей была взята из физиологии, ведь, как всем известно, прототип этой архитектуры был получен из примера функционирования человеческого мозга. Человеческая биологическая нейронная сеть была взята за основу. Ученые пытались смоделировать работу нашего мозга и протекающих в нем реакций, поэтому данный метод и нашел применение в основном в распознавании объектов, образов и в задачах прогнозирования. Существует несколько методов, взятых из научной литературы, для распознавания эмоций на лице человека:

1. Статичный метод. В данном методе классификатор, который используют авторы, распознает каждый кадр и сравнивает его с заявленными вначале эмоциями, и на основе каждого кадра делает вывод. Обычно в таких подходах используется Наивный Байесовский классификатор. В то же время, он использует строгое и нереалистичное предположение о том, что все признаки не взаимосвязаны друг с другом и являются новым классом.
2. Динамический метод. В этом методе, классификаторы принимают во внимание временный шаблон, полученный до этого, для отображения выражения лица. Скрытая Марковская модель наилучше подходит для такой реализации. СММ отслеживает сгенерированные до этого состояния, и на основе этого, достаточно оптимально выводит полученный результат. На основе данного алгоритма в работе получилось даже отследить долгий фрагмент и получить необходимый прогноз по отслеживанию эмоции. Работы по теме распознавания можно обычно разделить на два типа:

• Сравнение методов по обработке и определение эмоции

• Введение нового метода, полученного на основе предыдущих работ, который усовершенствует предыдущие результаты.

Иногда в последнее десятилетие можно выделить еще один тип, который описывает разницу в осуществлении разных подходах на разном типе машин и разном типе аппаратурной части (обучение на CPU и GPU [6]) .

4

**Приоритетные задачи**

Задача идентификации эмоций обычно решается в два этапа, каждый из которых бесспорно важен и определяет успешность результата, первый распознавание лица на изображении, второй распознавание лицевой экспрессии на основе полученного лица. В своей работе я буду рассматривать только второй этап, предполагая, что на изображении всегда есть лицо и будет рассмотрен метод, который будет распознавать эмоцию человека из поданного изображения лица при помощью построенной и обученной на выборке нейронной сети. Количество эмоций описанной в моей работе составляет две: лицо с улыбкой и без нее. Мой приоритет улучшить значение точности алгоритмов на относительно малом количестве выборки, увеличивать количество эмоций и достигнуть того, что компьютер решает проблему распознавания не хуже человеческого глаза. Самая основополагающая задача сгенерировать подход, который одинаково будет распознавать эмоции, вне зависимости от национальности и возраста.

Дополнительные задачи, которые необходимо выполнить:

• Необходимо изучить и исследовать нейронные сети в области детектирования экспрессий, и вообще в области распознавания

• Провести качественный обзор и сравнение уже существующих методов, с помощью которых уже была решена поставленная мною задача

• Выделить проблемы данной научной области и обозначить перспективы развития

Для дальнейшего прорыва в получении и создании искусственного интеллекта необходимо понимание особенностей работы мозга, в том числе распознавания объектов. Поэтому вектор данной работы очень востребован и популярен.

**1. Формализация задачи**

**1.1 Постановка задачи**

Данная задача была уже решена многими крупными компаниями, которые занимаются разработками систем по детектированию и распознаванию. Поэтому, в первую очередь данная работа носит в себе исследовательский характер. Поставлена цель создать алгоритм, изучить методы и средства разработки по идентификации двух эмоций. Конечный алгоритм, как и любой другой алгоритм по распознаванию лицевых экспрессий, будет разделен на три основных последовательных этапа:

1. Определения лица

Данный этап в работе требует тщательной работы, потому что только эта одна тема рассматривается как отдельная серьезная работа, над решением которой уже давно работают ученые. Это задача требует учета многих особенностей. Например, присутствует ли лицо на изображении или полученном кадре? Если был получен положительный ответ, то с помощью программных и аппаратных средств должен быть получен размер лица и его положение (то есть, его координаты в двухмерной плоскости) на изображении. Так же сюда может быть включен подпункт про отслеживание лица, то есть изменение его положения в плоскости по времени.

Однако, задача построения качественного детектора не будет решаться в работе. Поэтому, мы предполагаем, что лицо на фотографии существует.

5

1. Определение и описание особенностей лица человека

На данном этапе работы подход должен определять положение в плоскости значимых частей лица, которые будут задействованы в дальнейшем распознавании эмоций и лицевых экспрессий. В число этих важных частей лица могут входить: рот, глаза, брови, нос и так далее. Варьируются эти варианты в зависимости с выбранном или созданном алгоритме при определении. Для решения этой части необходимо построение архитектуры нейронной сети, которая будет генерировать матрицу весов для обучающей выборки, учитываю все эти особенности. Нейронная сеть - это некий черный ящик, то есть пользователь не имеет возможности узнать, какие ключевые элементы в основном учитывает построенная сеть.

1. Распознавание эмоции

Эта часть работы уже финальная и требует работы с полученными данными. На основе матрицы весов, полученной с помощью обучения сети, алгоритм должен сделать прогноз с помощью нейронной сети. Согласно информации, данной нам с первых двух частях работы, наш алгоритм должен выполнять свою первоначальную и незыблемую функцию - определить эмоцию на изображение.

* 1. **Обзор существующих алгоритмов**

Анализ распознавания лицевых экспрессий, или, как тривиальнее, эмоций, является сложной и комплексной задачей, поэтому существует огромное количество методов и алгоритмов, рассмотренных в работах разных авторов. Некоторые методы очень спорны и имеют существенные недостатки, так как авторы пытаются найти все более изощренный метод, который докажет свою состоятельность и выявит недостатки предыдущих алгоритмов. Есть некоторые особенности лица человека, из-за которых не существует на данный момент эталонного способа решения задачи. Часто многие методы не учитывают аксессуаров и лицевой растительности, типа бороды, усов, неравномерных бровей и объемных причесок, закрывающих или мешающих распознаванию. Плюс к тому, эмоция очень абстрактное выражение, которое очень изменяется от человека к человеку. Часто мы сталкиваемся с проблемой непонимания некоторой экспрессии, например, открытый рот не всегда означает удивление. Однако, основной недостаток, как было сказано выше, отсутствие эталонной обучающей выборки, на основе которой создается алгоритм. Методы, расписанные ниже, чаще применяются к детектированию лиц на изображении, но многие были применены и к эмоциональному распознаванию.

1. Распознавание эмоций в реальном времени на основе нечеткой логики

Данный алгоритм учитывает особенности нечеткой логики (Fuzzy Logic) . Авторы данного алгоритма показывают преимущества решения подобных плохо формализуемых задач, таких как определение эмоций. Методы адаптивного управления строят самообучающиеся распознающе-управляющие комплексы, которые одновременно решают множество задач, таких как:

* автоматическая классификация
* распознавание образов на изображении
* качественная оценка полученных значений (в данном случае определение эмоции)

6

* небольшая ошибка прогнозирования
* принятие окончательного решения

Данный метод доказал свою состоятельность и показал хорошие погрешности на тестируемых данных, которые привели ученых к выводу, что данный подход необходимо изучать и развивать. Из негативных сторон можно выделить то, что все эти задачи не всегда целесообразно и выгодно решать в одной системе. Иногда проще и конструктивнее создание комплексной системы, учитывающей все особенности. На Рис. 1 представлен пример определения эмоции с помощью этого метода.



Рис. 1: Модель эмоций на основе Fuzzy Logic

1. Метод главных компонент

Одним из самых популярных и проработанных подходов является метод главных компонент. Метод главных компонент (Principal Component Analysis, PCA) в основном применяется для сжатия или снижения признаков без существенной потери признаков. В основу этого был и положен метод определения лица человека, у которого многие несущественные признаки можно отбросить, то есть провести сжатие. Основной целью данного метода является уменьшение количества признаков таким образом, чтобы оно наилучшим способом имело возможность описать типичный эталон. Проделанная работа обычно сводится к тому, что вся наша обучающая выборка (в данном случае лиц) преобразовывается в одну общую матрицу данных, где каждое изображение кодируется в собственной строке. Для успешного выполнения этого алгоритма все изображения должны быть приведены в одному размеру и имели нормированную гистограмму.

Для каждого изображения путем вычисления получаются главные компоненты. Обычно таких ключевых значений от 5 до 200. Остальные компоненты берутся за мелкие различия лицами иной шум. Далее с помощью какой-либо метрики (обычно Евклидово расстояние) применяется сравнение главных компонент полученного изображения с компонентами других фрагментов. Этот метод достаточно оптимальный, но для его идеальной реализации необходимо, чтобы все входные данные были с одинаковой освещенностью, иначе возможно появление дополнительных ненужных компонент, которые могут затруднить определение лицевой экспрессии. Плюс к тому, данный подход очень чувствителен к появлению аксессуаров и растительности на лице. Но самый главный недостаток - высокая трудоемкость этого подхода. Иногда определение набора собственных векторов очень ресурсоемко и занимает большое количество времени, пример можно увидеть на Рис.

7

2. Существуют способы, чтобы распознавание и вычисление происходило быстрее.

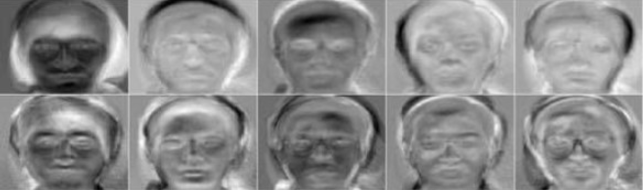


Рис. 2: Пример первых десяти собственных векторов (собственных лиц), полученных на обучаемом наборе лиц.

8

**2. Теоретические обоснования**

**2.1 Сверточные нейронные сети**

Нейросетевые методы - это методы, базирующиеся на применении различных типов нейронных сетей (НС). Основные направления применения различных НС для распознавания образов и изображений:

* применение для извлечение ключевых характеристик или признаков заданных образов,
* классификация самих образов или уже извлечённых из них характеристик (в первом случае извлечение ключевых характеристик происходит неявно внутри сети),
* решение оптимизационных задач.

Архитектура искусственных НС имеет некоторое сходство с естественными нейронными сетями. НС, предназначенные для решения различных задач, могут существенно различаться алгоритмами функционирования, но их главные свойства следующие.

НС состоит из элементов, называемых формальными нейронами, которые сами по себе очень просты и связаны с другими нейронами. Каждый нейрон преобразует набор сигналов, поступающих к нему на вход в выходной сигнал. Именно связи между нейронами, кодируемые весами, играют ключевую роль. Одно из преимуществ НС (а так же недостаток при реализации их на последовательной архитектуре) это то, что все элементы могут функционировать параллельно, тем самым существенно повышая эффективность решения задачи, особенно в обработке изображений. Кроме того, что НС позволяют эффективно решать многие задачи, они предоставляют мощные гибкие и универсальные механизмы обучения, что является их главным преимуществом перед другими методами [4,5] (вероятностные методы, линейные разделители, решающие деревья и т.п.). Обучение избавляет от необходимости выбирать ключевые признаки, их значимость и отношения между признаками. Но тем не менее выбор исходного представления входных данных (вектор в n-мерном пространстве, частотные характеристики, вэйвлеты и т.п.), существенно влияет на качество решения и является отдельной темой. НС обладают хорошей обобщающей способностью (лучше чем у решающих деревьев [5]), т.е. могут успешно распространять опыт, полученный на конечном обучающем наборе, на всё множество образов.

Сверточные сети представляют собой огромный прорыв в области информационных технологий в рамках изобретения компьютерного зрения и распознавания различного род образов. Такая модель сети используется в задачах широкого спектра различия, таких как распознавание человеческого голоса, обработка аудио сигналов, обработка временных рядов или анализа смыслового содержания текста. В последнее время сети данного вида стали применяться в процессе моделировании искусственного интеллекта для различных компьютерных приложений.

**Свёрточная нейронная сеть** (англ. Convolutional neural network, CNN) представляет собой специальную архитектуру, которая состоит из искусственных нейронных сетей. Впервые данная сеть была предложена Яном Лекуном. Целью такой сети стало эффективное распознавание изображений, поступающих на вход. Данная сеть является одной составляющей в технологии глубокого обучения. Эта технология создана на основе анализа принципов работы коры головного мозга, отвечающей за зрительное восприятие человеком окружающего мира. В этой области были открыты так называемые «простые

9

клетки», которые реагируют на прямые линии, идущие под разными углами, и «сложные клетки», реакция которых зависит от некоторого набора активных простых клеток. В обобщенном смысле, идея сверточных нейронных сетей заключается в чередовании сверточных и субдискретизирующих слоев (другими словами, под выборки). Сверточная нейросеть является особым видом нейросетей прямого распространения. Под прямым распространением понимается то, что переменные нейроны в этой сети разбиты на группы, называемые слоями. И когда такая слоистая нейронная сеть применяется к данным, то активация слоев — значение этих переменных— подсчитывается последовательно: сначала значение активации первого слоя, потом значение активации второго слоя, и так до последнего слоя. Активация последнего слоя и служит выходами нейронной сети. Ключевым моментом является то, что активации внутри некоторого определенного слоя могут подсчитываться параллельно в связи с тем, что они независимы друг от друга. Нейросети данного типа удобно и эффективно моделировать на современных высокоскоростных процессорах, в том числе и на графических сопроцессорах.

Из недостатков можно смело выделить неопределенность структуры, многие параметры необходимо настраивать под задачу, то есть эмпирическим путем, методом проб и ошибок, а обучение сети на огромных датасетах может длиться неделями. Поэтому для многих задач берется уже обученная модель с матрицей весов, так как намного более перспективно дообучать сеть на своих данных. Особенно это актуально при небольшом количестве изображений. На данный момент существует множество архитектур, построенных на основе сверточных нейронных сетей. Примерами являются: AlexNet, VGGNet, ResNet, GoogLeNet.



Рис. 3: Архитектура готовой сверточной модели нейронной сети VGG-16.

**Архитектура CNN**

Как было описано выше, у сверточной нейронной сети реализовано несколько слоев, которые преобразуют входные данные в вектор принадлежности к классам. Основные слои, которые используются в сверточной сети это слой свертки, слой линейной ректификации, слой субдискретизации и полносвязный слой. У каждого из этих своя функции и задача, но в итоге мы получаем сложно структурированную сеть с наилучшими результатами для классификации объектов на изображении. На Рис.9 подробно иллюстрирована архитектура CNN. Рассмотрим подробнее каждый из слоев:

1. Входной

Первый вход у любого типа нейронной сети – входной. Этот слой хранит необработанные значения пикселей входного изображения. В случае с изображением

10

это некоторое количество значений по ширине, высоте и глубине.

1. Слой свертки

Сверточный слой - это основной слой в CNN. Данный слой зиждется на операции свертки, которая заключается в том, что по всем входным данным «пробегается» матрица весов ограниченного и небольшого размера, создающая новое значение для некоторого числа пикселей. Такая матрица называется ядром свертки, ее особенностью является то, что весовые коэффициенты такой матрицы определяются во время обучения и заранее неизвестны. К необходимости использования этого слоя можно отнести то, что значительно уменьшает количество весов, созданных для нейронов, в отличии от полносвязной нейронной сети.

1. Слой с блоком линейной ректификации(ReLu)

Данный слой содержит функцию активацию после операции свертки. В сверточной нейронной сети используется вместо привычных гиперболической или сигмоидальной функции активации, функции f(x) = max(0,x). Эта активационная показала очень хорошие результаты на задаче классификации изображений. Более того, выбор был сделан на этой функции активации по причине того, что для большого объема данных она позволяет обучаться сети за адекватное время.

1. Слой субдискретизации

На данной этапе происходит пулинг полученных данных, то есть подвыборка. На этой стадии матрица характеристик, она же матрица со значениями пикселей уплотняется до матрицы меньшего размера. На примере изображений пулинг уменьшается размерность по ширине и высоте. Такое преобразование делается на уникальных и непересекающихся областях, обычно из прямоугольника 2х2 выбирается один пиксель, значение которого максимальное. Тем самым происходит удаление незначительных деталей и разрешение задачи переобучения. Существуют субдискретизации не только с функцией максимизации, но еще и функцией среднего или L2 нормирования.

1. Полносвязный слой(fullyconnected layer)

Данный слой представляет собой обычную нейронную сеть, где каждый нейрон связан со всем предыдущими. Предназначение такого слоя в том, чтобы вычислить уже по имеющимся данным вероятность принадлежности к классу. Значительно сократив объем данных по сравнению с исходными, наша сеть их передает в полносвязный нейронную сети, в которой может быть еще несколько слоев внутри.

11

**3. Реализация алгоритма**

**3.1 База изображений**

Первые проблемы, с которыми приходится столкнуться при обучении сети – большое количество шумов на изображениях, неоднородные данные или просто недостаточное количество картинок для полноценного обучения сверточной нейронной сети. Как отмечалось ранее, работа имеет исследовательский характер, поэтому база изображений собиралась в «ручную», тем самым, демонстрируя весь спектр работы этой области. Использована популярная база изображений «extended-cohn-kanade-images», где в следствии были отобраны и отсортированы на 2 класса эмоций.

 Размещено на три директории для «тренировки», «проверки» и «тестировании» нейронной сети. В каждой директории размещается два класса эмоций где «with» - портреты людей с улыбкой и «without» - портреты людей без нее.



Примеры изображений с улыбкой.



Примеры изображений без улыбки.

Количество обучающей выборки составило по 3000 изображений на каждый класс эмоций. Для задачи распознавания размер обучающей выборки достаточно скромен, для сравнения известная база изображений для классификации содержит в себе более 10000 изображений на класс.

Для пополнения базы изображения, в ходе работы использовалось keras.preprocessing.image «ImageDataGenerator», что позволило увеличить базу вдвое, благодаря ImageDataGenerator(rescale = 1. / 255, horizontal\_flip=True, fill\_mode='nearest') параметру «horizontal\_flip».

12

**3.2 Сверточная нейронная сеть**

В ходе исследовательской работы было проведено 6 экспериментов с использованием сверточной архитектуры. Каждый эксперимент имеет за собой собственную архитектур. С использованием методов проб и ошибок, в последнем эксперименте использовалась архитектура:

Слои выделяющие признаки:

3 слоя свертки и подвыборки с функцией активации ReLu:

Conv2D(32… 64… 128, (3, 3))

Conv2D(32… 64… 128, (3, 3))

MaxPooling2D(pool\_size=(2, 2))

Далее слой Flatten() => Позволяющий представлять любую матрицу в вектор.

Классификатор:

Состоящий из полносвязной сети и выходной слой с двумя нейронами.

Также после каждого слоя присутствует слой нормализации: BatchNormalization()

На рис. 4 представлено графики при обучения последнего экперимента.

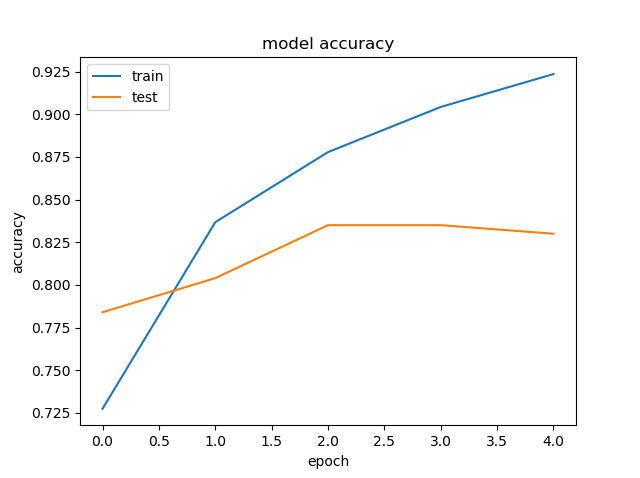
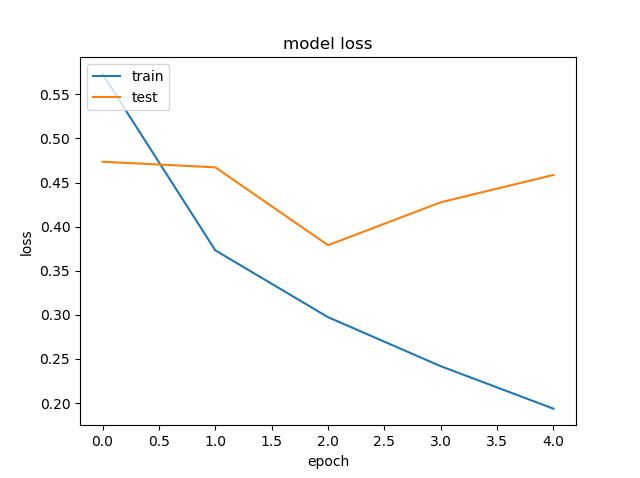


Рис. 4 – Графики изменения при обучении функции потери и точность.

Обучение производилось на ноутбуке с параметрами:

name: GeForce GT 710M major: 2 minor: 1 memoryClockRate(GHz): 1.55

totalMemory: 1.00GiB freeMemory: 823.06MiB

Длительность обучения составляло по 32-36 часов.

13

**3.3 Использование алгоритма**

Самой важной деталью программы и алгоритма является работа на реальных данных. Поэтому мной были взяты фотографии согласившихся на обработку данных студентов, которые должны были продемонстрировать на фото определенное проявление эмоций. Получился маленькая выборка, которая должна была проверить адекватность работы программы. Первое, что необходимо было сделать – это обработать полученную фотографию, с помощью препроцессора keras – ImageDataGenerator, который приводит полученное изображение в размер – (150, 150).

1. Тест:



0.21255456 - вероятность «с улыбкой»

0.7874454 - вероятность «без улыбки»

Как мы видим предсказание выполнено успешно!

1. Тест:



0.03484018 - вероятность «с улыбкой»

0.9651598 - вероятность «без улыбки»

Во второй раз модель предсказала неверно!

14

1. Тест:



0.66362244 - вероятность «с улыбкой»

0.33637765 - вероятность «без улыбки»

В третий раз модель показала правильный результат.

15

**4. Заключение**

В проделанной работе была реализована программа для распознавания 2-х эмоций на человеческом лице по изображению. Был придуман алгоритм обработки входных данных, а также была построена архитектура нейронной сеть, которая успешно была протестирована на тестовых данных. Более того, полученная нейронная сети идеально удовлетворяла соотношению время обучения/точность алгоритма. Так как обучение нейронной сети очень долгий и сложный процесс, поэтом данный фактор играл ключевую роль при выполнении работы. Самым трудным этапом в данной работе являлся эмпирический подбор параметров сети, который определял финальную точность и создание собственной базы изображений. Получение весов при обучении сверточной нейронной сети – процесс на подобии «черного ящика», этим обуславливалась критическая проблема при построении такой архитектуры. При выполнении практической и теоретической части работы были выполнены данные поставленные задачи:

1. Изучение такого инструмента машинного обучения как нейронные сети, которые в настоящее время превалируют в анализе данных и показывают наилучшие результаты со сложными задачами. На этом этапе так же были выделены и рассмотрены другие методы, с помощью которых решалась задача детектирования лицевых экспрессий, был проведен сравнительный анализ и выявлены достоинства и недостатки каждого из пару методов.

2. Смоделировал архитектуру сверточной нейронной сети, которая показывает один из наилучших результатов точности детектирования эмоций для небольшого объема обучающей выборки. В дополнении к этому, обучение на такой сети оказалось оптимальным для доступных машинных ресурсов, и при этом обучалась в более 20 раз быстрее, чем на процессоре.

3. Был успешно создан алгоритм по распознаванию эмоций на человеческом лице, а также программа была проверена в условиях реальной жизни. При несложных манипуляциях алгоритм такого типа можно встраивать в системы видеонаблюдения и лайв-детектирования.

4. В работе, в первую очередь, была поставлена задача показать высокую точность при детектировании при небольшом объеме обучающей выборки. Достигнутая точность оказалось достаточно высокой.

16

**4.1 Перспективы**

Завершая данную работу, было намечено огромное количество перспектив развития данной области и усовершенствования алгоритмов по распознаванию эмоциональных окрасок на человеческом лице при помощи средств машинного обучения. Прежде всего, нельзя забывать, что эмоции очень субъективная окраска человеческого лица, очень часто даже человек не может определить лицевую экспрессию без комплекса других деталей, которые мы можем отмечать во время диалога с собеседником. Часто появляется необходимость в детектировании таких деталей, как тембр голоса, телодвижения и ситуационная составляющая. Ситуация бесспорно влияет на наше восприятие ситуации, являясь зачастую источником этих самых изменений в эмоциональной окраске. Поэтому в перспективе распознавания эмоциональной составляющей человеческого лица – создание комплекса алгоритмов, которые бы учитывали все тонкости. Голосовое детектирование вместе с трекингом изменений окружающей среды, распознавание основных психологических телодвижений индивидуума и все это в совокупности с обычным эмоциональным детектором порядком увеличили бы точность любых программных средств в этой проблеме машинного обучения. Более того, уже сейчас начинают появляться комплексные алгоритмы, которые используют параллельные нейронные сети, для вычисления ключевых точек и обычную сверхточную нейтронную сеть на основе ResNet уже обученной на ImageNet. Данные матрицы весов, полученные с помощью этих сетей, являются еще одной тренировочный выборкой для объединяющей их обычный нейронной сети. Такие сложные и громоздкие вычисления производятся в огромных дата центрах, при использовании современных вычислительных машин с параллельными графическими процессорами нового поколения. Необходимость создания тенденции на интегрирование систем распознавания и систем видеонаблюдения уже становится понятной многим крупных компаниям, так как это влечёт за собой интерес государства и огромный финансовый успех на рынке систем безопасности. В дополнение ко всем, нельзя не упомянуть, что развитие таких систем неуклонно влечёт технологии и науку, в целом, к созданию искусственного интеллекта и любых его форм. Ведь умение понять, что чувствует собеседник или стопорный человек - одна из уникальных человеческих способностей, которая пока в полной мере недоступна компьютеру.

17

**Список литературы**

1. [Andrew\_Glassner] «Deep\_Learning. From\_Basic to Practics»
2. [Я. Гудфеллоу, И. Бенджио, А. Курвилль] «Глубокое обучение»
3. ВидеоЛекции А.Созыкина - <https://www.youtube.com/watch?v=GX7qxV5nh5o&list=PLtPJ9lKvJ4oiz9aaL_xcZd-x0qd8G0VN_>
4. <https://neurohive.io/ru/osnovy-data-science/glubokaya-svertochnaja-nejronnaja-set/>
5. <https://www.swmath.org/software/4974> - База Изображений
6. <https://habr.com/ru/post/133686/>
7. <https://medium.com/@neurohiveru/%D0%BA%D0%B0%D0%BA-%D0%BA%D0%BE%D0%B2%D0%B0%D1%80%D0%B8%D0%B0%D1%86%D0%B8%D1%8F-%D0%BF%D0%BE%D0%BC%D0%BE%D0%B3%D0%BB%D0%B0-%D1%83%D0%BB%D1%83%D1%87%D1%88%D0%B8%D1%82%D1%8C-%D1%80%D0%B5%D0%B7%D1%83%D0%BB%D1%8C%D1%82%D0%B0%D1%82-%D0%B2-%D1%80%D0%B0%D1%81%D0%BF%D0%BE%D0%B7%D0%BD%D0%B0%D0%B2%D0%B0%D0%BD%D0%B8%D0%B8-%D1%8D%D0%BC%D0%BE%D1%86%D0%B8%D0%B9-37a1a3e4869e>

18