

Министерство образования и науки Российской Федерации  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ  
Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра дискретной математики и  
информационных технологий

**Разработка программного обеспечения на основе сверточной нейронной  
сети для решения задач распознавания изображений**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

по направлению 09.04.01 «Информатика и вычислительная техника»  
студента 2 курса 271 группы  
факультета компьютерных наук и информационных технологий  
Куропатовой Анастасии Андреевны

Научный руководитель

к.ф.-м.н., доцент

\_\_\_\_\_

В.А. Поздняков

Заведующий кафедрой

к.ф.-м.н., доцент

\_\_\_\_\_

Л.Б. Тяпаев

Саратов 2018

**Введение.** В современном мире распознавание графических образов – это довольно популярная тема в области компьютерных технологий (распознавание штрих-кодов, изображений, лиц, жестов, опознавательных знаков технических объектов) [1]. Задачу распознавания образов решают разными способами с применением различных алгоритмов и методик.

Нейронные сети являются наиболее популярными моделями в области искусственного интеллекта. С их помощью решаются многие нетривиальные задачи, для которых невозможно вывести готовую формулу. Также нейронные сети используются для распознавания графических образов.

Одной из таких сетей является многослойный персептрон [2]. Эта сеть состоит из нескольких слоёв нейронов, в которых каждый нейрон одного слоя связан с каждым нейроном следующего. Эта сеть обучается с учителем, т.е. берётся заранее подготовленный массив объектов, для которых известен их класс. Обучение происходит с помощью алгоритма обратного распространения ошибки. Такая модель на данный момент наиболее сильно похожа на модель работы человеческого мозга. С помощью персептрона решаются задачи в области принятия решений, классификации, синтеза речи. Также он может использоваться для распознавания графических изображений, однако он не покажет положительных результатов, если на входе будет повёрнутое, масштабированное, искажённое или сильно зашумлённое изображение.

Ещё одной нейронной сетью, предназначенной для распознавания визуальных образов, являются сверточные нейронные сети. Они имеют большое количество слоев, по сравнению с классическим многослойным персептроном. За счет совместных весов, используемых сразу несколькими нейронами в каждом слое, удастся снизить общее количество обучаемых параметров сети и ускорить обучение. Также, в отличие от многослойного персептрона, сверточные сети восприимчивы к топологии входного изображения.

Цель выпускной квалификационной работы: разработка программного обеспечения на базе сверточной нейронной сети, позволяющей распознавать образы, используя метод обучения с учителем на малых данных.

Чтобы достичь поставленной цели нужно решить следующие задачи:

Изучить модель сверточной сети, выявить параметры, влияющие на качество распознавания.

Построить архитектуру сверточной сети.

Разработать модуль распознавания на базе сверточной сети.

Подготовить данные для обучения.

Произвести обучение сети.

Отладить и протестировать программное обеспечение.

За последние несколько лет работа с искусственным интеллектом стала одной из самых актуальных задач в индустрии. Нейросети внедряются во многие сферы прогнозирования, принятия решений, распознавание образов, оптимизацию и прочие. Искусственный интеллект используется в сфере распознавания изображений, что позволяет автоматизировать многие производственные процессы.

Работа состоит из четырех глав:

1. В первой главе рассмотрены основные понятия сверточных сетей, проанализированы архитектуры многослойных и сверточных сетей для выбора наиболее подходящей к решению поставленной задачи .
2. Во второй главе рассмотрены структуры сетей, слои и принципы их работы.
3. Третья глава описывает процесс распознавания в нейросетевой модели, принципы обучения сети а так же рассмотрены методы обучения с учителем.
4. Четвертая глава содержит решение поставленной задачи а так же тестирование и исследование модифицированной нейросетевой модели сверточной сети.

**Сверточные сети.** Машинное обучение – совокупность методов искусственного интеллекта, направленных не на прямое решение задачи, а на обучение в процессе применения решения множества схожих задач.

В данной работе необходимо применить методы машинного обучения для качественного распознавания изображений, найти критерии, влияющие на качество распознавания и произвести их модификацию. Сравнить результаты работы модифицированной и не модифицированной сети, подвести итог.

Наиболее современной на сегодняшний день является архитектура сверточных нейронных сетей. Эта архитектура, призванная эффективно детектировать изображения, входящая в состав технологии глубокого обучения.

Одной из разновидностей нейронных сетей являются многослойные нейронные сети (МНС). Эта сеть была предложена в работах Розенблатта и является наиболее часто используемой. Архитектура такой сети состоит из последовательно соединенных слоев, где нейрон каждого слоя своими входами связан со всеми нейронами предыдущего слоя, а выходами следующего [2]. Нейронная сеть, имеющая в своем составе два решающих слоя, может с любым значением точности аппроксимировать любую многомерную функцию. Нейронная сеть с одним решающим слоем весьма ограничена в круге задач, решаемых ею, так как способна формировать лишь линейные разделяющие поверхности.

Нейронная сеть, в состав которой входит нелинейная функция активации и два решающих слоя, позволяет производить формирование любых выпуклых областей в пространстве решений, а имея три решающих слоя – области любой сложности, включая невыпуклую. При этом многослойная нейронная сеть оставляет за собой способность к обобщению.

Особенность сверточной нейронной сети состоит в обеспечении частичной устойчивости к изменениям масштаба, поворотам, смене ракурса, смещениям и прочим искажениям. Сверточные нейронные сети объединяют

три архитектурных идеи, для обеспечения инвариантности к данным изменениям:

- локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов);
- общие синаптические коэффициенты (обеспечивают детектирование некоторых черт в любом месте изображения и уменьшают общее число весовых коэффициентов);
- иерархическая организация с пространственными подвыборками.

На сегодняшний день сверточная нейронная сеть и ее различные модификации считаются лучшими по точности и скорости алгоритмами нахождения объектов на сцене.

**Структура сетей.** Сверточный слой представляет собой слой, в котором нейрон соединен с ограниченным количеством нейронов предыдущего уровня, что отличает его от полносвязного слоя [2]. Сверточный слой аналогичен применению операции свертки, где используется лишь матрица весов небольшого размера (ядро свертки), которую двигают по всему обрабатываемому слою. Еще одной особенностью сверточного слоя является то, что он немного уменьшает размер изображения за счет краевых эффектов.

Еще одним видом слоя, который встречается в сверточной нейронной сети, является субдискретизирующий слой. Данный вид слоя выполняет уменьшение размерности обычно в несколько раз, используя, как правило, метод, основанный на выборе максимального элемента (max-pooling). При таком подходе карта признаков разделяется на множество ячеек, на которых в последствии выбираются максимальные значения. Рисунок 2.4 иллюстрирует пример работы субдискретизирующего слоя на основе метода выбора максимального элемента [2].

Работа полносвязного слоя сверточной нейронной сети представляет собой метод, в котором каждый нейрон соединен со всеми нейронами на предыдущем уровне таким образом, что у каждой связи имеется свой весовой

коэффициент [2].

Способность к распознаванию сложных иерархий признаков связана с тем, что все три вида слоев могут чередоваться в произвольном порядке, что позволяет составлять карты признаков из карт признаков. Более поздние слои извлекают более общие характеристики, меньше зависящие от искажений изображения. Стоит отметить, что в классической архитектуре сверточная нейронная сеть состоит из двух слоев: слоя свертки и слоя усреднения.

Таким образом, идея сверточных нейронных сетей заключается в чередовании сверточных и субдискретизирующих слоев, обеспечивающих частичную устойчивость к изменениям масштаба, смещениям, поворотам, искажениям. Для обучения такая сеть использует стандартные методы, чаще всего метод обратного распространения ошибки.

**Процесс распознавания в сверточной нейронной сети.** Распознавание в сверточной сети заключается в вычислении выходных значений со всех клеток уровень за уровнем. Результатом этого процесса является определение, к какому из обученных классов принадлежит входной образ.

При обработке вектора входных сигналов в выходной сигнал, искусственный нейрон умножает вектор входных сигналов на вектор, называемый вектором синаптических весов данного нейрона. Далее производит прибавление к скалярному произведению величины под названием порог данного нейрона, и применяет к результату функцию активации нейрона.

В поисках способа улучшения качества распознавания в данной работе было решено проверить работу сети, если в алгоритм расчёта значений клеток S-плоскостей будет добавлена сигмоидальная функция активации.

Обучение сверточных сетей происходит по алгоритму обратного распространения ошибки и с различными итерациями данного метода.

Он считается самым простым для понимания, и, несмотря на появление более новых методов, обладает некоторыми преимуществами.

Главной особенностью алгоритма обратного распространения ошибки

является вычисление вектора градиента поверхности ошибок. Вектор позволяет минимизировать ошибки посредством движения по кратчайшему пути из начальной точки. Последовательно передвигаясь по заданному пути, происходит приближение к минимуму того или иного типа. Главным вопросом в данной задаче является выбор оптимальной длины шага.

Эпоха – очередной шаг в обучении нейронных сетей. В ходе эпохи на вход нейронной сети поочередно подают заранее подготовленные обучающие данные, после чего выходные значения сравнивают с целевыми и производят вычисление значения ошибки. Этот прием позволяет скорректировать веса в сети, после чего наступает время следующей эпохи. Обучение завершается после прохождения заданного количества эпох или в случае, когда ошибка мала настолько, что дальнейшее прохождение эпох не позволит уменьшить ее значение. Первоначальные веса, как правило, задаются случайным образом.

Для обучения нейронной сети в течение одной эпохи могут быть использованы два различных подхода [3]. В первом варианте на вход нейронной сети последовательно подаются примеры из обучающей выборки. По прохождению каждого изображения веса сети обновляются. Этот метод называется стохастическим (онлайн) обучением. Во втором варианте на вход нейронной сети подается целый пакет обучающих примеров, после чего веса сети обновляются. Таким образом, внутри пакета ошибка накапливается на весах сети, чтобы затем произошло их обновление.

При решении поставленной задачи было принято решение использовать пакетный режим обучения, так как он обладает рядом преимуществ по сравнению со стохастическим:

- Условия сходимости хорошо изучены
- Большое количество техник ускорения обучения работают только с пакетным режимом
- Теоретический анализ динамики изменения весов и скорости сходимости проще

**Структура разработанного программного обеспечения.** Исходя из проведённого анализа предметной области и существующих решений, актуальной является разработка программного обеспечения распознавания образов на базе сверточной сети. Программное обеспечение должно обладать функциями подготовки нейронной сети, подготовки обучающего множества для нейронной сети, сохранения нейронной сети в файл и загрузки из файла, просмотра результатов распознавания.

В качестве языка разработки был выбран язык Python. Это основной язык разработки приложений для платформы .NET и входит в её начальную установку. Он сочетает в себе объектно-ориентированные и контекстно-ориентированные концепции, изначально ориентирован на написание безопасного кода. Язык имеет статическую типизацию и перегрузку операторов. Он обладает очень важными и полезными особенностями, необходимыми для эффективной разработки.

Разрабатываемая система для распознавания изображений будет состоять из таких компонентов как:

1. Нейронная сеть.

Задача: реализация сверточной нейронной сети в программной среде. Для реализации, объявления и обучения нейронной сети взята библиотека с открытым исходным кодом Keras.

2. Материал для обучения.

Задача: создание выборки для обучения нейронной сети, которая подходит под критерии: размер изображения, формат файлов. Для создания выборки был использован метод аугментации.

Таким образом, нейронная сеть была обучена за 100 итераций. Нейронная сеть обучалась до момента пока значения функции ошибки сети не стало наименьшим. Это было достигнуто на 86 итерации обучения.

С помощью модификации нейросети и примененного метода аугментации изображений, процент верно распознанных изображений достиг значения в 96%.

**Заключение.** Методы машинного обучения предоставляют большие возможности для решения задач распознавания изображений.

Использование предложенной модификации помогает улучшить качество распознавания изображений. Метод аугментации позволяет производить обучение нейросети на малых данных, что помогает повысить процент верно распознанных изображений.

Разработана сверточная нейронная сеть, состоящая из восьми слоев, способная решать задачу распознавания образов при использовании фильтров для извлечения признаков из изображений.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Теория распознавания образов [Электронный ресурс]. URL: [http://ru.wikipedia.org/wiki/Теория\\_распознавания\\_образов](http://ru.wikipedia.org/wiki/Теория_распознавания_образов) (Дата обращения: 11.03.2018).
- 2 Хайкин, С. Нейронные сети: полный курс, 2-е издание / С. Хайкин. – М.: Вильямс, 2006. – 1104 с.
- 3 Yoshua Bengio, Y. Le Cun, and D. Henderson. Globally trained handwritten word recognizer using spatial representation, space displacement neural networks and hidden Markov models. In Advances in Neural Information Processing Systems 6, San Mateo CA, 1994. Morgan Kaufmann.