**Слайд 1:**

Уважаемая Государственная экзаменационная комиссия. Позвольте представить Вашему вниманию выпускную работу бакалавра. Я, Вакулин Артём Александрович, Научные руководители — кандидат физико — математических наук, доцент кафедры цифровых технологий Крыловецкий Александр Абрамович и кандидат физико — математических наук, ассистент кафедры цифровых технологий Черников Игорь Сергеевич. Тема дипломной работы – «Применение сверточных нейронных сетей к задаче классификации трехмерных моделей».

**Слайд 2: (Цель)**

Целью моей работы была разработка математических методов и алгоритмов классификации трёхмерных объектов с помощью сверточных нейронных сетей.

**Слайд 3: (Задачи)**

Исходя из этого возникли следующие задачи:

1. Исследование возможных представлений(дескрипторов) трехмерных объектов

2. Исследование технологий сверточных нейронных сетей и построение архитектуры сети для решения поставленных задач

3. Создание програмной реализации системы распознавания трехмерных объектов с использованием сверточных нейронных сетей

**Слайд 4: (Схема системы)**

Для классификации моделей я предлагаю использовать систему, состоящую из четырех основных этапов: Формирование базы данных трехмерных объектов, Подготовка данных, Построение глобальных дескрипторов поверхности и Обучение сверточной нейронной сети.

Рассмотрим каждый из этих этапов.

**Слайд 5: (База данных)**

Для реализации системы классификации я взял модели, представленные в базе трехмерных моделей университета Princeton. А именно 180 объектов трех различных классов, представленных на слайде.

**Слайд 6: (Контроль разрешения объектов)**

Разрешение (или средняя длинна ребра) объектов в данной базе данных существенно различается. Для моделей с малым разрешением построение дескриптора, который бы корректно отражал форму поверхности невозможно. Для того чтобы контролировать разрешение моделей был реализован алгоритм, представленный в работе ссылка на которую будет в конце презентации.

В общих чертах алгоритм выглядит следующим образом: сначала создается динамическая очередь из ребер объекта. За тем первое ребро в очереди выталкивается из нее и обрабатывается. Если длина ребра больше заданного значения, ребро разбивается в его средней точке. Если длина ребра меньше заданной границы, оно сворачивается в точку.

**Слайд 7: (Спиновые изображения)**

Для того чтобы представить информацию о форме поверхности трёхмерных моделей в сжатом и ёмком виде используется глобальный дескриптор в виде интегрального спинового изображения.

Основная идея спиновых изображений - это сопоставление опорной точке (точке поверхности, для которой вычисляется локальный дескриптор) цилиндрической системы координат без учёта полярного угла. Опорная точка в этом случае принимается за начало системы координат, а ось z расположена вдоль вектора нормали к поверхности в опорной точке. Теперь поставим в соответствие координатам ρ и z цилиндрической системы координат две относительные координаты спинового изображения α и β, как показано на слайде.

Вершины с примерно одинаковыми значениями α и β разбивают на классы (корзины) и подсчитывают количество вершин поверхности, попавших в каждый класс.

Для построения глобального дескриптора поверхности предлагается использовать одно спиновое изображение, вычисленное для всей модели сразу.

**Слайд 8: (Примеры дескрипторов)**

На данном слайде вы можете видеть примеры спиновых изображений которые я получал в процессе создания системы классификации.

**Слайд 9: (СНС общее представление)**

Архитектура данного типа нейронных сетей схожа с некоторыми особенностями зрительной коры, в которой имеются простые клетки, реагирующие на прямые линии под разными углами, и сложные клетки, реакция которых связана с активацией определённого набора простых клеток.

Идея свёрточных нейронных сетей заключается в чередовании свёрточных слоев и слоев объединения. Данные слои формируют матрицы чисел - так называемые карты признаков. На последнем слое сети нейроны в полносвязной нейронной сети соединены со всеми картами активаций в предыдущем слое. Результатом её работы является вектор признаков в котором каждый элемент соответствует одному из распознаваемых классов.

Из преимуществ данного типа нейронных сетей можно выделить:

1. Удобное распараллеливание вычислений, а, следовательно, возможность реализации алгоритмов работы и обучения сети на графических процессорах.
2. Относительная устойчивость к повороту и сдвигу распознаваемого изображения.

**Слайд 10: (СНС свертка)**

Свёрточный слой является основным строительным блоком сверточной нейронной сети. Слой состоит из набора фильтров для обучения (или ядер свертки), которые имеют небольшое рецептивное поле, но простираются на всю глубину входной матрицы. Каждый фильтр осуществляет свертку, то есть передвигаясь по вводному изображению, он умножает значения фильтра на исходные значения пикселей изображения(поэлементно). Получившиеся значения суммируются. В итоге получается одно число которое занимает свое место в карте активации.

На слайде можно увидеть, что, если на вводном изображении есть форма, в общих чертах похожая на кривую, которую представляет этот фильтр, то результатом свертки будет большое значение. Таким образом можно говорить, что фильтр «активировался», т.е. нашел определенный признак на изображении.

**Слайд 11: (СНС объединение)**

При реализации архитектуры сверточных нейронных сетей, между последовательно идущими сверточными слоями обычно помещают объединяющие слои. Вся карта признаков разделяется на ячейки [2 х 2] элемента, из которых выбираются максимальные по значению.

Также стоит отметить, что первоначально значения весов сверточных слоев инициализированы случайным образом. Корректировку их значений в данной работе предлагается провести с помощью метода обратного распространения ошибки.

Далее рассмотрим реализацию системы для двух различных архитектур сверточной нейронной сети. Формирование базы данных и подготовка моделей для обоих систем существенно не отличались.

**Слайд 12: (Тест 1 архитектура)**

Итак, в первом эксперименте за основу архитектуры сети была принята модель сети, предназначенная для распознавания цветных изображений размером 32х32 пикселя. На слайде вы можете видеть, что данная архитектура использует 5 сверточных и 3 слоя объединения. Также на слайде указаны основные гиперпараметры данных слоев: шаг и отступ. Результатом работы сети является вектор признаков в котором каждый элемент соответствует одному из распознаваемых классов.

**Слайд 13: (Тест 1 обучение)**

На данном слайде представлен график изменения значений функции ошибки   
для данной реализации системы. Во время обучения сети, обучающая выборка делится на тренировочную(train) и проверочную(validation). После обновления фильтровых коэффициентов для тренировочной выборки, сеть тестируется проверочной выборкой чтобы оценить реальный результат обучения. На графике видно, что значение ошибки для проверочной выборки после некоторого момента перестает снижаться. Это говорит о том, что после этого сеть не продолжает реальный процесс обучения, а лишь затачивается под распознавание тренировочных изображений. Далее будет видно, что данная архитектура не показала хороших результатов, однако сам факт снижения ошибки говорит о том, что сеть способна выделять признаки у глобальных дескрипторов поверхности.

После некоторого количества экспериментов было решено упростить архитектуру сети и искусственно расширить обучающий набор данных путем поворота изображений на 180 градусов и их инвертирования (180 => 720 изображений для каждого класса, 2160 всего).

**Слайд 14: (Тест 2 архитектура)**

Во втором эксперименте за основу архитектуры сети была принята модель сети, предназначенная для распознавания рукописных символов. Размер изображений, передаваемых на вход сети, составляет 28х28 пикселей. На слайде вы можете видеть, что данная архитектура использует 4 сверточных и 2 слоя объединения. Также значительно увеличено количество ядер свертки на некоторых слоях.

**Слайд 15: (Тест 2 обучение)**

Уже на этапе обучения сети видно, что значение функции ошибки для данной реализации системы снижается для каждой эпохи анализа как для тренировочной, так и для проверочной выборок. Это свидетельствует о том, что сеть выделяет признаки из глобальных дескрипторов, а не просто «запоминает» тренировочные изображения.

**Слайд 16: (Результат)**

Оценим результаты проведенных тестов. Как я упоминал ранее, первая реализация системы не показала хороших результатов распознавания. В лучшем случае сеть смогла правильно классифицировать не более 40% передаваемых ей объектов. Вторая же реализация системы показала наилучшие результаты из всех проведенных тестов. Качество распознавания в этом случае достигает 95%.

**Слайд 17: (Итог презентации)**

Итак, в результате проведенной работы:

1. Разработаны математические методы построения дескрипторов трехмерных поверхностей для использования в системах глубокого обучения.

2. Проведено исследование возможных архитектур СНС и на основании проведенных вычислительных экспериментов предложены две возможные архитектуры для решения задачи распознавания трехмерных моделей.

3. Создана программная реализация системы распознавания трехмерных объектов с использованием нейронных сетей глубокого обучения и открытой базы данных 3D моделей The Princeton Shape Benchmark.

Спасибо за внимание.