

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра математической кибернетики и компьютерных наук

**ИНТЕРАКТИВНАЯ СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ С
ПРИМЕНЕНИЕМ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 5 курса 551 группы
направления 09.03.04 — Программная инженерия
факультета КНиИТ
Ранецкого Владимира Руслановича

Научный руководитель

к. ф.-м. н., доцент

Г. Г. Наркайтис

Заведующий кафедрой

к. ф.-м. н.

А. С. Иванов

Саратов 2020

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Нейронные сети	4
1.1 Общие сведения	4
1.2 Модель нейрона	4
1.3 Алгоритм обратного распространения ошибки	4
1.4 Программные библиотеки для обучения нейронных сетей.....	5
2 Сегментация.....	6
2.1 Общие сведения	6
2.2 Сверточные нейронные сети	6
2.3 Интерактивная сегментация	6
3 Практическая часть.....	7
3.1 Общие сведения	7
3.2 Используемые архитектуры нейронных сетей	7
3.3 Описание разработанного приложения	8
4 Исследование результатов обучения с применением интерактивной сегментации	9
4.1 Исследование FPN + ResNet50	9
4.2 Исследование FPN + ResNet101	10
ЗАКЛЮЧЕНИЕ	11
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	12

ВВЕДЕНИЕ

Человеческий мозг - невероятно сложная структура, способная обрабатывать данные нелинейно и параллельно, а также структурно самоорганизовываться, перестраивая нейронные связи. С самого рождения мозг обладает достаточной сложностью и потенциалом для того, чтобы начать выстраивать собственные правила поведения через то, что мы обычно называем опытом, адаптируясь к окружающему миру.

В процессе изучения функционирования нейронов и их связей удалось создать математические модели для проверки различных теорий. Выяснилось что эти модели не только повторяют функции мозга, но и способны выполнять задачи, имеющие собственную ценность. Это и стало фундаментом искусственных нейронных сетей.

Одной из задач нейронных сетей сегодня является сегментация изображений. Идея состоит в том, чтобы для любого входного изображения получить карту всех обнаруженных на этом изображении объектов (маску). Основная проблема данной задачи заключается в подготовке данных для обучения нейронной сети, так как на этапе подготовки требуется огромный человеческий труд. Необходимо, чтобы для каждого входного изображения, которое будет участвовать в обучении сети, была сделана максимально точная маска.

Однако, не так давно была выдвинута идея интерактивной сегментации [1]. Идея заключается в использовании уже предобученной нейронной сети для сегментации изображений, при этом от пользователя требуется лишь размещать положительные и отрицательные метки, на основе которых высчитывается карта расстояний, оптимизирующая предсказание используемой нейронной сети. Вскоре эта идея была улучшена несколькими исследователями из компании Samsung [2], позволяя оптимизировать не только человеческий труд, но и время работы данного подхода.

Задачи:

- изучить основы нейронных сетей.
- исследовать сверточные нейронные сети.
- реализовать программу, обучающую сверточную нейронную сеть с использованием интерактивной сегментации.
- сравнить результаты предсказаний с нейронными сетями, обученными на изображениях, сегментированных вручную.

1 Нейронные сети

1.1 Общие сведения

Нейронную сеть можно рассматривать как отлично распараллеленный и распределенный обработчик данных, который имеет склонность к накоплению эмпирических знаний и делает их доступными для использования [3].

Нейронные сети напоминают мозг по двум причинам:

1. Знания приобретаются сетью из внешнего мира во время обучения.
2. Сила связей между нейронами, которую еще называют синаптическими весами, используется для хранения полученных знаний.

1.2 Модель нейрона

Нейрон - основная единица обработки информации, являющаяся фундаментом для проектирования большого семейства нейронных сетей.

Выделяется три основных компонента нейронной модели:

1. Набор синапсов, или связей, каждая связь характеризуется весом (силой). Например, сигнал x_j , на входе синапса j , соединенного с нейроном k умножается на вес w_{jk} . Синаптический вес может принимать как положительные, так и отрицательные значения.
2. Сумматор - суммирующий входные сигналы с учетом весов.
3. Функция активации, также ее называют функцией сжатия - ограничивает (сжимает) входную сумму до допустимого диапазона значений.

1.3 Алгоритм обратного распространения ошибки

Настоящим толчком в развитии нейронных сетей послужило появление теоретически обоснованного алгоритма обучения многослойных сетей [4].

На начальном этапе обучения, всем весам сети должны быть присвоены начальные значения. После подготовительных действий следует следующий цикл обучения:

1. Производится выбор очередной обучающей пары из обучающего множества. Входной вектор пары подается на вход нейронной сети.
2. Вычисляется выход сети.
3. Вычисляется разность между выходом сети и требуемым выходом из обучающей пары.
4. Веса сети корректируются так, чтобы минимизировать ошибку.

Цикл повторяется на всем обучающем множестве до тех пор, пока не будет достигнута ошибка приемлемого уровня.

1.4 Программные библиотеки для обучения нейронных сетей

Основным структурным блоком библиотек машинного обучения является тензор. Фактически, тензор - объект, линейно преобразующий элементы одного линейного пространства, в элементы другого. В рамках библиотек, тензорами называют многомерные массивы, которые имеют дополнительный функционал, позволяющий им работать с гпу.

Следующей важной частью подобных библиотек является модуль, который позволяет автоматизированно выполнять вычисление градиента, для выполнения обучения.

Наконец, подобные библиотеки содержат уже готовые реализации различных слоев нейронных сетей, метрик, функций активации и потерь, а также различных оптимизаторов, делающих их использование очень удобным.

2 Сегментация

2.1 Общие сведения

Задача сегментации является одной из самых сложных задач компьютерного зрения. Под сегментацией понимают процесс выделения областей объектов на изображениях.

2.2 Сверточные нейронные сети

В нейронных сетях, состоящих из слоев нейронов, которые соединены друг с другом есть проблемы, которые делают их неэффективными при сегментации изображений. Во-первых, подобные сети имеют большое количество весов. Во-вторых, теряется информация о топологии объектов на нем, так как сеть анализирует соседние пиксели только по горизонтали. Сверточные нейронные сети спроектированы таким образом, что они могут учитывать пространственные входные структуры. Такие сети основываются на применении операций свертки к входным данным, из-за чего и получили свое название. Основной идеей сверточных сетей является создание таких слоев, признаками в которых могут служить ядра свертки. Обучение происходит по тому же принципу, что и обучение перцептронов: алгоритм обратного распространения применяется к ядрам свертки.

2.3 Интерактивная сегментация

В 2019 году на конференции IEEE двумя исследователями из Гарвардского и корейского университетов была предложена схема интерактивной сегментации для сверточных сетей, названная схемой обратного уточнения (англ. Backpropagating Refinement Scheme). Согласно этой схеме, пользователь делает положительные и отрицательные метки на сегментируемом изображении, затем выполняется прямой проход по сети с учетом этих меток, а после этого выполняется обратный проход с ограничениями в областях меток.

Она получилась эффективной и показала хорошие результаты, однако, нужны большие затраты с точки зрения вычислений за счет того, что обратное распространение ошибки выполняется итеративно. В 2020 году несколькими исследователями из компании Samsung было разработано улучшение этой схемы [2], благодаря этому улучшения обратное распространение необходимо выполнять только для части сети, что дает прирост скорости вычислений.

3 Практическая часть

3.1 Общие сведения

Для достижения целей данной работы было разработано приложение, позволяющее обучать сверточные нейронные сети и пользоваться интерактивной сегментацией на языке программирования python.

Для реализации обучения нейронных сетей была выбрана библиотека pytorch, обладающая высокой эффективностью вычислений и удобством. Также благодаря большому количеству пользователей, можно легко найти вопросы на интересующие вопросы и готовые примеры, выполненные с использованием pytorch.

С помощью реализованной программы можно выполнить следующий цикл обучения нейронной сети:

1. Разметка обучающей выборки
2. Обучение
3. Сегментирование части новых изображений на обученной сети
4. Сегментация плохо сегментированных изображений с помощью интерактивной утилиты
5. Повторение шагов 2 - 4 для новых изображений

Для исследования была выбрана уже размеченная обучающая выборка CamVid [5]. Эта обучающая выборка состоит из более чем семисот сегментированных изображений. Эти изображения были получены при съемке дорог в Кембридже, на каждом из этих изображений размечены различные объекты, начиная от неба, и заканчивая велосипедами. Для исследования был выбран класс машин.

3.2 Используемые архитектуры нейронных сетей

Для исследований была выбрана нейронная сеть с пирамидальной архитектурой (анг. Feature Pyramid Network). Подобные сети показывают хорошие результаты в различных приложениях. Главной особенностью такой архитектуры является то, что она устойчива к масштабу входных данных. Также пирамидальные сети являются достаточно небольшими, чтобы эффективно обучать их на среднем домашнем компьютере.

В дополнение к ней, для проведения исследований, была выбрана сверточная сеть ResNet [6]. Данная сверточная сеть показала хорошие результаты

во многих областях компьютерного зрения, в том числе и сегментации. Так же в ее архитектуре была применена революционная идея соединений быстрого доступа (анг. *shortcut connections*).

Основой соединений быстрого доступа было создание таких слоев, преобразование в которых, не меняли бы размер входных данных, а после выполнения преобразований полученный результат складывался бы с исходными данными. Таким образом, при обратном проходе сильно уменьшается вероятность затухания градиента и паралича сети.

В качестве функции потерь выбрана метрика Dice loss/ F1 score, которая схожа с коэффициентом Жаккара и часто используется в качестве функции потерь при обучении сверточных нейронных сетей.

3.3 Описание разработанного приложения

Первая вкладка реализованной программы предназначена для первоначальной настройки взаимодействия с ней. На этой вкладке необходимо задать пути до основных директорий операционной системы, которые будут использованы для работы. Также данная вкладка позволяет загрузить «CamVid» обучающую выборку и подготовить ее для работы с другими компонентами приложения.

На второй вкладке предусмотрены элементы управления для обучения нейронной сети на текущей обучающей выборке. Здесь можно настроить параметры обучения, а также посмотреть текущее состояние значений метрик.

После обучения нейронной сети можно пробовать сегментировать новые изображения с ее помощью. С помощью элементов управления, расположенных на третьей вкладке приложения, можно указывать количество (в процентном соотношении) изображений из исходной обучающей выборки, который будут подготовлены для сегментации, а затем сегментировать их, при этом в нейронную сеть будут загружены веса из последней сохраненной контрольной точки и начнется сегментация с ее помощью. Также, на этой вкладке можно открыть для просмотра сегментированные изображения, а также применить к ним интерактивную сегментацию.

4 Исследование результатов обучения с применением интерактивной сегментации

Цель создания программы, описанной в практической части - исследование обучения с применением интерактивной сегментации. Для этого было проведено два эксперимента, для того чтобы сравнить результаты для разных нейронных сетей.

Цель экспериментов - сравнить качество масок, предсказанных сетью обученной с применением интерактивной сегментации и обученной на выборке, сегментированной вручную. Для этого использовалась обучающая выборка CamVid, а в качестве объектов для сегментации были выбраны машины.

4.1 Исследование FPN + ResNet50

Исследование проводилось согласно алгоритму, описанному в практической части.

В качестве подготовки была загружена обучающая выборка CamVid. Затем с помощью функций реализованного приложения была подготовлена для использования. После этого в качестве начальной обучающей выборки для FPN + ResNet50 было подготовлено 34 изображения.

Для тестирования во время обучения использовалось 30% от начальной выборки, размер шага обучения была задан равным 8, количество эпох - 10.

После обучения было подготовлено 33 изображения для сегментации с помощью обученной сети.

После сегментации изображения были добавлены к текущей обучающей выборки.

На следующей итерации обучающая выборка содержит 67 изображений. Для тестирования во время обучения было использовано 15% изображений текущей выборки.

Во время каждого обучения веса нейронной сети сохранялись, если значение коэффициента Жаккара превышало последнее максимально значение. Данный цикл выполнялся до тех пор, пока улучшалось значение коэффициента Жаккара. Во время обучения процент изображений, используемых для тестирования корректировался таким образом, чтобы тестирующая выборка не превышала 15-20 изображений.

4.2 Исследование FPN + ResNet101

Исследование проводилось согласно алгоритму, описанному в практической части.

В качестве подготовки была загружена обучающая выборка CamVid. Затем с помощью функций реализованного приложения была подготовлена для использования. После этого в качестве начальной обучающей выборки для FPN + ResNet101 было подготовлено 34 изображения. ResNet101 содержит большее количество сверточных слоев, по сравнению с ResNet50.

Данная сеть обучилась за довольно короткое количество эпох, но показала результаты немного более плохие чем сеть из предыдущего исследования. При этом значения метрик сети, обученной на изображениях, сегментированных вручную, были очень близкими.

ЗАКЛЮЧЕНИЕ

Сегодня область компьютерного зрения популярна как никогда. Сегментация изображений, как составная ее часть, используется в большом количестве приложений. Она может быть применена и используется в сферах медицины, банковского дела, а также при распознавании объектов в реальном времени, например, для создания беспилотных автомобилей.

Исследования в области нейронных сетей также очень популярны, из-за того, что нейронные сети можно применять для решения задач компьютерного зрения. А также потому что в современном мире можно с легкостью получать большие массивы данных, обработка которых, помогает повысить качество предсказаний. При этом различные проводимые конкурсы показывают, что уже сейчас некоторые нейронные сети способны показывать результаты предсказаний лучше, чем это делает человек.

Было проведено два исследования. Результаты исследований показали, что с помощью интерактивной сегментации можно значительно ускорить подготовку обучающей выборки, однако, существует ряд проблем. Во-первых, интерактивная сегментация основывается на таком же предсказании с использованием нейронных сетей, а это означает, что для получения хороших результатов, необходимо обучить нейронную сеть, которая будет использоваться в дальнейшем. Здесь встает вопрос о ручной сегментации и о том, какую лучше архитектуру подобрать для получения наилучшего качества сегментирования. Во-вторых, маски, создаваемые интерактивно могут получаться не такими точными, как маски, сегментированные вручную.

Подводя итог можно отметить, что области компьютерного зрения, сегментации изображений и машинного обучения не утратят интерес исследователей в ближайшее время. С каждым годом можно видеть все более лучшие результаты, которые теперь находятся очень близко к возможностям человека по качеству, и значительно опережают его по быстродействию. Интерактивная сегментация может стать отличным инструментом, который поможет сократить большие трудозатраты на обучение нейронных сетей, однако пока он не обладает достаточно хорошей точностью для получения наилучших результатов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 *Jang, W.-D.* Interactive Image Segmentation via Backpropagating Refinement Scheme / W.-D. Jang, C.-S. Kim. — 2019. — URL: <https://vcg.seas.harvard.edu/publications/interactive-image-segmentation-via-backpropagating-refinement-scheme/paper> (Дата обращения 18.05.2020). Загл. с экр. Яз. рус.
- 2 *Sofiuk, K.* f-BRS: Rethinking Backpropagating Refinement for Interactive Segmentation [Электронный ресурс] / K. Sofiuk, I. Petrov, O. Varinova, A. Konushin. — 2020. — URL: <https://arxiv.org/abs/2001.10331> (Дата обращения 18.05.2020). Загл. с экр. Яз. англ.
- 3 *Haykin, S.* Neural Networks and Learning Machines / S. Haykin. — Pearson, 2009.
- 4 Learning internal representations by error propagation. [Электронный ресурс]. — URL: https://web.stanford.edu/class/psych209a/ReadingsByDate/02_06/PDPVol1Chapter8.pdf (Дата обращения 18.05.2020). Загл. с экр. Яз. англ.
- 5 Object recognition in video dataset [Электронный ресурс]. — URL: <http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamVid/> (Дата обращения 18.05.2020). Загл. с экр. Яз. англ.
- 6 Learning deep resnet blocks sequentially using boosting theory [Электронный ресурс]. — URL: <https://arxiv.org/pdf/1706.04964.pdf> (Дата обращения 18.05.2020). Загл. с экр. Яз. англ.