

Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

УДК 004.932.2

Сухов
Никита Юрьевич

**Решение задач обработки изображений с использованием алгоритмов
глубокого обучения**

АВТОРЕФЕРАТ
диссертации на соискание степени
магистра информатики и вычислительной техники

по специальности 1-40 81 04 – Обработка больших объемов информации

Научный руководитель
Калугина М.А.
к.физ-мат.н., доцент

Минск 2020

Работа выполнена на кафедре информатики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

Научный руководитель: **КАЛУГИНА Марина Алексеевна**,
кандидат физико-математических наук, доцент
кафедры информатики учреждения образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Рецензент: **ШЕШОЛКО Владимир Константинович**,
кандидат физико-математических наук, доцент
кафедры управления информационными
ресурсами учреждения образования «Академия
управления при Президенте РБ»

Защита диссертации состоится «23» июня 2020 г. года в 10⁰⁰ часов на заседании Государственной экзаменационной комиссии по защите магистерских диссертаций в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» по адресу: 220013, Минск, ул. Гикало, 9, копр. 4, ауд. 112, тел. 293-85-91, e-mail: inform@bsuir.by

С диссертацией можно ознакомиться в библиотеке учреждения образования «Белорусский государственный университет информатики и радиоэлектроники».

ВВЕДЕНИЕ

Глубокое обучение – одна из центральных компонент в современных интеллектуальных системах. На сегодняшний день многие компании используют глубокое обучение в своих проектах. Оно нашло применение в различных отраслях: медицине, банковской сфере, телекоммуникациях, рекламе. Глубокое обучение – это отрасль в программировании, использующая нейронные сети. За последние годы с их помощью разработки в этой сфере достигли рекордной точности, которая не только равна, но и превышает человеческую.

Системы глубокого обучения на данный момент не только выполняют рутинные действия (заменяя при этом человека), но и делают некоторые виды работ быстрее, качественнее, дешевле, чем люди. Примерами таких задач могут послужить классификация изображений, анализ текстов. Последние разработки в анализе текстов позволяют понять и извлечь полезную информацию из текста за время, которое человек может не успеть за всю свою жизнь.

Несмотря на всеобщую популярность, глубокое обучение не является панацеей от всех проблем и универсальным их решением. Внедрение этого метода в уже существующие системы требует некоторого уровня организованности данных, и, естественно, достаточного их количества.

Подход, использующий глубокие нейронные сети, в области распознавания и обработки изображений на текущий момент получил наибольшую популярность. Основной проблемой этого подхода является ограничение производительности компьютеров.

Большой прорыв в области распознавания изображений произошёл, когда были придуманы сверточные нейронные сети. Теперь не нужно человеку самому генерировать признаки и придумывать различные фильтры для обработки изображений, сверточные нейронные сети справляются с этой задачей самостоятельно: для этого нужно лишь большое количество данных и хорошее начальное приближение. На сегодняшний день с помощью сверточных нейронных сетей можно классифицировать изображения, сказать с точностью до пикселя, где находится определенный объект на изображении, генерировать изображения по определенному критерию, а также можно генерировать достаточно полные текстовые описания для различных изображений.

В работе произведен анализ и последующее решение задач классификации, детекции, сегментации медицинских изображений. Основные снимки, с которыми произведена работа, – это гистологические изображения толстой кишки и МРТ головного мозга.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цели и задачи исследования

В диссертационной работе излагаются результаты применения методов глубокого обучения для анализа изображений. Работа является логическим продолжением дипломной работы «Определение модели камеры по снимкам с использованием методов машинного обучения». В диссертационной работе продолжается исследование вопросов, как машинное обучение, а точнее глубокое, помогает решать задачи, связанные с анализом изображений, и как его заставить стать очень хорошим помощником для человека.

Целью данной работы является решение двух задач, связанных с анализом медицинских изображений. Первая задача – классификация проявления язвенного колита по гистологическим изображениям. Вторая задача – определение местоположения опухоли, отека и некроза на 3-D изображениях МРТ (магнитно-резонансная томография) головного мозга. Если вторая задача частично решена (опухоли сегментируются достаточно неплохо уже давно классическими алгоритмами компьютерного зрения, например, методом активного контура), то первая задача практически не имеет аналогов.

Решение задачи классификации язвенного колита сможет помочь врачам ставить диагнозы. Бывает, что не хватает врачей-специалистов, чтобы просмотреть все снимки и поставить диагноз, а также врач попросту может ошибиться из-за усталости, неопытности и т. д. Система является не заменой врачу, а помощником. Если разработанная система выделила на снимке подозрительные области и поставила активность, которая соответствует более тяжелому течению болезни, то специалист более подробно рассмотрит данное изображение.

Вторая задача имеет не менее важное значение для медицины, так как в последнее время всё чаще у детей обнаруживаются опухолевые образования в мозге, а также различные травмы головы.

Объектом исследования является разработка системы, которая способна автоматически обрабатывать медицинские изображения, выделяя на них структуры, с помощью которых врач сможет поставить диагноз пациенту и назначить соответствующее лечение.

Предметом исследования являются методы обнаружения и распознавания объектов на изображении с использованием глубоких нейронных сетей, а также алгоритмы анализа полученных данных.

Область исследования

Содержание диссертации соответствует образовательному стандарту высшего образования второй ступени (магистратуры) ОСВО 1-40 81 04-2020 специальности 1-40 81 04 – «Обработка больших объемов информации».

Теоретическая и методологическая основа исследования

В основу диссертации легли работы зарубежных ученых в области анализа изображений, а также анализ онлайн-сервисов и технической документации по рассматриваемой тематике.

Информационная база исследования сформирована на основе литературы, открытой информации, технических нормативно-правовых актов, сведений из электронных ресурсов, а также материалов научных конференций и семинаров.

Научная новизна

Научная новизна и значимость полученных результатов работы заключается в разработке алгоритмов, способных по гистологическим снимкам определять активность язвенного колита, а также методов, на основе которых можно на МРТ изображениях находить опухоль, некроз и отек.

Теоретическая значимость работы заключается в детальном анализе существующих алгоритмов обработки изображений с использованием методов глубокого обучения.

Практическая значимость диссертации заключается в разработке системы для обработки изображений.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались на 55-ой юбилейной научной конференции аспирантов, магистрантов и студентов БГУИР, 56-ой научной конференции аспирантов, магистрантов и студентов БГУИР и XI Международной научно-методической конференции (г. Минск).

Опубликованность результатов диссертации

По теме диссертации опубликованы 4 печатные работы, из них 3 работы в сборнике научной конференции аспирантов, магистрантов и студентов БГУИР, 1 работа в материалах XI Международной научно-методической конференции.

Структура и объем диссертации

Диссертация состоит из общей характеристики работы, введения, пяти глав, заключения, списка использованных источников, списка публикаций автора и приложения.

В первой главе представлен анализ предметной области. Рассматриваются теоретические сведения в области строения толстой кишки и головного мозга.

Во второй главе приведены принципы построения нейронных сетей и сверточных нейронных сетей, а также их преимущества в анализе изображений.

Третья глава дает краткую характеристику используемых технологий и описывает причины, по которым они были выбраны.

В четвертой главе рассматриваются детали реализации системы определения активности язвенного колита по гистологическим снимкам.

В пятой главе рассматриваются детали реализации системы сегментации опухоли, некроза и отека по снимкам МРТ головного мозга.

Общий объем работы составляет 89 страниц, из которых основного текста – 68 страниц, 32 рисунка на 28 страницах, список использованных источников из 39 наименований на 3 страницах и одно приложение на 21 странице.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** рассмотрен нейросетевой подход для решения современных задач, его преимущества и недостатки. Были приведены различные примеры использования глубокого обучения, особое внимание было уделено применению его в анализе изображений. Также описано обоснование актуальности темы.

В **общей характеристике работы** сформулированы цели и задачи диссертации, обозначена область исследований, научная и практическая значимость. Приведено количество опубликованных работ по теме диссертации, а также описана ее структура.

В **первой** главе были рассмотрены теоретические сведения, которые помогают понять предметную область, без знания которых невозможно приступить к построению систем высокого качества. Основное внимание было уделено:

- Строению толстой кишки.
- Признакам, по которым можно распознать язвенный колит.
- Симптомам, которые проявляются при заболевании язвенным колитом.
- Строению головного мозга.
- Классификации опухолей головного мозга.
- Типам сканирования МРТ.

Важным моментом этой главы является выделение признаков, с помощью которых производится классификация активности язвенного колита. Именно поиск этих признаков на изображении осуществляется в процессе решения задачи.

Зная, какие бывают и чем отличаются различные типы сканирования МРТ изображений, можно продумать алгоритм реализации системы, которая будет достаточно хорошо отличать опухоль головного мозга от отека.

Также в самом начале главы формулируются основные требования, которым должны удовлетворять разработанные системы.

Во **второй** главе рассмотрены нейронные сети как современный инструмент для решения задач обработки изображений. Были продемонстрированы примеры построения нейронных сетей, описан алгоритм их обучения – backpropagation, благодаря которому использование нейронных сетей стало возможным для большинства задач, так как с помощью этого метода довольно быстро обновляются параметры сети. Знания о видах функции активации дают возможность построения более разнообразных нейронных сетей.

Особое внимание уделено сверточным нейронным сетям. Изобретение сверточных нейронных сетей обеспечило большой прорыв в области обработки изображений и видео. Благодаря сверточным слоям стало возможным смотреть изнутри алгоритма на изображение в том виде, в котором оно есть, не преобразуя его в вектор, тем самым теряя информацию о местоположении объектов на нём. Также не нужно придумывать различные фильтры для работы с картинкой – сеть в процессе обучения сама формирует их. В конце главы приводится описание работы слоя объединения. Он помогает уменьшить количество вычислений в сети.

Использование нейронных сетей в работе очень сильно сократило время разработки систем, позволило добиться качества решения задач, не хуже чем человек. А также с помощью глубокого обучения удалось достигнуть высокой обобщающей способности алгоритмов.

В **третьей** главе рассмотрена краткая характеристика используемых технологий с обоснованием выбора. Для разработки системы был выбран язык Python. Выбор в сторону этого языка встал по причине его простоты. На нём можно очень сложную логику выразить всего в несколько строк кода. Но главным преимуществом выбора именно этого языка является богатство библиотек для обработки данных и фреймворков с готовыми алгоритмами машинного и глубинного обучения. Так, в Keras, можно всего за несколько минут написать код модели нейронной сети и ее обучения. Ещё Python лучшего всего подходит для развертывания обученных моделей, по сравнению, например, с Matlab.

Использование CUDA и CuDNN позволяет обучать нейронные сети на GPU. Благодаря этим технологиям и стало возможным обучать глубокие сети. Не используя возможности графического процессора, глубокую нейронную сеть на большом количестве данных приходилось бы обучать от нескольких недель до нескольких месяцев. Можно сделать вывод, что Python на сегодняшний день занимает ведущее место в качестве языка разработки систем, связанных с искусственным интеллектом.

В **четвертой** главе рассмотрено решение задачи прогнозирования активности язвенного колита. В самом начале главы рассматриваются примеры данных и их разметка. Было показано, что разметка данных является немаловажным фактором для решения задачи. Хорошая разметка позволяет сделать декомпозицию задачи, что позволяет находить более узкие места в ней и уделить этой проблеме больше внимания.

Далее рассматривается архитектура сегментационной сети Deeplabv3+, с помощью которой были решены многие подзадачи. Данная сеть хорошо справляется с изображениями в разных масштабах благодаря пространственному объединению в пирамиду. «Дырявые свертки» позволяют

увеличивать рецептивное поле, не уменьшая при этом пространственное разрешение изображения.

В ходе исследования было выявлено самое слабое место в алгоритме решения задачи – классификация клеток. В будущем планируется решать эту проблему путем разметки новых типов клеток на изображении.

Создание собственной метрики для классификации активности язвенного колита дало возможность более объективно оценить качество работы системы. Из-за того, что чем выше активность, тем более тяжело протекает заболевание и это обязательно нужно учитывать, стандартные метрики задач классификации оказались непригодными.

Лучший результат, которого удалось добиться – это 0.75 по разработанной метрике. Этот результат примерно является средним значением результатов оценки врачами-специалистами.

В **пятой** главе рассмотрено выделение областей, содержащих опухоль, отек и некроз на 3-D изображениях головного мозга. Для решения поставленной задачи использовалась нейронная сеть V-Net, реализованная с помощью фреймворка Python Keras. Выбор в сторону этой сети встал из-за того, что она умеет работать с трехмерными изображениями.

Для более точного прогнозирования отека и некроза внутри опухоли сначала решалась задача сегментации опухоли головного мозга, а потом уже в ней осуществлялся поиск некроза и отека. Использование разных режимов МРТ сканирования дало возможность добиться качества решения задачи 0.81 по метрике IOU. В целом, можно сделать вывод, что использование нейронных сетей с трехмерными свертками является лучшим подходом на сегодняшний день для сегментации 3-D изображений, по сравнению с применением алгоритма сегментации для каждого среза по отдельности, так как при этом теряется пространственная информация о расположении объектов.

ЗАКЛЮЧЕНИЕ

В диссертационной работе были затронуты темы классификации, сегментации, детекции изображений с использованием алгоритмов глубокого обучения и две определенные прикладные задачи: определение активности язвенного колита на основании гистологических снимков и сегментация опухоли, некроза и отека по МРТ изображениям головного мозга.

После исследования предметной области были разработаны две системы для решения поставленных задач. Качество работы системы по определению активности язвенного колита оценивалось с помощью специально разработанной метрики для этой задачи. Лучшего значения качества этой метрики удалось достигнуть с помощью декомпозиции исходной задачи на подзадачи. Оно составило 0.75 (или 75%) на тестовой выборке. Качество работы системы по сегментированию головного мозга на три структуры: опухоль, отек, некроз измерялось метрикой IOU. Также с помощью разделения данной задачи на несколько подзадач удалось достигнуть наилучшего качества, которое получилось равным 0.81.

Для выполнения работы выбран оптимальный язык программирования – Python, который используется в современном мире для решения задач, связанных с использованием нейронных сетей. Были выбраны такие библиотеки глубокого обучения, как Keras и Tensorflow.

Для решения поставленных задач были использованы сверточные нейронные сети. Это один из наиболее прогрессивных современных подходов к анализу изображений. Именно благодаря им было достигнуто такое высокое качество. Основное преимущество сверточных сетей – использование изображений в том виде, в котором они есть, без преобразования в одномерный вектор, а также снижение количества параметров модели, что позволяет использовать глубокие нейронные сети для анализа изображений с имеющимися на сегодняшний день мощностями вычислительной техники.

Поставленная задача диссертационной работы была достигнута. Получено хорошее качество решения задач. Однако, несмотря на достигнутые успехи решаемых задач, существуют некоторые методы, позволяющие улучшить точность моделей. Например, обучение более глубоких сетей, а также объединение моделей в ансамбли. Но это требует больших вычислительных мощностей. Также использование очень больших ансамблей моделей может являться критичным в условиях использования системы (пользователь может долго ждать, пока модель выдаст предсказания).

По мере развития сферы компьютерного зрения и применения накопленного опыта, полученные системы могут быть не только улучшены качественно (путем добавления новых значимых характеристик, которые будут

учитываться алгоритмом машинного обучения), но и распространены в смежные области.

Библиотека БГУИР

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

1-А. Козак А. В. Байесовская оптимизация / А. В. Козак, Н. Ю. Сухов // Компьютерные системы и сети: 55-я юбилейная научная конференция аспирантов, магистрантов и студентов, Минск, 22-26 апреля 2019 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2019. – С. 200 – 202.

2-А. Саскевич А. В. Алгоритмы машинного обучения в дистанционных информационных образовательных системах / Саскевич А. В., Сухов Н. Ю., Хотеев А. Л. // Дистанционное обучение – образовательная среда XXI века : материалы XI Международной научно-методической конференции, Минск, 12-13 декабря 2019 г. / редкол. : В. А. Прытков [и др.]. – Минск : БГУИР, 2019. – С. 276.

3-А. Сухов Н. Ю. Обнаружение рака лёгкого по снимкам с использованием алгоритмов машинного обучения / Н. Ю. Сухов // Компьютерные системы и сети: 56-я научная конференция аспирантов, магистрантов и студентов, Минск / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2020 (В печати).

4-А. Сухов Н. Ю. Распознавание лиц на изображениях плохого качества с использованием нейросетевых алгоритмов / Н. Ю. Сухов, Н.С. Латушкин, А.В. Козак // Компьютерные системы и сети: 56-я научная конференция аспирантов, магистрантов и студентов, Минск / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2020 (В печати).