

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

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

Кафедра дискретной математики и информационных технологий

**СВЕРТОЧНЫЕ НЕЙТРОННЫЕ СЕТИ И ИХ ПРИМЕНЕНИЕ В
ЗАДАЧАХ РАСПОЗНОВАНИЯ ИЗОБРАЖЕНИЙ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 421 группы
направления 09.03.01 — Информатика и вычислительная техника
факультета КНиИТ
Кабанова Даниила Дмитриевича

Научный руководитель

доцент, к. ф.-м. н.

И. Д. Сагаева

Заведующий кафедрой

доцент, к. ф.-м. н.

Л. Б. Тяпаев

Саратов 2024

ВВЕДЕНИЕ

Актуальность темы. Современные медицинские исследования активно используют методы машинного обучения и искусственного интеллекта для диагностики различных заболеваний. Спектроскопия, как неинвазивный метод анализа, позволяет получать детализированную информацию о биологических тканях, что делает её перспективным инструментом для диагностики кожных заболеваний.

Актуальность данной работы обусловлена необходимостью создания автоматизированных систем для раннего выявления и классификации кожных патологий, таких как меланома, базалиома и невусы, что может существенно повысить эффективность и точность диагностики.

Цель выпускной квалификационной работы – исследование точности дифференциации патологий по результатам АФ и КР спектроскопии с помощью нейросетевого алгоритма.

Объект исследования – процесс дифференциации образцов биоткани по результатам спектроскопии комбинационного рассеивания и флуоресценции.

Предмет исследования – искусственные нейронные сети.

Рассматривается проблема неинвазивной диагностики раковых заболеваний тканей кожи человека. Проводится анализ результатов спектроскопии и автофлуоресценции (АФ) и комбинационного рассеивания (КР), для выявления признаков дифференциации кожных патологий с использованием нейронной сети.

Методологические основы применения нейронных сетей для решения подобных проблем представлены в работах К.А. Сержантова [1], С.А. Тюляндина [2], В.С. Медведева [3], Л.В. Демидова [4], В.В. Тучина [5], В.Л. Быкова [6].

Практическая значимость бакалаврской работы. В ходе выполнения работы была создана система автоматической классификации кожных заболеваний. Протестированные модели показали высокую точность в классификации меланом, базалиом и невусов на основе спектроскопических данных.

Структура и объем работы. Работа состоит из введения, трех глав и заключения. Объем работы составляет 52 страницы, на которых размещены

30 рисунков и 2 приложения. При написании данной работы было использовано 20 литературных источников, в том числе 7 источников на иностранном языке.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении подчеркивается актуальность разработки неинвазивных методов диагностики кожных заболеваний, особенно раковых новообразований. Проблема ранней диагностики кожных раковых заболеваний важна из-за их высокой распространенности и потенциальной смертности. Основной целью работы является разработка и обучение сверточной нейронной сети для автоматической классификации кожных заболеваний на основе спектроскопических данных. Задачи работы включают анализ существующих методов диагностики, разработку нейронной сети и тестирование ее эффективности.

Первая глава посвящена анализу существующих методов спектроскопии и машинного обучения, используемых для диагностики кожных заболеваний.

Рассматриваются текущие методы диагностики, такие как автофлуоресценция и спектроскопия. Особое внимание уделяется спектроскопии как перспективному методу, позволяющему получать данные о кожных новообразованиях без инвазивного вмешательства. Обсуждаются преимущества и недостатки различных подходов, а также современные достижения и проблемы в области неинвазивной диагностики.

Описываются основные понятия и механизмы работы нейронных сетей, включая их архитектуру, функции активации, обучение и оптимизацию.

В разделе **”Виды нейронных сетей”** рассматриваются различные типы нейронных сетей, такие как полносвязные сети, сверточные нейронные сети (CNN), рекуррентные нейронные сети (RNN) и их модификации. Описаны их особенности, преимущества и области применения, а также приведены примеры успешного использования этих сетей в медицинских задачах.

Во второй главе описывается процесс разработки и обучения сверточной нейронной сети (CNN) для классификации кожных заболеваний.

В подразделе 2.1 описаны инструменты и библиотеки, использованные для разработки и обучения нейронной сети, такие как Keras, TensorFlow, Scikit-learn и другие. Приведены аргументы в пользу выбора этих инструментов и их возможности.

В подразделе 2.2 подробно рассматривается процесс подготовки данных для обучения сверточной нейронной сети. Сначала осуществляется сбор

спектральных данных кожных образцов с использованием различных методов спектроскопии, таких как Рамановская спектроскопия и автофлуоресценция. Эти данные собираются от пациентов с различными типами кожных новообразований, включая меланому, базальноклеточную карциному, невусы и кератомы.

Для устранения диспропорции между классами обучающая выборка балансируется. Используются методы форматирования данных, такие как вращение, масштабирование, чтобы увеличить количество образцов и предотвратить переобучение модели.

В подразделе 2.3 рассмотрены методы повышения точности классификации, включая форматирование данных, регуляризацию и использование различных архитектур нейронных сетей.

В подразделах 2.4 - 2.7 подробно описана архитектура сверточных нейронных сетей, включая их основные компоненты: сверточные слои, слои подвыборки (pooling), слои активации и полносвязные слои. Рассматриваются особенности работы сверточных сетей и их преимущества для задач классификации изображений.

В подразделе 2.8 рассматривается использование предобученных моделей CNN для задач классификации кожных заболеваний. Описаны методы переноса обучения (transfer learning) и адаптации предобученных моделей к новой задаче.

Подраздел 2.9 описывает процесс предобработки спектральных данных и создание различных наборов данных для обучения и тестирования моделей. Спектры нормализуются для приведения их к единому масштабу. Это позволяет устранить различия в интенсивности сигналов и сделать данные более сопоставимыми.

Спектральные данные могут иметь разное разрешение, поэтому проводится интерполяция и дублирование для приведения их к единому размеру. Используются методы, такие как линейная и сплайн-интерполяция, для создания равномерных наборов данных. На основе предобработанных спектров формируются обучающие, валидационные и тестовые наборы данных. Каждому набору присваиваются метки классов, соответствующие типам кожных заболеваний. Наборы данных структурируются таким образом, чтобы обеспечить равномерное представление всех классов.

Третья глава посвящена тестированию разработанных нейросетевых архитектур и оценке их эффективности в задаче классификации кожных заболеваний.

В подразделе 3.1 описаны методы дифференциации результатов автофлуоресценции и их применение для классификации кожных новообразований. Рассматриваются методы анализа спектральных данных и их интерпретация.

Используя 446 усредненных образца, включающих 223 образца кожи и 223 образца патологии, была разработана и обучена нейронная сеть. Обучающая выборка была случайным образом разделена на три набора: 60% использовались для обучения (Training), 20% для проверки достоверности результатов и предотвращения переобучения (Validation), и еще 20% для независимого тестирования сети (Test).

После процесса обучения была разработана регрессионная модель, отображающая выходные результаты нейронной сети, как это показано на рисунке 1.

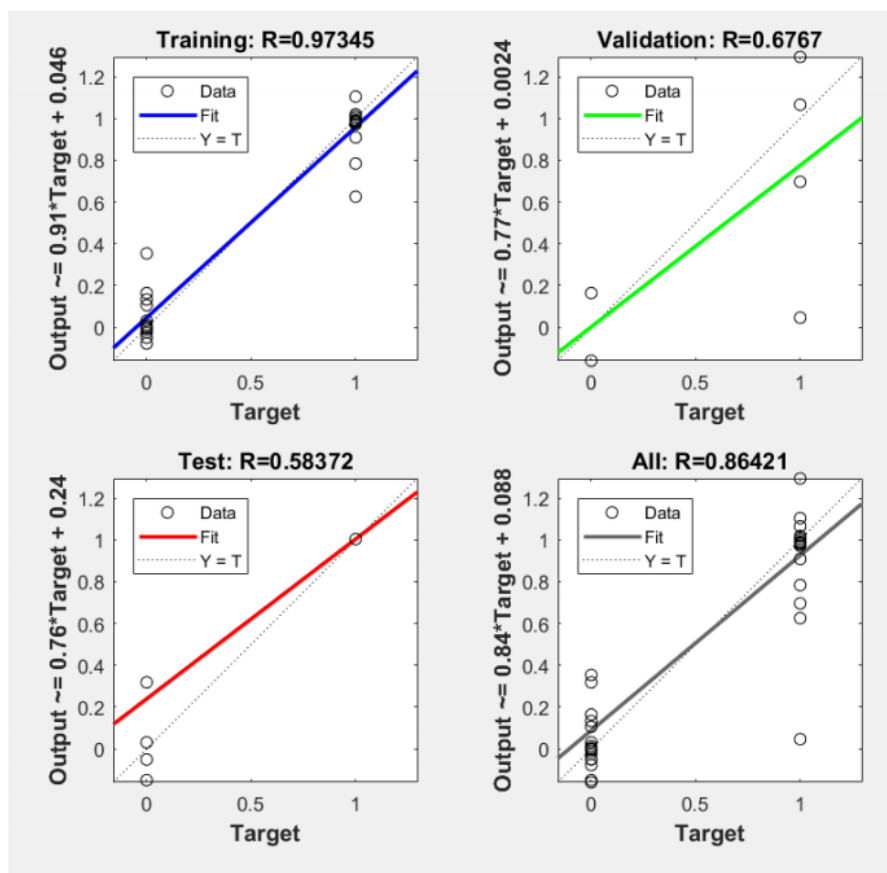


Рисунок 1 – Регрессионная модель нейронной сети.

В подразделе 3.2 приводятся результаты тестирования различных архитектур сверточных нейронных сетей на наборе данных кожных заболеваний.

Для каждой из этих задач из загруженных наборов данных извлекались выборки, необходимые для обучения нейронной сети. Выборкам присваивались метки классов 0 и 1, представленные в виде массива, соответствующие определенной задаче бинарной классификации.

В результате экспериментов показанных на рисунке 2 наибольшую точность продемонстрировала модель нейронной сети, обученная на наборе данных формата $72 \times 72 (32 \times 32 \times 4) _1 \times 1024$, которая в последствие стала основой для тонкой настройки.

Набор данных	Задача	Метод масштабирования	Размер выборки	Точность Test, %	Точность Training, %
dataSet_72x72(36x36x4)_2x648	м/б	Дублирование спектра	86/60	54	61
dataSet_72x72(36x36x4)_2x648	м/н	Дублирование спектра	86/110	54	60
dataSet_72x72(36x36x4)_2x648	б/н	Дублирование спектра	110/60	63	69
dataSet_72x72(36x36x4)_2x648	Д/з	Дублирование спектра	86+60/110	63	65
dataSet_72x72(32x32x4)_1x1024	м/б	Интерполяция + Дублирование спектра	86/60	55.56	60.45
dataSet_72x72(32x32x4)_1x1024	б/н	Интерполяция + Дублирование спектра	110/60	71,41	72
dataSet_72x72(32x32x4)_1x1024	Д/з	Интерполяция + Дублирование спектра	86+60/110	69,53	65.56
dataSet_72x72(32x32x4)_1x1024	м/н	Интерполяция + Дублирование спектра	86/110	71,51	65.10

Рисунок 2 – Результаты тестирования Xception.

Сравнивая результаты до и после этого процесса тонкой настройки, показанных на рисунке 3, увидели, что качество работы модели значительно возросло.

В подразделе 3.3 был разработан алгоритм, который объединяет результаты работы этих сетей и принимает решение. Этот алгоритм позволяет использовать информацию, полученную от нескольких моделей, для более точного прогнозирования.

Точность работы данного алгоритма составляет 70%, что означает, что в 70% случаев алгоритм правильно классифицирует данные. Специфичность

Название набора данных	Задача	Точность на тестовой выборке	Точность на обучающей выборке	Чувствительность, %	Специфичность, %
dataSet_72x72(32x32x4)_1x1024	б/н	83.33	89.84	73.33	98.15
dataSet_72x72(32x32x4)_1x1024	м/н	59.32	68.12	40.48	95.74
dataSet_72x72(32x32x4)_1x1024	м/б	65.31	75.05	90.43	57,16
dataSet_72x72(32x32x4)_1x1024	д/з	64.84	76.82	85.91	54,21

Рисунок 3 – Результаты тонкой настройки.

составляет 53%, что указывает на то, что алгоритм правильно идентифицирует отсутствие патологии у 53% здоровых пациентов. Чувствительность же равна 93,64%, что означает, что алгоритм правильно обнаруживает наличие патологии у 93,64% больных.

ЗАКЛЮЧЕНИЕ

Проведенные исследования сосредоточены на разработке нейронных сетей для автоматического анализа данных автофлуоресценции кожи. Основная цель заключалась в создании эффективных инструментов диагностики рака кожи, которые бы улучшили раннюю диагностику, снизили количество ложноположительных и ложноотрицательных результатов, повысили эффективность лечения и улучшили прогнозы выживаемости пациентов.

Были разработаны несколько моделей сверточных нейронных сетей, способных эффективно анализировать данные автофлуоресценции. Эти модели были настроены на выявление различных типов рака кожи и здоровых тканей. Проведенная серия экспериментов позволила определить оптимальные параметры и архитектуру сетей, обеспечивающие наилучшие результаты.

Проведено обширное тестирование разработанных архитектур на независимых наборах данных, чтобы убедиться в их точности и надежности. Результаты тестирования показали, что реализованные модели достигают высокой точности в дифференциации различных типов рака кожи.

Проект продемонстрировал результаты, сопоставимые с существующими коммерческими программами для анализа кожных образцов, такими как SkinVision. Однако, для дальнейшего повышения эффективности разработанных моделей машинного обучения необходимо собрать больше обучающих данных. С увеличением количества образцов и разнообразия ситуаций для обучения нейронной сети можно достичь более высокой точности и надежности в определении признаков рака кожи.

Таким образом, разработанный проект имеет значительный потенциал для улучшения диагностики рака кожи и может быть использован как альтернативный метод, способный обеспечить более раннее и точное выявление заболевания.

Основные источники информации:

- 1 Сержантов К.А. Автофлуоресцентный анализ кожных патологий с использованием нейросетевого алгоритма: науч. работа / К.А. Сержантов, М.Г. Лисовская, В.П. Захаров, А.А. Морятов, С.В. Козлов / сборник статей «Информационные технологии в моделировании и управлении: подходы, методы, решения», 2017. – 256-263 с.
- 2 Тюляндин С.А., Моисеенко В.М. Практическая онкология: учеб. посо-

- бие / С.А. Тюляндин, В.М. Моисеенко – СПб.: Центр ТОММ, 2014. – 23 с.
- 3 Медведев В.С., Потемкин В.Г. Нейронные сети. МАТЛАВ 6 / В.С. Медведев, В.Г. Потемкин – Москва: ДИАЛОГ-МИФИ, 2013. – 496 с.
- 4 Демидов Л.В. Меланома кожи: стадирование, диагностика и лечение: науч. работа / Л.В. Демидов – Москва, 2013. – 658-665 с.
- 5 Тучин В.В. Оптическая биомедицинская диагностика: учеб. пособие / В.В. Тучин – Москва: ФИЗМАТЛИТ, 2015. – 13 с.
- 6 Быков В.Л. Цитология и общая гистология: учеб. пособие / В.Л. Быков – СПб.: СОТИС, 2015. – 520 с.