

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

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

Кафедра дискретной математики и информационных технологий

**РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ СЕГМЕНТИРОВАНИЯ  
ИЗОБРАЖЕНИЙ В ФОРМАТЕ NIFTI НА ОСНОВЕ  
СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ**

**АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ**

студента 4 курса 421 группы  
направления 09.03.01 — Информатика и вычислительная техника  
факультета КНиИТ  
Соколова Вадима Павловича

Научный руководитель  
доцент, к. э. н.

\_\_\_\_\_

Г. Ю. Чернышова

Заведующий кафедрой  
доцент, к. ф.-м. н.

\_\_\_\_\_

Л. Б. Тяпаев

Саратов 2024

## ВВЕДЕНИЕ

Цифровые изображения играют важную роль в медицинских исследованиях. Для представления изображений могут применяться разнообразные форматы, причем существуют специфические форматы, в частности представляющие собой набор изображений, формирующих трехмерные образы. Задача обработки изображений и применения эффективных методов для анализа визуальной информации приобретают особую значимость. Для решения задач обработки изображений существует широкий спектр методов как компьютерного зрения, так и методов машинного обучения. С учетом разнообразия вариантов решения задачи, подбор конкретных методов является важным этапом.

Целью бакалаврской работы является разработка приложения для сегментации изображений МРТ-снимков брюшной полости. Для достижения цели требуется решить следующие задачи:

- классифицировать методы предобработки и сегментации изображений;
- осуществить анализ инструментальных средств для реализации предобработки изображений;
- осуществить анализ инструментальных средств для сегментации изображений;
- разработать приложение для сегментации печени на медицинских изображениях формата NIfTI.

Предметом данной бакалаврской работы является применение методов обработки изображений в рамках анализа МРТ-снимков брюшной полости человека. Объектом данной бакалаврской работы являются инструментальные средства для реализации обработки изображений.

Бакалаврская работа содержит три раздела. Первый раздел включает в себя обзор современных методов анализа и обработки изображений. Второй раздел посвящен анализу инструментальных средств для разработки приложения. Третий раздел содержит результаты разработки приложения для анализа МРТ-снимков в формате NIfTI на основе сверточных нейронных сетей.

Данная работа содержит семнадцать рисунков, одну таблицу и двадцать используемых источников.

В качестве материалов в бакалаврской работе использовались научные

издания, учебники, учебные пособия, документация по отдельным библиотекам на платформе Python, статистические данные в виде набора NIfTI-файлов, содержащих результаты МРТ.

## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

### Первый раздел

Изображение представляет собой набор пикселей, каждый из которых имеет определенное значение яркости и цвета. При работе с изображениями важную роль играет понятие канала. Например, RGB – цветовая модель, представление цвета в которой задается совокупностью трех цветовых каналов: красного, зеленого и синего. Каждый цветовой канал представляет собой матрицу 8-разрядных двоичных или десятичных значений в диапазоне от 0 до 255 [1], соответствующих интенсивности некоторого цвета для данного пикселя. Изображения в формате .PNG имеют дополнительный альфа-канал, значения в котором отвечают за прозрачность соответствующего пикселя изображения.

Канал изображения представляет собой его отдельные составляющие, такие, как цвет, прозрачность, яркость или более сложные признаки.

В отличие от большинства форматов визуализации, файлы NIfTI хранят изображения не в двумерном, а в трехмерном виде. Также помимо самих изображений, файлы формата NIfTI хранят и дополнительную информацию. Эта информация хранится в виде заголовка, который содержит набор полей, каждое из которых имеет уникальный идентификатор (код) и значение [2].

Таким образом, структура файла NIfTI состоит из двух частей – заголовка и непосредственно данных изображения. Данные изображения содержат сами изображения в виде многомерного массива пикселей.

Выбор методов и последовательности их применения зависит от конкретной задачи. Существуют следующие методы предварительной обработки изображений:

- преобразование к оттенкам серого;
- нормализация яркости;
- аффинные преобразования;
- обрезка фона;
- удаление шумов.

Операция сегментации изображения – это процесс разделения изображения на отдельные сегменты путем присвоения каждому пикселю изображения метки некоторого класса и заполнения бинарных масок, соответствующих выделяемым сегментам. Методы сегментации делятся на структурные,

стохастические и смешанные (гибридные) [3].

Метод пороговой обработки или thresholding заключается в преобразовании изображения из многоканального цветного в одноканальное в градациях серого. Таким образом, каждый пиксель изображения будет иметь значение на промежутке от 0 до 255, где 0 – это полностью черный пиксель, 255 – белый. На вход методу подается пороговое значение и максимальное. Пиксели, имеющие значение ниже порогового, приравниваются к 0, остальные – к максимальному заданному.

Из методов сегментации в машинном обучении выделяют методы кластеризации и классификации, а именно  $k$ -средних, иерархическую кластеризацию и  $k$ -ближайших соседей, метод опорных векторов, логистическую регрессию соответственно.

$k$ -средних выполняет два основных шага – присвоение и пересчет, при первом шаге каждый пиксель изображения присваивается к ближайшему центру кластеров, а при втором – центры пересчитываются как среднее значение пикселей в каждом кластере.

Иерархическая кластеризация строит иерархию кластеров, начинает с определенного количества отдельно расположенных пикселей и постепенно объединяет их в кластеры.

Методы глубокого обучения, в частности архитектуры сверточных нейронных сетей, нашли широкое применение в задачах сегментации изображений [4].

Процесс свертки изображения – это математическая операция, которая заключается в применении ядра (фильтра) к каждому пикселю изображения. Само ядро свертки представляет собой матрицу коэффициентов, имеющую размеры  $n \times m$ , где  $n$  и  $m$  должны быть кратны ширине и высоте исходного изображения. Фильтр должен иметь то же количество слоев, что и изображение, к которому оно применяется.

Результатом применения фильтра является двумерное изображение, каждый пиксель которого является суммой соответствующих пикселей, полученных при применении фильтра. Такое изображение называется картой активации изображения, оно показывает, насколько фильтр похож на участок изображения, к которому он был применен [5].

Операция подвыборки производится следующим образом: карта акти-

вадий разбивается на непересекающиеся прямоугольники, и из каждого такого прямоугольника по определенному принципу выбирается один пиксель. Есть множество принципов, по которым может выбираться пиксель, например *max-pooling* (выбор максимального значения), *min-pooling* (выбор минимального значения), *average-pooling* (вычисление среднего значения), однако, в большинстве случаев используется именно *max-pooling*.

Работа сверточных нейронных сетей заключается в последовательном применении к изображению блоков сверток и подвыборок, в процессе чего размеры получаемых карт активации уменьшаются (могут достигать единицы), но увеличивается их количество – количество каналов изображения. В итоге получается набор карт активации, содержащих абстрактные признаки из исходного изображения. Полученные данные (многоканальное изображение) отправляются в полносвязные слои в сверточных нейронных сетях.

К методам глубокого обучения применительно к задаче сегментации относится сегментация при помощи архитектуры сверточной сети U-Net. Данная архитектура представляет собой сверточную нейронную сеть, предназначенную для применения к изображениям, в частности, широко применяется в области медицины. В отличие от остальных сверточных архитектур, U-Net не имеет полносвязного слоя и является полностью сверточной сетью.

Сеть U-Net состоит из двух основных частей:

- сверточная часть или *encoder* состоит из нескольких сверточных блоков, которые постепенно увеличивают глубину сети и уменьшают размерность входного изображения. Каждый сверточный блок состоит из двух сверточных слоев с ядрами размера  $3 \times 3$ , за которыми следует слой нормализации и слой активации с функцией активации ReLU:  $f(x) = \max(0, x)$  [6]. После каждого сверточного блока применяется слой пулинга, который уменьшает размерность и разрешение изображения вдвое, тем самым вдвое увеличивает количество каналов изображения;
- расширяющаяся часть или *decoder* состоит из нескольких расширяющих блоков, которые постепенно уменьшают глубину сети (вдвое за каждый *up-conv*) и увеличивают размерность и разрешение выходного изображения. Каждый расширяющий блок состоит из слоя транспонированной свертки, за которым следует слой нормализации и слой

активации. Затем происходит конкатенация выходного тензора расширяющего блока с соответствующим тензором из сверточной части сети. Это позволяет сети использовать информацию о контексте изображения при сегментации.

К метрикам оценки качества моделей сегментации изображений относят точность (accuracy), чувствительность (sensitivity), специфичность (specificity) и метрику Дайса [7].

Метрика Дайса (Dice Coefficient) – это метрика, измеряющая сходство между сегментированным и истинным изображениями. Данный метод учитывает оба типа ошибок, как ложные положительные, так и ложные отрицательные. Полученное значение 1 означает идеальную сегментацию. Метрика Дайса выражается через формулу:

$$Dice = \frac{2 * TrueNegatives}{2 * TruePositives + FalsePositives + FalseNegatives}, \quad (1)$$

или

$$Dice = \frac{2 * |A \cap B|}{|A| * |B|}, \quad (2)$$

где  $A$  – предсказанная маска изображения,  $B$  – истинная маска изображения [8].

Таким образом, получается функция потерь Дайса или Dice loss:

$$L(a) = 1 - Dice, \quad (3)$$

где  $a$  – сегментационная модель.

Далее в работе будут реализованы такие методы, как метод пороговой обработки изображений и сегментация при помощи архитектуры глубокого обучения U-Net.

## Второй раздел

На сегодняшний день существует множество технологий, предоставляющих возможность реализации модулей сегментации изображений.

Широкое применение в решении задач искусственного интеллекта, в частности, компьютерного зрения нашел язык программирования высокого уровня Python. Существует множество готовых библиотек, реализующих как

методы, относящиеся к обычному компьютерному зрению (такие библиотеки, как OpenCV, scikit-image, PIL), так и методы машинного обучения. Библиотеки языка Python предоставляют не только готовые модели (TensorFlow, PyTorch), но так же и готовые функции и слои, с помощью которых можно спроектировать уникальную архитектуру под конкретные требования и задачу (Keras, Monai). Кроме того, Python-код легко отлаживать и модифицировать, что позволяет быстро вносить изменения в приложение.

Эффективным средством решения задач компьютерного зрения является высокоуровневый язык программирования C++.

Одним из самых главных недостатков данного подхода является то, что процесс написания и отладки кода на C++ занимает гораздо больше времени, чем написание такого же по функционалу модуля на языке Python.

Существуют и другие подходы, в частности еще одним возможным средством решения задачи сегментации изображений является высокоуровневый интерпретируемый язык программирования MatLab с соответствующим пакетом для решения задач компьютерного зрения Image Processing Toolbox и собственной интегрированной средой разработки.

Однако, MatLab является дорогостоящим коммерческим продуктом, и его применение в медицинских учреждениях весьма ограничено.

Для дальнейшей разработки был выбран язык высокого уровня Python, так как он является open-source проектом и содержит большое количество библиотек, что ускоряет процесс тестирования и разработки приложения и делает его более гибким.

В качестве вспомогательных инструментов для разработки модуля сегментации были выбраны такие библиотеки, как OpenCV, Monai и Pytorch.

В контексте применения на платформе Python Monai – это библиотека, предназначенная для построения, обучения и применения моделей глубокого обучения в области медицинских технологий. Отличительной чертой данной библиотеки является предоставляемая ей возможность работы с файлами медицинского формата, такими, как DICOM и NIfTI [9].

OpenCV, в отличие от других библиотек, реализующих работу с изображениями в Python (matplotlib и scikit-image), предоставляет не только стандартные методы для загрузки, сохранения и изменения цветовой палитры изображения, но так же и методы пороговой обработки и поиска контуров.

В качестве инструмента для реализации серверной части был выбран фреймворк Flask – библиотека для создания веб-приложений на языке Python, которая позволяет организовать серверную часть приложения путем настройки HTTP-сервера, обработки HTTP-запросов и отправки ответов [10].

Для разработки клиентской части был выбран фреймворк Vue.js 3 для языка JavaScript. Его преимуществами перед другими фреймворками и библиотеками, например такими, как Angular, React и JQuery являются:

- условная отрисовка – компоненты Vue не встраиваются в DOM-дерево ровно до того момента, пока они не будут вызваны [11];
- гибкость – благодаря своей гибкости, реактивности данных и множеству вспомогательных директив Vue.js позволяет сэкономить время на разработке клиентской части;
- Vue имеет наименьший вес среди остальных JavaScript-фреймворков.

Vue является наиболее подходящим решением для минималистичного веб-приложения, не теряя при этом возможности реализации функционала.

### **Третий раздел**

Первоначальные данные, поступившие из клинической больницы, имели неподходящий формат (png и webm вместо NIfTI, используемого в медицине повсеместно). Кроме того, данные были неразмеченными, что сужает перечень моделей, способных решить поставленную задачу. Таким образом, на имевшихся данных был опробован метод пороговой обработки, относящийся к методам компьютерного зрения.

Данный метод показал вполне неплохой визуальный результат при ручной настройке, однако данные не были размечены, соответственно, для оценки результата, полученного при помощи метода пороговой обработки, нельзя подобрать численной метрики.

Помимо этого, OpenCV находит все контуры, присутствующие на изображении, и не всегда нужный контур печени будет являться самым большим на изображении, поэтому придется просматривать все контуры для поиска необходимого.

Было решено использовать модель сверточной нейронной сети архитектуры U-Net. Для обучения, настройки и опробации данной модели необходимо было добыть размеченные МРТ-изображения брюшной полости с масками, соответствующими печени. Для решения данной задачи был исполь-

зован The Cancer Imaging Archive, находящийся в открытом доступе в сети Интернет. Ресурс предоставляет наборы данных, собранных Национальным институтом онкологии, США, при помощи американских медицинских клиник, центров и институтов радиологии. Архив содержит анонимизированные КТ- и МРТ-изображения брюшной полости [12]. Для реализации модели были отобраны T1-взвешенные МРТ-снимки в артериальной фазе.

Первым шагом предобработки трёхмерных данных формата NIfTI для их использования моделями глубокого обучения является изменение порядка величин массива, хранящего размеры изображения. В данном случае на первое место ставится количество каналов изображения.

Затем все изображения приводятся к одному значению пространственного расширения. В данном случае значение берется равным (2x2x1.5), так как при таком значении пространственного разрешения модель показывает оптимальный результат по времени обработки одного NIfTI-файла.

Далее выбирается пространственная ориентация изображения – в данном случае это "RAS". Переориентация изображений перед их подачей на вход модели является необходимой, так как изначально снимки могут иметь разную пространственную ориентацию.

Верхняя граница интенсивности пикселей задана значением 400, нижняя граница яркости пикселей задана значением -100. Таким образом, отдельные органы и ткани больше контрастируют друг с другом и становятся визуально различимыми.

Далее изображение обрезается вокруг переднего плана, оставляя только области, соответствующие телу пациента. Полученное изображение приводится к одному размеру (в данном случае это 128x128x64 пикселей) и преобразуется в Torch-тензор, над которым будут производиться вычисления внутри модели глубокого обучения.

Обучение модели производилось на батчах размерности 1 в течении 600 эпох. Данное количество эпох было выбрано исходя из полученного значения функции потерь (Dice loss) - 0.21, что является хорошим значением для решаемой задачи. Соответственно, полученная точность модели равняется 0,79 по метрике Дайса.

Далее был разработан модуль, реализующий сегментацию новых изображений при помощи уже обученной модели. Функция "segment\_liver" дан-

ного модуля принимает на вход трехмерный МРТ-снимок пользователя в формате NIfTI.

Поступившее изображение покадрово обрабатывается настроенной моделью, и то изображение, маска которого содержит наибольшее количество ненулевых (содержащих интересующий орган) пикселей, возвращается из функции вместе с маской и площадью органа в пикселях.

Архитектура приложения представляет собой модель "клиент-сервер".

Серверная часть приложения определяет две конечные точки – адреса, по которым можно обратиться к приложению. Первая конечная точка реализует схему вида "{scheme}://{host}:{port}/" и представляет собой корневой адрес, который по HTTP запросу с методом "POST" ожидает файл в формате .nii.gz, который он сохранит, обработает при помощи уже обученной модели и сохранит полученные результаты. В ответ пользователю отправляются пути, по которым клиент может сделать запрос для получения среза и маски исходного изображения, на котором орган имеет максимальную площадь.

Второй эндпоинт имеет вид "{scheme}://{host}:{port}/<path:path>". В качестве параметра "path" клиент указывает переданное ему ранее название файла, которое хранится в директории для статических файлов на серверной части приложения.

Для разработки визуального интерфейса используются стандартные средства разработки веб-приложений, такие, как HTML, CSS и JavaScript. Для более быстрой и гибкой разработки был выбран фреймворк Vue.js 3.

Было решено сделать 2 раздела приложения – раздел загрузки данных и раздел отображения сегментированных снимков. Также интерфейс содержит шапку веб-страницы, на которой расположены кнопки для перехода по разделам.

При нажатии на кнопку "Загрузить NIfTI-файл" пользователю открывается диалоговое окно, предлагающее выбрать файл из файловой системы для последующей его обработки приложением. При успешной загрузке файла происходит его обработка. "Данные не загружены", а сам пользователь может запросить отображение результатов по нажатии кнопки "Показать изображение".

При нажатии на кнопку "Показать изображения" открывается раздел показа изображений, отображающий срез исходного изображения, на кото-

ром интересующий орган имеет максимальную площадь в пикселях, а также маску, соответствующую данному срезу. Под описанными изображениями располагается сопутствующая информация о проекции среза и размерах органа.

Для рассмотренного примера площадь наибольшей проекции (аксиальной) [13] составляет 1311 пикселей.

Разработанное приложение позволяет более объективно сравнить размеры органа для двух различных снимков. Это может выполняться как для двух отдельных пациентов, так и для отдельных периодов выбранного пациента.

## ЗАКЛЮЧЕНИЕ

В ходе данной бакалаврской работы были опробованы разные подходы к сегментации медицинских NIfTI-изображений, в частности, методы глубокого обучения – модель сверточной нейронной сети U-Net, а так же метод пороговой обработки, относящийся к методам классического компьютерного зрения. Для сегментации МРТ-снимков брюшной полости была осуществлена предварительная обработка данных, затем были использованы традиционные методы компьютерного зрения.

Разработанное web-приложение состоит из двух главных модулей. Первый модуль представляет собой серверную часть приложения, реализованную с использованием платформы Python. Серверная часть приложения принимает NIfTI-файлы посредством HTTP-запросов, производит их обработку и сегментацию, затем возвращает обработанные данные. Клиентская часть, реализованная при помощи фреймворка Vue.js для языка JavaScript, предоставляет пользователю web-интерфейс для взаимодействия с приложением.

Функционал разработанного приложения позволяет загрузить МРТ-снимки в формате NIfTI, произвести их предварительную обработку, сегментацию и получить кадр, на котором печень имеет самый большой размер в аксиальной проекции, соответствующую ему маску и размер органа на данном кадре в пикселях.

Для настройки модели использовались данные из открытого источника – базы данных медицинских изображений The Cancer Imaging Archive, основанного при поддержке американского национального института онкологии. Из всего массива данных, представляющих собой анонимизированные КТ- и МРТ-изображения брюшной полости, были выбраны именно МРТ-изображения и соответствующие им бинарные маски печени, поскольку данный тип исследований лучше показывает мягкие ткани, а также исходные данные являлись МРТ-снимками.

Возможности практического использования разработанного приложения заключаются в применении врачами-специалистами для работы как с текущими, так и архивными снимками в распространенном формате NIfTI.

## Основные источники информации:

- 1 Савельева, М.Г., Урбанович П.П. Метод стеганографического преобразования web-документов на основе растровой графики и модели RGB : [Электронный ресурс] - URL: <https://cyberleninka.ru/article/n/metod-steganograficheskogo-preobrazovaniya-web-dokumentov-na-osnove-rastrovoy-grafikii-modeli-rgb> (дата обращения 02.11.2023). - Загл. с экрана. - Яз. рус.
- 2 Документация формата медицинских изображений NIfTI : [Электронный ресурс] - URL: <https://nifti.nih.gov/nifti-1/> (дата обращения 02.11.2023). - Загл. с экрана. - Яз. англ.
- 3 Джо К.О., Гергет О.М. Методы и алгоритмы сегментации изображений // Автоматизация и моделирование в проектировании и управлении. - 2020. - №2. - С. 11-16.
- 4 Bai, W. [Электронный ресурс] Wiley : An ENet Semantic Segmentation Method Combined with Attention Mechanism / W. Bai . URL: <https://onlinelibrary.wiley.com/doi/10.1155/2023/6965259> (дата обращения 19.12.2023). - Загл. с экрана. Яз. англ.
- 5 Saleem, M. Application of Deep Convolution Neural Network for Image : [Электронный ресурс] - URL: [https://www.researchgate.net/publication/373988922\\_Application\\_of\\_Deep\\_Convolution\\_Neural\\_Network\\_for\\_Image](https://www.researchgate.net/publication/373988922_Application_of_Deep_Convolution_Neural_Network_for_Image) (дата обращения 08.01.2024) - Загл. с экрана. - Яз англ.
- 6 Chou, A., Li, W., Roman, E. GI Tract Image Segmentation with U-Net and Mask R-CNN : [Электронный ресурс] - URL: <https://cs231n.stanford.edu/reports/2022/pdfs/164.pdf> (дата обращения 11.01.2024) - Загл. с экрана. - Яз. англ.
- 7 Minaee, S., Бойков, Ю., Porikli, F., Plaza, A. Image Segmentation Using Deep Learning:A Survey : [Электронный ресурс] - URL: [https://www.researchgate.net/publication/338644066\\_Image\\_Segmentation\\_Using\\_Deep\\_Learning\\_A\\_Survey](https://www.researchgate.net/publication/338644066_Image_Segmentation_Using_Deep_Learning_A_Survey) (дата обращения 11.01.2024) - Загл. с экрана. - Яз. англ.
- 8 Черненький, И.М., Черненький, М.М., Фиев, Д.Н., Сирота, Е.С. Сегментация почечных структур по изображениям контрастной компьютерной томографии с помощью сверточной нейронной сети : [Электронный ресурс] - URL: <https://cyberleninka.ru/article/n/segmentatsiya-pochechnyh->

struktur-po-izobrazheniyam-kontrastnoy-kompyuternoy-tomografii  
-s-pomoschyu-svertochnoy-neyronnoy-seti/viewer

(дата обращения 18.01.2024) - Загл. с экрана. - Яз. рус.

- 9 Cardoso, M.J., Li, W., Brown, R., Ma, N. MONAI: An open-source framework for deep learning in healthcare : [Электронный ресурс] - URL: <https://arxiv.org/abs/2211.02701> (дата обращения 23.01.2024) - Загл. с экрана. - Яз. англ.
- 10 Документация фреймворка Flask [Электронный ресурс] - URL: <https://flask.palletsprojects.com/en/3.0.x/> (дата обращения 08.04.2024) - Загл. с экрана. - Яз. англ.
- 11 Документация фреймворка Vue.js [Электронный ресурс] - URL: <https://vuejs.netlify.app/guide/introduction.html> (дата обращения 11.04.2024) - Загл. с экрана - Яз. англ.
- 12 The Cancer Genome Atlas Liver Hepatocellular Carcinoma Collection : [Электронный ресурс] - URL: <https://www.cancerimagingarchive.net/collection/tcga-lihc/> (дата обращения 02.02.2024) - Загл. с экрана. - Яз.
- 13 Anderson, M.W., Fox, M.G. Sectional Anatomy by MRI and CT. - 4 изд. - Amsterdam: Elsevier, 2016. - 580 с.