

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 WEB-ПРИЛОЖЕНИЯ ДЛЯ ВЕКТОРИЗАЦИИ
ИЗОБРАЖЕНИЙ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 271 группы
направления 09.04.01 — Информатика и вычислительная техника
факультета КНиИТ
Кузнецова Теймура Мохлюдовича

Научный руководитель

доцент, к. э. н.

Г. Ю. Чернышова

Заведующий кафедрой

доцент, к. ф.-м. н.

Л. Б. Тяпаев

Саратов 2025

ВВЕДЕНИЕ

Современные технологии обработки изображений, особенно в медицине, играют ключевую роль в диагностике заболеваний и планировании лечения. Одним из важнейших направлений является векторизация изображений — процесс преобразования растровых данных в структурированные векторные представления. В медицинской диагностике, например, при работе с магнитно-резонансной томографией (МРТ) и компьютерной томографией (КТ), векторизация позволяет выделять значимые области, такие как органы, ткани и патологические изменения, что критически важно для анализа и принятия врачебных решений. Однако существующие методы сталкиваются с проблемами высокой вариативности данных, наличия шумов и сложных анатомических структур, а также с недостатком размеченных данных. Разработка эффективных инструментов для сегментации медицинских изображений на основе современных методов глубокого обучения, таких как U-Net, является актуальной задачей, способной значительно улучшить качество диагностики и автоматизировать процесс анализа медицинских данных.

Целью магистерской работы является разработка web-приложения для сегментации медицинских изображений на основе современных методов векторизации.

В соответствии с поставленной целью определены следующие задачи:

- провести анализ существующих подходов к векторизации изображений, в частности методов сегментации;
- исследовать инструментальные средства и алгоритмы, применяемые для сегментации медицинских изображений;
- реализовать модель для построения сегментированного изображения и интегрировать ее в web-приложение;
- провести оценку качества работы модели на медицинских данных.

В работе использовались данные из набора Liver Dataset из Medical Segmentation Decathlon, включающего 3D-изображения в формате NIfTI и соответствующие маски сегментации. Для реализации модели применялись библиотеки Python, такие как MONAI, PyTorch и фреймворки Flask и Vue.js для разработки web-приложения.

Магистерская работа состоит из введения, трёх разделов, заключения, списка использованных источников и одного приложения. Первый раздел по-

священ теоретическим основам сегментации изображений, включая методы векторизации, анализу классических и нейросетевых подходов к сегментации, а также критериям оценки качества для валидации результатов. Второй раздел включает в себя анализ основных инструментов и библиотек Python, применяемых для решения задач сегментации, включая специализированные фреймворки для работы с медицинскими данными. Третий раздел посвящён разработке web-приложения на Flask и Vue.js для сегментации МРТ-изображений печени с использованием модели U-Net.

Общий объём работы 65 страниц, из них 54 страница – основное содержание, включая 21 рисунок и 2 таблицы, список использованных источников информации – 43 наименования.

Научная новизна работы заключается в разработке web-приложения, интегрирующего модель U-Net для сегментации медицинских изображений, с использованием современных инструментов и методов глубокого обучения. Практическая значимость работы состоит в создании инструмента, который может быть использован в медицинской диагностике для автоматизации анализа МРТ-снимков, что способствует повышению точности диагностики и сокращению времени обработки данных.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Классификация основных методов и подходов к сегментации изображений» посвящён анализу ключевых подходов к обработке изображений, включая векторизацию и сегментацию, с акцентом на их применение в медицинской диагностике. Особое внимание уделяется переходу от классических алгоритмов к современным нейросетевым архитектурам, а также методам оценки качества сегментации.

Описывается процесс преобразования растровых данных в структурированное математическое представление, основанное на контурах, графах и геометрических примитивах. Этот этап критически важен для последующего анализа, поскольку позволяет перейти от плотного пиксельного описания к компактным параметрическим моделям.

Основной задачей векторизации указывается подготовка данных для алгоритмической обработки, включая выделение границ объектов, снижение размерности и создание промежуточного представления, пригодного для машинного обучения. В работе подробно разбираются следующие методы:

- пороговая обработка (Thresholding), являющаяся простейшим подходом, основанным на бинаризации изображения по заданному уровню яркости;
- дискретное вейвлет-преобразование (DWT) [1]. Данный метод является частотным анализом, применяемым для сжатия данных и выделения текстурных особенностей;
- методы локальных дескрипторов (SIFT, HOG) [2, 3], обеспечивающий инвариантность к масштабу и повороту, что важно для задач сопоставления изображений;
- свёрточные нейронные сети (CNN) [4], являющиеся наиболее современным подходом, автоматически извлекающими иерархические признаки.

Отмечается, что векторизация сама по себе не решает задачу семантического анализа, но создаёт основу для последующей сегментации.

Проводится сравнение традиционных алгоритмов и современных архитектур глубокого обучения.

Классические методы сегментации включают:

- пороговые алгоритмы (глобальные и адаптивные), которые разделяют пиксели по интенсивности;
- кластеризацию (K-Means, Mean Shift), группирующую пиксели по схожести признаков;
- методы на основе границ (операторы Собеля, Кэнни), выделяющие контуры объектов;
- морфологические операции (эрозия, дилатация), улучшающие качество бинарных масок.

Показано, что классические методы (например, реализованные в OpenCV) требуют ручной настройки параметров и плохо масштабируются на 3D-данные. В то же время свёрточные нейронные сети (CNN) демонстрируют более высокую точность благодаря способности автоматически извлекать признаки и анализировать сложные пространственные зависимости.

Особое внимание уделяется следующим нейросетевым подходам:

- U-Net архитектура с кодировщиком-декодировщиком и skip-connections, ставшая стандартом в медицинской сегментации;
- transformer-модели (Swin-UNet, UNetR), обеспечивающие глобальный контекстный анализ;
- гибридные методы (CNN + топологический анализ), которые комбинируют

преимущества разных подходов.

Отдельно обсуждаются проблемы малых размеченных датасетов и способы их решения через аугментацию данных и transfer learning.

Детально разбирается структура сети, включающая кодировщик (encoder), представляющий последовательность свёрточных слоёв и max-pooling, уменьшающих разрешение и выделяющих семантические признаки, декодировщик (decoder), включающий транспонированные свёртки и skip-connections, восстанавливающие пространственную информацию, и 3D-модификации (V-Net) для адаптации к обработке объёмных медицинских изображений.

Подчёркивается, что U-Net особенно эффективна для задач с ограниченными данными благодаря сохранению деталей через skip-connections [5, 6].

Анализируются метрики, используемые для валидации результатов:

- Dice Similarity Coefficient (DSC) как основной показатель перекрытия масок.
- Normalized Surface Distance (NSD) для оценка точности границ.
- чувствительность и специфичность, используемые для баланса между ложными положительными и отрицательными срабатываниями.

В разделе систематизированы методы векторизации и сегментации, показан переход от классических алгоритмов к нейросетевым подходам. Доказано, что U-Net и её модификации являются наиболее перспективными для медицинской сегментации, обеспечивая высокую точность даже при малом объёме данных.

Второй раздел «Инструментальные средства для сегментации изображений» рассматривает комплекс программных решений для сегментации изображений, доступных на различных платформах и языках программирования. Особое внимание уделяется сравнительному анализу их функциональных возможностей, производительности и применимости для решения специализированных задач, в частности – обработки медицинских изображений.

Описываются ключевые программные продукты для сегментации изображений, представленные в различных экосистемах разработки. Анализируются их преимущества и ограничения в контексте практического применения.

MATLAB Image Processing Toolbox выделяется среди кроссплатформенных решений, предлагая оптимизированные алгоритмы глубокого обучения, включая современную Segment Anything Model (SAM). Его отличительной осо-

бенностью является наличие интерактивного приложения Image Segmenter, позволяющего выполнять разметку без необходимости написания кода, что значительно ускоряет процесс обработки изображений.

Для разработчиков, работающих с C++, представляет интерес библиотека `mlprack`, которая обеспечивает высокую производительность за счет использования `Armadillo` для линейной алгебры. Однако её применение сопряжено с рядом сложностей, включая проблемы совместимости в Windows-средах, необходимость ручной настройки BLAS/LAPACK и ограниченную документацию по задачам сегментации. На практике для C++ проектов чаще выбирают OpenCV с модулями `dnn` и `imgproc`, которые предоставляют более широкие возможности для работы с изображениями.

В Java-экосистеме ключевым инструментом является ImageJ/Fiji, расширяемый многочисленными плагинами, такими как Trainable Weka Segmentation для классификации пикселей. Также стоит отметить DeepJavaLibrary (DJL) - фреймворк для инференса моделей PyTorch и TensorFlow с поддержкой GPU-ускорения.

Облачные сервисы, такие как Google Cloud Vision API и Amazon Rekognition, предлагают готовые решения для сегментации через REST API, включая функции постобработки и выделения объектов. Однако их использование может быть ограничено требованиями к конфиденциальности данных, особенно в медицинских приложениях.

Python выделяется среди других платформ наличием высокоуровневых API в библиотеках типа MONAI [7], глубокой интеграцией с фреймворками PyTorch и TensorFlow, развитой экосистемой для обработки данных и широким сообществом разработчиков и обширной документацией.

Подробно рассматриваются специализированные библиотеки Python, предназначенные для решения задач сегментации изображений. Основное внимание уделяется их функциональным возможностям и особенностям применения в различных предметных областях.

Фреймворки TensorFlow и Keras предоставляют мощные инструменты для проектирования и обучения нейросетевых архитектур. Они включают готовые реализации популярных моделей (DeepLabV3+, U-Net), оптимизированные конвейеры обработки данных (`tf.data.Dataset`) и средства аугментации (`tf.image`). Однако для работы с медицинскими изображениями более подходящей оказы-

вается специализированная библиотека MONAI (Medical Open Network for AI), расширяющая возможности PyTorch.

Особого внимания заслуживает реализация 3D U-Net в MONAI, которая позволяет анализировать целые объемы медицинских изображений, а не отдельные срезы. Библиотека также предоставляет специализированные функции потерь (DiceCELoss) и метрики качества (DiceScore, HausdorffDistance), критически важные для медицинских приложений.

В заключении подраздела обосновывается выбор MONAI в качестве основного инструмента для данной работы.

Третий раздел «Разработка модели для сегментирования изображения» посвящён процессу создания и оптимизации web-приложения для сегментации медицинских изображений с использованием нейронных сетей. Особое внимание уделяется архитектурным решениям, выбору технологического стека и методам повышения эффективности модели.

Описывается процесс создания web-приложения для взаимодействия пользователя с моделью семантической сегментации. Анализируются преимущества выбранного технологического подхода и особенности реализации ключевых функциональных модулей.

Для реализации backend-части приложения был выбран микрофреймворк Flask, который обеспечивает необходимую гибкость при отсутствии избыточных компонентов. Этот выбор обусловлен минималистичной архитектурой, позволяющей реализовать только требуемый функционал, простотой интеграции с библиотеками обработки медицинских изображений (nibabel, numpy) и легкостью настройки маршрутизации для основных операций.

Приложение реализует следующий функциональный набор:

- загрузку медицинских изображений в формате NIfTI через интерфейс Drag-and-drop;
- визуализацию оригинальных и сегментированных данных с навигацией по срезам;
- отображение статистики по классам (фон, здоровая ткань, пораженные области);
- возможность скачивания результатов обработки.

Подробно рассматривается процесс создания и оптимизации модели сегментации на основе архитектуры U-Net. Особое внимание уделяется работе с

медицинскими данными и методам решения специфических проблем.

Для обучения модели использовался Liver Dataset из Medical Segmentation Decathlon [8], содержащий:

- 131 тренировочное изображение с соответствующими масками;
- 70 тестовых изображений для валидации;
- данные в формате NIfTI с разрешением $0.7 * 0.7 * 1.5$ мм.

Предобработка данных включала комплекс из стандартизации ориентации (RAS), нормализации интенсивностей пикселей, обрезки фоновых областей и приведения к единому размеру.

Результаты работы модели:

- средний Dice score был 0.67 (после оптимизации);
- точность сегментации здоровых тканей показала 0.867 ± 0.05 ;
- точность сегментации опухолей составила 0.032-0.735;
- время обучения составило около 12 часов (200 эпох на CPU).

Разработанное решение демонстрирует стабильную работу. Но несмотря на значимые результаты, проведённое исследование выявило ряд ограничений и перспективных направлений для дальнейшего развития системы. Наиболее критичным является вопрос качества данных - существующая выборка требует существенного расширения, особенно в части редких патологий, что позволит повысить точность сегментации аномальных участков. Вторым ключевым направлением является повышение вычислительной эффективности за счёт перехода на GPU-ускорение с использованием технологии CUDA, что критически важно для обработки полных объёмов медицинских изображений без потери диагностически значимой информации.

Особый потенциал для развития системы лежит в области функционального расширения. В первую очередь это касается реализации поддержки мультимодальных данных, включая КТ и УЗИ исследования, что значительно расширит область применения разработки. Не менее важной задачей является интеграция с медицинскими стандартами, в частности с DICOM-форматом, и существующими информационными системами лечебных учреждений. Перспективным направлением представляется развитие интеллектуальных функций автоматизированной диагностической поддержки, включая количественный анализ характеристик выявленных патологий и генерацию структурированных заключений.

Реализация указанных направлений развития позволит создать комплексное решение, соответствующее актуальным требованиям современной цифровой медицины и способное существенно повысить эффективность работы врачей-диагностов. Особое внимание при этом должно быть уделено вопросам валидации системы на расширенных клинических данных и обеспечению необходимого уровня объяснимости принимаемых алгоритмом решений для их успешного внедрения в клиническую практику.

ЗАКЛЮЧЕНИЕ

В ходе выполнения магистерской работы был проведен комплексный анализ современных методов сегментации медицинских изображений, в результате которого архитектура U-Net была идентифицирована как оптимальное решение для обработки МРТ-данных благодаря своей способности сохранять пространственные взаимосвязи между срезами и эффективно работать с ограниченными размеченными данными. На основе этого исследования было разработано решение, включающее web-приложение на Flask 3.1.0 и Vue.js с поддержкой формата NIfTI и интерактивными функциями анализа, а также оптимизированную модель 3D U-Net, достигшую среднего Dice score 0.9 на тренировочной и 0.7 на тестовой выборках.

Для реализации решения был обоснован и применен технологический стек на основе Python 3.9 с библиотекой MONAI 1.1.0, обеспечившей специализированную обработку медицинских изображений, включая механизмы компенсации дисбаланса классов и аугментации данных. Разработанное приложение поддерживает ключевые функции для клинического применения: загрузку и визуализацию медицинских изображений, интерактивный просмотр срезов, отображение сегментированных областей и 3D-реконструкцию.

Несмотря на достигнутые результаты, исследование выявило ряд перспективных направлений для дальнейшего развития. Основные усилия следует сосредоточить на расширении обучающей выборки, особенно для редких патологий, что позволит повысить точность сегментации патологических изменений. Существенный прирост производительности может быть достигнут за счет перехода на GPU-ускорение с использованием технологии CUDA, что обеспечит обработку полных объемов данных без компромиссов в качестве. Дополнительные перспективы включают расширение функциональности для поддержки

мультимодальных данных (КТ, УЗИ), интеграцию с DICOM-стандартом и медицинскими информационными системами, а также развитие функций автоматизированной диагностической поддержки.

Основные источники информации:

- 1 Wavelet Integrated CNNs for Noise-Robust Image Classification [Text] / Li, Q., Shen, L., Guo, S., and Lai, Z. — 2020.
- 2 Krasnabayeu, E. A. Comparison of binary feature points descriptors of images under distortion conditions [Text] / Krasnabayeu, E. A., Chistabayeu, D. V., and Malyshev, A. L. // Computer Optics. — 2019. — June. — Vol. 43, no. 3.
- 3 Южаков, Г. Б. Алгоритм быстрого построения дескрипторов изображения, основанных на технике гистограмм ориентированных градиентов [Текст] // ТРУДЫ МФТИ. — 2013. — Т. 5, № 3. — С. 84–91.
- 4 Review of deep learning: concepts, CNN architectures, challenges, applications, future directions [Text] / Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., and Farhan, L. // Journal of Big Data. — 2021. — Mar. — Vol. 8, no. 1.
- 5 nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation [Text] / Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., and Maier-Hein, K. H. // Nature Methods. — 2020. — Dec. — Vol. 18, no. 2. — P. 203–211.
- 6 Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation [Text] / Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., and Wang, M. — 2021. — 2105.05537.
- 7 MONAI Team. — MONAI Documentation [Text] : 2023. — Medical Open Network for AI. Access mode: <https://docs.monai.io> (online; accessed: June 6, 2025).
- 8 A large annotated medical image dataset for the development and evaluation of segmentation algorithms [Text] / Simpson, A. L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., van Ginneken, B., Kopp-Schneider, A., Landman, B. A., Litjens, G., Menze, B., Ronneberger, O., Summers, R. M., Bilic, P., Christ, P. F., Do, R. K. G., Gollub, M., Golia-Pernicka, J., Heckers, S. H., Jarnagin, W. R., McHugo, M. K., Napel, S., Vorontsov, E., Maier-Hein, L., and Cardoso, M. J. — 2019. — 1902.09063.