

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА

На правах рукописи

Пенкин Максим Александрович

**Гибридные методы анализа и повышения качества
медицинских изображений**

Специальность 1.2.2
Математическое моделирование, численные методы и комплексы
программ

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
кандидата физико-математических наук

Москва — 2025

Диссертация подготовлена на кафедре математической физики факультета вычислительной математики и кибернетики Московского государственного университета имени М.В. Ломоносова

Научный руководитель: **Крылов Андрей Серджевич**
доктор физико-математических наук, профессор

Официальные оппоненты: **Визильтер Юрий Валентинович**,
доктор физико-математических наук,
профессор РАН,
Государственный научно-исследовательский
институт авиационных систем, директор
по направлению – руководитель научного
комплекса «Искусственный интеллект и
техническое зрение»
Сулимов Владимир Борисович,
доктор физико-математических наук, доцент,
Московский государственный университет
имени М. В. Ломоносова,
Научно-исследовательский вычислительный
центр, заведующий лабораторией вычислительных
систем и прикладных технологий
программирования
Турлапов Вадим Евгеньевич,
доктор технических наук, доцент,
Национальный исследовательский Нижегородский
государственный университет
имени Н. И. Лобачевского, профессор кафедры
высокопроизводительных вычислений и
системного программирования

Защита диссертации состоится «25» февраля 2026 г. в 15 часов 30 минут
на заседании диссертационного совета МГУ.012.1
Московского государственного университета имени М.В. Ломоносова по адресу:
119991, г. Москва, ул. Ленинские горы, д. 1, стр. 52, факультет ВМК, ауд. 685.
E-mail: ds@cs.msu.ru

С диссертацией можно ознакомиться в отделе диссертаций научной библиотеки
МГУ имени М.В. Ломоносова (Ломоносовский просп., д. 27) и на портале:
<https://dissovet.msu.ru/dissertation/3734>

Автореферат разослан «__» ____ 20__ г.

Учёный секретарь

диссертационного совета МГУ.012.1,
доктор физико-математических наук,
член-корреспондент РАН



А.В. Ильин

Общая характеристика работы

Актуальность темы

Современная медицина активно использует методы визуализации для диагностики, мониторинга заболеваний и планирования лечения. Однако рост объёма и сложности медицинских изображений требует более эффективных методов математического моделирования и проблемно-ориентированных комплексов программ для их анализа и обработки с применением современных компьютерных технологий. Создание таких методов и эффективных алгоритмов их реализации, является актуальной научно-технической проблемой в области математического моделирования, решение которой позволит повысить точность, интерпретируемость и надежность систем компьютерной диагностики. Ключевой задачей при этом является синтез методологий, позволяющий интегрировать классические математические модели с подходами глубокого обучения, извлекающими знания из данных.

Актуальность темы обусловлена необходимостью повышения точности диагностики. Традиционные методы обработки изображений (например, пороговая сегментация, фильтрация) часто недостаточно эффективны при работе с зашумлёнными или низкоконтрастными медицинскими снимками. Гибридные методы, сочетающие классические математические модели с моделями машинного обучения (свёрточные нейронные сети, трансформеры), позволяют улучшить качество распознавания патологий, например, объединяя аппарат дифференциальных уравнений и анализ и обработку больших данных.

Более того, классические численные методы хорошо справляются лишь с определёнными типами искажений, часто требуя ручной настройки параметров и не всегда адаптируясь к сложным случаям. Методы машинного обучения, наоборот, демонстрируют высокую обобщающую способность и эффективность в задачах шумоподавления, повышения разрешения и улучшения контраста, но могут быть чрезмерно сложными или требовать больших вычислительных ресурсов, затрудняя проведение вычислительных экспериментов. Гибридный подход, сочетающий преимущества классических моделей (интерпретируемость, низкие вычислительные затраты) и методов машинного обучения (адаптивность, высокая точность), позволяет преодолеть эти ограничения.

Внедрение гибридных методов анализа и обработки изображений ведёт к автоматизации и ускорению медицинской диагностики. Ручная обработка изображений трудоёмка и подвержена субъективным ошибкам. Комбинация алгоритмов предварительной обработки и нейросетевой классификации сокращает время диагностики и снижает нагрузку на врачей. Появление новых архитектур нейронных сетей (например, vision transformers) и методов объяснимого ИИ (XAI) открывает возможности для создания интерпретируемых гибридных методов, критически важных в медицине.

Таким образом, разработка и совершенствование методов анализа и обработки медицинских изображений представляет собой актуальную научно-практическую задачу, решение которой способствует повышению качества диагностики, снижению временных затрат и улучшению интерпретируемости.

Объектом исследования в данной диссертационной работе является процесс компьютерного анализа медицинских изображений различных модальностей. Предметом исследования – гибридные методы повышения качества, классификации и сегментации медицинских изображений.

Цель работы

Основная цель диссертационной работы заключается в разработке новых гибридных методов для обработки и анализа медицинских изображений различных модальностей, таких как изображения магнитно-резонансной томографии, ультразвуковые изображения, гистологические полнослайдовые изображения и колоноскопические изображения, а также в реализации соответствующего комплекса программ для проведения вычислительных экспериментов.

Научная новизна

В данной диссертационной работе разработаны:

- метод повышения качества изображений магнитно-резонансной томографии,
- масштабно-инвариантный метод подавления осцилляций Гиббса и уменьшения аддитивного гауссовского шума на изображениях магнитно-резонансной томографии,
- проекционный метод сегментации опухолей на ультразвуковых, гистологических и колоноскопических изображениях,
- метод выбора масштаба классификации полнослайдовых изображений.

Теоретическая и практическая ценность

Комбинирование классических методов математического моделирования и современных методов машинного обучения расширяет теоретическую базу для создания более эффективных гибридных методов. Разработка новых гибридных методов способствует интеграции подходов из разных областей и позволяет находить баланс между точностью и сложностью модели, например, применяя нейросетевые аппроксимации трудоёмких классических моделей. Предлагаемые гибридные методы более интерпретируемы и устойчивы, чем чёрные ящики глубокого обучения, что важно для врачебного принятия решений.

Разработанные в диссертационной работе методы обработки и анализа медицинских изображений могут применяться как независимо при проведении медицинских исследований, так и могут быть реализованы в виде связных модулей системы медицинской компьютерной диагностики.

Применение разработанных методов обработки и анализа медицинских изображений имеет существенную практическую значимость. Внедрение разработанных решений повышения качества медицинских данных способно привести к уменьшению количества повторных исследований. Улучшение качества визуализации способствует более точному обнаружению патологий, включая

ранние стадии опухолей. Сокращение времени диагностики ускоряет постановку диагноза и начало лечения, что критично для онкологических и острых состояний, а снижение нагрузки на врачей за счёт автоматизации рутинных задач (например, разметки полипов на колоноскопических снимках) позволяет перераспределить ресурсы на сложные случаи. Таким образом, результаты могут быть востребованы в клиниках и диагностических центрах, в разработке медицинского ПО, а также в научных исследованиях.

Степень разработанности темы

Проблема анализа и обработки медицинских изображений является междисциплинарной и находится на стыке математики, компьютерных наук и медицины. Степень её разработанности характеризуется наличием двух крупных исторически сложившихся направлений: классических методов, основанных на математическом моделировании и численных методах, и методов машинного (глубокого) обучения. В последние годы наблюдается устойчивая тенденция к их конвергенции в рамках гибридных подходов.

Классические методы математического моделирования и численные методы обработки изображений, такие как методы фильтрации, вейвлет-преобразования, вариационные методы и методы решения обратных задач, имеют солидную теоретическую базу и хорошо изучены. Эти методы отличаются высокой интерпретируемостью, предсказуемостью поведения и относительно низкими вычислительными затратами. Однако, как отмечено в актуальности, их эффективность часто ограничивается специфическими типами искажений и необходимостью ручной настройки параметров для каждого конкретного случая, что затрудняет их применение для обработки разнородных медицинских данных с комплексными артефактами.

В настоящее время большую популярность приобрели методы глубокого обучения в области математических методов обработки изображений и медицинской визуализации. Так, свёрточные нейронные сети, в частности архитектуры U-Net, SegNet, DeepLab и их модификации, продемонстрировали выдающиеся результаты в задачах сегментации, классификации и повышения качества изображений. Более поздние архитектуры, такие как трансформеры и диффузионные сети, расширили возможности подхода. Исследования в этой области показывают, что данные методы обладают высокой адаптивностью и способностью извлекать сложные признаки непосредственно из данных. Тем не менее, им присущи недостатки: они часто рассматриваются как чёрные ящики, требуют больших объёмов размеченных данных и значительных вычислительных ресурсов, а также могут быть неустойчивы к изменению распределения входных данных.

Таким образом, в научном сообществе сформировалось направление гибридных методов, которое активно развивается в последнее десятилетие. Данное направление можно условно разделить на несколько потоков.

Первый поток предполагает использование классических методов для предобработки данных для нейронных сетей. Этот подход хорошо изучен, но часто носит характер простой композиции, без глубокой интеграции методов.

Более глубокий уровень синтеза демонстрирует второй поток, связанный с прямым встраиванием знаний предметной области (англ. domain knowledge) в архитектуры нейронных сетей. Сюда относятся такие техники, как применение физически информированных нейронных сетей и интеграция известных преобразований (например, Фурье) в виде специальных слоёв.

Третий поток направлен на использование нейронных сетей для аппроксимации трудоёмких этапов классических алгоритмов. Этот подход позволяет сохранить интерпретируемость классической модели, значительно ускорив её работу. Исследования в этой области, например, работы по нейронным операторам, показывают перспективность интеграции, но их применение к анализу и повышению качества медицинских изображений остаётся недостаточно развитым.

Проведённый анализ степени разработанности темы подтверждает актуальность и научную новизну диссертационного исследования. Существующий зазор между классическими методами математического моделирования и современными методами глубокого обучения создаёт пространство для разработки новых, более эффективных гибридных подходов, предлагаемых в данной работе. Комбинация строгости математического моделирования и адаптивности машинного обучения позволяет сформулировать новые научные решения для актуальных задач медицинской диагностики.

Методология и методы исследования

Методологической основой исследования являются методы математического моделирования. Практическая реализация включает вычислительные эксперименты с использованием искусственных и реальных данных в рамках задач машинного обучения и анализа изображений. Для решения поставленных задач применяется гибридный подход, интегрирующий классические математические модели с архитектурами глубокого обучения.

Степень достоверности результатов

Достоверность результатов проведённых исследований обеспечивается опорой на теоретическую базу, воспроизводимыми вычислительными экспериментами и тестированием алгоритмов на искусственных и реальных данных.

Апробация работы

Основные результаты работы докладывались на:

1. 32-ой международной конференции по компьютерной графике и зрению «ГрафиКон'2022», (Москва, 2022);
2. 8-ой международной конференции по биомедицинской визуализации и обработке сигналов «ICBSP'2023», (Сингапур, 2023);
3. 34-ой международной конференции по компьютерной графике и зрению «ГрафиКон'2024», (Омск, 2024);

4. 9-ой международной конференции по биомедицинской визуализации и обработке сигналов «ICBSP'2024», (Гонконг, 2024);
5. 1-ой международной школе-конференции по тензорным методам в математике и задачах искусственного интеллекта «SMBU'2024», (Шэньчжэнь, 2024).

Публикации

По теме исследования опубликовано 7 работ, из них 3 работы в рецензируемых научных изданиях, рекомендованных для защиты в докторской совете МГУ по специальности и отрасли наук. Список опубликованных работ приведён в конце докторской работы.

Личный вклад

Все результаты работы получены автором лично под научным руководством д.ф.-м.н., проф. А.С. Крылова. В работах, написанных в соавторстве, вклад автора докторской работы в результаты является определяющим.

Основные положения, выносимые на защиту

1. Метод, объединяющий классический математический метод поиска оптимальных субпиксельных сдвигов и свёрточную нейронную сеть, позволил осуществить подавление осцилляций Гиббса на изображениях магнитно-резонансной томографии головного мозга.
2. Нейронный оператор Фурье, обученный на результатах численных расчётов классическими математическими методами, показал эффективность для масштабно-инвариантного уменьшения осцилляций Гиббса и аддитивного гауссовского шума на изображениях магнитно-резонансной томографии головного мозга.
3. Проекционные сети Колмогорова-Арнольда, используемые в качестве основы метода сегментации, позволили получить устойчивый метод сегментации опухолей на ультразвуковых изображениях молочной железы, гистологических изображениях слизистых желёз и колоноскопических изображениях.
4. Автоматический метод выбора масштаба гистологических полнослайдовых изображений показал эффективность в классификации изображений стенок желудка предобученными нейронными сетями.
5. Разработанный программный комплекс по обучению гибридных методов обработки и анализа медицинских изображений ориентирован на проведение вычислительных экспериментов.

Структура и объём работы

Докторская работа состоит из введения, четырёх глав, заключения, списков литературы, публикаций автора по теме исследования, рисунков и таблиц. Полный объём докторской работы составляет 103 страницы, включая 32 рисунка и 6 таблиц. Список литературы содержит 111 наименований.

Основное содержание работы

В **введении** обосновывается актуальность работы, изложены её цель, научная новизна и практическая ценность, даны основные характеристики работы, сформулированы положения, выносимые на защиту, личный вклад автора, представлен отчёт об апробации работы и публикациях, содержащих основные результаты.

В **первой** главе рассматривается проблема повышения качества изображений магнитно-резонансной томографии головного мозга. Магнитно-резонансная томография формирует изображение $I(\mathbf{x})$ через обратное преобразование Фурье k -пространства $K(\mathbf{k})$:

$$I(\mathbf{x}) = \mathcal{F}^{-1}(K(\mathbf{k})), \quad (1)$$

где $\mathbf{x} = (x, y)$ и $\mathbf{k} = (k_x, k_y)$. В практических условиях измеряется лишь конечная область k -пространства, то есть истинный спектр $K(\mathbf{k})$ подвергается усечению оператором обрезки высоких частот:

$$\mathcal{R} = \begin{cases} 1, & |k_x| \leq r_{max}, |k_y| \leq r_{max}, \\ 0, & \text{иначе.} \end{cases}$$

Тогда математическая модель реконструированного изображения $I^0(\mathbf{x})$, согласно теореме о свёртке для обратного преобразования Фурье, имеет вид:

$$I^0(\mathbf{x}) = \mathcal{F}^{-1}(\mathcal{R} \cdot K(\mathbf{k})). \quad (2)$$

Таким образом, естественной моделью искажений на изображениях МРТ являются осцилляции Гиббса:

$$I^0(\mathbf{x}) = I(\mathbf{x}) + \varepsilon_G(\mathbf{x}), \quad (3)$$

где осцилляционный компонент $\varepsilon_G(\mathbf{x})$ представляет собой колебательные добавки, локализованные в областях больших градиентов, например, на границах тканей мозга. Осцилляции Гиббса не исчезают при увеличении числа гармоник и концентрируются в окрестностях контуров на изображениях.

В главе предложен гибридный метод для подавления осцилляций Гиббса. Метод относится к классу вычислительных методов реконструкции изображений и объединяет классический поиск оптимальных субпиксельных сдвигов с обучаемой свёрточной нейронной сетью. Такой подход позволяет получить решение задачи устранения артефактов при существенно меньшей вычислительной сложности по сравнению с существующими глубокими нейронными сетями.

Разработанная гибридная архитектура DGAS9-CNN содержит два параллельных канала кодирования признаков изображений: I^0 – реконструкции по усечённому Фурье-спектру и $I^{K\uparrow}$ – интерполированного результата классического метода Кельнера. Оба канала описаны последовательностью свёрточных

операторов, вводящих отображения вида:

$$\mathbf{x}_l = \mathbf{x}_{l-1} + \Phi_l(\mathbf{x}_{l-1}), l = 1, 2, \dots, L, \quad (4)$$

где L – число остаточных блоков, а $\Phi_l(\mathbf{x}_{l-1}) = \sigma(\mathbf{W}_l * \mathbf{x}_{l-1})$ – l -й, свёрточный блок, обрабатывающий признаковое описание \mathbf{x}_{l-1} . Признаковые описания агрегируются в модуле декодирования посредством свёртки предварительно конкатенированных тензоров. Архитектура сети завершается модулем реконструкции, который преобразует ставшее результатом фильтрации признаковое описание в изображение целевого цветового пространства – оттенки серого (англ. greyscale).

Предлагаемый метод обучен с учителем на сформированном наборе данных с синтетическими артефактами Гиббса, полученными путём обрезки 1/9 центральной части Фурье-спектра. В ходе обучения была использована следующая функция потерь:

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N \left\| \mathcal{F}(I_i^0; \theta) - I_i^1 \right\|_1 + \lambda \cdot \|\theta\|_2^2 \rightarrow \min_{\theta \in \Theta}, \quad (5)$$

где I_i^0 – входное изображение с осцилляциями Гиббса; I_i^1 – референсное изображение без артефактов; \mathcal{F} – гибридный метод с параметрами θ ; λ – коэффициент регуляризации (10^{-4}). Обучение было проведено на графическом процессоре с применением пакетной обработки данных и полиномиальной схемы уменьшения темпа обучения:

$$\eta_k = \eta_0 \cdot \left(1 - \frac{k}{K_{\max}}\right)^{\gamma}, \quad (6)$$

где η_k – темп обучения на итерации k ($\eta_0 = 10^{-4}$); K_{\max} — максимальное количество итераций (1000); $\gamma > 0$ — показатель степени (0.3), определяющий скорость уменьшения темпа обучения.

Разработанный гибридный метод сравнивается с глубокой свёрточной нейронной сетью GAS-CNN и классическим методом Кельнера. Алгоритм Кельнера основан на поиске оптимальных субпиксельных сдвигов с целью минимизации функционала, оценивающего величину осцилляций. В качестве такого функционала выступает полная вариация.

Проведённое исследование демонстрирует, что гибридный метод превосходит по качеству реконструкции алгоритм Кельнера и базовую нейронную сеть GAS-CNN. Согласно Таблице 1, метод DGAS9-CNN достигает наивысшего значения PSNR (29.57 dB) и величины полной вариации, наиболее близкой к референсным изображениям, при этом обеспечивая время выполнения на центральном процессоре Intel(R) Core(TM) i7-8700 на 30–40 % меньшее, чем у GAS-CNN. Результаты первой главы рекомендуют метод DGAS9-CNN как эффективный вычислительный инструмент для подавления осцилляций Гиббса на изображениях магнитно-резонансной томографии головного мозга (см. Рис. 1).

Таблица 1 — Средние значения PSNR, полной вариации и времени выполнения методов улучшения качества изображений МРТ головного мозга на тестовом наборе данных из 2617 изображений.

Метод	PSNR (дБ) \uparrow	TV	Время (с) \downarrow
I^0	20.70	659.68	—
I^1	—	601.75	—
Алгоритм Кельнера	21.68	540.18	0.23
GAS-CNN	29.46	620.09	1.03
DGAS9-CNN	29.57	611.25	0.68

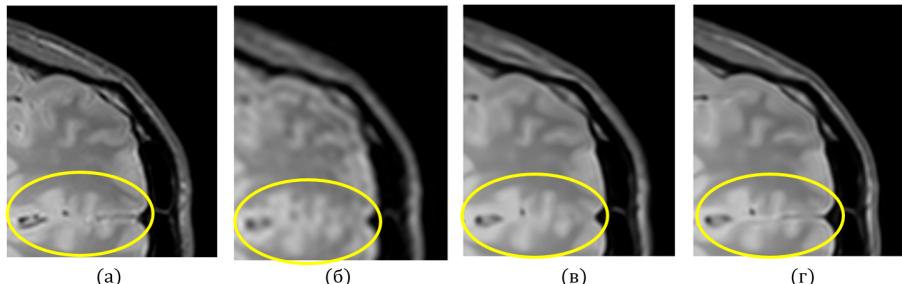


Рисунок 1 — Сравнение разработанного гибридного метода DGAS9-CNN улучшения качества изображений МРТ головного мозга с нейронной сетью GAS-CNN. (а) – референсное изображение без артефактов, (б) – изображение с осцилляциями Гиббса, (в) – результат нейронной сети GAS-CNN, (г) – результат гибридного метода DGAS9-CNN.

В **второй** главе исследуются методы повышения качества медицинских изображений и автоматической семантической сегментации патологических областей на основе нейронных операторов.

Первая часть главы посвящена применению нейронного оператора Фурье (англ. Fourier neural operator, FNO) для решения задачи повышения качества изображений магнитно-резонансной томографии головного мозга. Теоретической основой архитектуры FNO является интегральное представление решения краевых задач, используя функцию Грина. Предлагаемый в главе метод заключается в обучении нейронного оператора Фурье аппроксимации решения уравнения анизотропной диффузии, использованном в модели Перона-Малика для подавления шума на изображениях:

$$\frac{\partial I}{\partial t} = \operatorname{div}(c(\mathbf{x}, t) \cdot \nabla I), \quad (7)$$

$$I(\mathbf{x}, 0) = I^0(\mathbf{x}), \frac{\partial I(\mathbf{x}, t)}{\partial n} \Big|_{\mathbf{x} \in \partial \Omega} = 0, \quad (8)$$

где $I(\mathbf{x}, t)$ – изображение на итерации t ; $c(\mathbf{x}, t)$ – коэффициент диффузии; ∇I – градиент изображения; div – оператор дивергенции; $\partial\Omega$ – контуры изображения; n – нормаль.

Коэффициент диффузии выбирается так, чтобы уменьшать диффузию в областях с большим градиентом:

$$c(|\nabla I|) = e^{-(\frac{|\nabla I|}{k})^2}, \quad (9)$$

где k – параметр, управляющий чувствительностью к контурам.

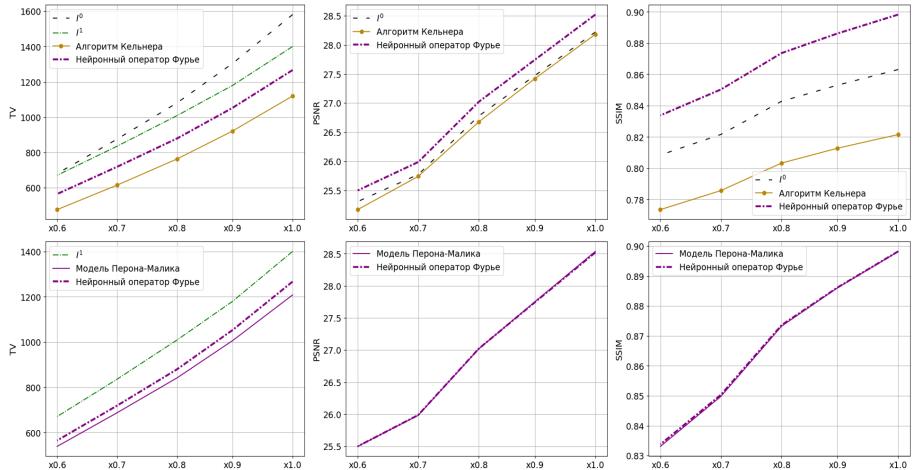


Рисунок 2 — Средние значения TV, PSNR, SSIM методов улучшения качества изображений МРТ головного мозга по тестовому набору данных IXI из 2617 изображений. Оценка устойчивости методов к изменению входного разрешения и регуляризующей способности нейронного оператора Фурье, обученного аппроксимации модели Перона-Малика на масштабе $\times 1.0$. Масштаб $\times 1.0$ соответствует равномерной пиксельной сетке 145×145 . Остальные масштабы получены линейной интерполяцией.

Эксперименты на наборе данных IXI показали (см. Рис. 2), что обученный нейронный оператор Фурье на результатах численных расчётов трёх итераций дифференциальной модели анизотропной диффузии, с временным шагом $\Delta t = 0.1$ и параметром чувствительности к контурам $k = 0.1$, превосходит по качеству реконструкции классический алгоритм Кельнера. Важнейшим результатом является демонстрация свойства масштабно-инвариантности метода на основе нейронного оператора Фурье. Благодаря глобальной обработке информации в частотной области, метод способен корректно обрабатывать изображения с пространственными характеристиками, не представленными в обучающей выборке. Это означает, что оператор обобщается не только на новые данные из того же

распределения, но и на данные, чей масштаб или разрешение отличаются от тех, что использовались при обучении. Данное свойство критически важно для практического применения метода в клинических условиях, где изображения могут быть получены на сканерах с различными протоколами и разрешениями.

Несмотря на то что пиковое отношение сигнала к шуму (PSNR), рассчитываемое на основе среднеквадратической ошибки и выражаемое в децибелах, является стандартной количественной метрикой различия изображений, его применение ограничено слабой корреляцией с воспринимаемым человеком качеством. В связи с этим, для оценки результатов повышения качества магнитно-резонансных изображений показатель PSNR дополнен структурным индексом подобия (SSIM) и величиной полной вариации (TV). Контроль величины полной вариации, в свою очередь, позволяет количественно оценивать амплитуду артефактов Гиббса и минимизировать риск избыточного сглаживания диагностически значимых элементов.

Во второй части главы разработан новый гибридный метод сегментации на основе обобщения сетей Колмогорова-Арнольда на задачи анализа и обработки изображений. Несмотря на аппроксимационные свойства сетей Колмогорова-Арнольда в задачах многомерной аппроксимации, их прямое применение к обработке изображений имеет фундаментальные ограничения. Ключевая проблема заключается в том, что сети Колмогорова-Арнольда рассматривают входные данные как скалярные величины, игнорируя двумерную геометрическую структуру, присущую визуальным данным.

В главе выдвинута и экспериментально подтверждена гипотеза о возможности представления непрерывных функционалов $f(\chi_1, \dots, \chi_n)$ на H^n в виде композиции линейных непрерывных функционалов, непрерывных функций одной переменной и операции сложения:

$$f(\chi_1, \dots, \chi_n) \rightsquigarrow \sum_j \zeta_j \left(\sum_i \varphi_{ji}(\chi_i) \right), \quad (10)$$

где $\chi_i \in H$; $\varphi_{ji} \in H^*$ и $\zeta_j: \mathbb{R} \rightarrow \mathbb{R}$. На основе этой гипотезы предлагается архитектура функциональной сети Колмогорова-Арнольда (*англ. functional Kolmogorov-Arnold network, FunKAN*), в которой двумерные карты признаков χ_i трактуются как элементы гильбертова пространства H , а их обработка в первом слое осуществляется проекционным методом по функциям Эрмита $\{\psi_k\}_{k=1}^r$ – собственным функциям интегрального оператора Фурье, согласно выражению:

$$\chi_{l,j} = \sum_{i=1}^n \omega_{l,j} \left(\sum_{k=1}^r \langle \varphi_l(\chi_{l-1,i}), \psi_{l,k}(\chi_{l-1,i}) \rangle \psi_{l,k}(\chi_{l-1,i}) \right), \quad (11)$$

где $\varphi_l \in \mathbb{R}^{n \times h \times w}$ – обучаемые внутренние функции; $\psi_l \in \mathbb{R}^{n \times r \times h \times w}$ – базисные функции Эрмита и $W_l \in \mathbb{R}^{1 \times 1 \times n \times m}$ – матрица свёрточных весов $\omega_{l,j}$.

Разработанный метод был применён как для повышения качества изображений МРТ, так и для семантической сегментации, будучи интегрированным в U-образную архитектуру (U-FunKAN).

Таблица 2 — Сравнение качества современных методов сегментации в трёх клинически различных медицинских сценариях. Результаты включают средние значения IoU и значения F1.

Метод	BUSI		GlaS		CVC	
	IoU ↑	F1 ↑	IoU ↑	F1 ↑	IoU ↑	F1 ↑
U-Net	57.22	71.91	86.66	92.79	83.79	91.06
Att-Unet	55.18	70.22	86.84	92.89	84.52	91.46
U-Net++	57.41	72.11	87.07	92.96	84.61	91.53
U-NeXt	59.06	73.08	84.51	91.55	74.83	85.36
Rolling-UNet	61.00	74.67	86.42	92.63	82.87	90.48
U-Mamba	61.81	75.55	87.01	93.02	84.79	91.63
UKAGNet	63.45	77.64	87.31	93.23	76.85	86.91
U-KAN	63.38	76.40	87.64	93.37	85.05	91.88
U-FunKAN	68.49	77.37	88.02	93.50	85.93	91.42

Таблица 3 — Сравнение вычислительной эффективности методов сегментации. Результаты получены с помощью инструмента THOP Python для профилирования моделей PyTorch.

Метод	Gflops ↓	Params (M) ↓
U-Net	524.2	34.53
Att-Unet	533.1	34.9
U-Net++	1109	36.6
U-NeXt	4.58	1.47
Rolling-UNet	16.82	1.78
U-Mamba	2087	86.3
U-KAN	14.02	6.35
U-FunKAN	4.35	3.6

Повышению качества изображений МРТ гибридный метод был обучен с учителем с функцией потерь среднеквадратической ошибки:

$$\mathcal{L}_{enh} = \frac{1}{N} \sum_{i=1}^N \|I_i^* - I_i^1\|_2^2, \quad (12)$$

где I^* – предсказываемые изображения, а I^1 – референсные изображения.

Сегментации гибридный метод был обучен также с учителем с помощью взвешенной комбинации бинарной перекрестной энтропии и коэффициента Сёренсена:

$$\mathcal{L}_{segm} = \frac{1}{N} \sum_{i=1}^N 0.1 \cdot CE(I_i^*, I_i^1) + Dice(I_i^*, I_i^1), \quad (13)$$

где I^* – предсказываемые изображения, а I^1 – референсные изображения.

Результаты показали (см. Таблицу 2), что U-FunKAN устанавливает новый уровень точности по метрике IoU на всех трёх наборах данных: BUSI (ультразвук молочной железы), GlaS (гистологические срезы) и CVC-ClinicDB (колоноскопические изображения).

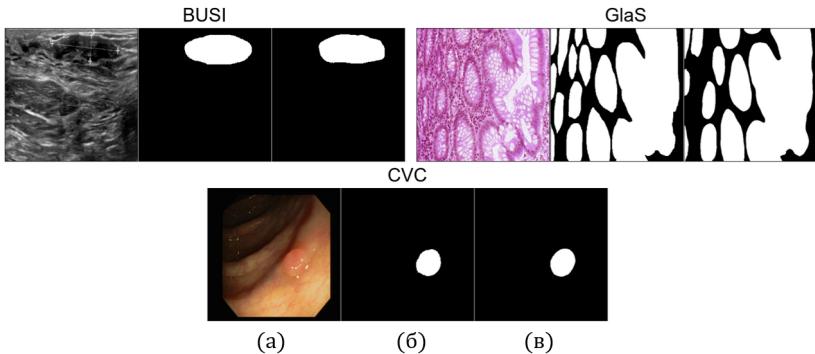


Рисунок 3 — Качественные результаты работы U-FunKAN для трёх разнородных медицинских сценариев. (а) – входные изображения, (б) – референсные маски, (в) – результаты метода.

Наиболее значительный прирост точности продемонстрирован на наборе данных BUSI. Значение метрики IoU для метода U-FunKAN составило 68.49 %, что на 5.04–13.31 процентных пункта выше результатов всех сравниваемых методов. По метрике F1, учитывающей баланс между точностью и полнотой, U-FunKAN показал результат, сравнимый с наилучшим 77.64 %, и уверенно пре-взошёл классические и современные аналоги.

На наборе данных GlaS предложенный метод также показал наивысшую точность, достигнув значения IoU, равного 88.02 %, и значения F1, равного 93.50 %. Преимущество над ближайшими конкурентами U-KAN и UKAGNet является значимым, хотя и менее выраженным, чем на BUSI, что указывает на высокую эффективность сравниваемых методов на данном типе гистологических изображений.

На наборе данных CVC-ClinicDB наблюдается схожая картина: U-FunKAN демонстрирует наилучший показатель IoU, равный 85.93 %, опережая U-KAN на 0.88 процентных пункта и U-Mamba на 1.14 процентных пункта. По метрике F1 результат U-FunKAN оказался несколько ниже наилучшего значения, показанного методом U-KAN, однако разница находится в пределах возможной статистической погрешности.

Более того, U-FunKAN демонстрирует существенно более высокую вычислительную эффективность (см. Таблицу 3), требуя в 3 раза меньше операций с плавающей точкой и на 43 % меньше обучаемых параметров по сравнению с указанными аналогами. Качественные результаты сегментации, полученные предложенным проекционным методом представлены на Рис. 3. Дополнительным преимуществом является интерпретируемость гибридного метода (см. Рис. 4), обеспечиваемая возможностью анализа спектрального распределения энергии признаков функций Эрмита. Точность, вычислительная экономичность и интерпретируемость делают метод ценным для задач медицинской диагностики.

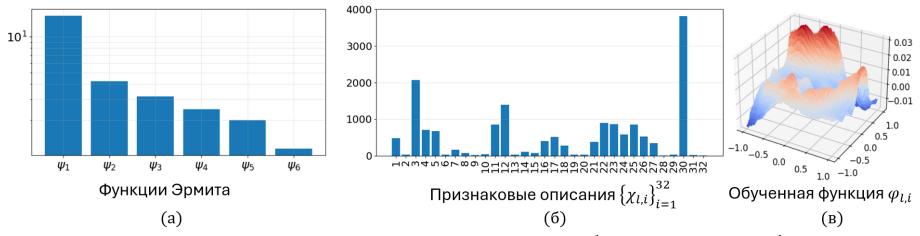


Рисунок 4 — Интерпретируемость слоя FunKAN, обученного на наборе данных IXI для улучшения качества изображений МРТ. (а) – агрегированный вклад базисных функций Эрмита по всем обученным внутренним функциям слоя, (б) – характеристика важности признаков, вычисленная как спектральная энергия, (в) – визуализация одной из обученных внутренних функций.

В третьей главе рассматривается проблема в области компьютерного анализа гистологических изображений, а именно проблема низкой устойчивости предобученных свёрточных нейронных сетей к изменению масштаба (оптического увеличения) входных данных. Широкое внедрение технологии полнослайдовых изображений (англ. whole slide imaging, WSI) в цифровую патологию, пришедшей на смену световой микроскопии, создаёт потребность в эффективных методах автоматизированной обработки. Однако эффективность нейросетевых классификаторов, таких как DenseNet121, ResNet50, VGG19, критически зависит от соответствия масштаба анализируемого изображения тому масштабу, на котором метод был обучен (см. Рис. 5). Эмпирически показано (см.

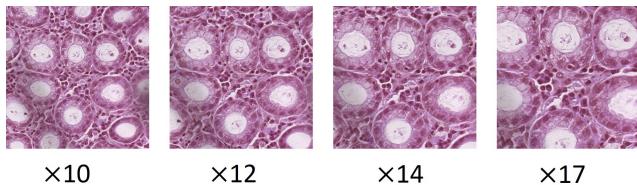


Рисунок 5 — Участок полнослайдового гистологического изображения в различных оптических увеличениях (масштабах).

Таблицу 4), что отклонение от оптимального масштаба приводит к значительному падению точности нейросетевой классификации (на 17 % в целом и на 67 % для отдельных классов тканей). Поэтому, задача автоматического выбора масштаба анализа для заданной предобученной нейронной сети является ключевой для построения систем компьютерной диагностики на основе WSI.

В главе предложен метод выбора масштаба обработки полнослайдового изображения на основе геометрического смысла нейронного классификатора. Метод основан на интерпретации последнего линейного слоя классификатора как набора разделяющих гиперплоскостей в признаковом пространстве.

Таблица 4 — Точность (%) классификации фрагментов (224×224) полнослайдового изображения стенки желудка в зависимости от выбранного масштаба входных данных. Acc – общая точность, Acc@AT – точность распознавания участков подслизистой основы, Acc@BG – точность распознавания фона, Acc@LP – точность распознавания неизменённых участков собственной пластиинки слизистой оболочки, Acc@MM – точность распознавания неизменённых участков мышечной пластиинки слизистой оболочки, Acc@TUM – точность распознавания участков adenокарциномы.

Масштаб	Acc ↑	Acc@AT ↑	Acc@BG ↑	Acc@LP ↑	Acc@MM ↑	Acc@TUM ↑
$\times 10$	73.01	87.80	99.84	61.74	17.47	98.19
$\times 20$	76.04	84.42	99.58	82.41	15.74	98.04
$\times 40$	59.04	78.85	99.46	15.01	7.35	94.52

Суть метода заключается в следующем:

1. Для анализируемого изображения случайным образом выбирается множество точек.
2. Для каждой точки извлекается пирамида фрагментов на различных масштабах из заданного диапазона.
3. Каждый многомасштабный фрагмент пропускается через сеть, и для него вычисляется расстояние в признаковом пространстве до каждой из разделяющих гиперплоскостей.
4. Для каждого фрагмента выбирается масштаб, на котором достигается максимальное расстояние до гиперплоскости своего (референсного) класса.
5. Итоговый масштаб для всего слайда определяется как мода (наиболее частый выбор) по всем проанализированным точкам.

Алгоритм не требует априорной информации об оптимальном масштабе или исходном увеличении сканирования и использует только внутренние параметры самой нейронной сети.

Предлагаемый метод экспериментально проверен на наборе данных PATH-DT-MSU, содержащем аннотированные полнослайдовые изображения стенок желудка. В качестве тестовой нейронной сети использовалась предобученная свёрточная сеть DenseNet121, обученная классификации фрагментов на 5 классов: подслизистая основа и мышечные слои (AT), фон (BG), собственная пластиинка слизистой (LP), мышечная пластиинка слизистой (MM) и adenокарцинома (TUM). Набор масштабов для поиска включал 9 уровней: $\{\times 10, \dots, \times 43\}$.

Результаты подтверждают эффективность метода. В частности (см. Рис. 6), для класса LP более 85 % прогнозов метода попадает в интервал: $\{\times 12, \dots, \times 25\}$, – в окрестности истинного масштаба $\times 20$, на котором метод был обучен.

В четвёртой главе рассматривается проблема обеспечения эффективности, воспроизводимости и управляемости вычислительных экспериментов при разработке комплекса программ, реализующих гибридные методы анализа и повышения качества медицинских изображений. Рост масштабов исследований в области искусственного интеллекта выявил ряд сложностей на стыке теории и

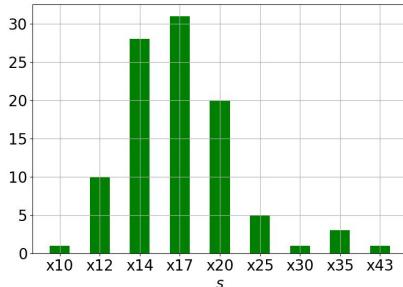


Рисунок 6 — Гистограмма по множеству предложенных методом масштабов для 100 случайных фрагментов класса LP.

программной реализации, решение которых является необходимым условием для достижения научной строгости и практической эффективности. Действительно, отсутствие единой системы конфигурации параметров эксперимента приводит к трудностям в точном воспроизведении полученных результатов, ставя под сомнение их достоверность и затрудняя накопление научного знания. А экспоненциальный рост числа гиперпараметров, архитектурных решений и вариантов предобработки данных делает непрактичным и подверженным ошибкам управление экспериментами через традиционные интерфейсы командной строки.

В связи с этим, в рамках главы были поставлены и решены следующие задачи:

1. Разработать модульную архитектуру программного комплекса, основанную на принципах разделения ответственности и инверсии управления, для гибкой интеграции новых методов.
2. Внедрить систему автоматизированного контроля качества кода, позволяющую поддерживать соответствие стандартам программирования в условиях активной разработки.
3. Создать гибридную систему конфигурации вычислительных экспериментов, сочетающую преимущества декларативного описания параметров и гибкости командной строки, для гарантии их полной воспроизводимости и управляемости.
4. Апробировать разработанный программный комплекс на реальных задачах анализа и обработки изображений для подтверждения его практической эффективности.

Разработанный программный комплекс написан на Python 3 и построен на принципах SOLID с использованием фреймворка PyTorch Lightning. Для повышения гибкости и сопровождаемости кода разработан механизм внедрения зависимостей, позволяющий динамически конфигурировать компоненты системы через внешние конфигурационные файлы, упрощая проведение экспериментов.

Для обеспечения единого качества кода программного комплекса внедрена система статического анализа на основе современного инструмента Ruff.

Предложена и реализована гибридная система конфигурации вычислительных экспериментов, которая сочетает достоинства декларативного подхода (конфигурационные файлы) и императивного подхода (интерфейс командной строки для часто изменяемых параметров). Внедрение модульного подхода к организации конфигурационных файлов с помощью директивы `#include` позволило создать переиспользуемую иерархическую структуру. Для работы с этой структурой была реализована фабрика объектов, выполняющая десериализацию компонентов системы непосредственно из их декларативного описания.

Разработанный программный комплекс был апробирован в серии вычислительных экспериментов по обучению гибридных методов анализа и обработки медицинских изображений различных типов, что подтвердило его эффективность и масштабируемость.

Заключение

В **заключении** приведены основные результаты работы, которые заключаются в следующем:

1. Разработан гибридный метод подавления осцилляций Гиббса на изображениях МРТ головного мозга.
2. Предложен масштабно-инвариантный метод повышения качества изображений МРТ.
3. Создан проекционный метод сегментации опухолей на ультразвуковых изображениях молочной железы, гистологических и колоноскопических изображениях с использованием сетей Колмогорова-Арнольда.
4. Разработан метод автоматического выбора масштаба классификации полнослайдовых гистологических изображений.
5. Реализован комплекс проблемно-ориентированных программных средств, обеспечивающий воспроизводимость вычислительных экспериментов.

Публикации автора по теме диссертации

Научные статьи, опубликованные в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ по специальности и отрасли наук

1. Hybrid Method for Gibbs-ringing Artifact Suppression in Magnetic Resonance Images / M.A. Penkin, A.S. Krylov, A.V. Khvostikov // Programming and Computer Software — 2021. — vol. 47. No. 3. — pp. 207–214. EDN: EUGNNS. Импакт-фактор 0,212 (SJR) [0.35 / 0.46] п.л.

Автором разработан гибридный метод уменьшения осцилляций Гиббса на изображениях магнитно-резонансной томографии головного мозга. Выполнена программная реализация, обучение и тестирование предлагаемого метода на наборе данных IXI.

2. Automated Method for Optimum Scale Search when using Trained Models for Histological Image Analysis / M.A. Penkin, A.V. Khvostikov, A.S. Krylov // Programming and Computer Software — 2023. — vol. 49. No. 3. — pp. 172-177. EDN: QRCFPG. Импакт-фактор 0,212 (SJR) [0.2 / 0.35] п.л.

Автором предложен численный метод определения масштаба обработки и анализа полнослайдовых гистологических изображений предобученными свёрточными нейронными сетями. Выполнена программная реализация и тестирование алгоритма на гистологическом наборе данных стенок желудка PATH-DT-MSU.

3. Adaptive Method for Selecting Basis Functions in Kolmogorov–Arnold Networks for Magnetic Resonance Image Enhancement / M.A. Penkin, A.S. Krylov // Programming and Computer Software — 2025. — vol. 51. No. 3. — pp. 167–172. EDN: GLUUTO. Импакт-фактор 0,212 (SJR) [0.3 / 0.4] п.л.

Автором разработан масштабно-инвариантный метод подавления осцилляций Гиббса и аддитивного гауссовского шума на медицинских изображениях с использованием предложенной схемы адаптивного выбора базисных функций для аппроксимации функций Колмогорова-Арнольда в нейронных сетях Колмогорова-Арнольда. Выполнена программная реализация, обучение и тестирование метода.

Иные публикации

4. How to Choose Adaptively Parameters of Image Denoising Methods? / A. Krylov, M. Penkin, N. Mamaev, A. Khvostikov // 2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA) — 2019. — pp. 1–6. DOI: 10.1109/IPTA.2019.8936123 (WoS, Scopus) [0.1 / 0.35] п.л.
5. Attention-based Convolutional Neural Network for MRI Gibbs-ringing Artifact Suppression / M. Penkin, A. Krylov, A. Khvostikov // CEUR Workshop Proceedings — 2020. — vol. 2744. — pp. 1–12. DOI: 10.51130/graphicon-2020-2-3-34 Импакт-фактор 0,166 (SJR) [0.5 / 0.7] п.л.
6. Medical Image Joint Deringing and Denoising Using Fourier Neural Operator / M. Penkin, A. Krylov // Proceedings of the 2023 8th International Conference on Biomedical Imaging, Signal Processing — 2023. — pp. 40–45. DOI: 10.1145/3634875.3634881 (ACM International Conference Proceeding Series, Scopus) Импакт-фактор 0,191 (SJR) [0.2 / 0.35] п.л.
7. Kolmogorov-Arnold Networks as Deep Feature Extractors for MRI Reconstruction / M. Penkin, A. Krylov // Proceedings of the 2024 9th International Conference on Biomedical Imaging, Signal Processing — 2024. — pp. 92–97. DOI: 10.1145/3707172.3707186 (ACM International Conference Proceeding Series, Scopus) Импакт-фактор 0,191 (SJR) [0.25 / 0.35] п.л.

Пенкин Максим Александрович

Гибридные методы анализа и повышения качества медицинских изображений

Автореф. дис. на соискание ученой степени канд. физ.-мат. наук

Подписано в печать _____._____._____. Заказ № _____

Формат 60×90/16. Усл. печ. л. 1. Тираж 100 экз.

Типография _____