

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

ДИССЕРТАЦИЯ
на соискание учёной степени
кандидата физико-математических наук

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

Москва — 2025

Оглавление

	Стр.
Введение	4
Глава 1. Методы повышения качества изображений	
магнитно-резонансной томографии	11
1.1 Классический математический метод поиска оптимальных субпиксельных сдвигов для подавления осцилляций Гиббса на изображениях головного мозга	16
1.2 Нейросетевой свёрточный метод повышения качества изображений головного мозга	18
1.3 Гибридный метод подавления осцилляций Гиббса на изображениях головного мозга	21
1.4 Эксперименты и результаты	25
1.5 Выводы	27
Глава 2. Нейронные операторы в обработке медицинских изображений	28
2.1 Нейронный оператор Фурье аппроксимации математических методов, используемых для повышения качества изображений магнитно-резонансной томографии на произвольном разрешении .	33
2.2 Проекционные сети Колмогорова-Арнольда для повышения качества изображений магнитно-резонансной томографии и сегментации ультразвуковых, гистологических и колоноскопических изображений	40
2.3 Эксперименты и результаты	50
2.4 Выводы	59
Глава 3. Методы выбора масштаба анализа гистологических	
изображений предобученными нейронными сетями	60
3.1 Нейронная классификация полнослайдовых изображений стенок желудка	63

3.2	Метод автоматического определения масштаба обработки полнослайдовых изображений на основе уверенности нейронного классификатора	66
3.3	Эксперименты и результаты	68
3.4	Выводы	70
 Глава 4. Программный комплекс гибридных методов анализа и повышения качества медицинских изображений		
4.1	Архитектура программного комплекса	74
4.2	Технология статического анализа программного кода	78
4.3	Система конфигурации вычислительных экспериментов	81
Заключение		85
Список литературы		86
Список рисунков		98
Список таблиц		103

Введение

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 надёжность 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 [1].

Более того, классические численные методы хорошо справляются лишь с определёнными типами искажений, часто требуя ручной настройки параметров [2] и не всегда адаптируясь к сложным случаям. Методы машинного обучения, напротив, демонстрируют высокую обобщающую способность и эффективность в задачах шумоподавления [3], повышения разрешения и улучшения контраста [4], но могут быть чрезмерно сложными или требовать больших вычислительных ресурсов [5], 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 [6]) и методов объяснимого ИИ (XAI) [7] 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.

Принципиально значимым является взаимное обогащение между классическим математическим аппаратом, включающим математические модели и численные методы, и технологиями искусственного интеллекта. Так, методы искусственного интеллекта не просто добавляются к традиционным подходам, но и сами развиваются под влиянием фундаментальной математики [8]. Классические методы, в свою очередь, вносят вклад в повышение надёжности решений, полученных с помощью нейронных сетей, а также стимулируют появление новых подходов в глубоком машинном обучении. Иллюстрацией такого синтеза служит эволюция многосеточных алгоритмов, концепции которых легли в основу построения иерархических архитектур нейронных сетей [9]. Это взаимодействие ярко проявляется в гибридных методах анализа и повышения качества медицинских изображений.

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 связанных 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 [10].

Классические методы математического моделирования и численные методы обработки изображений, такие как методы фильтрации [11], вейвлет-преобразования [12], вариационные методы [13] и методы решения обратных задач [14], 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.

Методы машинного и глубокого обучения совершили революцию в области компьютерного зрения, в том числе и в медицинской визуализации. Свёрточные нейронные сети, в частности архитектуры U-Net [15], SegNet [16], DeepLab [17] и их модификации, продемонстрировали выдающиеся результаты в задачах сегментации, классификации и повышения качества изображений. Более поздние архитектуры, такие как трансформеры [6] и диффузионные сети [18], 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 преобработки данных для нейронных сетей. Этот подход хорошо изучен, но часто носит характер простой композиции [19], 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 [20; 21] 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 [22; 23], 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 диссертационном совете МГУ по специальности и отрасли наук. Список опубликованных работ приведён в конце диссертационной 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 диссертационной работы в полученные 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

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

Глава 1. Методы повышения качества изображений магнитно-резонансной томографии

Качество изображений магнитно-резонансной томографии (МРТ) может снижаться из-за шумов, движения, низкого разрешения или недостаточной контрастности. Для улучшения их качества применяют классические методы математического моделирования, современные методы машинного обучения и гибридные подходы.

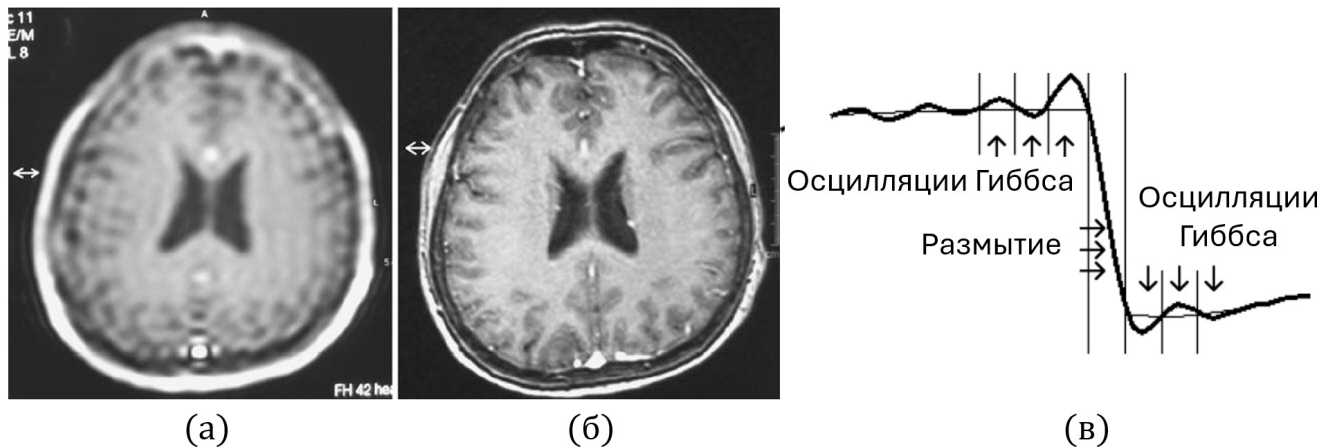


Рисунок 1.1 — Проявление артефактов Гиббса на изображениях магнитно-резонансной томографии. (а) – T1-взвешенное изображение МРТ мозга в низком разрешении (64×64) с осцилляциями Гиббса в виде ярких и тёмных линий, параллельных краям резких изменений интенсивности, (б) – T1-взвешенное изображение МРТ мозга в высоком разрешении (128×128), (в) – пример осцилляций Гиббса на ступенчатой функции, моделирующей резкий контур на изображении.

Типичным дефектом на изображениях магнитно-резонансной томографии являются осцилляции Гиббса [24] (см. Рис. 1.1). Осцилляции Гиббса представляют собой волны возле контрастных перепадов яркости на изображениях, например, на границах головного мозга и спинномозговой жидкости в желудочках мозга и субарахноидальном пространстве. Данный артефакт возникает вследствие усечения ряда Фурье при реконструкции изображения. Наличие этих осцилляций существенно затрудняет анализ тканей и снижает диагностическую ценность изображений.

В магнитно-резонансной томографии осцилляции Гиббса возникают из-за ограниченного числа высокочастотных компонент в разложении сигнала. В процессе МРТ-сканирования измеряются интегральные величины, определяющие

коэффициенты Фурье, таким образом, пространственное изображение формируется посредством применения обратного преобразования Фурье к полученным данным. В результате, получается изображение с ограниченным разрешением и характерными осцилляциями Гиббса возле резких контуров. Число осцилляций Гиббса является бесконечным, тем не менее, на реальных изображениях обычно заметно не более 1-2 осцилляций. Незначительное искажение изображений человек может не заметить, тогда как при больших потерях информации артефакт приводит к видимым окаймлениям (см. Рис. 1.1).

Феномен Гиббса, впервые описанный Генри Уилбрахамом в 1848 году и позднее переоткрытый Дж. Уиллардом Гиббсом в 1898 году, длительное время не привлекал значительного внимания математического сообщества. Интенсивное изучение связанных с ним осцилляций началось лишь в конце 1970-х годов [25]. Последующее увеличение числа публикаций по данной теме [26; 27] было во многом обусловлено развитием смежных технологических областей. В частности, импульс исследованиям могли придать достижения в области магнитно-резонансной томографии, методологическая основа которой была заложена П. Лотербургом в 1973 году и впоследствии усовершенствована П. Мэнсфилдом, что в 2003 году было отмечено Нобелевской премией. Примечательно, что как феномен Гиббса, так и метод МРТ характеризуются множественным авторством, а их признание сопровождалось дискуссиями о приоритете. Эта ситуация иллюстрирует распространённый в истории науки феномен одновременных и независимых открытий, который подробно анализировал социолог Роберт Мертон. Классическими примерами служат: создание математического анализа Ньютоном и Лейбницем (XVII в.), формулировка закона Бойля–Мариотта (XVII в.), открытие кислорода Шееле, Пристли и Лавуазье (XVIII в.), построение неевклидовой геометрии Лобачевским, Бойяи и Гауссом (XIX в.), а также разработка теории эволюции Дарвином и Уоллесом (XIX в.).

Для демонстрации причин осцилляций Гиббса с математической точки зрения рассмотрим следующую вещественную периодическую функцию:

$$\xi(t) = \begin{cases} a, & t \in [-\tau/2, \tau/2] \\ 0, & t \in [-T/2, T/2] \setminus [-\tau/2, \tau/2] \end{cases}, \quad (1.1)$$

$$\xi(t) = \xi(t + T),$$

где a – величина скачка; T – период функции; τ – ширина импульса.

Используя преобразование Фурье в комплексной форме и предполагая $T = 2\tau$, (1.1) может быть записано в виде:

$$\xi(t) = \sum_{k=-\infty}^{+\infty} d_k e^{i\omega_k t}, \quad (1.2)$$

где $d_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \xi(t) e^{-i\omega_k t} dt$, $\omega_k = \Omega k$, $\Omega = 2\pi/T$.

$$\begin{aligned} \xi(t) &= 2 \cdot \frac{a}{2} \sum_{k=0}^{+\infty} (-1)^k \cdot \frac{1}{(2k+1)\pi} \cdot \cos(2k+1) \cdot \frac{2\pi}{T} t \\ &= \frac{a}{\pi} (\cos \Omega t - \frac{1}{3} \cos 3\Omega t + \dots). \end{aligned} \quad (1.3)$$

На практике часто используется не полный Фурье-спектр сигнала, а лишь его часть для минимизации времени сбора данных, что, как было отмечено выше, приводит к осцилляциям Гиббса, характерным для аппроксимации тригонометрическим рядом функций с разрывами (см. Рис. 1.2). Важно заметить, что

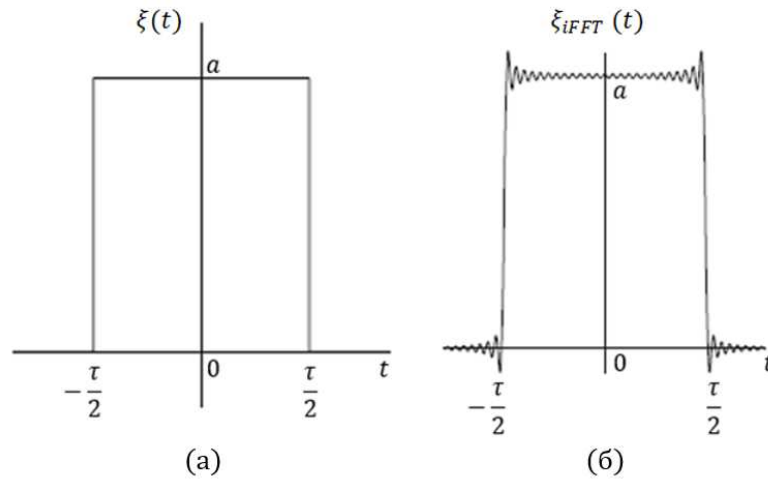


Рисунок 1.2 — Осцилляции Гиббса на примере конечного разрыва. (а) – ступенчатая функция $\xi(t)$, (б) – результат восстановления сигнала $\xi(t)$ по усечённому ряду Фурье, демонстрирующий осцилляции в окрестностях точек разрыва.

максимальная амплитуда осцилляций Гиббса постоянна и не зависит от выбранной частоты обрезки ряда [12].

Задачу уменьшения осцилляций Гиббса на изображениях можно решать, например, вариационным путём [28], оптимизируя целевой функционал (1.4) на выбранном компакте соответствующего функционального пространства.

$$J(I) = \frac{1}{2} \|I - I^0\|^2 + \lambda \int_D |\nabla I(\mathbf{x})| d\mathbf{x} \rightarrow \min_I, \quad (1.4)$$

где I^0 – входное изображение с осцилляциями Гиббса; I – искомое изображение среди изображений, принадлежащих выбранному компакт; D – область задания изображений; $\mathbf{x} = (x, y)$ – пространственные координаты; λ – параметр регуляризации.

В работе [29] предложен подход, основанный на теории разреженных представлений – математическом аппарате, который позволяет эффективно описывать сигналы с помощью ограниченного числа значимых коэффициентов в оптимально подобранном базовом наборе функций. Данная методика особенно эффективна для обработки медицинских изображений, где важные диагностические признаки часто могут быть представлены в виде небольшого числа информативных компонент.

Среди классических методов наиболее эффективным алгоритмом является алгоритм Кельнера [30], в котором осцилляции Гиббса уменьшаются методом поиска оптимальных субпиксельных сдвигов.

В настоящее время большую популярность приобрели методы глубокого обучения в области математических методов обработки медицинских изображений. Так, свёрточные нейронные сети успешно решают массу прикладных задач обработки медицинских изображений, таких как сегментация [15; 31–33], уменьшение уровня шума [34; 35], и, конечно, подавление осцилляций Гиббса на изображениях магнитно-резонансной томографии головного мозга [36]. Более того, свёрточные нейронные сети успешно осуществляют локальную обработку по построению, устойчивы к сдвигам по построению, эффективны на малых данных благодаря предобучению и интерпретируемы (например, с помощью алгоритма GradCAM [37]). К недостаткам свёрточных сетей относят низкую устойчивость к изменению масштаба входных данных [38] и фиксированный размер рецептивного поля, определяемый архитектурой сети.

Трансформерные сети приобрели взрывную популярность [39–41], главное преимущество которых заключается в наличии глобального контекста (*англ.* self-attention), отсутствии заикленности на локальных паттернах и в масштабируемости на изображения высокого разрешения. Тем не менее, трансформерные сети требуют огромного количества данных для обучения, в то время как медицинские наборы данных обычно достаточно ограничены в объёме и вариативности.

Улучшать качество результатов нейронных сетей можно различными способами. Нарращивание ансамбля слоёв является одним из часто используемых подходов в машинном обучении [42]. Однако, увеличивая число слоёв в ансамбле,

резко растёт и число параметров, требующих оптимизации, увеличивается трудоёмкость вычисления и оптимизации модели, более того, появляется значительный риск переобучения нейронной сети под тренировочную выборку. В качестве альтернативы выступает гибридный подход [43; 44].

Гибридные методы представляют сочетание методов глубокого обучения с классическими математическими моделями обработки изображений. Цель гибридных методов заключается в уменьшении числа обучаемых параметров, упрощении архитектуры, уменьшая тем самым риск переобучения, в сохранении обобщающей способности модели и в синергии и эмерджентности полученной системы. Известным примером гибридного метода является метод сегментации опухолей мозга [45] с помощью комбинации U-Net [15] и методов активных контуров [46]. Синергия гибридного алгоритма в данном случае достигается тем, что U-Net быстро находит область интереса (*англ.* region of interest, ROI), а метод активных контуров уточняет границы. Классическая модель компенсирует недостатки нейронной сети (неточности в мелких деталях), а нейронная сеть избавляется от ручной инициализации контура. Эмерджентность заключается в возникновении у системы способности точно сегментировать опухоли даже при слабых границах (например, при глиомах на МРТ), чего не может сделать ни одна модель по отдельности. В результате, повышается точность, уменьшается время обработки и повышается интерпретируемость, так как врачи видят не только *чёрный ящик* нейронной сети, но и физически обоснованную обработку.

В данной главе будет представлен разработанный гибридный метод подавления артефактов Гиббса на изображениях МРТ головного мозга, объединяющий классическую математическую модель Кельнера поиска оптимальных субпиксельных сдвигов и обучаемую свёрточную нейронную сеть.

1.1 Классический математический метод поиска оптимальных субпиксельных сдвигов для подавления осцилляций Гиббса на изображениях головного мозга

В современной обработке медицинских изображений особое место занимает неитеративный математический метод Кельнера [30], разработанная для подавления осцилляций Гиббса в том числе на изображениях магнитно-резонансной томографии головного мозга.

Алгоритм Кельнера основан на численном поиске оптимальных субпиксельных сдвигов с целью минимизации функционала, оценивающего величину осцилляций. В качестве такого функционала используется полная вариация [47], находящая своё применение в задачах регуляризации и восстановления изображений. В непрерывном случае полная вариация изображения $I(\mathbf{x})$, $\mathbf{x} = (x, y) \in D$ определяется как:

$$TV(I; D) = \int_D |\nabla I(\mathbf{x})| d\mathbf{x}. \quad (1.5)$$

Метод Кельнера характеризуется двумя ключевыми параметрами Δ_1 и Δ_2 , которые определяют размер окна для локальной оптимизации. В рамках данного подхода реализована двухэтапная процедура минимизации полной вариации: на первом этапе выполняется численная оптимизация субпиксельных сдвигов в ортогональных направлениях, на втором этапе осуществляется синтез полученных результатов по следующей формуле:

$$I^K = \mathcal{F}^{-1} \{ \mathcal{F}(I^{K_x}) \cdot G_x + \mathcal{F}(I^{K_y}) \cdot G_y \}, \quad (1.6)$$

где I^K – итоговый результат алгоритма Кельнера; I^{K_x} – результат одномерного алгоритма Кельнера вдоль оси OX ; I^{K_y} – результат одномерного алгоритма Кельнера вдоль оси OY ; \mathcal{F} – прямое дискретное преобразование Фурье; \mathcal{F}^{-1} – обратное дискретное преобразование Фурье; G_x, G_y – весовые функции:

$$G_x(t_x, t_y) = \frac{1 + \cos t_y}{(1 + \cos t_x) + (1 + \cos t_y)}, \quad (1.7)$$

$$G_y(t_x, t_y) = \frac{1 + \cos t_x}{(1 + \cos t_x) + (1 + \cos t_y)}, \quad (1.8)$$

где $t_x \in [-\pi, \pi]$, $t_y \in [-\pi, \pi]$.

Рассмотрим подробнее одномерный алгоритм Кельнера с математической точки зрения. Пусть I – исходный сигнал и C_k – k -я комплексная амплитуда. Тогда исходный сигнал можно представить, разложив его в дискретный ряд Фурье:

$$I(x) = \frac{1}{N} \sum_{k=0}^{N-1} C_k \cdot e^{i2\pi \frac{k}{N} x}. \quad (1.9)$$

С помощью теоремы о сдвиге получим $2M$ копий сигнала для дальнейшего выбора оптимальных субпиксельных сдвигов:

$$I_s(x) = \frac{1}{N} \sum_{k=0}^{N-1} (C_k \cdot e^{-i2\pi \frac{k}{N} \cdot \frac{s}{2M}}) \cdot e^{i2\pi \frac{k}{N} x}, \quad (1.10)$$

где $s \in [-M, M - 1], s \in \mathbb{Z}$.

Алгоритм Кельнера проводит поиск оптимальных сдвигов вдоль положительного и отрицательного направлений независимо:

$$TV^\pm(I_s) = \sum_{\delta=\Delta_1}^{\Delta_2} |I_s(x \pm \delta) - I_s(x \pm (\delta - 1))|, \quad (1.11)$$

$$r^\pm(x) = \arg \min_s TV^\pm(I_s). \quad (1.12)$$

Из сдвигов $r^+(x)$, $r^-(x)$ выбирается $r(x)$, обеспечивающий наименьшую полную вариацию.

В заключение, значения в соответствующих субпикселях $x - r(x)/(2M)$ вычисляются с помощью линейной интерполяции. Для большинства рассматриваемых прикладных случаев (см. Рис. 1.3) оптимальными параметрами метода являются: $\Delta_1 = 1, \Delta_2 = 3$.

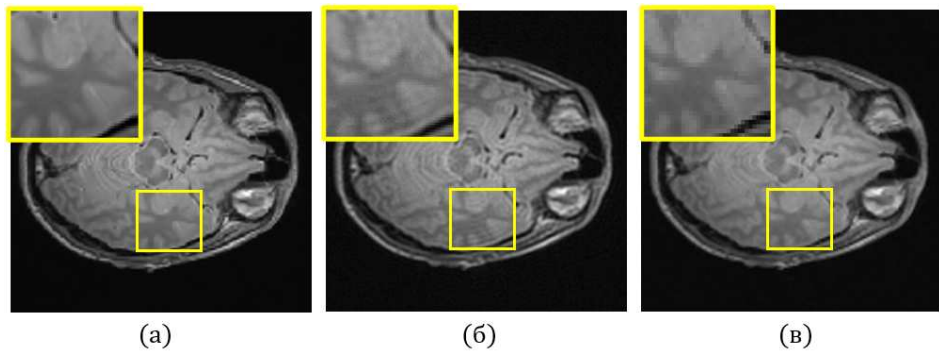


Рисунок 1.3 — Пример улучшения качества изображения МРТ головного мозга алгоритмом Кельнера. (а) – референсное изображение без артефактов, (б) – изображение с осцилляциями Гиббса, (в) – результат алгоритма Кельнера.

1.2 Нейросетевой свёрточный метод повышения качества изображений головного мозга

Одной из актуальных свёрточных нейронных сетей для подавления осцилляций Гиббса на изображениях МРТ головного мозга является нейронная сеть GAS-CNN (*англ.* Gibbs-ringing artifact suppression convolutional neural network) [36]. Она реализована в виде трёх последовательных модулей: модуль извлечения признаков, нелинейного отображения и реконструкции.

Основными отличительными особенностями GAS-CNN являются:

- отказ от слоёв уменьшения пространственной размерности, например, свёртки с шагом 2 или слоя субдискретизации с функцией максимума (*англ.* max pooling), в силу локальности осцилляций Гиббса,
- уменьшение размерности признакового пространства,
- добавление пространственных связей (*англ.* skip connection), подобных используемым в U-Net [15] между кодировщиком и декодировщиком.

С помощью модуля извлечения признаков GAS-CNN отображает входное изображение в признаковое пространство размерности 64, в котором затем осуществляется фильтрация с помощью 32-ух свёрточных остаточных блоков (*англ.* residual block) [48], составляющих модуль нелинейного отображения:

$$\mathbf{x}_l = \mathbf{x}_{l-1} + \Phi_l(\mathbf{x}_{l-1}), l = 1, 2, \dots, L, \quad (1.13)$$

где L – число остаточных блоков, а $\Phi_l(\mathbf{x}_{l-1}) = \max[\mathbf{x}_{l-1} * \mathbf{w}_{l,1} + \mathbf{b}_{l,1}, 0] * \mathbf{w}_{l,2} + \mathbf{b}_{l,2}$ – l -й свёрточный блок, обрабатывающий тензор признаков \mathbf{x}_{l-1} . Архитектура сети завершается модулем реконструкции, который преобразует признаковое описание в изображение целевого цветового пространства.

В качестве функции потерь используется взвешенная комбинация MAE (*англ.* mean average error) и SSIM (*англ.* structural similarity index measure) [49]:

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N (\|\mathcal{F}(I_i^0; \theta) - I_i^1\|_1 + \lambda \cdot SSIM(\mathcal{F}(I_i^0; \theta), I_i^1)) \rightarrow \min_{\theta \in \Theta}, \quad (1.14)$$

где I_i^0 – входное изображение с осцилляциями Гиббса; I_i^1 – референсное изображение без артефактов; \mathcal{F} – нейронная сеть GAS-CNN с совокупностью обучаемых параметров θ , оптимизация которых ведётся модифицированным численным методом стохастического градиентного спуска [50] Adam [51] ($\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$), по набору данных $\mathcal{D} = \{(I_i^0, I_i^1)\}_{i=1}^N$.

Визуально и численно, по метрикам PSNR и SSIM оценки качества изображений [52], оценить результаты применения обученной GAS-CNN для улучшения качества изображений MPT мозга можно на Рис. 1.4.

Пиковое отношение сигнала к шуму (*англ.* peak signal-to-noise ratio, PSNR) является показателем для количественной оценки качества реконструкции изображений. Данная величина широко применяется в задачах обработки медицинских изображений, включая восстановление, сжатие и подавление артефактов. PSNR определяется как:

$$PSNR = 10 \cdot \log_{10} \frac{MAX^2}{MSE}, \quad (1.15)$$

где MAX – максимально возможное значение пикселя; MSE – среднеквадратическая ошибка между референсным и восстановленным изображениями.

Индекс структурного сходства (*англ.* structural similarity index measure, SSIM) был разработан для преодоления ограничений PSNR через моделирование человеческого восприятия изображений. SSIM оценивает три компоненты: яркость (l), контраст (c) и структуру (s).

$$SSIM(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y), \quad (1.16)$$

$$\text{где } l(x, y) = \frac{2\mu_x\mu_y+c_1}{\mu_x^2+\mu_y^2+c_1}, \quad c(x, y) = \frac{2\sigma_x\sigma_y+c_2}{\sigma_x^2+\sigma_y^2+c_2}, \quad s(x, y) = \frac{\sigma_{xy}+c_3}{\sigma_x\sigma_y+c_3}.$$

Хотя PSNR, вычисляемый через среднеквадратическую ошибку [53] и выражаемый в децибелах, остаётся стандартным инструментом количественной оценки различий между изображениями, его главным ограничением является слабая корреляция с перцептивным качеством. Поэтому в задаче повышения качества изображений MPT этот показатель обычно дополняют индексом SSIM или величиной полной вариации TV. В отличие от PSNR, индекс SSIM демонстрирует лучшую согласованность с субъективной оценкой экспертов, что особенно важно при анализе медицинских изображений, где критическое значение имеет сохранение морфологических деталей и чёткости границ между тканями. Контроль полной вариации позволяет отслеживать амплитуду остаточных колебаний Гиббса и предотвращать избыточное сглаживание значимых деталей. Таким образом, комбинированный анализ PSNR, SSIM и TV обеспечивает наиболее полную оценку качества обработки, учитывая как глобальные, так и локальные особенности, критически важные для медицинской диагностики.

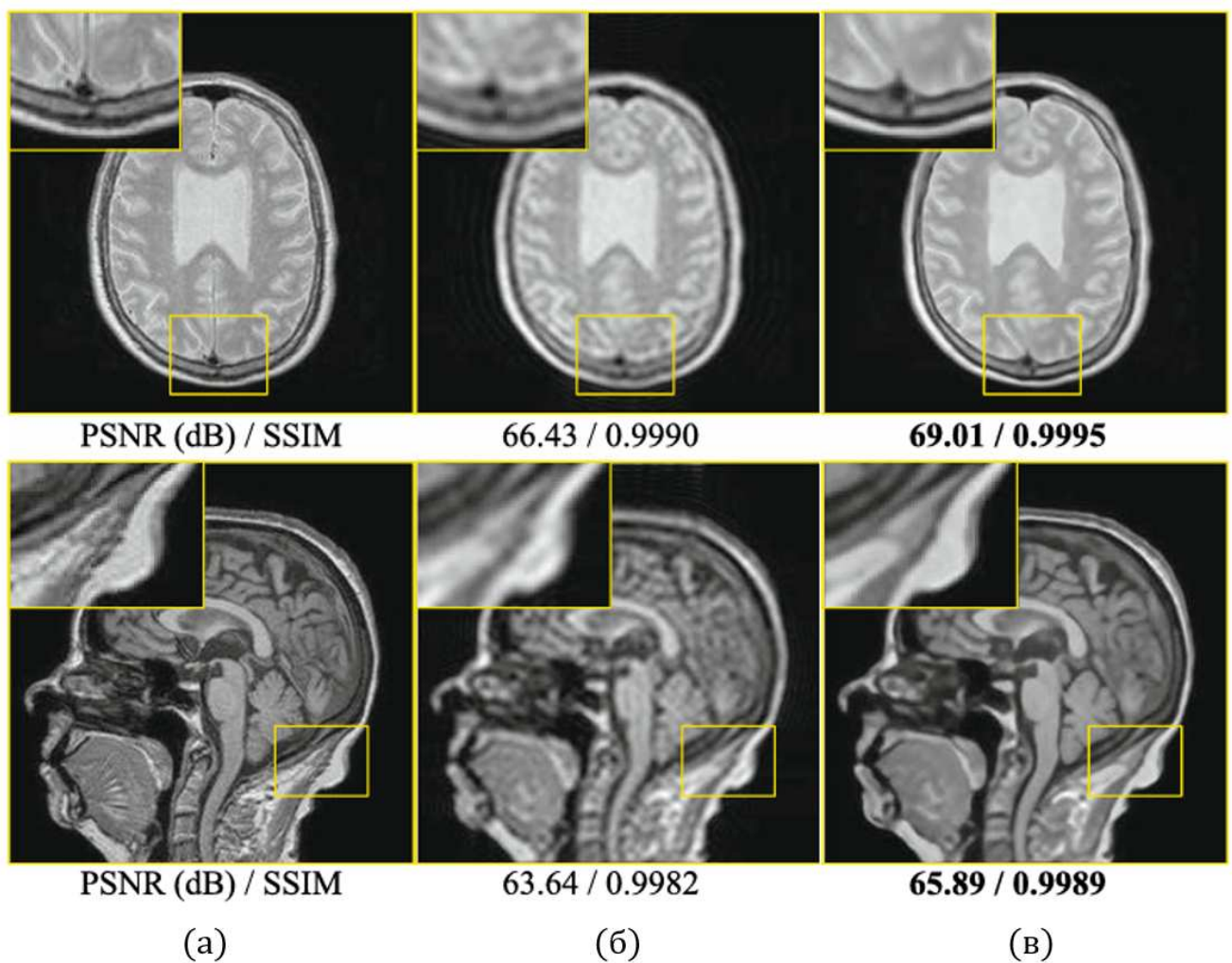


Рисунок 1.4 — Примеры улучшения качества изображений МРТ головного мозга нейронной сетью GAS-CNN. (а) – референсные изображения без артефактов, (б) – изображения с осцилляциями Гиббса, (в) – результаты нейронной сети GAS-CNN.

1.3 Гибридный метод подавления осцилляций Гиббса на изображениях головного мозга

В данном разделе описывается разработанный гибридный метод подавления осцилляций Гиббса на изображениях магнитно-резонансной томографии головного мозга, объединяющий обучаемую нейронную сеть и классический математический метод. В качестве нейронной сети используется облегчённая модификация свёрточной сети GAS-CNN [36], исходная глубина которой оказалась избыточной, а в качестве классического метода выступает рассмотренный ранее алгоритм Кельнера [30] поиска оптимальных субпиксельных сдвигов.

Обучаемая часть предлагаемого гибридного метода состоит всего из 9 остаточных свёрточных блоков, вместо 32-ух у GAS-CNN, однако, благодаря использованию вспомогательных признаков от предварительно обработанных алгоритмом Кельнера входных изображений, разработанный метод сохраняет обобщающую способность на тестовом наборе данных IXI¹.

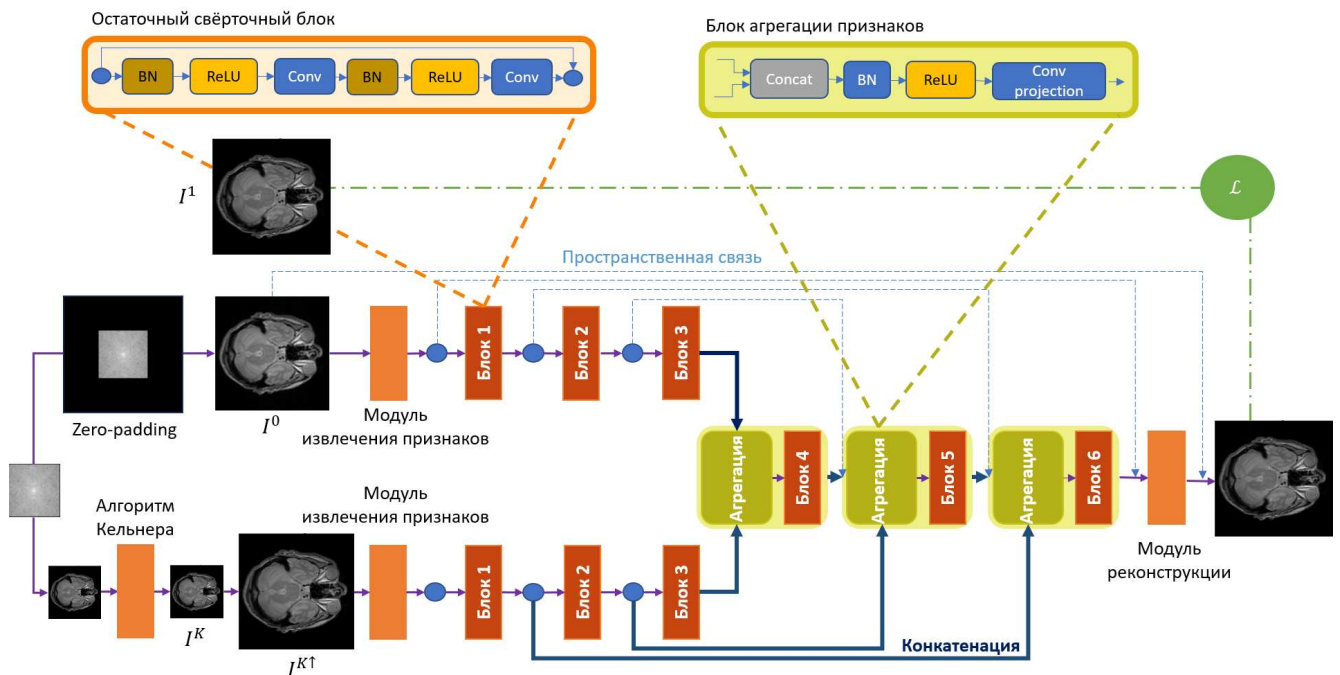


Рисунок 1.5 — Архитектура разработанного гибридного метода DGAS9-CNN подавления осцилляций Гиббса на изображениях МРТ головного мозга.

Схема разработанного метода, обозначаемого как DGAS9-CNN (*англ.* dual Gibbs-ringing artifact suppression convolutional neural network), представлена на Рис. 1.5.

¹<http://brain-development.org/ixi-dataset/>

Предлагаемый гибридный метод имеет следующую структуру:

1. Вход: центральная (1/9) часть частотного спектра Фурье.
2. Изображение I^0 (256×256): результат обратного преобразования Фурье к входному спектру, дополненному нулями по периметру (*англ.* zero padding).
3. Изображение I^K (85×85): результат алгоритма Кельнера.
4. Изображение $I^{K\uparrow}$ (256×256): интерполированный методом ближайшего соседа результат алгоритма Кельнера до соответствия размерам референсного изображения I^1 .
5. Модуль извлечения признаков из изображения I^0 (признаковое пространство размерности 64).
6. Модуль извлечения признаков из изображения $I^{K\uparrow}$ (признаковое пространство размерности 64).
7. Модуль кодирования признаков изображения I^0 .
8. Модуль кодирования признаков изображения $I^{K\uparrow}$.
9. Модуль декодирования признаков восстанавливаемого изображения с агрегацией признаков изображений I^0 и $I^{K\uparrow}$.
10. Модуль реконструкции восстанавливаемого изображения.

Итого, предлагаемый гибридный метод включает в себя модули извлечения признаков, кодировщики, параллельно кодирующие признаки изображений I^0 , $I^{K\uparrow}$ в шести свёрточных остаточных блоках, декодировщик, агрегирующий соответствующие признаковые описания в трёх свёрточных остаточных блоках, и модуль реконструкции.

Модули извлечения признаков описаний изображений I^0 и $I^{K\uparrow}$ состоят из свёртки с ядром 3×3 с последующей пакетной нормализацией (*англ.* batch normalization) [54] и активацией ReLU.

В отличие от GAS-CNN используемый остаточный блок в модуле нелинейной фильтрации использует пред-активацию [55] (см. Рис. 1.5) и содержит слои пакетной нормализации, позволяющие более эффективно вести обучение глубоких нейронных сетей, выступая в роли регуляризатора [56; 57]. Особенностью модуля декодирования является блок агрегации признаков описаний (см. Рис. 1.5), реализованный свёрткой с ядром 3×3 предварительно конкатенированных и нормализованных признаков описаний изображений I^0 и $I^{K\uparrow}$.

Заключительный модуль реконструкции воплощён в виде двух последовательных свёрточных слоёв с ядрами 3×3 , первый из которых осуществляет

промежуточное проецирование признаков описаний размерности 64 на подпространство размерности 16, а второй отображает полученный тензор в целевое цветовое пространство.

Обучение предлагаемого гибридного метода осуществлялось с применением численного метода итеративной оптимизации Adam, предложенного в работе [51]. Данный алгоритм представляет собой подход стохастической оптимизации, который сочетает преимущества методов моментов и адаптивного изменения темпа обучения (*англ.* learning rate). Алгоритм осуществляет адаптацию величины шага обучения каждого параметра модели индивидуально посредством использования оценок первых и вторых моментов градиентов целевой функции с коэффициентами, соответственно, $\beta_1=0.9$, $\beta_2=0.999$.

Гибридный метод был обучен на графическом ускорителе GPU NVIDIA GeForce RTX 2080 Ti по пакету (*англ.* batch) из 20 случайно вырезанных и аугментированных вращениями, на углы кратные 90° , и отражениями фрагментов (*англ.* patch) размера 48×48 с полиномиальным уменьшением темпа обучения, начиная с 10^{-4} .

В качестве функции потерь использовалась устойчивая к выбросам MAE (*англ.* mean average error) [58] с l_2 регуляризацией весов ($\lambda = 10^{-4}$):

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N \|\mathcal{F}(I_i^0; \theta) - I_i^1\|_1 + \lambda \cdot \|\theta\|_2^2 \rightarrow \min_{\theta \in \Theta}, \quad (1.17)$$

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.

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

Объединение метода глубокого обучения и классического математического метода в разработанном гибридном методе проявляется в улучшении качества обработки как визуально в более точном восстановлении деталей снимка

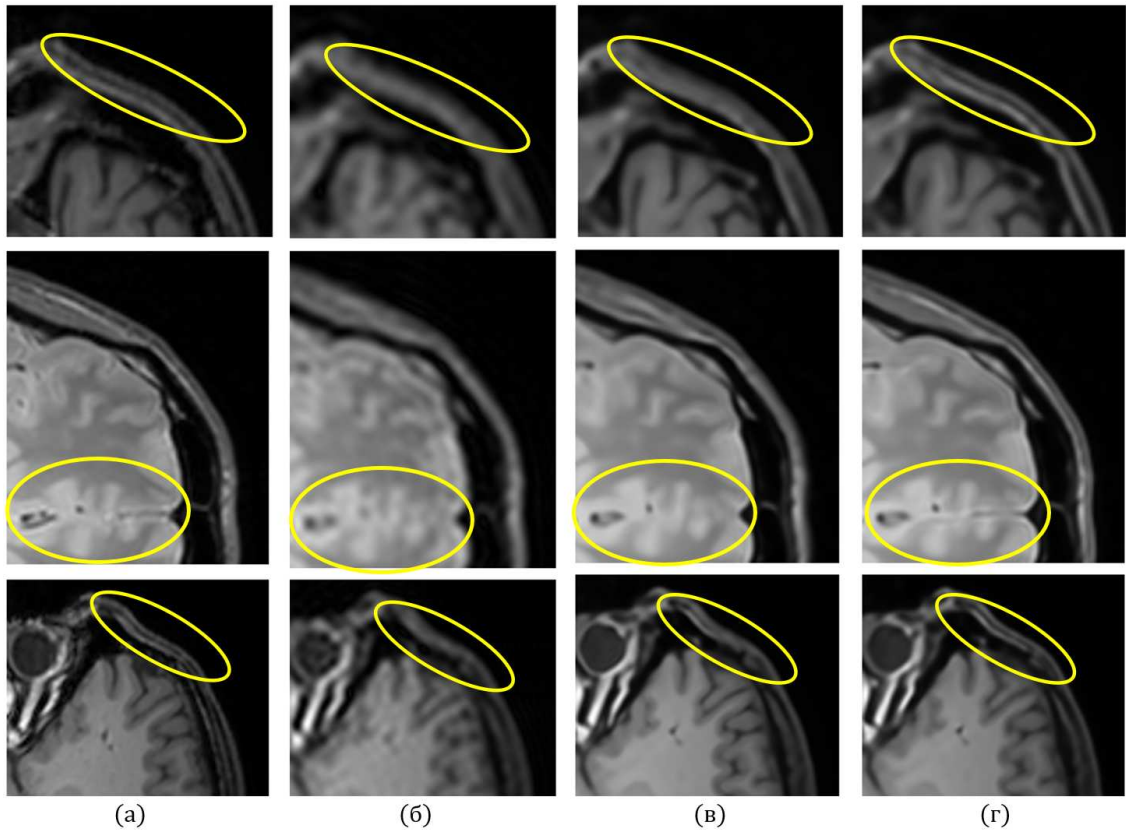


Рисунок 1.6 — Сравнение разработанного гибридного метода DGAS9-CNN улучшения качества изображений МРТ головного мозга с нейронной сетью GAS-CNN. (а) – референсные изображения без артефактов, (б) – изображения с осцилляциями Гиббса, (в) – результаты нейронной сети GAS-CNN, (г) – результаты гибридного метода DGAS9-CNN.

(см. Рис. 1.6), так и по метрике оценки качества изображений PSNR (см. Таблицу 1) и величине полной вариации TV, оценивающей степень оставшихся осцилляций. Для численного сравнения (см. Таблицу 1) разработанного метода с алгоритмом Кельнера изображения были приведены к размеру изображений I^K . Полная вариация изображений, обработанных гибридным методом, наиболее близка к полной вариации референсных изображений I^1 , в то время как алгоритм Кельнера, являясь самым быстрым, существенно размывает изображения, занижая значение полной вариации. Наконец, время выполнения гибридного метода DGAS9-CNN на центральном процессоре Intel(R) Core(TM) i7-8700 в полтора раза меньше времени выполнения свёрточной нейронной сети GAS-CNN при сопоставимом и даже лучшем качестве.

1.4 Эксперименты и результаты

Предлагаемый гибридный метод подавления осцилляций Гиббса на изображениях магнитно-резонансной томографии головного мозга был обучен и протестирован на наборе данных IXI [36].

Датасет IXI представляет собой общедоступный мультимодальный медицинский набор данных, содержащий около 600 высококачественных МРТ-сканов здоровых пациентов, полученных на томографах Philips и GE Healthcare с магнитной индукцией 1.5Т и 3Т. В состав датасета входит 581 объём (*англ.* volume) T1-взвешенных изображений, 578 объёмов T2-взвешенных изображений и 578 объёмов PD-изображений, сохранённых в формате NIFTI.

При формировании набора данных были выполнены следующие шаги:

1. Определено пересечение всех модальностей, в результате чего отобрано 577 объёмов, содержащих все три модальности (T1, T2 и PD).
2. Произведено разделение на три выборки:
 - первые 400 объёмов для создания обучающей выборки,
 - следующие 100 объёмов составили тестовую выборку,
 - оставшиеся данные вошли в валидационную выборку.

Подготовка изображений к обработке включала в себя следующие шаги:

1. Конверсия NIFTI в NumPy массивы.
2. Удаление 25 крайних срезов с каждого конца объёмов.
3. Отбор каждого десятого среза для формирования пар изображений (с артефактами и без артефактов).
4. Нормализация интенсивностей.
5. В результате сформированы:
 - обучающая выборка из 10473 пар изображений,
 - валидационная выборка из 2016 пар изображений,
 - тестовая выборка из 2617 пар изображений.

Протокол генерации изображений с артефактами Гиббса:

1. Загрузка очередного изображения.
2. Применение прямого преобразования Фурье.
3. Обрезка центральной 1/9 области частотного спектра.
4. Применение обратного преобразования Фурье.

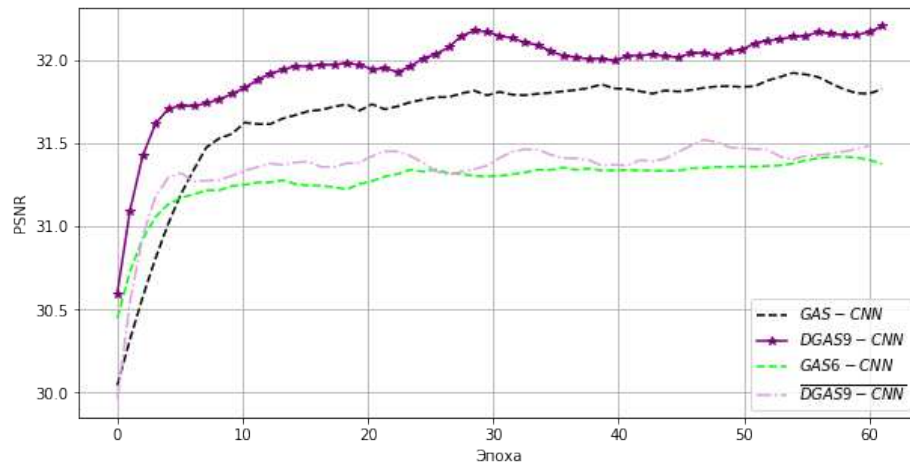


Рисунок 1.7 — Динамика средних значений PSNR на валидационном наборе данных из 2016 изображений в зависимости от эпохи. Оценка влияния вспомогательных признаков описаний алгоритма Кельнера на качество предлагаемого гибридного метода DGAS9-CNN.

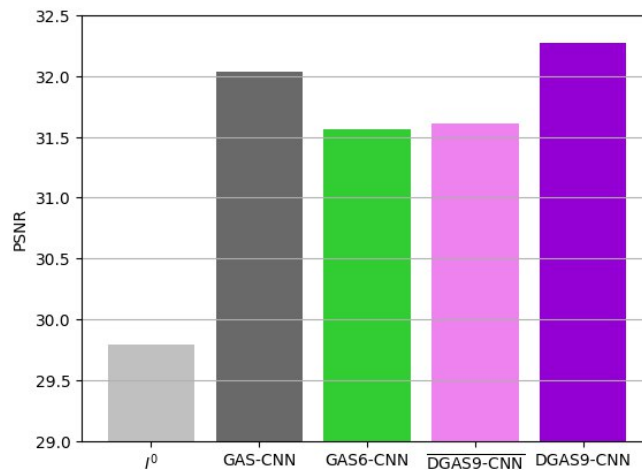


Рисунок 1.8 — Средние значения PSNR по тестовому набору данных из 2617 изображений. Оценка влияния вспомогательных признаков описаний алгоритма Кельнера на качество предлагаемого гибридного метода DGAS9-CNN.

Для доказательства того, что вспомогательные признаки от предварительно обработанного алгоритмом Кельнера входного изображения оказывают положительное влияние на обобщающую способность предлагаемого гибридного метода DGAS9-CNN, были дополнительно обучены следующие вариации предлагаемой архитектуры:

- GAS6-CNN – архитектура, полностью повторяющая предлагаемую, однако без использования признаков описаний $I^{K\uparrow}$;
- $\overline{\text{DGAS9-CNN}}$ – архитектура, полностью повторяющая предлагаемую, но с пропущенной обработкой изображения алгоритмом Кельнера.

Анализ результатов обучения выше указанных методов (см. Рис. 1.7 и Рис. 1.8) демонстрирует, что нейронная сеть GAS6-CNN теряет обобщающую способность на тестовом наборе данных на 0.5 дБ по сравнению с базовой GAS-CNN. В то же время добавление к GAS6-CNN дополнительных признаков, от обработанных алгоритмом Кельнера изображений, приводит к предлагаемой гибридной сети DGAS9-CNN и возвращает обобщающую способность на уровень нейронной сети GAS-CNN и даже превосходит её на 0.2 дБ, причём последняя имеет в 3.5 раза большее число свёрточных остаточных блоков в нелинейном отображении. Важно также заметить, что сама по себе обработка алгоритмом Кельнера существенна, так как метод $\overline{\text{DGAS9-CNN}}$, в котором алгоритм Кельнера пропущен, демонстрирует схожие с GAS6-CNN результаты. Представленная столбчатая диаграмма свидетельствует, что гибридный метод DGAS9-CNN показывает наивысшее качество реконструкции со значением PSNR 32.27 дБ.

1.5 Выводы

Проведенное исследование методов повышения качества изображений магнитно-резонансной томографии позволяет заключить, что гибридный метод DGAS9-CNN демонстрирует наилучшую эффективность в подавлении артефактов.

Действительно, DGAS9-CNN в полтора раза быстрее базовой нейронной сети GAS-CNN и при этом обеспечивает лучшие показатели используемых метрик оценки качества изображений (см. Таблицу 1). Таблица 1 показывает, что классический метод Кельнера, существенно подавляя осцилляции Гиббса, приводит к чрезмерно низкому среднему значению полной вариации (меньше референсного значения) и, соответственно, к потере диагностически значимых деталей снимков, а также к низкому среднему значению PSNR. В то же время, разработанный гибридный метод DGAS9-CNN, объединяющий классический метод поиска оптимальных субпиксельных сдвигов и свёрточную нейронную сеть GAS-CNN, предлагает полную вариацию наиболее близкую к референсным изображениям и наибольшую среднюю величину PSNR на тестовом наборе данных IXI из 2617 изображений.

Глава 2. Нейронные операторы в обработке медицинских изображений

Нейронные операторы – это класс глубоких нейронных сетей, обучаемых по результатам численных расчётов классическими математическими методами и, соответственно, способные аппроксимировать нелинейные отображения между пространствами $\mathcal{A} \rightarrow \mathcal{U}$, в том числе между функциональными:

$$\mathcal{F} : \mathcal{A} \times \Theta \rightarrow \mathcal{U}, \quad (2.1)$$

или, в эквивалентной записи:

$$\mathcal{F}_\theta : \mathcal{A} \rightarrow \mathcal{U}, \quad (2.2)$$

где \mathcal{A} – входное пространство; \mathcal{U} – целевое пространство; Θ – параметры.

В конечномерных пространствах нейронные операторы обобщают по построению нейросетевые модели в смысле инвариантности к различным дискретизациям входных данных, сохраняя идею параметризации сложных нелинейных отображений. Так, полносвязные сети можно рассматривать как частный случай нейронных операторов на арифметических пространствах, а свёрточные сети – как дискретный аналог интегральных операторов с разностными ядрами. Эта обобщающая концепция позволяет создавать новые методы, устойчивые к передискретизации, что является фундаментальным требованием при решении ряда задач обработки изображений.

Архитектура нейронного оператора обычно имеет вид ансамбля из нейронных модулей:

$$\mathcal{F} : Q \circ \Phi_L \circ \Phi_{L-1} \circ \dots \circ \Phi_1 \circ P, \quad (2.3)$$

где P – модуль извлечения признаков (кодировщик); Q – модуль реконструкции (декодировщик); Φ_l – очередной блок модуля нелинейной фильтрации.

Различные архитектуры блоков Φ_l задают ключевые архитектуры нейронных операторов:

- нейронный оператор Фурье [22],
- сети Колмогорова-Арнольда [59; 60],
- DeepONet [61],
- сети неявного представления [62; 63].

Пусть, например, нас интересует отображение, определяющее некоторый рентгеновский снимок грудной клетки в оттенках серого, иными словами отображение компакта D арифметического пространства \mathbb{R}^2 в компакт цветового пространства (в качестве $D \subset \mathbb{R}^2$ обычно выступает двумерный единичный куб $[0, 1]^2$). Зададим и сравним три конечномерных нейронных оператора в форме (2.3), аппроксимирующих это отображение, выбрав в качестве $P: \mathbb{R}^2 \rightarrow \mathbb{R}^{128}$ и $Q: \mathbb{R}^{128} \rightarrow \mathbb{R}$ линейные преобразования, а в качестве $\Phi_l: \mathbb{R}^{128} \rightarrow \mathbb{R}^{128}$ – три варианта нелинейных преобразований, отличающихся только функциями активации, а именно: ReLU, функции Габора, синусы.

Проведём обучение заданных нейронных операторов:

- $\mathcal{F}_\theta^{ReLU}: \mathbb{R}^2 \rightarrow \mathbb{R}$,
- $\mathcal{F}_\theta^{WIRE}$ [63]: $\mathbb{R}^2 \rightarrow \mathbb{R}$,
- $\mathcal{F}_\theta^{SIREN}$ [62]: $\mathbb{R}^2 \rightarrow \mathbb{R}$, –

на наборе данных $\mathcal{D} = \{(\mathbf{x}_i, I(\mathbf{x}_i))\}_{i=1}^N$, где $I(\mathbf{x}_i) \in \mathbb{R}$ – значение интенсивности пикселя в оттенках серого, отвечающего очередной координате $\mathbf{x}_i \in \mathbb{R}^2$ на изображении I , заданном в области D . Тогда функция потерь примет вид:

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N \|\mathcal{F}_\theta(\mathbf{x}_i) - I(\mathbf{x}_i)\|^2 \rightarrow \min_{\theta \in \Theta}. \quad (2.4)$$

Минимизация рассматриваемого функционала осуществляется с применением стандартных численных методов оптимизации первого порядка. В качестве базового подхода используется стохастический градиентный спуск [50]. Для ускорения сходимости и повышения устойчивости процесса оптимизации применяются современные модификации метода, в частности, алгоритм Adam [51], сочетающий преимущества адаптивной оценки моментов градиента и инерционности обновления параметров.

Динамика обучения трёх заданных конечномерных нейронных операторов приведена на Рис. 2.1. Архитектуры нейронных операторов отличаются только выбранной нелинейностью. Используется 2 блока в модуле нелинейного отображения. Набор данных \mathcal{D} состоит из 2048 пар, в которых точки \mathbf{x}_i получены из равномерного целочисленного распределения.

Визуально качество реконструкции рентгеновского снимка грудной клетки нейронными операторами с различными нелинейностями можно оценить на Рис. 2.2. Важно, что рассматриваемые нейронные операторы по построению способны параллельно обрабатывать произвольное число точек, варьируя размер

пакета (*англ.* batch size). Так, в данном случае (см. Рис. 2.2), для восстановления изображения размера 224×224 , не ограничивая общности, пакет точек состоял из равномерной пиксельной сетки из 50176 точек, однако вывод доступен на произвольной дискретизации, а также градиент и лапласиан такого отображения, применяя аппарат автоматического дифференцирования.

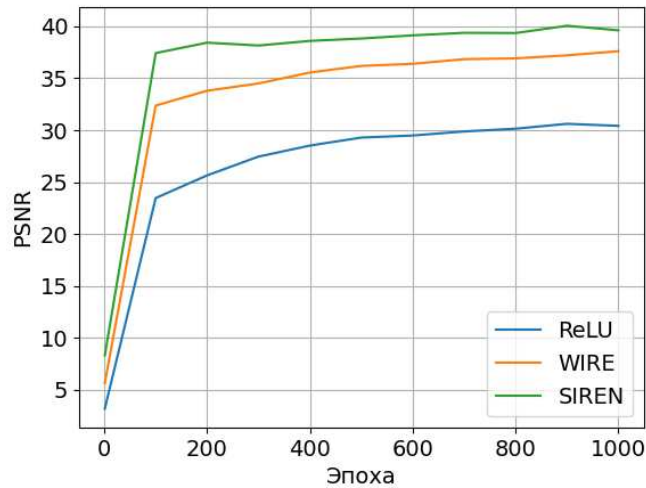


Рисунок 2.1 — Мониторинг качества реконструкции изображения рентгеновского снимка грудной клетки размера 224×224 тремя сравниваемыми нейронными операторами, отличающимися видом нелинейности: ReLU, функции Габора (WIRE [63]), синусы (SIREN [62]).

Результаты демонстрируют, что выбор нелинейности оказывает критическое влияние на качество реконструкции. Периодические функции активации намного лучше подходят для представления сложных естественных сигналов в медицинской визуализации. Количественные оценки подтверждают, что использование периодических активаций позволяет достичь улучшения PSNR на 2-3 дБ. Эти результаты имеют важное мотивирующее значение для разработки новых методов на основе нейронных операторов.

Недостатком конечномерных нейронных операторов является их *переобучение* под конкретное изображение. Этот же недостаток свойственен подходу физически-информированных нейронных сетей (*англ.* physics-informed neural networks, PINNs) [64], где для каждого уравнения конкретное решение $u(x)$ ищется, обучая нейронную сеть отображению из координат x в значения решения $u(x)$. Так, например, если в уравнении Пуассона мы изменим правую часть или начальное условие, то процесс обучения придётся начинать сначала. Частично разрешить этот недостаток возможно с помощью добавления дополнительного

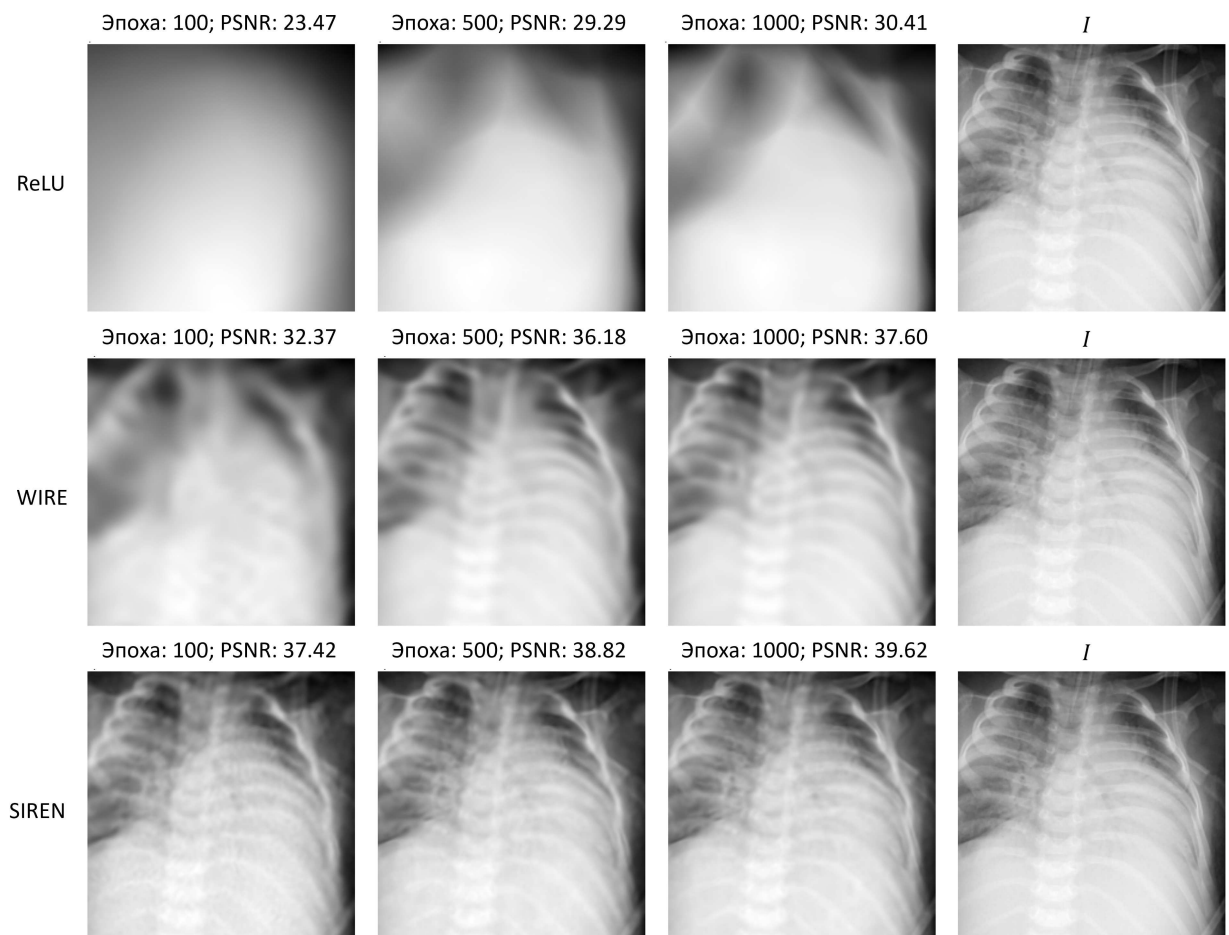


Рисунок 2.2 — Примеры восстановления изображения рентгеновского снимка грудной клетки размера 224×224 тремя нейронными операторами, отличающимися видом нелинейности: ReLU, функции Габора (WIRE [63]), синусы (SIREN [62]). *I* – референсное изображение из набора данных PneumoniaMNIST [<https://medmnist.com/>].

кодировщика [65], предоставляющего алгоритму признаки с информацией о реконструируемом изображении, однако более комплексное решение заключается в построении нейронных операторов, действующих непосредственно в функциональных пространствах [22]. Для перехода в функциональные пространства нам понадобятся некоторые факты анализа [66].

Рассмотрим изображение *I* как функцию, заданную на пиксельной сетке (равномерной дискретизации двумерного единичного куба $[0, 1]^2$) и принимающую конечное число значений. Действительно, если значения пикселей изображения закодированы восемью битами, то они принимают конечный набор целых значений от 0 до 255, если шестнадцатью, то от 0 до 65535 и так далее. Иными словами, перед нами оказывается линейная комбинация индикаторов. Множество функций, принимающих лишь конечное число значений, как известно, всюду

плотно в банаховом пространстве суммируемых функций L_1 . То есть, с одной стороны, изображение – это функция из пространства $L_1(D; \mathbb{R}^{d_I})$, заданная на компакте $D \subset \mathbb{R}^d$ и принимающая значения из \mathbb{R}^{d_I} . С другой стороны, часто бывает удобно рассматривать изображение как функцию с интегрируемым квадратом из гильбертова пространства $L_2(D; \mathbb{R}^{d_I}) \subseteq L_1(D; \mathbb{R}^{d_I})$. Таким образом, в общем случае, нейронные операторы в бесконечномерных функциональных пространствах аппроксимируют нелинейные отображения между банаховыми пространствами: $\mathcal{F}_\theta : L_p \rightarrow L_q$.

Используя фундаментальное свойство всюду плотной вложенности непрерывных функций C в пространствах L_p , мы получаем доступ к широкому спектру теорем аппроксимации, позволяющих рассматривать различные полные системы функций в C (а значит, и в L_p) в качестве потенциальных архитектурных решений блока модуля нелинейной фильтрации. В частности, из данной теоретической базы следует возможность использования ортогональных полиномов (Лежандра, Чебышева), применимость тригонометрических систем Фурье для аппроксимации периодических компонент сигнала и эффективность вейвлет-разложений при работе с локальными особенностями медицинских изображений. Самым известным примером полной системы функций в L_p является совокупность: $1, x, x^2, \dots$. Полнота этой системы следует из теоремы Вейерштрасса о равномерном приближении непрерывной функции $I(x)$ на сегменте последовательностью многочленов $P_n(x)$ и из всюду плотности C в L_p . Причём в качестве таких многочленов выступают свёртки целевой функции $I(x)$ с дельта-образной функциональной последовательностью ядер $Q_n(x)$: $\int_0^1 I(t)Q_n(t-x)dt$, где $Q_n(x)$ задаётся, например, как $c_n(1-x^2)^n$, $n \in \mathbb{N}$.

Данная глава посвящена исследованию нейронных операторов улучшения качества изображений магнитно-резонансной томографии головного мозга, а также сегментации опухолей на ультразвуковых, гистологических и колоноскопических изображениях.

2.1 Нейронный оператор Фурье аппроксимации математических методов, используемых для повышения качества изображений магнитно-резонансной томографии на произвольном разрешении

Нейронный оператор Фурье (*англ.* Fourier neural operator, FNO) [22] – это модель глубокого обучения, предназначенная для аппроксимации нелинейных операторов в бесконечномерных функциональных пространствах и обучаемая по результатам численных расчётов классическими математическими методами. Данный подход обеспечивает устойчивость к изменению разрешения входных данных по построению. В основе FNO (см. Рис. 2.3) лежит сочетание спектральных и линейных преобразований.

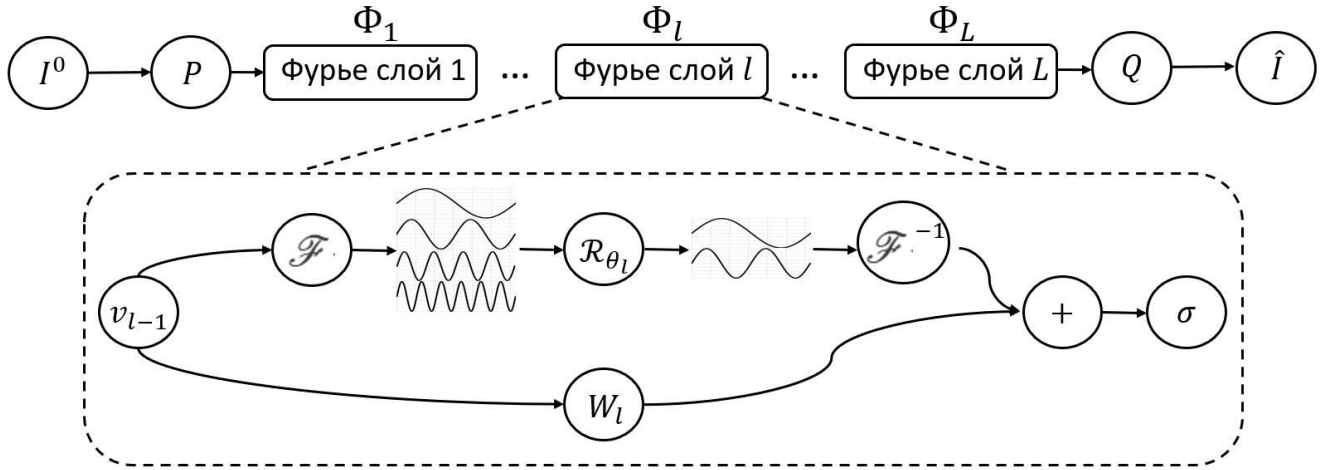


Рисунок 2.3 — Архитектура нейронного оператора Фурье. I^0 – входное изображение, требующее улучшения качества, \hat{I} – изображение, обработанное нейронным оператором Фурье, \mathcal{F} – прямое преобразование Фурье, \mathcal{F}^{-1} – обратное преобразование Фурье, \mathcal{R}_θ – параметризованный ядром оператор обрезки высоких частот, W – линейное преобразование, встроенное в пространственную связь (*англ.* skip connection), σ – нелинейность (например, ReLU).

Блок Φ_l модуля нелинейной фильтрации FNO имеет следующий вид:

$$\Phi_l(v_{l-1})(\mathbf{x}) = \sigma(W_l v_{l-1}(\mathbf{x}) + \mathcal{F}^{-1}(\mathcal{R}_{\theta_l} \cdot \mathcal{F}(v_{l-1}))(\mathbf{x})), \quad (2.5)$$

где \mathcal{F} – прямое преобразование Фурье; \mathcal{F}^{-1} – обратное преобразование Фурье; \mathcal{R}_{θ_l} – параметризованный ядром оператор обрезки высоких частот, сохраняющий лишь низкие частоты $|r| < r_{max}$; W_l – линейное преобразование; σ – нелинейность (например, ReLU).

Нейронный оператор Фурье был предложен для решения уравнений в частных производных на равномерной сетке произвольного масштаба, где в сравнении с традиционными свёрточными нейронными сетями продемонстрировал большую эффективность. Так, нейронный оператор Фурье успешно применяется для решения задачи переноса тепла и задач машинного обучения [67].

Архитектура нейронного оператора Фурье во многом мотивирована общим видом решения краевой задачи для линейного обыкновенного дифференциального уравнения второго порядка. Рассмотрим такую краевую задачу, предполагая, что выполнена редукция к однородным краевым условиям:

$$a_0(x)u''(x) + a_1(x)u'(x) + a_2(x)u(x) = f_1(x), 0 \leq x \leq 1, \quad (2.6)$$

$$\alpha_1 u'(0) + \beta_1 u(0) = 0, \alpha_2 u'(1) + \beta_2 u(1) = 0, \quad (2.7)$$

где функции $a_i(x) \in C[0, 1]$, $i = 0, 1, 2$, $f_1(x) \in C[0, 1]$ и постоянные $\alpha_1, \alpha_2, \beta_1, \beta_2$ заданы, причём $a_0(x) \neq 0$, а постоянные такие, что $\alpha_i^2 + \beta_i^2 > 0$. Почленно разделим уравнение (2.6) на $a_0(x)$, а затем умножим на $p(x) = \exp(\int_0^x \frac{a_1(s)}{a_0(s)} ds)$. Выделяя полную производную, уравнение (2.6) сводится к операторному виду:

$$\mathcal{L}u = f, 0 \leq x \leq 1, \quad (2.8)$$

$$\mathcal{L} = \frac{d}{dx}(p(x)\frac{d}{dx}) - q(x), \quad (2.9)$$

где $p(x) > 0$; $q(x) = -\frac{p(x)a_2(x)}{a_0(x)}$; $f(x) = \frac{p(x)f_1(x)}{a_0(x)}$; \mathcal{L} – оператор Штурма-Лиувилля.

Известно [68], что если однородная краевая задача имеет только нулевое решение, то решение краевой задачи (2.8) с граничными условиями (2.7) существует и единственно и задаётся формулой:

$$u(x) = \int_0^1 G(x, \xi) f(\xi) d\xi, 0 \leq x \leq 1, \quad (2.10)$$

где $G(x, \xi)$ – функция Грина, то есть функция, определённая и непрерывная в квадрате $[0, 1] \times [0, 1]$, удовлетворяющая однородному уравнению $\mathcal{L}G = 0$ и однородным краевым условиям по x , а также дважды непрерывно дифференцируемая по переменной x на множестве $[0, \xi) \cup (\xi, 1]$ для $\forall \xi \in (0, 1)$ с частной производной при $x = \xi$ такой, что $G_x(\xi + 0, \xi) - G_x(\xi - 0, \xi) = \frac{1}{p(\xi)}$. В физике $G(x, \xi)$ описывает поле в точке x , созданное точечным источником в ξ , например, потенциал точечного заряда в электростатике.

Можно показать [68], что формула (2.10) даёт решение также и в случае нелинейной краевой задачи:

$$u''(x) + a^2 u(x) = F(x, u(x)), 0 \leq x \leq 1, \quad (2.11)$$

$$u(0) = u(1) = 0, \quad (2.12)$$

принимая вид:

$$u(x) = \int_0^1 G(x, \xi) F(\xi, u(\xi)) d\xi, 0 \leq x \leq 1. \quad (2.13)$$

Рассматривая уравнение (2.13) как интегральное уравнение в паре банаховых пространств, например, $C[0, 1]$, применим метод последовательных приближений:

$$v_l(x) = \int_0^1 G(x, \xi) F(\xi, v_{l-1}(\xi)) d\xi, 0 \leq x \leq 1, \quad (2.14)$$

который даст в пределе $\{v_l(x)\}_{l=1}^{\infty}$ единственное решение (неподвижную точку), если интегральный оператор есть сжимающее отображение.

Общий вид поиска решения методом последовательных приближений (2.14) является ключом к последовательной архитектуре нейронного оператора (2.3), а вид интегрального преобразования (2.14) – к архитектуре блока (2.5) модуля нелинейной фильтрации. Действительно, предположив, что функция Грина имеет вид разностного ядра, то есть $G(x, \xi) = G(x - \xi)$, интеграл (2.14) определяет ни что иное как свёртку, которая сводится к перемножению соответствующих Фурье-образов признакового описания и ядра оператора \mathcal{R}_{θ_l} , оставляя лишь несколько низких частот. Обрезкой высоких частот достигается устойчивость, во-первых, к входному разрешению, во-вторых, к шуму в признаковом пространстве, а также учёт глобальных закономерностей через низкочастотные моды. Пространственная связь, с линейным преобразованием W_l , снижает риск затухания градиента [55] в ходе обучения. Остаётся заметить, что если оператор \mathcal{L} трансляционно инвариантен, то есть если \mathcal{L} имеет постоянные коэффициенты, то функция Грина, действительно, может быть выбрана в виде конволюционного оператора.

Решая аналитически дифференциальные уравнения в частных производных второго порядка, например, параболические (уравнение теплопроводности) или гиперболические (волновое уравнение), используется метод Фурье (метод разделения переменных), применяя который задача сводится вновь к решению

обыкновенных дифференциальных уравнений второго порядка, в частности к решению задачи на собственные значения с оператором Штурма-Лиувилля (2.9).

Таким образом, нейронный оператор Фурье параметризует функцию Грина в неявном, спектральном, виде в ядрах операторов $\{\mathcal{R}_{\theta_i}\}_{i=1}^L$. Архитектура нейронного оператора Фурье мотивирована общим видом решения линейной краевой задачи для дифференциального уравнения второго порядка, однако ввиду обобщающей способности методов машинного обучения и будучи обучаемым исключительно на данных (*англ.* data-driven) численных расчётов классическими математическими методами, FNO позволяет искать решения также и нелинейных дифференциальных уравнений со сложными граничными условиями в привычном виде (2.10). Так, сильными сторонами нейронного оператора Фурье являются: универсальность, вычислительная эффективность и обобщаемость на произвольные разрешения без необходимости повторного обучения или тонкой настройки.

Применим нейронный оператор Фурье для улучшения качества изображений магнитно-резонансной томографии головного мозга на произвольном разрешении. МРТ является одним из наиболее распространённых и информативных методов, позволяющих получить детализированные изображения внутренних структур организма. Однако качество изображений МРТ может варьироваться в зависимости от множества факторов, таких как параметры сканирования, движение пациента и шумы, возникающие в процессе получения данных. Типичными дефектами на таких изображениях являются осцилляции Гиббса и шум. В качестве аппроксимируемой классической модели, используемой для улучшения качества изображений, выберем модель анизотропной (нелинейной) диффузии Перона-Малика [69], которая сглаживает слабые градиенты (шум, мелкие осцилляции), сохраняя сильные (резкие контуры):

$$\frac{\partial I}{\partial t} = \text{div}(c(\mathbf{x}, t) \cdot \nabla I), \quad (2.15)$$

$$I(\mathbf{x}, 0) = I^0(\mathbf{x}), \quad \left. \frac{\partial I(\mathbf{x}, t)}{\partial n} \right|_{\mathbf{x} \in \partial \Omega} = 0, \quad (2.16)$$

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:

$$c(|\nabla I|) = \frac{1}{1 + \left(\frac{|\nabla I|}{k}\right)^2} \quad (2.17)$$

или альтернативный вариант, используемый в исследовании:

$$c(|\nabla I|) = e^{-(\frac{\nabla I}{k})^2}, \quad (2.18)$$

где k – параметр, управляющий чувствительностью к контурам.

Для генерации набора данных воспользуемся моделью аддитивного гауссовского шума. В рамках данной модели наблюдаемое изображение I^0 формируется как сумма референсного изображения I^1 и статистически независимой случайной помехи η :

$$I^0(\mathbf{x}) = \eta(\mathbf{x}) + \mathcal{F}^{-1}(\mathcal{R} \cdot \mathcal{F}(I^1))(\mathbf{x}), \quad (2.19)$$

где I^1 – референсное изображение высокого разрешения 255×255 из набора данных IXI¹; I^0 – изображение с артефактами Гиббса низкого разрешения 145×145 и аддитивным гауссовским шумом $\eta \sim \mathcal{N}(0, \sigma^2)$ с $\sigma = 0.01$; \mathcal{F} – прямое преобразование Фурье; \mathcal{F}^{-1} – обратное преобразование Фурье; \mathcal{R} – оператор обрезки высоких частот, сохраняющий лишь центральные 25 % низких частот, см. Рис. 2.4.

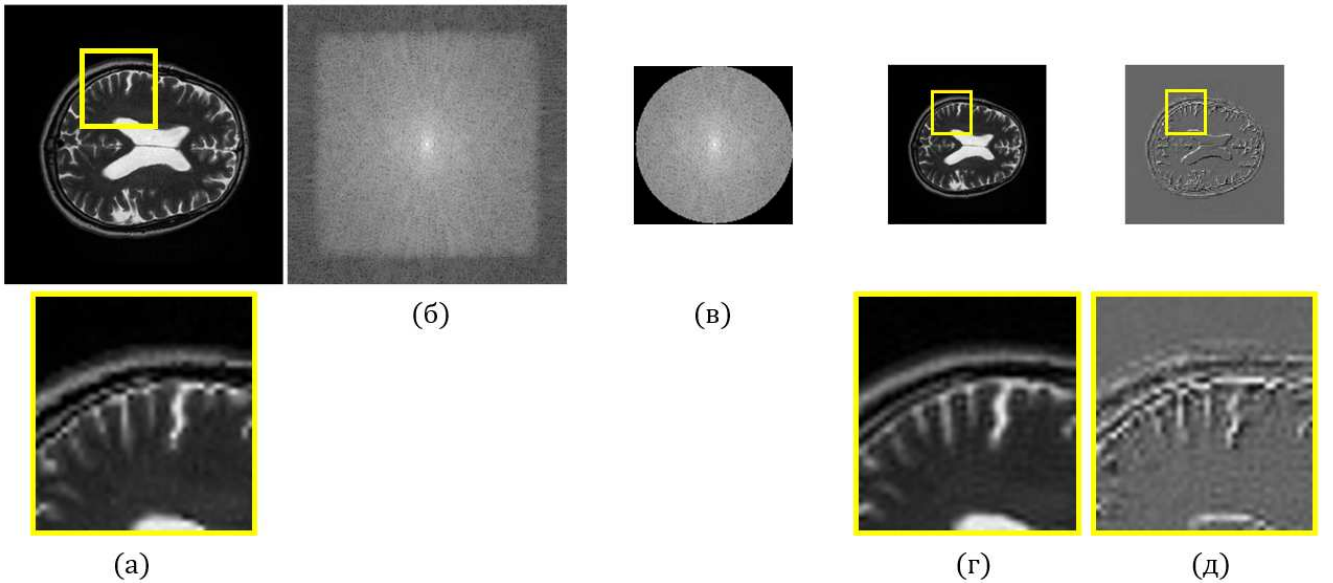


Рисунок 2.4 — Генерация изображения магнитно-резонансной томографии головного мозга с осцилляциями Гиббса и аддитивным гауссовским шумом. (а) – референсное изображение I^1 высокого разрешения 255×255 , (б) – Фурье-спектр изображения I^1 , (в) – результат обрезки высоких частот оператором \mathcal{R} , (г) – изображение I^0 низкого разрешения 145×145 , требующее улучшения качества, (д) – остаточный кадр $|I^0 - I^1|$, демонстрирующий артефакты на снимке.

¹<http://brain-development.org/ixi-dataset/>

Нейронный оператор Фурье в данном случае аппроксимирует нелинейное отображение в функциональном пространстве $L_2([0, 1]^2; \mathbb{R})$ в себя, соответственно, начального условия I^0 в результат трёх итераций дифференциальной модели анизотропной диффузии (2.15) с временным шагом $\Delta t = 0.1$ и параметром чувствительности к контурам $k = 0.1$. Используется 4 блока Φ_l в модуле нелинейной фильтрации (см. Рис. 2.3). Размерность признакового пространства равна 64. Параметризованные ядрами операторы обрезки высоких частот $\{\mathcal{R}_{\theta_l}\}_{l=1}^4$ сохраняют низкие частоты в окне 32×32 , отбрасывая остальные.

Обучение велось на графическом ускорителе GPU NVIDIA RTX A6000 на случайных пакетах (*англ.* batch) из 20 изображений, минимизируя среднюю квадратичную ошибку численным методом оптимизации Adam [51] с темпом обучения (*англ.* learning rate) 10^{-3} в течение 100 эпох. Adam, будучи адаптивным численным методом оптимизации, автоматически адаптирует темп обучения для каждого параметра модели, что важно для устойчивого обучения методов, оперирующих спектральными представлениями, например, FNO, градиенты которого могут иметь разный порядок в разных частотных компонентах.

Качество обработки изображений МРТ головного мозга нейронным оператором Фурье можно проанализировать на Рис. 2.5 и визуально, и численно. Заметим, что полная вариация в данном случае не используется в качестве регуляризирующей компоненты функции потерь, а применяется как показатель присутствия осцилляций Гиббса на изображениях. Осцилляции Гиббса, проявляющиеся в виде высокочастотных колебаний вблизи резких контуров, приводят к локальному росту градиентов, что повышает TV. Полная вариация чувствительна и к другим факторам, поэтому интерпретация требует осторожности и дополнительного контроля метриками оценки качества изображений, например, PSNR и SSIM. Так, в сравнении с классическим алгоритмом Кельнера [30], существенно размывающим изображение, что численно отражается в низком значении полной вариации, нейронный оператор Фурье повышает качество реконструкции тонких деталей снимка (см. Рис. 2.5), приближая значение TV к референсному, одновременно демонстрируя рост показателей PSNR и SSIM (см. Рис. 2.6).

Нейронный оператор Фурье по построению, благодаря оператору \mathcal{R}_{θ} , обобщаем на равномерные сетки произвольного масштаба. Так, будучи обученным на масштабе $\times 1.0$, FNO демонстрирует лучшую обобщающую способность на прочие масштабы, чем другие алгоритмы (см. Рис. 2.6). Более того, нейронный оператор Фурье проявляет регуляризирующее свойство, приближая значения

полной вариации на всех рассматриваемых разрешениях к референсным, при этом повторяя значения PSNR и SSIM модели анизотропной диффузии Перона-Малика (см. Рис. 2.6), на численных расчётах которой он обучался.

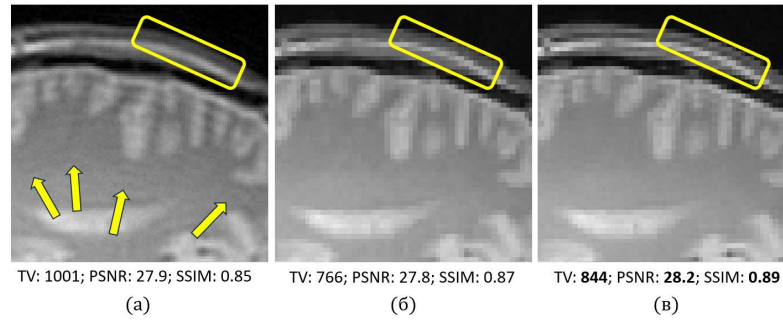


Рисунок 2.5 — Обработка изображения МРТ головного мозга нейронным оператором Фурье и классическим методом Кельнера. (а) – входное изображение с осцилляциями Гиббса (отмечены жёлтыми стрелками) и шумом, (б) – результат обработки алгоритмом Кельнера, (в) – результат обработки нейронным оператором Фурье, аппроксимирующим решение уравнения анизотропной диффузии, использованном в модели Перона-Малика.

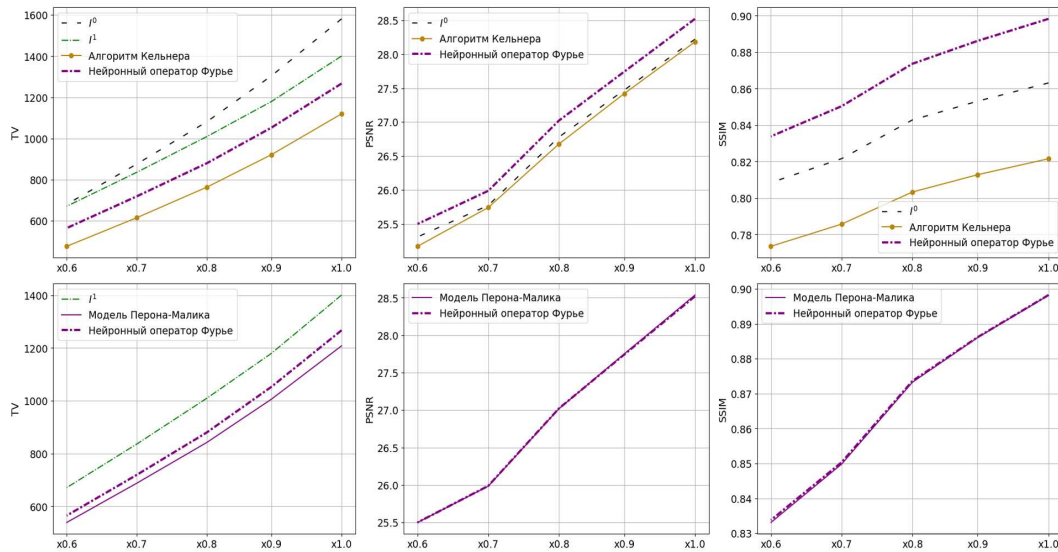


Рисунок 2.6 — Средние значения TV, PSNR, SSIM методов улучшения качества изображений МРТ головного мозга по тестовому набору данных IXI из 2617 изображений. Оценка устойчивости методов к изменению входного разрешения и регуляризующей способности нейронного оператора Фурье, обученного аппроксимации решения уравнения анизотропной диффузии, использованном в модели Перона-Малика на масштабе $\times 1.0$. Масштаб $\times 1.0$ соответствует равномерной пиксельной сетке 145×145 . Остальные масштабы получены линейной интерполяцией.

2.2 Проекционные сети Колмогорова-Арнольда для повышения качества изображений магнитно-резонансной томографии и сегментации ультразвуковых, гистологических и колоноскопических изображений

В данном разделе главы рассматривается разработанный гибридный метод анализа и обработки медицинских изображений, объединяющий свёрточную нейронную сеть с недавно предложенными сетями Колмогорова-Арнольда (*англ.* Kolmogorov-Arnold networks, KANs) [59], адаптированными в диссертационной работе для задач обработки изображений.

Сети Колмогорова-Арнольда представляют перспективную альтернативу многослойным перцептронам (*англ.* multi-layer perceptrons, MLPs) и мотивированы одноимённой теоремой [70] из анализа действительного переменного и теории приближений, которая утверждает, что многомерную непрерывную функцию можно записать в виде конечной композиции одномерных непрерывных функций и бинарной операции сложения:

$$f(\mathbf{x}) = f(x_1, \dots, x_n) = \sum_{j=1}^{2n+1} \zeta_j \left(\sum_{i=1}^n \phi_{ji}(x_i) \right), \quad (2.20)$$

где $\phi_{ji}: [0, 1] \rightarrow \mathbb{R}$ и $\zeta_j: \mathbb{R} \rightarrow \mathbb{R}$ – одномерные непрерывные функции Колмогорова-Арнольда.

Подобно теореме Цыбенко [71], утверждающей всюду плотность множества искусственных нейронных сетей с одним скрытым слоем и сигмоидной функцией активации в пространстве непрерывных функций:

$$\forall \varepsilon > 0 \exists M > 0 : \exists \{\alpha_j, \mathbf{w}_j, b_j\}_{j=1}^M : \left| \sum_{j=1}^M \alpha_j \phi(\mathbf{w}_j^T \mathbf{x} + b_j) - f(\mathbf{x}) \right| < \varepsilon, \quad (2.21)$$

$$\phi(x) = \frac{1}{1 + e^{-x}}, \quad (2.22)$$

теорема Колмогорова-Арнольда (2.20) утверждает, что нейронные сети с $2n + 1$ нейронами в скрытом слое и одномерными непрерывными функциями активации, сколь угодно точно аппроксимируют непрерывные функции n -переменных на компакте. Важно заметить, что в теореме Цыбенко ничего не сказано о числе нейронов в скрытом слое, а сигмоидная функция активации является глобальной для всех нейронов, в то время как теорема Колмогорова-Арнольда задаёт конкретное число нейронов в скрытом слое, равное $2n + 1$, а функции активации

определяются индивидуально для каждого нейрона. В каком-то смысле, Колмогоров и Арнольд показали, что единственная истинная функция многих переменных – это сложение, поскольку все другие функции можно записать с использованием функций одной переменной и сложения.

Оборотной стороной теоремы Колмогорова-Арнольда является отсутствие эффективного алгоритма для нахождения одномерных функций ϕ_{ji} и ζ_j , а также отсутствие гарантий их *гладкости*. Действительно, на практике функции Колмогорова-Арнольда могут оказаться, например, фрактальными и, соответственно, не обучаемыми стандартными алгоритмами [50]. Так, известным примером непрерывной функции, нигде не имеющей производной, является функция Вейерштрасса, заданная на всей вещественной прямой единым выражением:

$$\phi(x) = \sum_{k=0}^{\infty} b^k \cos(a^k \pi x), \quad (2.23)$$

где a – произвольное нечётное число, не равное единице, а b – положительное число, меньшее единицы. Функция была предложена Карлом Вейерштрассом в 1872 году как контрпример к интуитивному представлению о том, что непрерывные функции должны быть дифференцируемыми «почти везде». Этот функциональный ряд мажорируется сходящимся числовым рядом $\sum_{k=0}^{\infty} b^k$, поэтому функция $\phi(x)$ определена и непрерывна при всех вещественных x . Тем не менее, эта функция не имеет производной по крайней мере при $ab > \frac{3}{2}\pi + 1$. Это связано с сильными осцилляциями на всех масштабах. График функции Вейерштрасса обладает самоподобием и является фракталом. Функция Вейерштрасса сыграла ключевую роль в развитии анализа, показав, что непрерывность не гарантирует *гладкость*. Позже были построены другие примеры, например, функция Ван дер Вардена.

Из-за подобной возможной патологической природы одномерных функций ϕ_{ji} и ζ_j теорема Колмогорова-Арнольда не использовалась в машинном обучении, считаясь теоретически обоснованной, но практически бесполезной [72] вплоть до 2024-го года [59]. В работе [59] была предложена глубокая нейронная сеть (2.24), состоящая из слоёв Колмогорова-Арнольда (2.26). Гладкость функций Колмогорова-Арнольда была обеспечена их представлением в виде линейных комбинаций дифференцируемых базисных функций – В-сплайнов, при этом, подобно многослойным персептронам, глубина сети была обобщена с исходных двух слоёв до L , а ширина с исходных $2n + 1$ нейронов в скрытом слое до произвольных m , где L и m – гиперпараметры сети, требующие настройки.

Итак, сеть Колмогорова-Арнольда задаётся следующим образом:

$$KAN(\mathbf{x}) = (\Phi_L \circ \Phi_{L-1} \circ \dots \circ \Phi_1)(\mathbf{x}), \quad (2.24)$$

где $\{\Phi_l\}_{l=1}^L$ – слои Колмогорова-Арнольда, вычисляющие значение очередного нейрона как:

$$x_{l,j} = \sum_{i=1}^n \phi_{l,ji}(x_{l-1,i}) \quad (2.25)$$

$$\begin{bmatrix} x_{l,1} \\ x_{l,2} \\ \vdots \\ x_{l,m} \end{bmatrix} = \begin{bmatrix} \phi_{l,11}(\cdot) & \phi_{l,12}(\cdot) & \dots & \phi_{l,1n}(\cdot) \\ \phi_{l,21}(\cdot) & \phi_{l,22}(\cdot) & \dots & \phi_{l,2n}(\cdot) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_{l,m1}(\cdot) & \phi_{l,m2}(\cdot) & \dots & \phi_{l,mn}(\cdot) \end{bmatrix} \cdot \begin{bmatrix} x_{l-1,1} \\ x_{l-1,2} \\ \vdots \\ x_{l-1,n} \end{bmatrix}, \quad (2.26)$$

где $\phi_{l,ji}$ – функции Колмогорова-Арнольда (2.20) (см. Рис. 2.7).

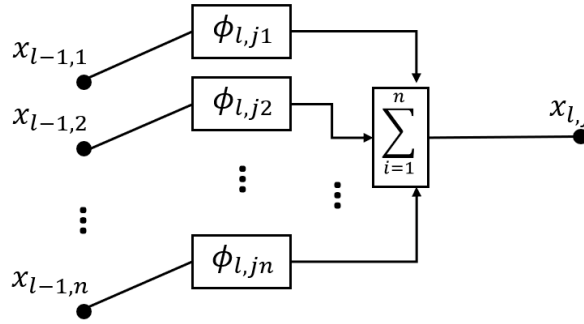


Рисунок 2.7 — Схема вычисления очередного нейрона $x_{l,j}$ слоем Φ_l сети Колмогорова-Арнольда.

Определение сети KAN схоже с многослойным персептроном:

$$MLP(\mathbf{x}) = (W_L \circ \sigma \circ W_{L-1} \circ \dots \circ \sigma \circ W_1)(\mathbf{x}), \quad (2.27)$$

где $\{W_l\}_{l=1}^L$ – линейные слои, вычисляющие значение очередного нейрона как:

$$x_{l,j} = \sum_{i=1}^n \omega_{l,ji} x_{l-1,i} \quad (2.28)$$

$$\begin{bmatrix} x_{l,1} \\ x_{l,2} \\ \vdots \\ x_{l,m} \end{bmatrix} = \begin{bmatrix} \omega_{l,11} & \omega_{l,12} & \dots & \omega_{l,1n} \\ \omega_{l,21} & \omega_{l,22} & \dots & \omega_{l,2n} \\ \vdots & \vdots & \vdots & \vdots \\ \omega_{l,m1} & \omega_{l,m2} & \dots & \omega_{l,mn} \end{bmatrix} \cdot \begin{bmatrix} x_{l-1,1} \\ x_{l-1,2} \\ \vdots \\ x_{l-1,n} \end{bmatrix}, \quad (2.29)$$

где $\omega_{l,ji}$ – веса (2.21), связывающие нейроны на l -ом и $(l-1)$ -ом слоях.

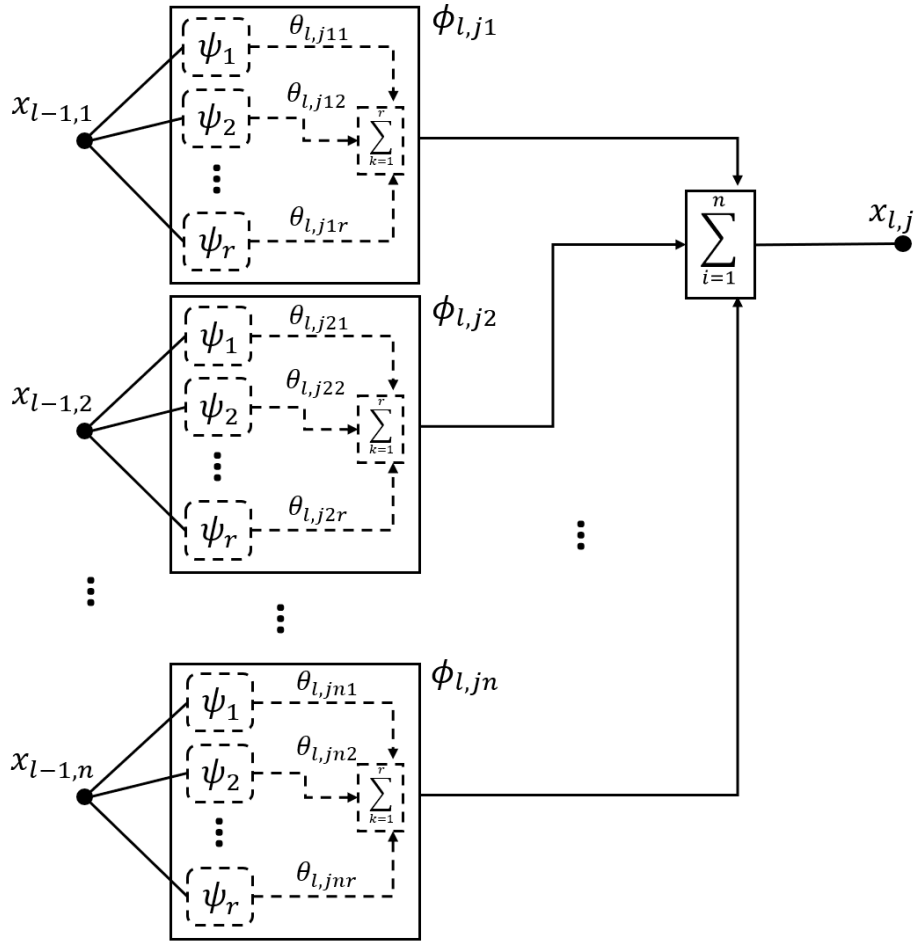


Рисунок 2.8 — Схема вычисления очередного нейрона $x_{l,j}$ слоем Φ_l сети Колмогорова-Арнольда через представление функций Колмогорова-Арнольда в виде линейных комбинаций дифференцируемых базисных функций $\{\psi_k\}_{k=1}^r$ с коэффициентами $\theta_{l,jik}$, формирующими матрицу $\Theta_l \in \mathbb{R}^{m \times n \times r}$ весов l -го слоя, где m — число нейронов в l -ом слое, n — число нейронов в $(l - 1)$ -ом слое, r — число базисных функций.

Необходимость применения градиентного спуска при обучении метода накладывает ограничение дифференцируемости на представление функций Колмогорова-Арнольда, так что на практике они строятся в виде линейной комбинации В-сплайнов [59] или других дифференцируемых базисных функций [73] $\{\psi_k\}_{k=1}^r$ (см. Рис. 2.8):

$$x_{l,j} = \sum_{i=1}^n \sum_{k=1}^r \theta_{l,jik} \psi_k(x_{l-1,i}), \quad (2.30)$$

где коэффициенты $\theta_{l,jik}$ образуют матрицу весов $\Theta_l \in \mathbb{R}^{m \times n \times r}$ очередного слоя.

В качестве $\{\psi_k\}_{k=1}^r$ обычно выбирают первые r функций некоторой полной системы в пространстве непрерывных функций, например, первые r полино-

мов Чебышева [74] или r функций Эрмита. Причём выбор функций Эрмита для анализа сигналов не случаен [75]. При исследовании и решении операторных уравнений математической физики (запишем формально такое уравнение как $Az = u$), большую информацию, соответствующую физике изучаемых явлений, содержат собственные функции оператора A (или оператора A^*A в случае несамосопряжённости оператора) для рассматриваемого функционального пространства. Это и обуславливает использование функций Эрмита, являющихся собственными функциями преобразования Фурье в L_2 , в методах обработки и анализа сигналов, во многих случаях основанных на использовании аппарата преобразования Фурье.

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.



Рисунок 2.9 — Признаковое описание изображения МРТ головного мозга, состоящее из n функций, значения каждой из которых известны на сетке $h \times w$.

Адаптировать сети Колмогорова-Арнольда под задачи обработки изображений возможно разными способами: во-первых, выпрямляя (*англ.* flatten) признаковые тензора перед их подачей в KAN [76]; во-вторых, используя свёрточные версии KAN [77; 78]; в-третьих, и это тот путь, который мы выберем, обобщая сети Колмогорова-Арнольда на бесконечномерные функциональные пространства. Теорема Колмогорова-Арнольда формулируется (2.20) для непрерывных функций многих переменных $f(x_1, \dots, x_n)$, $x_i \in \mathbb{R}$. В этом разделе выдвигается гипотеза об обобщении исходной теоремы на случай непрерывных функционалов $f(\chi_1, \dots, \chi_n)$, заданных на H^n , где каждый χ_i является элементом гильбертова пространства H . Хотя на данный момент это обобщение не имеет математического доказательства, оно проверено экспериментально на медицинских данных

различных модальностей, демонстрируя наилучшие результаты на сегодняшний день. Потенциальная теоретическая справедливость открывает перспективы новым гибридным методам на основе сетей Колмогорова-Арнольда.

Предположим, что если f – непрерывный функционал на H^n , тогда он может быть представлен в виде композиции линейных непрерывных функционалов, непрерывных функций одной переменной и операции сложения:

$$f(\chi_1, \dots, \chi_n) \rightsquigarrow \sum_j \zeta_j \left(\sum_i \varphi_{ji}(\chi_i) \right), \quad (2.31)$$

где H – гильбертово пространство; $\chi_i \in H$; $\varphi_{ji} \in H^*$ и $\zeta_j: \mathbb{R} \rightarrow \mathbb{R}$.

В отличие от оригинальной сети KAN [59], предлагаемый подход обладает естественной совместимостью со свёрточными нейронными сетями по построению, рассматривая каждую двумерную карту признаков как элемент гильбертова пространства, дискретизированный на пространственной сетке размера $h \times w$ (см. Рис. 2.9). Формирование карты признаков $\chi_{l,j}$ очередного слоя осуществляется посредством выражения:

$$\chi_{l,j} = \sum_{i=1}^n \varphi_{l,ji}(\chi_{l-1,i}), \quad (2.32)$$

где $\varphi_{l,ji} \in H^*$.

Для построения вычислительно эффективного метода воспользуемся теоремой Рисса о представлении [66]. Устанавливаемый ею изоморфизм между гильбертовым пространством H и его сопряжённым H^* позволяет перейти от элементов сопряжённого пространства $\varphi_{l,ji}$ к их представителям $\varphi_{l,j}(\chi_{l-1,i})$ в самом пространстве H . Далее эти внутренние функции параметризуются дифференцируемым способом посредством разложения в ряд Фурье по первым r функциям Эрмита $\{\psi_k\}_{k=1}^r$. Поскольку функции Эрмита являются собственными функциями интегрального преобразования Фурье в $L_2(\mathbb{R})$, применяемая стратегия спектрального усечения позволяет сохранить наиболее информативные моды, что соответствует принципу частотной фильтрации в нейронных операторах Фурье [22]. Таким образом, карта признаков $\chi_{l,j}$ очередного слоя определяется следующим выражением:

$$\chi_{l,j} = \sum_{i=1}^n \sum_{k=1}^r \langle \varphi_{l,j}(\chi_{l-1,i}), \psi_k \rangle \psi_k. \quad (2.33)$$

Путём вынесения индекса j за пределы внутренней суммы предложенное представление естественным образом приводит к свёртке 1×1 , параметризованной весами $W_l = \{\omega_{l,j}\}_{j=1}^m$, $\omega_{l,j} \in \mathbb{R}^{1 \times 1 \times n}$. Такая факторизация позволяет получить итоговую вычислительную формулу слоя функциональной сети Колмогорова-Арнольда (англ. functional Kolmogorov-Arnold network, FunKAN) (см. Рис. 2.10):

$$\chi_{l,j} = \sum_{i=1}^n \omega_{l,j} \left(\sum_{k=1}^r \langle \varphi_l(\chi_{l-1,i}), \psi_{l,k}(\chi_{l-1,i}) \rangle \psi_{l,k}(\chi_{l-1,i}) \right). \quad (2.34)$$

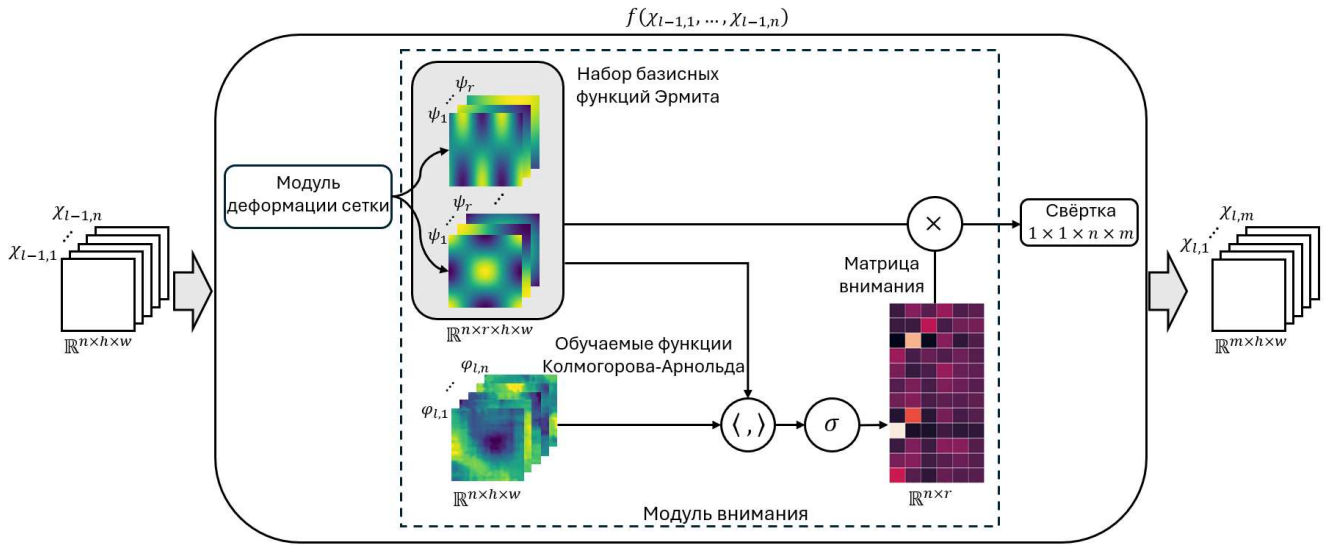


Рисунок 2.10 — Схема слоя функциональной сети Колмогорова-Арнольда.

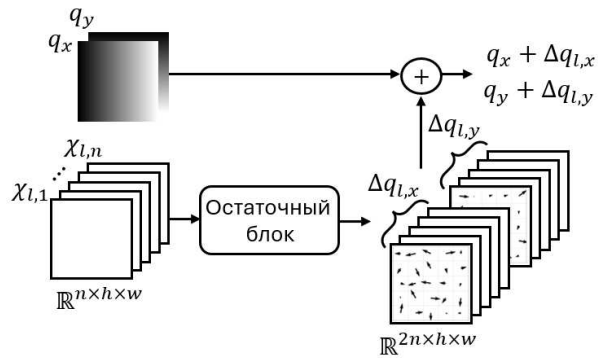


Рисунок 2.11 — Архитектура модуля деформации сетки, иллюстрирующая поле смещений $\{\Delta q_{l,x}, \Delta q_{l,y}\}$, сгенерированное обученным остаточным свёрточным блоком. Предсказанные смещения аддитивно комбинируются с равномерной сеткой $\{q_x, q_y\}$, чтобы создать деформированную сетку для вычисления базисных функций Эрмита.

Вдохновившись современными неявными нейронными сетями [79], вычисление базисных функций Эрмита реализовано (2.34) на адаптивной сетке

с динамической коррекцией координат. Ключевым модулем подхода является модуль деформации (см. Рис. 2.11). Пространственные координаты модифицируются через обучаемое векторное поле смещений $\Delta q_l = \{\Delta q_{l,x}, \Delta q_{l,y}\}$. Поле генерируется остаточным блоком [48]. Деформированная сетка вычисляется как $q + \Delta q_l$, где $q = \{q_x, q_y\}$ равномерная сетка.

Таким образом, структура очередного слоя FunKAN (см. Рис. 2.10) определяется тремя компонентами: тензором обучаемых внутренних функций $\phi_l \in \mathbb{R}^{n \times h \times w}$, множеством базисных функций Эрмита $\psi_l \in \mathbb{R}^{n \times r \times h \times w}$ и матрицей свёрточных весов $W_l \in \mathbb{R}^{1 \times 1 \times n \times m}$.

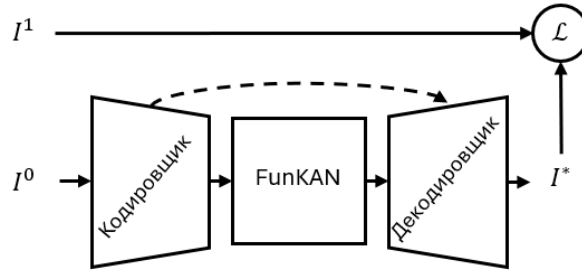


Рисунок 2.12 — FunKAN как многоцелевой метод обработки и анализа медицинских изображений. Обучение обработке входных изображений I^0 контролируется функцией потерь $\mathcal{L}(I^*, I^1)$ и референсными изображениями I^1 .

Функциональная сеть Колмогорова-Арнольда применяется в составе архитектуры (см. Рис. 2.12), во-первых, для задач улучшения качества изображений магнитно-резонансной томографии, во-вторых, для задач сегментации опухолей на ультразвуковых, гистологических и колоноскопических изображениях.

Для задачи подавления осцилляций Гиббса и аддитивного гауссовского шума на изображениях МРТ головного мозга (см. Рис. 2.13) была использована следующая конфигурация:

1. Проецирование: Свёртка 5×5 , преобразующая входное изображение в 16-мерное пространство признаков.
2. Модуль кодирования: Свёртка 3×3 , отображающая предварительно активированные через ReLU признаки в 32-мерное пространство с сохранением пространственного разрешения.
3. Модуль нелинейной фильтрации: Последовательность из трёх блоков FunKAN с пространственными связями в 32-мерном пространстве признаков ($n = 32$), где каждый блок кодирует спектральное представление внутренних функций через первые шесть базисных функций Эрмита ($r = 6$).

4. Модуль декодирования: Свёртка 3×3 , преобразующая предварительно активированные через ReLU признаки в 16-мерное пространство с сохранением пространственного разрешения.
5. Восстановление: Свёртка 1×1 , отображающая предварительно активированные через ReLU признаки в целевое цветовое пространство.

Метод обучался с учителем с функцией потерь среднеквадратической ошибки:

$$\mathcal{L}_{enh} = \frac{1}{N} \sum_{i=1}^N \|I_i^* - I_i^1\|_2^2, \quad (2.35)$$

где N – размер пакета, равный 8.

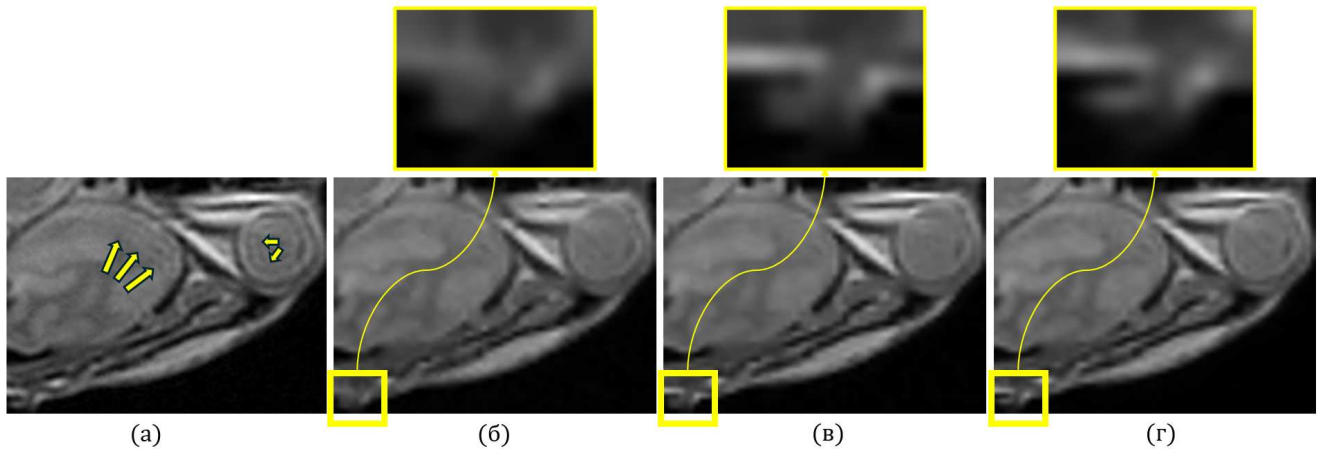


Рисунок 2.13 — Обработка изображения МРТ головного мозга гибридным и классическими методами. (а) – входное изображение с осцилляциями Гиббса (отмечены жёлтыми стрелками) и шумом, (б) – результат алгоритма Кельнера, (в) – результат модели Перона-Малика, (г) – результат гибридного метода, осуществляющего фильтрацию глубоких признаков описаний функциональной сетью Колмогорова-Арнольда по первым шести функциям Эрмита.

Для задачи медицинской сегментации опухолей FunKAN была интегрирована в U-образную сегментационную архитектуру:

1. Проецирование: Свёртка 3×3 , преобразующая входное изображение в 16-мерное пространство признаков.
2. Модуль кодирования: Четыре последовательных остаточных блока в стиле U-Net [15] с прогрессивным увеличением количества фильтров: $32 (C_1) \rightarrow 64 (C_2) \rightarrow 128 (C_3) \rightarrow 128$, где каждый блок уменьшает пространственное разрешение вдвое с помощью свёртки 3×3 с шагом 2 (*англ.* strided convolution).

3. Модуль нелинейной фильтрации: Последовательность из трёх блоков FunKAN с пространственными связями в 128-мерном пространстве признаков ($n = 128$), реализующих спектральное кодирование каждой внутренней функции через первые шесть базисных функций Эрмита ($r = 6$).
4. Модуль декодирования: Четыре последовательных остаточных блока в стиле U-Net с постепенным уменьшением количества фильтров: $128 (C_3) \rightarrow 64 (C_2) \rightarrow 32 (C_1) \rightarrow 16$, где каждый блок выполняет $\times 2$ повышение разрешения методом ближайшего соседа с последующей свёрткой 3×3 для уточнения признаков и пространственной связью с соответствующим блоком кодировщика.
5. Восстановление: Свёртка 1×1 , преобразующая предварительно активированные через ReLU признаки в вероятности.

Метод также обучался с учителем с помощью взвешенной комбинации бинарной перекрестной энтропии (*англ.* binary cross-entropy) и коэффициента Сёренсена (*англ.* Dice):

$$\mathcal{L}_{segm} = \frac{1}{N} \sum_{i=1}^N 0.1 \cdot CE(I_i^*, I_i^1) + Dice(I_i^*, I_i^1), \quad (2.36)$$

где N – размер пакета, равный 8.

Использование базисных функций Эрмита обусловлено их свойством вычислительной локализации, которое следует из их фундаментальной роли собственных функций интегрального преобразования Фурье [80]. Это свойство обеспечивает одновременную локализацию в пространственной и частотной областях, естественное описание локальных особенностей изображений и эффективное представление граничных структур. Количество функций ($r = 6$) определено экспериментально, а также проведён сравнительный анализ трёх базисов: В-сплайны, полиномы Чебышева и функции Эрмита.

2.3 Эксперименты и результаты

Апробация предложенного проекционного метода с использованием сетей Колмогорова-Арнольда была проведена на четырёх наборах медицинских изображений: один набор данных (IXI) повышения качества изображений магнитно-резонансной томографии и три набора данных (BUSI, GlaS, CVC-ClinicDB) сегментации опухолей. Подбор данных осуществлялся с целью проверки устойчивости метода к различным типам анатомических структур (неврологическим, онкологическим, гистологическим и эндоскопическим) и модальностям визуализации (МРТ, ультразвук, гистопатология и колоноскопия).

Экспериментальная платформа была реализована на Python 3.12 с использованием фреймворка PyTorch 2.5. Все вычисления выполнялись с полной точностью (*англ.* floating-point arithmetic) на графическом ускорителе NVIDIA RTX A6000. Программный стек включал PyTorch Lightning 2.5.1, CUDA 11.8 и cuDNN 9. Для обеспечения полной воспроизводимости результатов была реализована фиксация начальных значений генераторов случайных чисел (*англ.* seed) и система управления параметрами экспериментов через YAML-конфигурации. Процесс обучения рассматриваемых методов с нуля (*англ.* from scratch) до сходимости осуществлялся с применением численного метода стохастической оптимизации Adam [51] ($\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$) с ручным планированием темпа обучения (*англ.* learning rate) по схеме: 10^{-4} , $5 \cdot 10^{-5}$, 10^{-5} . Для улучшения обобщающей способности использовались следующие методы аугментации: добавление гауссовского шума ($\sigma = 0.01$) для данных МРТ и случайные отражения, повороты и транспозиции (вероятность 0.5) для наборов сегментации.

Оценка повышения качества изображений МРТ проводилась по метрикам оценки качества изображений PSNR (пиковое отношение сигнал/шум) и TV (полная вариация). Метрика PSNR количественно оценивает близость восстановленного изображения к эталону в среднеквадратичном смысле. Функционал полной вариации косвенно оценивает наличие артефактов, таких как осцилляции Гиббса, контролируя гладкость изображения. Для задач сегментации опухолей применялись IoU, вычисляющий отношение площади пересечения к площади объединения, и F1, оценивающий баланс между точностью и полнотой метода. IoU чувствительнее к перекрытию в пикселях, а F1 – к ошибкам классификации.

Таблица 2 — Сравнительный анализ модулей нелинейной фильтрации в рамках одной свёрточной нейронной сети для повышения качества изображений МРТ. Результаты отражают среднее пиковое отношение сигнал/шум и полную вариацию для 2617 тестовых изображений (145×145) из набора данных IXI. I^0 , I^1 обозначают изображения с артефактами и без артефактов соответственно.

Метод	IXI			
	PSNR (дБ) \uparrow	TV	Gflops \downarrow	Params (M) \downarrow
I^0	31.33	1476.55	–	–
I^1	–	1255.40	–	–
Kellner	31.09	1120.05	–	–
MLP	37.96	1145.57	0.19	0.01
KAN	38.10	1161.63	0.12	0.04
ChebyKAN	38.01	1156.56	0.12	0.03
HermiteKAN	38.04	1161.31	0.12	0.03
FunKAN	39.05	1174.86	3.11	2.2

Набор данных IXI [36] представляет собой общедоступную мультимодальную коллекцию нейровизуализационных данных, включающую примерно 600 высококачественных МРТ сканов здоровых людей, полученных на томографах Philips и GE Healthcare. Датасет содержит 581 объём T1-взвешенных изображений, 578 объёмов T2-взвешенных изображений и 578 объёмов PD-изображений, сохранённых в формате NIFTI – стандартный формат файлов, широко используемый для хранения и обмена данными.

В ходе подготовки данных, для обеспечения согласованности, сначала было выделено пересечение всех модальностей, в результате чего получено 577 объёмов, содержащих полный набор модальностей (T1, T2 и PD). Затем данные были разделены на три подвыборки: обучающую (первые 400 объёмов), тестовую (следующие 100 объёмов) и валидационную (оставшиеся 77 объёмов). В итоге были получены обучающая (10473), валидационная (2016) и тестовая (2617) выборки.

Проведённое количественное сравнение различных архитектур модуля нелинейной фильтрации (см. Таблицу 2) в рамках единой свёрточной структуры для улучшения изображений МРТ на наборе данных IXI выявило следующие закономерности. Замена традиционного многослойного перцептрона (MLP) на сеть Колмогорова-Арнольда (KAN) с последующим переобучением всей сети даёт незначительное улучшение качества – всего 0.1 дБ по метрике PSNR. Напротив, предложенная функциональная сеть Колмогорова-Арнольда (FunKAN) демонстрирует существенно более высокий прирост качества, превосходя стандартные сети KAN на 1 дБ. Это преимущество обусловлено способностью FunKAN учитывать геометрические взаимосвязи в визуальных данных по построению.

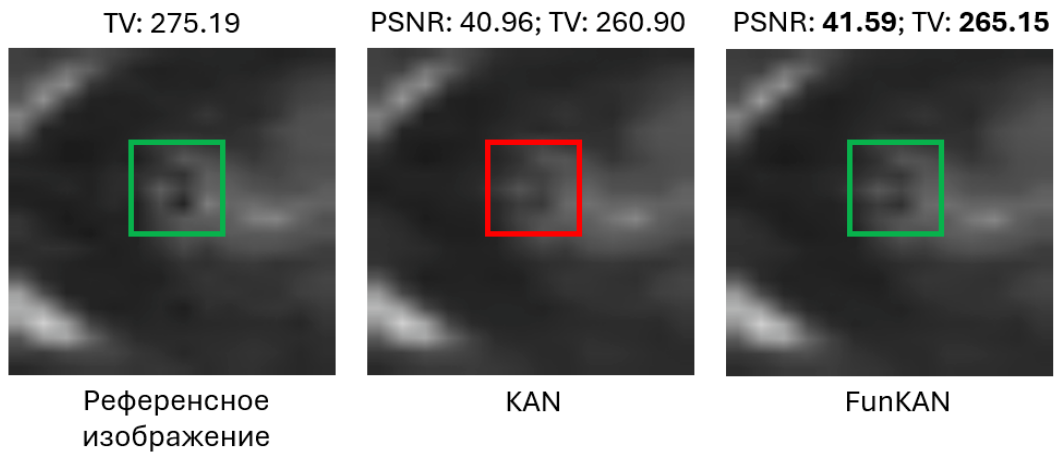


Рисунок 2.14 — Качественное сравнение KAN и FunKAN в рамках одной свёрточной нейронной сети. Выделены тонкие анатомические структуры.

Особое значение имеет способность FunKAN сохранять диагностически важные детали изображений при одновременном эффективном подавлении артефактов, что подтверждается как количественными метриками, так и визуальной оценкой результатов реконструкции (см. Рис. 2.14). Проведённое качественное сравнение выявило, что стандартные сети Колмогорова-Арнольда размывают мелкие структуры, в то время как предложенная функциональная версия сетей Колмогорова-Арнольда сохраняет резкие контуры и высокочастотные детали изображения. Клиническая значимость этих результатов заключается в том, что размытое изображение может маскировать диагностически важные особенности, такие как ранние стадии патологических изменений, микроструктурные аномалии и тонкие морфологические особенности. Таким образом, предложенный подход вносит вклад в развитие методов вычислительной диагностики, предлагая математически обоснованный инструмент для построения точных и интерпретируемых методов повышения качества изображений.

Набор данных BUSI [81] представляет собой общедоступную коллекцию ультразвуковых изображений молочной железы, предназначенную для задач сегментации опухолей. Он включает 780 двумерных изображений в формате PNG, полученных от 600 пациенток в возрасте от 25 до 75 лет. Данные разделены на три категории: 133 изображения без видимых опухолей (нормальные), 437 изображений с доброкачественными образованиями, 210 изображений со злокачественными опухолями (подтвержденными гистологически). Для экспериментов были отобраны 647 изображений (доброкачественные и злокачественные случаи), масштабированные до размера 256×256 пикселей.

Таблица 3 — Сравнение качества методов сегментации в трёх клинически различных медицинских сценариях. Результаты включают средние значения IoU и F1 со стандартным отклонением для трёх случайных запусков обучений.

Метод	BUSI		GlaS		CVC	
	IoU \uparrow	F1 \uparrow	IoU \uparrow	F1 \uparrow	IoU \uparrow	F1 \uparrow
U-Net [15]	57.22 \pm 4.74	71.91 \pm 3.54	86.66 \pm 0.91	92.79 \pm 0.56	83.79 \pm 0.77	91.06 \pm 0.47
Att-Unet [31]	55.18 \pm 3.61	70.22 \pm 2.88	86.84 \pm 1.19	92.89 \pm 0.65	84.52 \pm 0.51	91.46 \pm 0.25
U-Net++ [32]	57.41 \pm 4.77	72.11 \pm 3.90	87.07 \pm 0.76	92.96 \pm 0.44	84.61 \pm 1.47	91.53 \pm 0.88
U-NeXt [84]	59.06 \pm 1.03	73.08 \pm 1.32	84.51 \pm 0.37	91.55 \pm 0.23	74.83 \pm 0.24	85.36 \pm 0.17
Rolling-UNet [85]	61.00 \pm 0.64	74.67 \pm 1.24	86.42 \pm 0.96	92.63 \pm 0.62	82.87 \pm 1.42	90.48 \pm 0.83
U-Mamba [86]	61.81 \pm 3.24	75.55 \pm 3.01	87.01 \pm 0.39	93.02 \pm 0.24	84.79 \pm 0.58	91.63 \pm 0.39
UKAGNet [78]	63.45	77.64	87.31	93.23	76.85	86.91
U-KAN [76]	63.38 \pm 2.83	76.40 \pm 2.90	87.64 \pm 0.32	93.37 \pm 0.16	85.05 \pm 0.53	91.88\pm0.29
U-FunKAN	68.49\pm0.62	77.37 \pm 0.58	88.02\pm0.24	93.50\pm0.12	85.93\pm0.72	91.42 \pm 0.61

Таблица 4 — Сравнение вычислительной эффективности методов сегментации. Результаты получены с помощью инструмента THOR Python для профилирования моделей PyTorch.

Метод	Gflops \downarrow	Params (M) \downarrow
U-Net [15]	524.2	34.53
Att-Unet [31]	533.1	34.9
U-Net++ [32]	1109	36.6
U-NeXt [84]	4.58	1.47
Rolling-UNet [85]	16.82	1.78
U-Mamba [86]	2087	86.3
U-KAN [76]	14.02	6.35
U-FunKAN	4.35	3.6

Набор данных GlaS [82] является широко используемым специализированным ресурсом для задач сегментации желёз на гистологических изображениях. Он включает 165 RGB изображений гистологических препаратов, окрашенных гематоксилином и эозином (H&E). Все изображения были масштабированы до разрешения 512×512 пикселей. Снимки в данном наборе данных были получены с помощью аппаратов Olympus Q160AL и Q165L, в комплекте с видеообработывающим устройством Extra II.

Набор данных CVC-ClinicDB [83] предназначен для задач сегментации полипов на колоноскопических изображениях. Он включает 612 высококачественных RGB кадров, полученных из 29 видео-последовательностей, которые характеризуются различными условиями освещения, наличием зеркальных бликов, разнообразием текстур слизистой оболочки. Все изображения были масштабированы до разрешения 256×256 пикселей для обеспечения единообразия обработки.

Проведённые эксперименты (см. Таблицу 3) демонстрируют, что разработанный гибридный метод U-FunKAN устанавливает новый стандарт точности сегментации медицинских изображений различных типов, включая ультразвуковые, гистологические и колоноскопические изображения, обеспечивая также и вычислительную эффективность (см. Таблицу 4). Особое внимание было уделено надёжности оценок – все метрики рассчитывались на основе трёх независимых запусков с различными значениями генераторов случайных чисел (50, 100, 150), что обеспечило воспроизводимость и статистическую достоверность результатов.

Действительно, на трёх медицинских наборах данных метод показал наивысшие показатели по метрике IoU, подтверждённые статистически значимыми результатами, усреднёнными по последним 50 эпохам обучения. По показателю F1 разработанный метод продемонстрировал наилучший результат на наборе данных GlaS, предназначенном для сегментации желёз на гистологических изображениях, при этом незначительно уступив методу UKAGNet на ультразвуковых данных BUSI. В задачах колоноскопической сегментации на данных CVC-ClinicDB предложенный метод показал несколько меньшие значения F1 по сравнению с U-KAN.

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 превзошёл 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 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.

Таблица 5 — Исследование влияния масштабирования каналов в U-FunKAN на качество сегментации и эффективность. Приведены наилучшие значения метрик IoU и F1.

U-FunKAN			BUSI			
C_1	C_2	C_3	IoU \uparrow	F1 \uparrow	Gflops \downarrow	Params (M) \downarrow
32	64	128	69.11	77.95	4.35	3.6
64	96	128	69.94	78.42	10.84	4.1
128	160	256	69.49	78.39	40.42	15.7
256	320	512	70.62	79.31	161.43	62.4

Наблюдаемый профиль метрик IoU и F1, сочетающий высокий IoU на всех трёх наборах данных с относительно сниженной F1-мерой, свидетельствует о специфическом балансе ошибок. Такое соотношение метрик указывает на минимизацию ложноположительных детекций (что непосредственно отражается в IoU) в ущерб некоторому росту ложноотрицательных (что учитывается F1-мерой). Более глубокие версии U-FunKAN позволяют ослабить данный компромисс, демонстрируя лучшие результаты как по IoU, так и по F1 одновременно, например, на ультразвуковых данных BUSI (см. Таблицу 5).

Ключевое преимущество проекционных сетей Колмогорова-Арнольда проявляется при анализе вычислительной эффективности (см. Таблицу 4): U-FunKAN требует в 3 раза меньше операций с плавающей точкой и содержит на 43 % меньше обучаемых параметров по сравнению с аналогами. Это делает U-FunKAN практически применимым методом, потенциально адаптируемым для развертывания на оборудовании с ограниченными ресурсами.

На Рис. 2.15 представлены результаты исследования, оценивающего влияние количества базисных функций Эрмита на качество сегментации. Для обеспечения корректного сравнения все методы обучались с нуля со скоростью обучения 10^{-4} до сходимости на наборе ультразвуковых данных BUSI. Результаты показывают, что обе метрики, IoU и F1, растут при увеличении числа базисных функций со значения по умолчанию, равного 6, до 8 и 10. Однако это дополнительное улучшение точности достигается за счёт увеличения времени вывода, что детально показано на Рис. 2.16. На Рис. 2.16 представлен график, количественно характеризующий вычислительную сложность предложенного метода сегментации в зависимости от размера используемого функционального базиса. По оси абсцисс отложено количество базисных функций Эрмита, по оси ординат – медианное время вывода (*англ.* inference) на процессоре Intel Core i7-14700HX для одного изображения из набора данных BUSI, измеренное в миллисекундах.

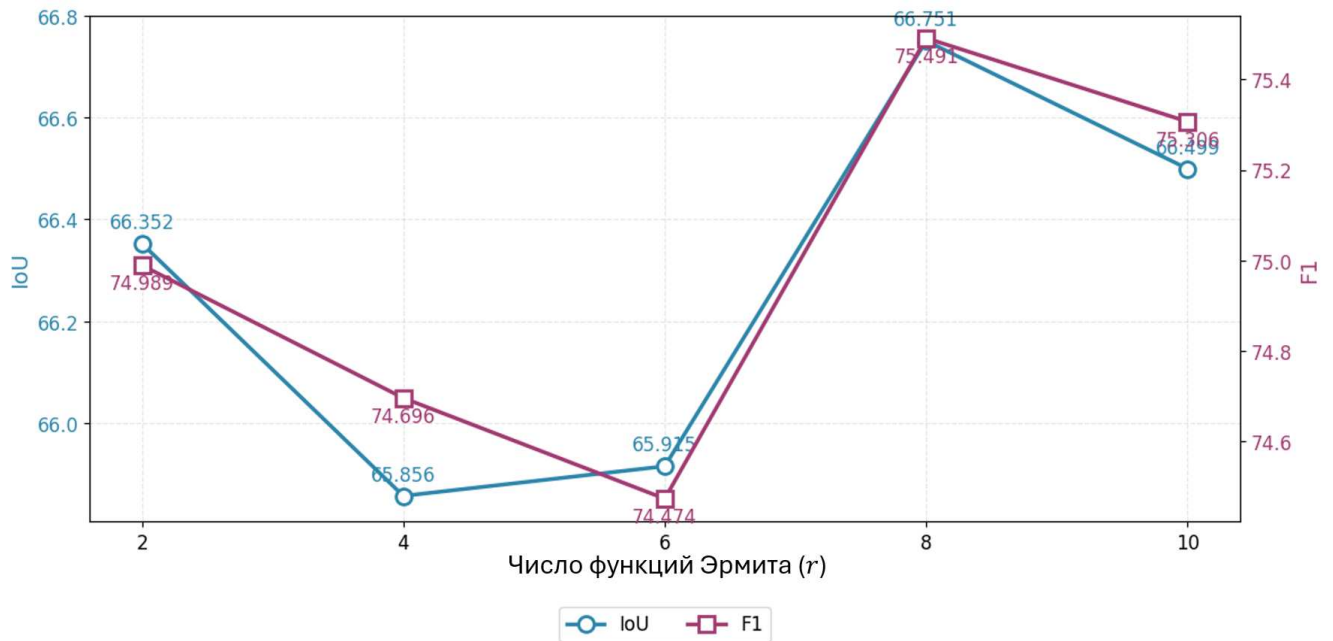


Рисунок 2.15 — Исследование влияния количества базисных функций Эрмита в U-FunKAN на качество сегментации рака молочной железы на ультразвуковом наборе данных BUSI. Для каждого эксперимента приведены усреднённые значения метрик IoU и F1 за последние 50 эпох обучения.

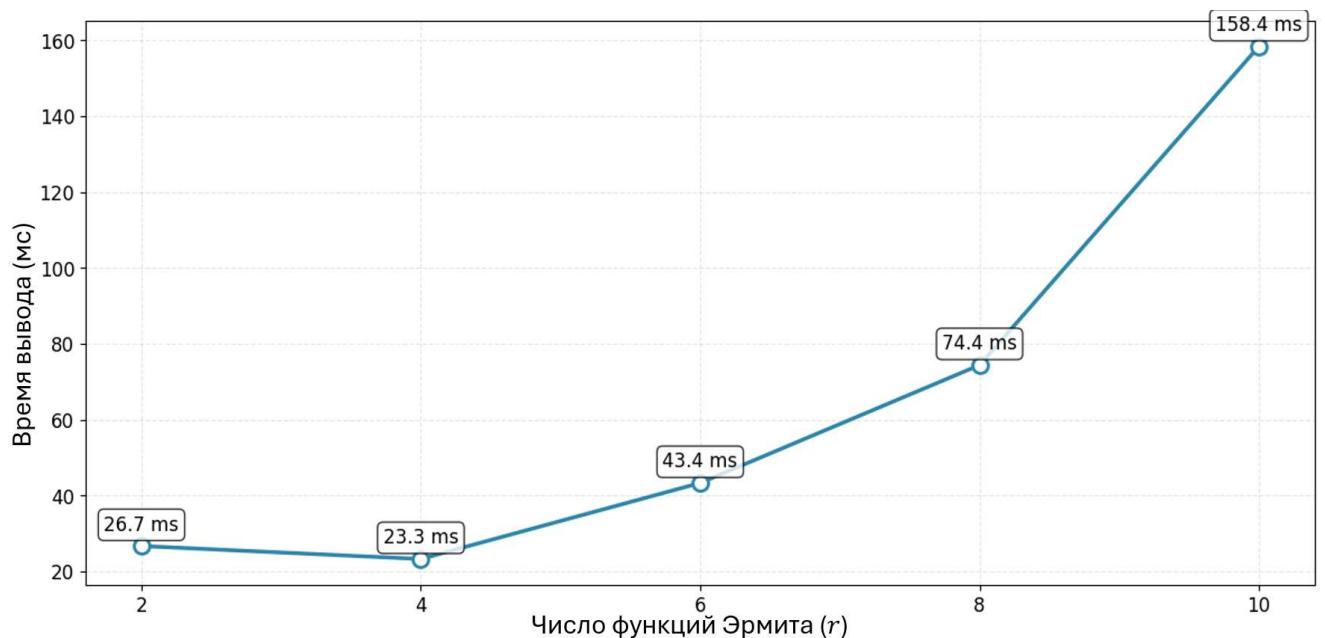


Рисунок 2.16 — Исследование влияния количества базисных функций Эрмита в U-FunKAN на скорость вывода (эффективность). Для каждого эксперимента приведено медианное время вывода на процессоре Intel Core i7-14700HX, усреднённое по 100 запускам для входного изображения размером 256×256 пикселей.

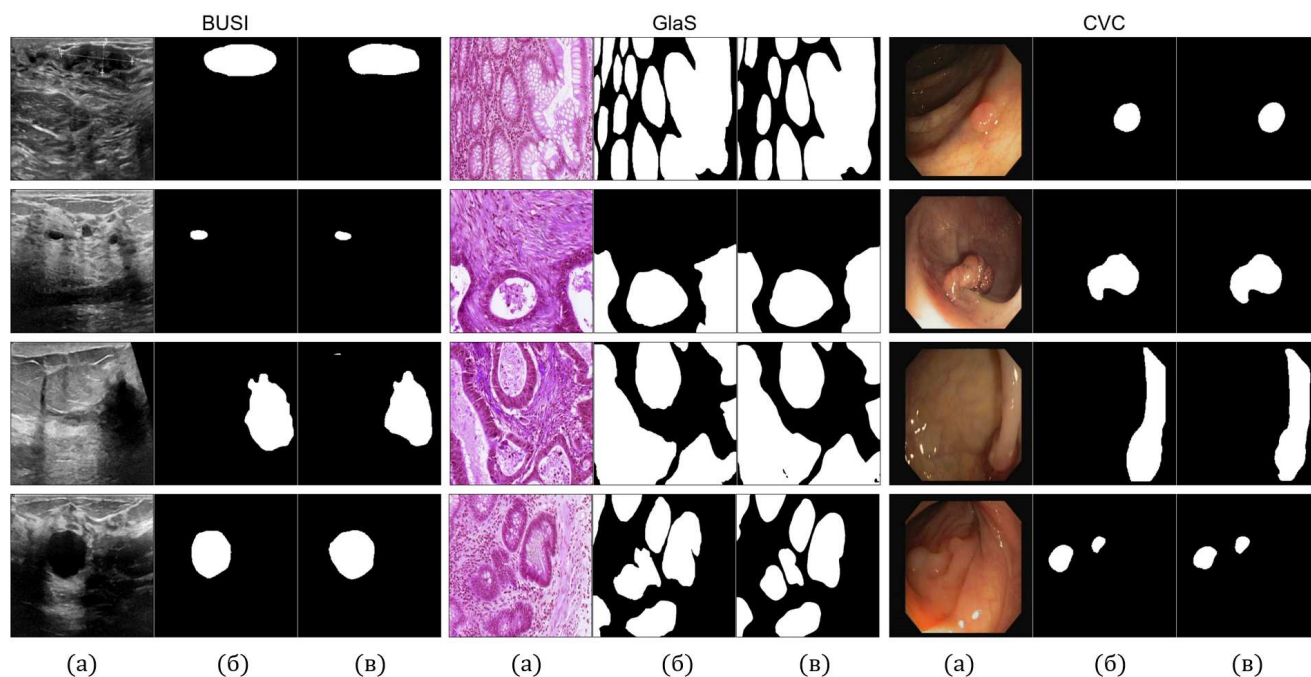


Рисунок 2.17 — Качественные результаты работы U-FunKAN для трёх разнородных медицинских сценариев. (а) – входные изображения, (б) – референсные маски, (в) – результаты предлагаемого метода.

Наблюдается нелинейный рост времени обработки с увеличением числа базисных функций. При переходе от 6 к 8 базисным функциям время вывода увеличивается в 1.7 раза, а от 8 к 10 – более чем в 2 раза. Начальный участок графика демонстрирует нетипичное поведение: время вывода при четырёх функциях Эрмита немного меньше, чем при двух. Данный эффект может быть объяснён особенностями оптимизации вычислительных процессов в используемых библиотеках. График показывает ключевой компромисс между точностью метода и его вычислительной эффективностью. Выбор шести функций Эрмита, принятый по умолчанию, представляет собой точку разумного баланса: время вывода (43.4 мс, что соответствует частоте обработки 23 кадра в секунду) является приемлемым для задач, не требующих строго реального времени вывода (*англ.* real-time inference), при этом обеспечивается удовлетворительное качество сегментации. Использование десяти функций Эрмита, хотя и дает наивысшую точность, приводит к почти четырёхкратному увеличению времени вывода, что может быть неприемлемо для систем, работающих с большими объёмами данных. Полученная зависимость является критически важной для практической реализации алгоритма. Она позволяет оценить требования к вычислительным ресурсам (производительности CPU/GPU) в зависимости от требуемой точности.

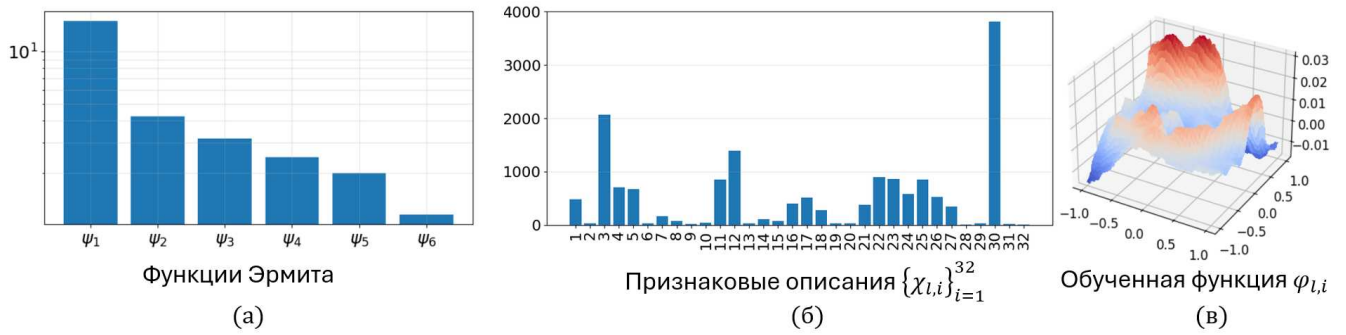


Рисунок 2.18 — Интерпретируемость слоя FunKAN, обученного на наборе данных IXI для улучшения качества изображений МРТ. (а) – агрегированный вклад базисных функций Эрмита по всем обученным внутренним функциям слоя, (б) – характеристика важности признаков, вычисленная как спектральная энергия, (в) – визуализация одной из обученных внутренних функций.

На Рис. 2.17 представлены качественные результаты сегментации, полученные предложенной проекционной сетью, для трёх разнородных медицинских сценариев: обнаружения рака молочной железы на ультразвуковых снимках, сегментации желёз на гистологических изображениях и сегментации полипов на изображениях колоноскопии.

С точки зрения интерпретируемости (см. Рис. 2.18), разложение признаков по базису Эрмита по своей сути даёт спектральную характеристику их сложности. Концентрация спектральной энергии в коэффициентах низкого порядка Эрмита указывает на гладкость внутренних функций, что, как показывают эксперименты, коррелирует с устойчивыми и обобщаемыми признаками. И наоборот, значительная энергия в коэффициентах высокого порядка свидетельствует о склонности к переобучению и повышенной уязвимости к атакам [87].

2.4 Выводы

Результаты демонстрируют, что нейронные операторы, в частности, нейронный оператор Фурье и проекционные сети Колмогорова-Арнольда, устанавливают новый стандарт в обработке медицинских изображений, обеспечивая баланс между подавлением артефактов, сохранением диагностически ценной информации и точностью распознавания анатомических структур. Такие характеристики делают предложенные гибридные методы особенно ценными для клинической практики, где точность визуализации напрямую влияет на качество диагностики. Важно, что достижение конкурентоспособных показателей сопровождается существенным снижением требований к аппаратному обеспечению. Это делает нейронные операторы перспективным решением для интеграции в системы медицинской диагностики.

Глава 3. Методы выбора масштаба анализа гистологических изображений предобученными нейронными сетями

Значительные технологические достижения привели к появлению инновационных методов цифровой визуализации в гистопатологии – разделе патологической анатомии, изучающем структурные и функциональные изменения тканей при различных заболеваниях под микроскопом для диагностики, классификации и понимания природы заболеваний.

Поиск решений известных проблем световой микроскопии таких как: затраты на подготовку и хранение препаратов, потеря качества препаратов со временем, сложность организации дистанционной работы, – вызвал стремительную эволюцию взаимодействия с гистологическими препаратами, и за последние 100 лет был осуществлён переход от световой микроскопии к цифровой [88; 89].

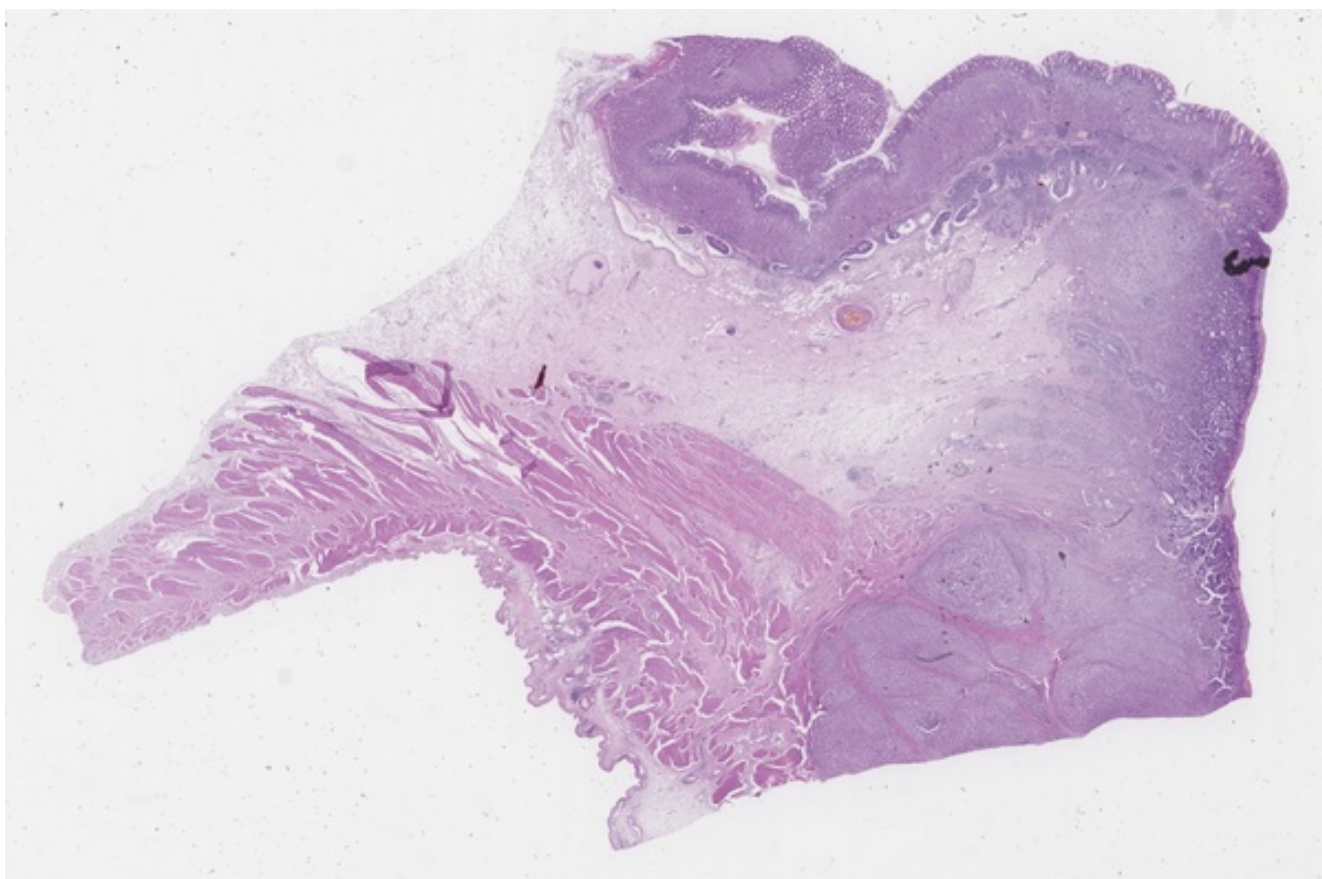


Рисунок 3.1 — Полнослайдовое гистологическое изображение стенки желудка из набора данных PATH-DT-MSU [<https://imaging.cs.msu.ru/en/research/histology/path-dt-msu>]. Снимок сделан на оптическом увеличении $\times 40$ с разрешением 65809×99600 пикселей.

На данный момент одним из передовых решений является сбор и анализ гистологических данных в виде полнослайдовых изображений (*англ.* whole slide imaging, WSI) [90; 91] (см. Рис. 3.1), представляющих цифровую копию всего гистологического или цитологического стекла, полученную с помощью специальных сканеров. В отличие от традиционной микроскопии, где врач изучает препарат через окуляр, технология WSI позволяет анализировать весь образец на компьютере с возможностью масштабирования и автоматизации обработки.

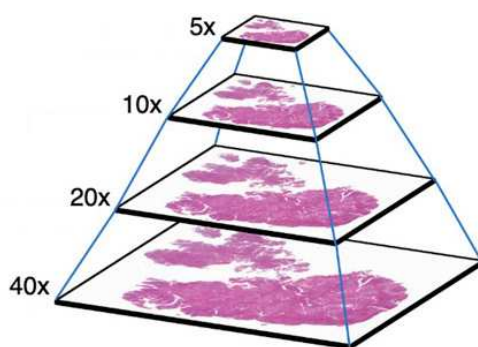


Рисунок 3.2 — Визуализация многомасштабной структуры полнослайдового изображения, включающего несколько уровней оптического увеличения.

В большинстве случаев полнослайдовое изображение – это многомасштабная пирамида из нескольких уровней (см. Рис. 3.2). Нижний уровень содержит изображение образца в максимальном оптическом увеличении. Очередной уровень пирамиды хранит последовательно уменьшенную копию нулевого уровня, вплоть до небольшого обзорного снимка. Каждый уровень, включая нулевой, представляет собой композицию, составленную из небольших фрагментов квадратной формы, видимые границы между которыми отсутствуют. Такая структура представления обеспечивает высокую скорость доступа к любому участку полнослайдового изображения на желаемом оптическом увеличении (масштабе). Хранятся [92] полнослайдовые изображения в различных форматах, например, в форматах SVS и BigTIFF.

Несмотря на перспективность технологии WSI, её внедрение в гистопатологическую практику сталкивается с рядом ограничений и проблем [93]. Основным недостатком WSI по сравнению со световой микроскопией является критическая необходимость в компьютерной автоматизированной обработке, без которой практическое использование полнослайдовых изображений в ручном режиме может оказаться неэффективным, дорогостоящим и требующим неприемлемо больших временных затрат. По сравнению с традиционной микроскопией могут также возникать артефакты сканирования, такие как размытие, нерезкость и

блики (особенно при автоматическом фокусировании). В некоторых странах, в частности в США, полнослайдовые цифровые микроскопические исследования ограничены в использовании для первичной диагностики онкологии и применяются лишь для вторичного подтверждения диагноза. Тем не менее, технология WSI продолжает набирать популярность [94] среди патологоанатомов в диагностических, образовательных и исследовательских целях.

В настоящее время в области компьютерной обработки гистологических изображений большое распространение приобрели нейронные сети [95]. Так, примером успешного применения свёрточных сетей для классификации полнослайдовых изображений мозга является работа [96], в которой точность классификации нейронными сетями ResNet50 [97] и VGG19 [98] достигла 90 %. Эффективность таких методов во многом зависит от выбора масштаба обработки гистологических срезов, по построению характеризующихся высокой вариативностью размеров структур и разрешений. Слишком мелкий масштаб может привести к потере контекстной информации, а слишком крупный – к избыточной детализации и повышению шума. В то же время, свёрточные нейронные сети известны низкой устойчивостью к изменению масштаба входных данных [38].

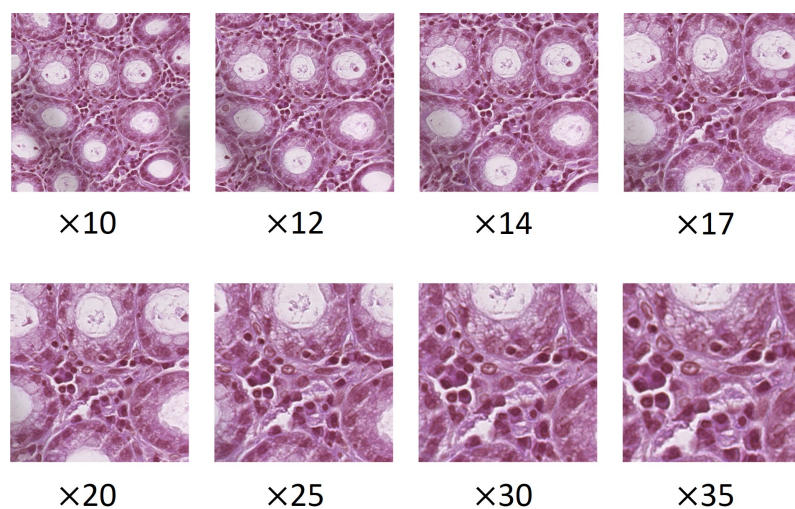


Рисунок 3.3 — Участок полнослайдового гистологического изображения в различных оптических увеличениях (масштабах).

Таким образом, задача выбора масштаба (см. Рис. 3.3) анализа и обработки гистологических изображений, оптимального для выбранной предобученной нейронной сети, важна как для автоматизации анализа полнослайдовых изображений, так и для разработки комплексных систем [99].

3.1 Нейронная классификация полнослайдовых изображений стенок желудка

Классификация полнослайдовых изображений стенок желудка проводится с учётом гистологических, патологических и молекулярных особенностей тканей.

На практике решаются следующие задачи классификации стенок желудка.

1. По гистологическим признакам.

- Нормальная ткань:
 - слизистая оболочка (эпителий, железы, ямки),
 - подслизистая основа,
 - мышечный слой,
 - серозная оболочка.
- Воспалительные изменения:
 - хронический гастрит,
 - острый гастрит,
 - лимфоцитарный/эозинофильный гастрит.
- Предраковые изменения:
 - атрофия слизистой,
 - кишечная метаплазия (полная/неполная),
 - дисплазия (низкая/высокая степень).

2. По онкологическим патологиям.

- Аденокарцинома желудка:
 - кишечный тип,
 - диффузный тип,
 - смешанный тип.
- Редкие опухоли:
 - нейроэндокринные (карциноиды),
 - лимфомы (MALT-лимфома),
 - мезенхимальные опухоли (GIST).
- Предраковые изменения:
 - атрофия слизистой,
 - кишечная метаплазия (полная/неполная),
 - дисплазия (низкая/высокая степень).

Высокое разрешение полнослайдовых изображений затрудняет применение методов глубокого обучения *напрямую*, подобно классификации обычных изображений, к примеру, нейронными сетями ResNet50 [97], VGG19 [98], InceptionNet [100], DenseNet121 [101]. Поэтому типичный нейросетевой классификатор опухолей желудка осуществляет так называемую *по-патчевую* классификацию, в ходе которой очередное полнослайдовое изображение разделяется на небольшие квадратные фрагменты, каждому из которых нейронный классификатор присваивает метку соответствующего класса. Затем полученный набор классифицированных фрагментов склеивается в единое изображение (см. Рис. 3.4). С вычислительной точки зрения, эта процедура представляет собой алгоритм скользящего окна. Таким образом, получается грубая карта *сегментации*, точность которой можно увеличить, уменьшая сдвиг окна (*англ. stride*) и усредняя предсказания. Данная методика, по сути, решает задачу семантической сегментации опосредованно, через классификацию отдельных участков.

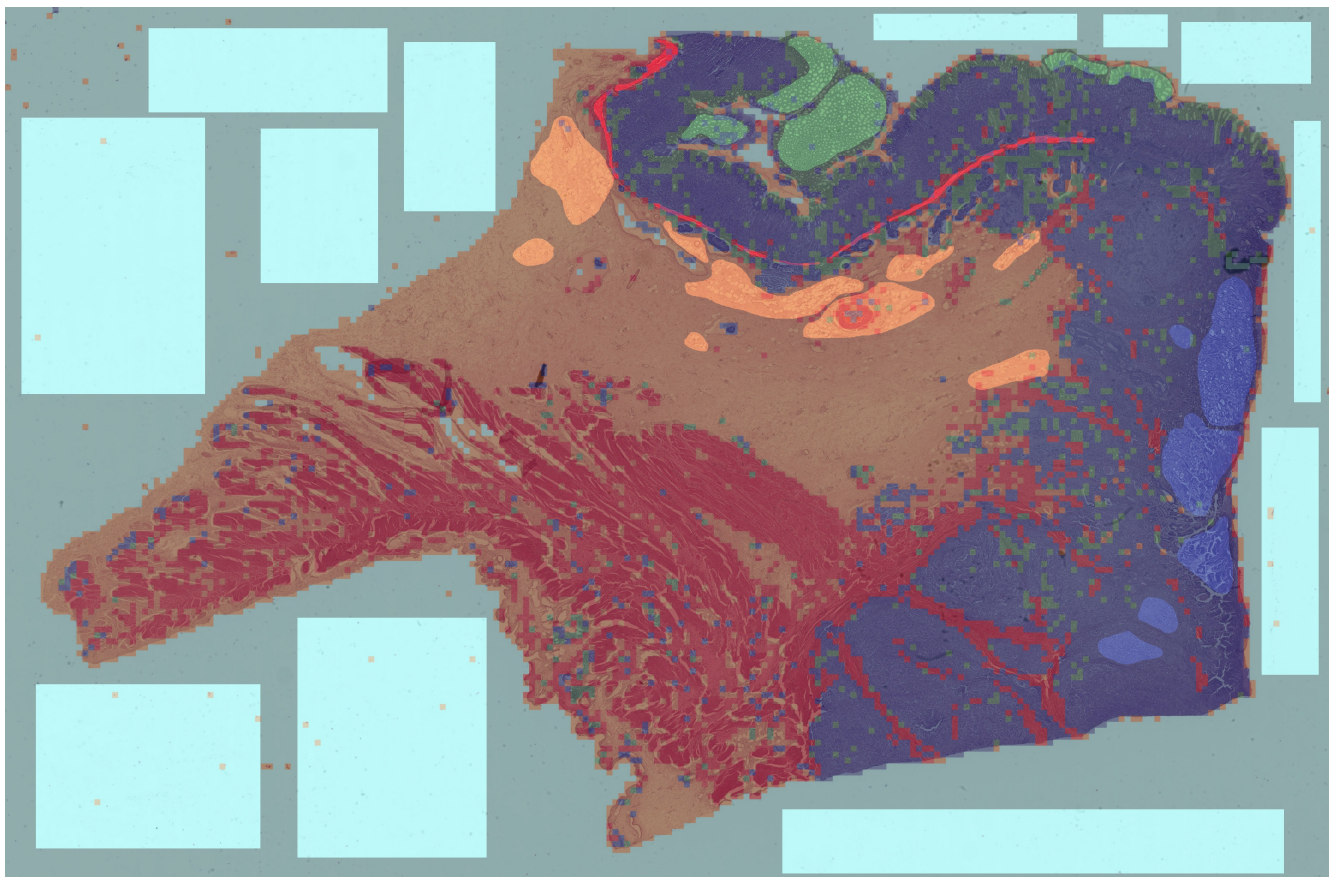


Рисунок 3.4 — Результат классификации полнослайдового гистологического изображения стенки желудка свёрточной нейронной сетью DenseNet121. Увеличенной яркостью выделены области, размеченные специалистами.

В данном разделе будет представлен разработанный метод автоматического выбора масштаба классификации полнослайдовых изображений стенок желудка предобученным нейронным классификатором. В качестве такого нейронного классификатора выбрана свёрточная нейронная сеть DenseNet121 [101], предобученная классификации пяти типов тканей, имеющихся в наборе данных PATH-DT-MSU (WSS1, WSS2):

- участки подслизистой основы (AT),
- фон изображения (BG),
- неизменённые участки собственной пластинки слизистой оболочки (LP),
- неизменённые участки мышечной пластинки слизистой оболочки (MM),
- участки аденокарциномы желудка (TUM).

Обучающая и валидационная выборки для предобучения нейронного классификатора были сформированы из фрагментов (224×224) полнослайдовых изображений на масштабе $\times 20$.

В Таблице 6 приведены значения точности классификации на оптическом увеличении $\times 20$, на котором велось обучение, на меньшем увеличении – $\times 10$ и на большем увеличении – $\times 40$. При выборе масштаба $\times 10$, отличного от оптимального для выбранного метода, наблюдается значительное снижение точности распознавания класса LP (на 20.7 %) и, как следствие, падение общей точности классификации (на 3 %). А при выборе масштаба $\times 40$, отличного от оптимального, падение точности классификации ещё более существенно и наблюдается уже на четырёх классах: AT (на 5.6 %), LP (на 67.4 %), MM (на 8.4 %) и TUM (на 3.5 %), – что приводит к падению общей точности классификации на 17 %.

Таблица 6 — Точность (%) классификации фрагментов (224×224) полнослайдового изображения стенки желудка в зависимости от выбранного масштаба входных данных. Асс – общая точность, Асс@AT – точность распознавания участков подслизистой основы, Асс@BG – точность распознавания фона, Асс@LP – точность распознавания неизменённых участков собственной пластинки слизистой оболочки, Асс@MM – точность распознавания неизменённых участков мышечной пластинки слизистой оболочки, Асс@TUM – точность распознавания участков аденокарциномы.

Оптическое увеличение	Асс ↑	Асс@AT ↑	Асс@BG ↑	Асс@LP ↑	Асс@MM ↑	Асс@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

3.2 Метод автоматического определения масштаба обработки полнослайдовых изображений на основе уверенности нейронного классификатора

На практике, в отличие от рассмотренного примера, информация об оптимальном масштабе классификации для выбранной нейронной сети может быть недоступна, или же может быть утеряна информация об оптическом увеличении, на котором было получено анализируемое полнослайдовое изображение.

Нейронный классификатор можно разделить на два структурных модуля:

- модуль нелинейного отображения Φ фрагментов $\varepsilon \times \varepsilon$,
- модуль разделяющих гиперплоскостей $W \in \mathbb{R}^{d \times C}$.

Решение о выборе масштаба обработки полнослайдового изображения предлагается принимать, анализируя расстояния до разделяющих гиперплоскостей нейронного классификатора при различных масштабах входных данных (см. Алгоритм 1). Метод позволяет количественно оценить уверенность метода в классификации при разных уровнях увеличения и выбрать масштаб, обеспечивающий наиболее достоверное предсказание.

На вход предлагаемому методу подаются полнослайдовое изображение, предобученный нейронный классификатор, предназначенный для анализа фрагментов гистологических изображений, а также набор масштабов $\mathcal{S} = \{s_1, \dots, s_N\}$, среди которых осуществляется выбор масштаба обработки, где s_1 соответствует наименьшему оптическому увеличению, обеспечивающему обзор на уровне ткани, а s_N – наибольшему, позволяющему детально анализировать структуры.

На полнослайдовом изображении I^{WSI} случайным образом выбирается P точек $\mathcal{X} = \{x_1, \dots, x_P\}$. Вокруг очередной точки x_k на каждом рассматриваемом масштабе s_i вырезается квадратный фрагмент $U^\varepsilon(x_k(s_i))$ формы $\varepsilon \times \varepsilon$. Соответствующие признаковые описания $\{\Phi(U^\varepsilon(x_k(s_i)); \theta)\}_{i=1}^N$ многомасштабной окрестности точки x_k образуют матрицу объектов-признаков $F_k \in \mathbb{R}^{N \times d}$. Обработка очередной точки x_k завершается вычислением матрицы расстояний $D_k \in \mathbb{R}^{N \times C}$ от многомасштабных фрагментов до каждой из C разделяющих гиперплоскостей нейронного классификатора. Масштаб, отвечающий наибольшему расстоянию, сохраняется в множество предлагаемых масштабов $\hat{\mathcal{S}}$, и алгоритм переходит к анализу многомасштабных фрагментов вокруг следующей точки множества \mathcal{X} .

Вход : I^{WSI} – полнослайдовое изображение.

$\Phi(\cdot; \theta)$ – модуль нелинейного отображения.

$W \in \mathbb{R}^{d \times C}$ – матрица разделяющих гиперплоскостей.

```

1  $\mathcal{X} \leftarrow \{x_k\}_{k=1}^P: x_k \in I^{WSI}$ 
2  $\mathcal{S} \leftarrow \{s_i\}_{i=1}^N: s_i \in \mathbb{Q}^+$ 
3  $\hat{\mathcal{S}} \leftarrow \{\}$ 
4 for  $k \leftarrow 1$  to  $P$  do
5    $F_k \leftarrow \{\}$ 
6   for  $i \leftarrow 1$  to  $N$  do
7      $F_k \leftarrow F_k \cup \Phi(U^\epsilon(x_k(s_i)); \theta)$ 
8    $F_k \in \mathbb{R}^{N \times d} \leftarrow \text{Concat}(F_k)$ 
9    $D_k \in \mathbb{R}^{N \times C} \leftarrow \frac{F_k \cdot W}{\|W\|}$ 
10   $\hat{\mathcal{S}} \leftarrow \hat{\mathcal{S}} \cup \text{argmax}(\max(D_k, \text{axis}=-1))$ 
11  $\hat{s} \leftarrow Mo[\hat{\mathcal{S}}]$ 

```

Выход: \hat{s} – предлагаемый масштаб обработки изображения I^{WSI} .

Алгоритм 1 — Определение масштаба обработки полнослайдового гистологического изображения.

Решение о масштабе обработки полнослайдового изображения I^{WSI} принимается по моде, то есть голосованием, по множеству предлагаемых масштабов $\hat{\mathcal{S}}$.

Разработанный метод мотивирован геометрическим смыслом линейного классификатора $\langle \omega, x \rangle$. Геометрически линейный классификатор соответствует гиперплоскости с вектором нормали ω . Величина скалярного произведения $\langle \omega, x \rangle$ пропорциональна расстоянию от гиперплоскости до точки x , а его знак показывает, с какой стороны от гиперплоскости находится данная точка. Таким образом, линейный классификатор разделяет пространство на две части, и при этом одно полупространство он относит к положительному классу, а другое – к отрицательному. Нетрудно показать [102], что расстояние от точки x до разделяющей гиперплоскости ω вычисляется как $\frac{|\langle \omega, x \rangle|}{\|\omega\|}$. В нейронные классификаторы встроены C разделяющих гиперплоскостей, так что вектор ω превращается в матрицу W , столбцы которой задают нормали этих гиперплоскостей, а скалярное произведение – в матричное умножение.

3.3 Эксперименты и результаты

Представленный метод выбора масштаба анализа и обработки гистологических изображений предобученными нейронными сетями реализован на языке программирования Python 3 с использованием библиотеки для работы с полнослайдовыми изображениями OpenSlide.

Метод протестирован на задаче классификации полнослайдовых гистологических изображений стенок желудка из набора данных PATH-DT-MSU (WSS1, WSS2) предобученной на масштабе $\times 20$ свёрточной нейронной сетью DenseNet121. PATH-DT-MSU является специализированным набором данных для задач автоматического анализа гистопатологических изображений, в частности, для выявления и классификации опухолей. Он содержит аннотированные изображения срезов тканей, полученные с помощью микроскопии, с разметкой различных типов опухолей. Датасет включает данные различного разрешения, что позволяет использовать его для алгоритмов компьютерного зрения как на уровне отдельных клеток, так и на уровне тканевых структур. PATH-DT-MSU также содержит клинико-морфологические метаданные, что делает его полезным для исследований на стыке цифровой патологии и предиктивной аналитики. Этот набор данных был создан в Московском государственном университете имени М.В. Ломоносова и находит применение в разработке методов глубокого обучения для поддержки диагностики в нейроонкологии.

Параметры метода (см. Алгоритм 1) в данном случае принимают следующие значения: $d = 1024$, $\varepsilon = 224$, $C = 5$. Множество $\mathcal{S} = \{\times 10, \times 12, \times 14, \times 17, \times 20, \times 25, \times 30, \times 35, \times 43\}$ составлено так, что $\frac{s_{i+1}}{s_i} \approx 1.2$, $i = \overline{1..N-1}$.

Демонстрация применения метода к пяти различным многомасштабным фрагментам, относящихся к пяти различным классам (AT, BG, LP, MM, TUM) и вырезанных из полнослайдового изображения (Рис. 3.1), представлены на Рис. 3.5. Для всех пяти фрагментов наибольший отступ (*англ.* margin) порождается классификатором именно от гиперплоскости, отвечающей референсному классу. Причём наблюдается заметная разница между величинами отступов от референсной гиперплоскости и от всех остальных. Низкая устойчивость распознавания класса LP к изменению масштаба, отмеченная в Таблице 6, находит своё отражение и на Рис. 3.5 (в): достигая экстремума на масштабе $\times 17$ (близком к

истинному масштабу $\times 20$, на котором велось обучение), кривая расстояний до референсной гиперплоскости существенно сближается с кривой расстояний до гиперплоскости TUM при обработке фрагмента на большем или меньшем масштабе в сравнении с $\times 17$.

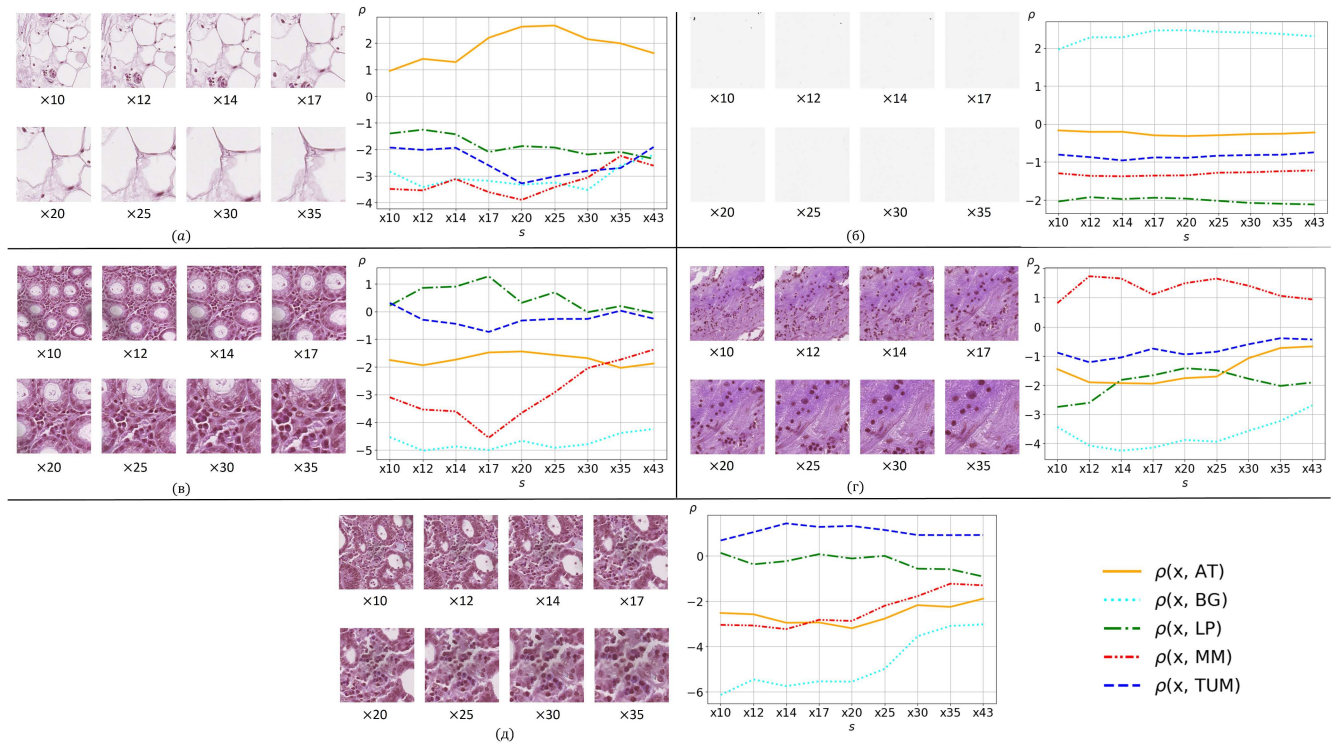


Рисунок 3.5 — Результаты применения метода к различным многомасштабным фрагментам полнослайдового изображения. На графиках показаны зависимости расстояний (ρ) до разделяющих гиперплоскостей в зависимости от масштаба (s) обрабатываемого фрагмента. (а) – фрагмент класса AT, (б) – фрагмент класса BG, (в) – фрагмент класса LP, (г) – фрагмент класса MM, (д) – фрагмент класса TUM.

В зависимости от специфики текстур фрагмента и ширины множества поиска варьируется характер представленных на Рис. 3.5 зависимостей:

- выпуклость (Рис. 3.5 (а, в)) – метод способен сделать однозначный вывод о масштабе обработки с высокой уверенностью, ширина множества \mathcal{S} достаточна;
- монотонность (Рис. 3.5 (г, д)) – метод способен выбрать масштаб обработки, однако следует увеличить ширину множества \mathcal{S} ;
- почти константность (Рис. 3.5 (б)) – масштабы равнозначны.

На Рис. 3.6 представлена гистограмма по множеству предложенных масштабов \hat{S} для 100 случайных фрагментов полнослайдового изображения, размеченных специалистами как LP-участки. Класс LP выбран ввиду особо низкой устойчивости сети DenseNet121 к распознаванию LP-фрагментов в различных

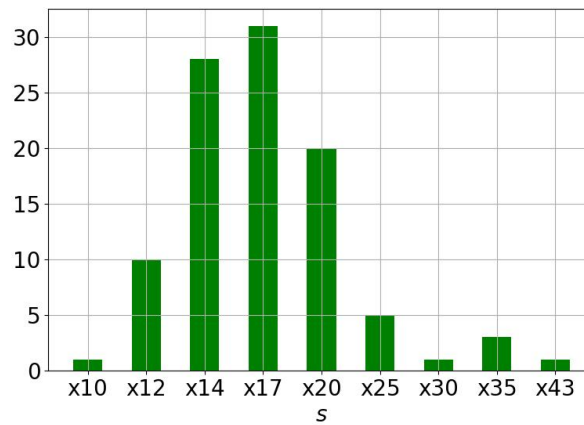


Рисунок 3.6 — Гистограмма по множеству предложенных методом масштабов \hat{S} для 100 случайных фрагментов LP.

масштабах (см. Таблицу 6), а значит высокой актуальности предлагаемого метода для анализа многомасштабных LP-фрагментов изображения. Топ-3 предлагаемых масштаба: $\times 17$, $\times 14$ и $\times 20$, соответственно. В доверительный интервал истинного масштаба ($\times 17 \leftarrow \times 20 \rightarrow \times 25$) попало 56 % всех прогнозов из \hat{S} . В более широкий доверительный интервал истинного масштаба ($\times 12 \leftarrow \times 20 \rightarrow \times 25$) попало более 85 % всех прогнозов.

3.4 Выводы

Данная глава обращает внимание на низкую устойчивость свёрточных нейронных сетей к изменению масштаба входных данных и предлагает автоматический метод выбора масштаба обработки, что особенно актуально при анализе многомасштабных полнослайдовых гистологических изображений.

Условно разделив свёрточную нейронную сеть на два модуля: модуль нелинейного отображения и модуль разделяющих гиперплоскостей, — данная глава представляет разработанный метод выбора масштаба на основе анализа расстояний до разделяющих гиперплоскостей.

Предлагаемый метод протестирован на задаче классификации полнослайдовых изображений стенок желудка из набора данных PATH-DT-MSU (WSS1, WSS2) предобученной свёрточной нейронной сетью DenseNet121.

Глава 4. Программный комплекс гибридных методов анализа и повышения качества медицинских изображений

Гибридные методы требуют не только разработки новых эффективных архитектур, но и создания надёжных, *гибких* инструментов их обучения. Традиционные подходы, основанные на *чистом* PyTorch, зачастую приводят к значительному дублированию кода, сложностям в воспроизведении вычислительных экспериментов и проблемам при масштабировании на распределённые системы. В данной главе будет представлен разработанный программный комплекс, реализованный на языке программирования Python на базе высокоуровневого фреймворка глубокого обучения PyTorch Lightning, решающий указанные проблемы и предоставляющий унифицированный *гибкий* [103] дизайн инструментария для обучения предлагаемых в диссертации гибридных методов анализа и обработки медицинских изображений. Это позволяет существенно сократить объём шаблонного кода и стандартизировать процесс проведения вычислительных экспериментов.

Дизайн программной системы – это абстрактная концепция. Он касается общего вида и структуры программы, а также детального вида и структуры каждого модуля, класса и метода. Дизайн можно представить в различных формах, но окончательным его воплощением является исходный код. В конечном итоге исходный код и является дизайном.

Хороший дизайн программы – это архитектура и структура кода, которые делают программное обеспечение удобным для разработки, поддержки, масштабирования и использования. Гибкое проектирование – это процесс, а не разовое событие. Это постоянное применение определённых паттернов и методик для улучшения структуры и понятности программы.

К признакам плохого дизайна [104] относят:

1. Жёсткость: дизайн трудно поддаётся изменению.
2. Хрупкость: дизайн легко разрушается.
3. Косность: дизайн трудно использовать повторно.
4. Вязкость: трудно добиться желаемого.
5. Ненужную сложность: избыточное проектирование.
6. Ненужные повторения: чрезмерное использование копирования.
7. Непрозрачность: плохо выраженная цель.

Дизайн называют жёстким, если единственное изменение провоцирует целый каскад других изменений в зависимых модулях. Чем больше модулей приходится менять, тем жёстче дизайн.

Дизайн является хрупким, если программа повреждается во многих местах при внесении единственного изменения. То есть новые проблемы возникают в частях, не имеющих концептуальной связи с той, что была изменена, а исправление одних проблем ведёт к появлению новых.

Дизайн называют косным, если он содержит части, которые могли бы оказаться полезны в других системах, но усилия и риски, сопряжённые с попыткой отделить эти части от оригинальной системы, слишком велики.

Вязкость программы проявляется в ситуации, когда внести изменения с сохранением дизайна оказывается значительно труднее, чем без сохранения.

Говорят, что дизайн обладает ненужной сложностью, если дизайн засоряется конструкциями, которые никогда не будут востребованы.

Дублирующийся код требует значительных усилий при рефакторинге, ведь ошибки, обнаруженные в повторяющемся блоке, должны быть исправлены во всех его копиях. Более того, поскольку повторения немного отличаются друг от друга, то и исправления будут разными.

Наконец, под непрозрачностью понимают трудность модуля для понимания. Действительность такова, что код, эволюционирующий со временем, постепенно становится всё более и более непрозрачным. Таким образом, разработчик должен примерять на себя роль читателя и прилагать сознательные усилия к рефакторингу кода таким образом, чтобы тот был понятен читателям.

Известно несколько принципов объектно-ориентированного проектирования, позволяющих избавиться от признаков плохого дизайна и создать наилучший дизайн для данного набора функций.

1. Принцип единственности обязанности (*англ.* single-responsibility principle, SRP) [105].
2. Принцип открытости/закрытости (*англ.* open/closed principle, OCP) [106].
3. Принцип подстановки Лисков (*англ.* Liskov substitution principle, LSP) [107].
4. Принцип инверсии зависимости (*англ.* dependency-inversion principle, DIP) [108].
5. Принцип разделения интерфейсов (*англ.* interface segregation principle, ISP) [109].

Данные принципы, сформулированные в работе [110], являются результатом многолетней эволюции и накопленного опыта в области программной инженерии. Их создание стало возможным благодаря коллективному интеллектуальному вкладу множества практиков и теоретиков. Вместе с тем, слепая догматическая приверженность даже проверенным постулатам может привести к негативным последствиям, таким как избыточное усложнение архитектуры, что противоречит самой цели применения этих правил.

Разработанный программный комплекс для обучения гибридных методов анализа и повышения качества медицинских изображений опубликован¹ на облачной платформе GitHub для хостинга и совместной разработки IT-проектов.

¹<https://github.com/MaksimPenkin/MedicalKAN>

4.1 Архитектура программного комплекса

Разработанный программный комплекс реализует концепцию разделения ответственности между компонентами системы, что соответствует принципам SOLID-программирования. Архитектура системы построена таким образом, что изменение одной функциональной части не требует модификации других, что обеспечивает масштабируемость и устойчивость к изменениям требований. Кроме того, подобная организация кода способствует повторному использованию компонентов в других исследовательских задачах.

Принцип единственности обязанности, гласящий, что у класса должна быть только одна причина для изменения, находит отражение в применяемом высокоуровневом фреймворке глубокого обучения PyTorch Lightning в разделении задач между компонентами, а именно:

1. `LightningModule` – ядро системы, инкапсулирующее:
 - архитектуру нейронной сети,
 - логику прямого распространения (*англ.* forward pass),
 - алгоритм обучения, валидации и тестирования.
2. `LightningDataModule` – абстракция, обеспечивающая:
 - загрузку и предобработку данных,
 - разделение на выборки,
 - генерацию пакетов (*англ.* batch).
3. `Callbacks` – механизм обратных вызовов для:
 - сохранения чекпоинтов,
 - мониторинга метрик.
4. `Trainer` – центральный класс, автоматизирующий:
 - распределение обучения по ускорителям (GPU, HPU, TPU, IPU),
 - логирование (TensorBoard, Wandb),
 - прочую гибкую настройку (максимальное число эпох и т.д.).

За поддержку модульности дополнительно отвечает разработанный механизм внедрения зависимости (*англ.* dependency injection, DI), являющийся своеобразной формой инверсии управления (*англ.* inversion of control, IoC) [111]. В сложных системах машинного обучения, где эксперименты включают множество компонентов, жёсткое связывание кода усложняет тестирование, модификацию и повторное использование. Предлагаемый механизм внедрения

зависимости решает эти проблемы, позволяя динамически управлять зависимостями между компонентами на основе конфигураций. В полном соответствии с принципом единственной ответственности объект (например, `CommonLitModel`, см. Листинг 4.1) отдаёт заботу о построении требуемых ему зависимостей внешнему, предназначенному для этого общему механизму (см. Листинг 4.2). Так, объект `CommonLitModel` делегирует получение доступа к архитектуре метода, функции потерь и оптимизатору (см. строки 11, 13, 15 Листинга 4.1). Механизм внедрения зависимости позволяет легче создавать альтернативные реализации, задавая их, например, в конфигурационных файлах, без внесения изменений в код объектов. Тем самым способствуя чистому коду, модульности, тестируемости и значительно упрощая управление сложными вычислительными экспериментами машинного обучения.

Листинг 4.1: Базовый класс `CommonLitModel`.

```

5  from . import nets
    from ..nn import losses, optimizers
    from src.utils.torch_utils import split_loss_logs
    from lightning import LightningModule

class CommonLitModel(LightningModule):
    def __init__(self, network, criterion=None, optimizer=None):
        super(CommonLitModel, self).__init__()

        self._model = nets.get(network)
        if criterion is not None:
            self._criterion = losses.get(criterion)
        if optimizer is not None:
            self._optimizer = optimizers.get(optimizer, partial=True)

    def forward(self, x, **kwargs):
        if isinstance(x, dict):
            try:
                y_pred = self._model(**x, **kwargs)
            except:
                y_pred = self._model(x, **kwargs)
        elif isinstance(x, (list, tuple)):
            try:
                y_pred = self._model(*x, **kwargs)
            except:
                y_pred = self._model(x, **kwargs)

```

```

    else:
        y_pred = self._model(x, **kwargs)
30     return y_pred

def compute_loss(self, y_pred, y):
    if isinstance(y, dict):
        try:
35             value = self._criterion(y_pred, **y)
        except:
            value = self._criterion(y_pred, y)
        elif isinstance(y, (list, tuple)):
            try:
40                 value = self._criterion(y_pred, *y)
            except:
                value = self._criterion(y_pred, y)
        else:
            value = self._criterion(y_pred, y)
45     return split_loss_logs(value)

def configure_optimizers(self):
    return self._optimizer(params=self.parameters())

```

Листинг 4.2: Реализация механизма внедрения зависимости на примере src/models.

```

from lightning import LightningModule
from .base_model import CommonLitModel
from .mri_enhancement_model import MRIEnhancementModel

5 from src.utils.serialization_utils import create_object

def get(identifier, **kwargs):
    obj = create_object(identifier,
10         module_objects={
            "CommonLitModel": CommonLitModel,
            "MRIEnhancementModel": MRIEnhancementModel},
            **kwargs)

15 if isinstance(obj, LightningModule):
    return obj
    raise ValueError(f"Could not interpret model instance: {obj}.")

```

Класс `CommonLitModel` открыт для расширения в соответствии с принципом открытости/закрытости. Так, часто меняющиеся методы, определяющие тренировочный, валидационный и тестовый шаги, реализованы в дочерних классах. Пример реализации тренировочного шага для обучения методов улучшения качества изображений магнитно-резонансной томографии приведён в Листинге 4.3. Важно осознанно применять абстракции только к тем фрагментам программы, которые часто изменяются. Отказ от преждевременного абстрагирования столь же важен, как и само абстрагирование.

Листинг 4.3: Класс `MRIEnhancementModel`, используемый для обучения методов улучшения качества изображений магнитно-резонансной томографии.

```

from .base_model import CommonLitModel

class MRIEnhancementModel(CommonLitModel):
5   def __init__(self, *args, **kwargs):
        super(MRIEnhancementModel, self).__init__(*args, **kwargs)

    def training_step(self, batch, batch_idx):
10        x, y = batch
        y_pred = self(x)
        loss, logs = self.compute_loss(y_pred, y)
        logs = {"train/" + k: v for k, v in logs.items()}
        self.log_dict(logs, on_step=True, on_epoch=True)
        return loss

```

4.2 Технология статического анализа программного кода

Статический анализ кода – это процесс анализа программного обеспечения без его фактического выполнения. Для Python существует множество инструментов статического анализа, которые помогают выявлять ошибки, проблемы дизайна и потенциальные уязвимости.

С помощью статического анализа обычно выявляют:

1. Ключевые признаки необходимости рефакторинга.
2. Потенциальные уязвимости.
3. Устаревшие конструкции.
4. Несоответствие кода стандартам.
5. Потенциальные ошибки.

В рамках разработки данного программного комплекса использовался современный инструмент Ruff, что позволило обеспечить соответствие кода стандарту PEP 8, выявить потенциальные логические ошибки и поддерживать единый стиль написания. Интеграция этого инструмента в процесс разработки существенно повысила читаемость и поддерживаемость кодовой базы. Традиционные линтеры, flake8, pylint и pycodestyle, обладают рядом ограничений, таких как высокое время выполнения и фрагментированность экосистемы, выраженная в необходимости использования нескольких инструментов для разных проверок.

Ruff – это современный высокопроизводительный линтер, объединяющий функциональность множества инструментов в единое решение с минимальными накладными расходами. Будучи написанным на Rust, Ruff работает от 10 до 100 раз быстрее других линтеров, предлагая низкоуровневую оптимизацию парсинга и анализа AST (*англ.* abstract syntax tree), параллельную обработку файлов и минимальные зависимости в сравнении с Python-инструментами. Настраивается Ruff с помощью конфигурационного файла (см. Листинг 4.4).

Листинг 4.4: Пример конфигурационного файла ruff.toml для линтера Ruff.

```
# Exclude a variety of commonly ignored directories.
exclude = [
    ".ruff_cache",
    ".mypy_cache",
5    ".pytest_cache",
    ".git",
    "__pycache__",
```

```

    "__init__.py",
    "build",
10    "venv",
    "*.egg-info",
    "*.pyi",
    "src/models/nets"]

15 extend-include = ["*.ipynb"]
    # extend-select = ["I"] # makes ruff sort imports

    # Same as Black.
    line-length = 150
20 indent-width = 4

    # Assume Python 3.11
    target-version = "py311"

25 [lint]
    # Enable Pyflakes (`F`) and a subset of the pycodestyle (`E`) codes.
    select = ["E4", "E7", "E9", "F"]
    ignore = ["E401"]

30 # Allow fix for all enabled rules (when `--fix`) is provided.
    fixable = ["ALL"]
    unfixable = []

    # Allow unused variables when underscore-prefixed.
35 dummy-variable-rgx = "^(_+|(_+[a-zA-Z0-9_]*[a-zA-Z0-9]+?))$"

    # Enable per-file disabling rules.
    [lint.per-file-ignores]
    "src/models/base_model.py" = ["E722"]

40 [format]
    # Like Black, use double quotes for strings.
    quote-style = "double"
    # Like Black, indent with spaces, rather than tabs.
45 indent-style = "space"
    # Like Black, respect magic trailing commas.
    skip-magic-trailing-comma = false
    # Like Black, automatically detect the appropriate line ending.
    line-ending = "auto"

```

Использование Ruff на протяжении всего жизненного цикла разработки программного комплекса позволило достичь нескольких ключевых целей, напрямую влияющих на эффективность научно-исследовательской деятельности. Ruff, выступая в роли формализованного арбитра, обеспечил строгое и автоматическое соблюдение единых стандартов оформления, минимизируя риск фрагментации стиля кода, упрощая чтение, анализ и модификацию. Наиболее значимым вкладом Ruff оказалось прямое выявление потенциальных источников ошибок на этапе написания кода, до запуска вычислительных экспериментов. В контексте численного моделирования, где выполнение одного расчёта может занимать часы или сутки, выявление подобной ошибки на стадии кодирования сэкономило значительные вычислительные и временные ресурсы.

К недостаткам статических анализаторов кода можно отнести ложные срабатывания. Например, при реализации специализированных численных алгоритмов разработчик может сознательно отступать от общих правил в пользу производительности или точности. Анализатор, не обладающий знаниями о предметной области, может интерпретировать такие решения как нарушения.

4.3 Система конфигурации вычислительных экспериментов

Современные исследования в области машинного обучения и глубоких нейронных сетей, используемых в гибридных методах, требуют проведения многочисленных вычислительных экспериментов с различными архитектурами, гиперпараметрами и наборами данных. Для обеспечения воспроизводимости, управляемости и масштабируемости вычислительных экспериментов необходима система конфигурации. Двумя основными подходами здесь являются императивный интерфейс командной строки (*англ.* command line interface, CLI) и декларативная конфигурация, например, с помощью JSON или YAML файлов.

Несмотря на преимущества классического интерфейса командной строки, такие как оперативность внесения изменений и удобную интеграцию с bash-циклами поиска оптимальных параметров, недостатки перевешивают, когда речь идёт о разработке больших программных комплексов, требующих длительной поддержки. Плоская структура параметров быстро приводит к информационной перегрузке интерфейса, а также к экспоненциальному росту числа флагов при увеличении сложности эксперимента. Следствием плоской структуры является отсутствие семантической группировки, смешение разноуровневых параметров и невозможность передачи сложных структур.

Использование декларативной системы конфигурации экспериментов на основе конфигурационных файлов разрешает описанные выше проблемы, обеспечивая научную строгость (воспроизводимость) и инженерную надёжность (разделение логики и конфигурации). Более того, конфигурационные файлы служат не только для настройки эксперимента, но и для его документирования. Действительно, они позволяют добавлять комментарии и метаданные, связывают результаты обучения с конкретными параметрами и упрощают сравнение экспериментов за счёт чёткой фиксации изменений. Примером современной мощной декларативной системы конфигурации экспериментов является Hydra. Однако она не лишена недостатков: во-первых, при большом количестве параметров конфигурационные файлы становятся громоздкими и сложными для навигации, во-вторых, Hydra не позволяет использовать динамические вычисления внутри конфигурационных файлов, которые часто оказываются удобными, например, для задания переменных окружения.

В данном разделе описывается разработанная гибридная система конфигурации вычислительных экспериментов, задающая очередной эксперимент YAML файлом (см. Листинг 4.5), причём настройка наиболее часто изменяемых параметров вынесена в CLI (см. Листинг 4.6).

Листинг 4.5: Корневая конфигурация одного из вычислительных экспериментов улучшения качества изображений магнитно-резонансной томографии головного мозга. Задано обучение метода из семейства нейронных операторов на графическом ускорителе GPU на наборе данных IXI. Вложенные спецификации гиперпараметров выбранного метода и загрузчиков данных скрыты в соответствующих YAML файлах и добавляются динамически во время загрузки.

```

trainer: !include "./trainers/gpu.yaml"
model: !include "./models/neuraloperator/model.yaml"
data:
  train_dataloaders:
5  - !include "./data/ixi/train_dataloader.yaml"
  val_dataloaders:
    - !include "./data/ixi/val_dataloader.yaml"

```

Листинг 4.6: Пример использования системы конфигурации экспериментов.

```
python train.py --config ./configs/train.yaml --limit_train_batches 0.1
```

Стандарт оформления корневого конфигурационного файла отображён в Листинге 4.5 и включает в себя следующие обязательные секции: `trainer`, `model`, `data`, – задающие конфигурации процесса обучения, метода и набора данных для обучения, валидации и тестирования.

Директива `!include` не включена в синтаксис YAML и реализована отдельно. Использование этой директивы позволяет повысить читаемость объёмных конфигурационных файлов, скрывая спецификации некоторых логических компонент в отдельных YAML файлах, а также способствует повторному использованию общих параметров в разных экспериментах и управлению версионностью отдельных блоков. Так, например, конфигурация метода повышения качества изображений магнитно-резонансной томографии включает в себя несколько логических компонент (см. Листинг 4.7): версия архитектуры нейронной сети задаётся в поле `network`, спецификация функции потерь в поле `criterion` и оптимизатора в поле `optimizer`. Эти логические компоненты естественным образом образуют словарь именованных параметров `config`, который подаётся инициализатору целевого объекта в фабрике объектов через `kwargs` (см. строку 13

Листинга 4.8). Задание целевого объекта осуществляется по двум полям: название модуля `module`, содержащего целевой класс, и, собственно, имя целевого класса `class_name`. Если название модуля не указано явно, то фабрика объектов будет искать целевой класс в реестре `module_objects`. Фабрика объектов поддерживает частичное (*англ.* `partial`) создание объектов, фиксируя часть аргументов и сохраняя возможность передачи оставшихся аргументов в дальнейшем. Такой сценарий реализуется при создании оптимизатора (см. Листинг 4.1).

Листинг 4.7: Конфигурация метода улучшения качества изображений магнитно-резонансной томографии.

```

module: src.models
class_name: MRIEnhancementModel
config:
    network: !include "./nets/funkan.yaml"
5   criterion: !include "../..../losses/mse.yaml"
    optimizer: !include "../..../optimizers/adam.yaml"

```

Листинг 4.8: Фабрика объектов.

```

def create_object_from_config(config, module_objects=None, partial=False):
    if config is None:
        return None

5   cls_module, cls_name, cls_config =
        config.get("module"), config["class_name"], config["config"]
    if not cls_module:
        cls = module_objects[cls_name]
    else:
10      cls = getattr(import_module(cls_module), cls_name)

    if not partial:
        return cls(**cls_config)
    else:
15      return functools.partial(cls, **cls_config)

```

Таким образом, осуществляется *десериализация* объектов из их иерархического представления в виде вложенной структуры YAML файлов. Важно отметить, что имена ключей словаря `config` заданы не системой конфигурации, как в единственном корневом конфигурационном файле (см. Листинг 4.5), а наследуются из кода, то есть определяются параметрами инициализатора целевого объекта.

Это обеспечивает дополнительную гибкость, прозрачность и сохранение единства именования одних и тех же сущностей в коде и файлах конфигураций, снижая когнитивную нагрузку на пользователя.

В рамках диссертационной работы разработанная гибридная система конфигурации экспериментов использовалась для:

- обучения и валидации предлагаемых гибридных методов,
- выбора гиперпараметров через автоматизированные запуски,
- фиксации условий вычислительных экспериментов для последующего воспроизведения.

Заключение

Диссертационная работа посвящена разработке гибридных методов обработки и анализа медицинских изображений различных типов, включая изображения магнитно-резонансной томографии, ультразвуковые, гистологические, полнослайдовые и колоноскопические изображения.

В ходе проведённого исследования полностью достигнута основная цель работы, разработан комплекс новых гибридных методов и соответствующий программный инструментарий для проведения вычислительных экспериментов.

Результаты работы 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.

Список литературы

1. *Li, S.* Exploring Neural Ordinary Differential Equations as Interpretable Healthcare Classifiers [Текст] / S. Li // arXiv preprint arXiv:2503.03129. — 2025.
2. *Mamaev, N.* Finding the Parameters of a Nonlinear Diffusion Denoising Method by Ridge Analysis [Текст] / N. Mamaev, D. Yurin, A. Krylov // Computational Mathematics and Modeling. — 2018. — Т. 29, № 3. — С. 334—343.
3. *Farooq, Y.* Noise Removal from the Image using Convolutional Neural Networks-based Denoising Auto Encoder [Текст] / Y. Farooq, S. Savaş // Journal of Emerging Computer Technologies. — 2024. — Т. 3, № 1. — С. 21—28.
4. Diffusion models, Image Super-resolution, and Everything: A Survey [Текст] / B. B. Moser [и др.] // IEEE Transactions on Neural Networks and Learning Systems. — 2024.
5. Neural Network Quantization with AI Model Efficiency Toolkit (AIMET) [Текст] / S. Siddegowda [и др.] // arXiv preprint arXiv:2201.08442. — 2022.
6. Transformers for Vision: A Survey on Innovative Methods for Computer Vision [Текст] / V. Hassija [и др.] // IEEE Access. — 2025.
7. *Mumuni, F.* Explainable Artificial Intelligence (XAI): from Inherent Explainability to Large Language Models [Текст] / F. Mumuni, A. Mumuni // arXiv preprint arXiv:2501.09967. — 2025.
8. *Abdulkadirov, R.* Survey of Optimization Algorithms in Modern Neural Networks [Текст] / R. Abdulkadirov, P. Lyakhov, N. Nagornov // Mathematics. — 2023. — Т. 11, № 11. — С. 2466.
9. Deep Stacked Hierarchical Multi-patch Network for Image Deblurring [Текст] / H. Zhang [и др.] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. — 2019. — С. 5978—5986.
10. Advances in Medical Image Segmentation: A Comprehensive Review of Traditional, Deep Learning and Hybrid approaches [Текст] / Y. Xu [и др.] // Bioengineering. — 2024. — Т. 11, № 10. — С. 1034.
11. *Petrou, M. M.* Image Processing: The Fundamentals [Текст] / M. M. Petrou, C. Petrou. — John Wiley & Sons, 2010.

12. *Малла, С.* Вейвлеты в Обработке Сигналов [Текст] / С. Малла. — 2005.
13. *Vese, L. A.* Variational Methods in Image Processing [Текст] / L. A. Vese, C. Le Guyader. — CRC Press Boca Raton, 2016.
14. *Ribes, A.* Linear Inverse Problems in Imaging [Текст] / A. Ribes, F. Schmitt // IEEE Signal Processing Magazine. — 2008. — Т. 25, № 4. — С. 84—99.
15. *Ronneberger, O.* U-Net: Convolutional Networks for Biomedical Image Segmentation [Текст] / O. Ronneberger, P. Fischer, T. Brox // Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. — Springer. 2015. — С. 234—241.
16. *Badrinarayanan, V.* SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation [Текст] / V. Badrinarayanan, A. Kendall, R. Cipolla // IEEE transactions on pattern analysis and machine intelligence. — 2017. — Т. 39, № 12. — С. 2481—2495.
17. Rethinking Atrous Convolution for Semantic Image Segmentation [Текст] / L.-C. Chen [и др.] // arXiv preprint arXiv:1706.05587. — 2017.
18. Diffusion Models for Image Restoration and Enhancement: A Comprehensive Survey [Текст] / X. Li [и др.] // International Journal of Computer Vision. — 2025. — С. 1—31.
19. *Jin, Q.* A Review of an Old Dilemma: Demosaicking First, or Denoising First? [Текст] / Q. Jin, G. Facciolo, J.-M. Morel // proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. — 2020. — С. 514—515.
20. Neural Ordinary Differential Equations [Текст] / R. T. Chen [и др.] // Advances in neural information processing systems. — 2018. — Т. 31.
21. Physics-Informed Neural Networks for PDE Problems: A Comprehensive Review [Текст] / K. Luo [и др.] // Artificial Intelligence Review. — 2025. — Т. 58, № 10. — С. 1—43.
22. Fourier Neural Operator for Parametric Partial Differential Equations [Текст] / Z. Li [и др.] // arXiv preprint arXiv:2010.08895. — 2020.

23. On the Spectral Bias of Neural Networks [Текст] / N. Rahaman [и др.] // International conference on machine learning. — PMLR. 2019. — C. 5301—5310.
24. *Wilton, J.* The Gibbs Phenomenon in Fourier-Bessel Series [Текст] / J. Wilton. — 1928.
25. *Gottlieb, D.* Numerical Analysis of Spectral Methods: Theory and Applications [Текст] / D. Gottlieb, S. A. Orszag. — SIAM, 1977.
26. *Pinsky, M. A.* Fourier Inversion for Piecewise Smooth Functions in Several Variables [Текст] / M. A. Pinsky // Proceedings of the American Mathematical Society. — 1993. — T. 118, № 3. — C. 903—910.
27. *Pinsky, M. A.* Pointwise Fourier Inversion in Several Variables [Текст] / M. A. Pinsky // Notices of the AMS. — 1995. — T. 42, № 3. — C. 330—334.
28. *Sitdikov, I.* Variational Image Deringing using Varying Regularization Parameter [Текст] / I. Sitdikov, A. Krylov // Pattern Recognition and Image Analysis. — 2015. — T. 25. — C. 96—100.
29. *Umnov, A.* Sparse Approach to Image Ringing Detection and Suppression [Текст] / A. Umnov, A. Krylov // Pattern Recognition and Image Analysis. — 2017. — T. 27. — C. 754—762.
30. Gibbs-Ringing Artifact Removal based on Local Subvoxel-Shifts [Текст] / E. Kellner [и др.] // Magnetic resonance in medicine. — 2016. — T. 76, № 5. — C. 1574—1581.
31. Attention U-Net: Learning Where to Look for the Pancreas [Текст] / O. Oktay [и др.] // arXiv preprint arXiv:1804.03999. — 2018.
32. UNet++: A Nested U-Net Architecture for Medical Image Segmentation [Текст] / Z. Zhou [и др.] // International workshop on deep learning in medical image analysis. — Springer. 2018. — C. 3—11.
33. *Sinha, A.* Multi-Scale Self-Guided Attention for Medical Image Segmentation [Текст] / A. Sinha, J. Dolz // IEEE journal of biomedical and health informatics. — 2020. — T. 25, № 1. — C. 121—130.
34. *Tripathi, P. C.* CNN-DMRI: A Convolutional Neural Network for Denoising of Magnetic Resonance Images [Текст] / P. C. Tripathi, S. Bag // Pattern Recognition Letters. — 2020. — T. 135. — C. 57—63.

35. A Loss-based Patch Label Denoising Method for Improving Whole-Slide Image Analysis using a Convolutional Neural Network [Текст] / M. Ashraf [и др.] // Scientific reports. — 2022. — Т. 12, № 1. — С. 1392.
36. Gibbs-Ringing Artifact Suppression with Knowledge Transfer from Natural Images to MR Images [Текст] / X. Zhao [и др.] // Multimedia Tools and Applications. — 2020. — Т. 79. — С. 33711—33733.
37. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization [Текст] / R. R. Selvaraju [и др.] // International journal of computer vision. — 2020. — Т. 128. — С. 336—359.
38. In What Ways are Deep Neural Networks Invariant and How Should We Measure This? [Текст] / H. Kvinge [и др.] // Advances in Neural Information Processing Systems. — 2022. — Т. 35. — С. 32816—32829.
39. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows [Текст] / Z. Liu [и др.] // Proceedings of the IEEE/CVF international conference on computer vision. — 2021. — С. 10012—10022.
40. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation [Текст] / J. Chen [и др.] // arXiv preprint arXiv:2102.04306. — 2021.
41. Medformer: A Multi-Granularity Patching Transformer for Medical Time-series Classification [Текст] / Y. Wang [и др.] // arXiv preprint arXiv:2405.19363. — 2024.
42. Wu, Z. Wider or Deeper: Revisiting the ResNet Model for Visual Recognition [Текст] / Z. Wu, C. Shen, A. Van Den Hengel // Pattern recognition. — 2019. — Т. 90. — С. 119—133.
43. Hybrid Method for Biomedical Image Denoising [Текст] / A. Krylov [и др.] // Proceedings of the 2019 4th International Conference on Biomedical Imaging, Signal Processing. — 2019. — С. 60—64.
44. GraFMRI: A Graph-based Fusion Framework for Robust Multi-Modal MRI Reconstruction [Текст] / S. Ahmed [и др.] // Magnetic Resonance Imaging. — 2025. — Т. 116. — С. 110279.

45. *Nguyen, D.-T.* Attention U-Net with Active Contour based Hybrid Loss for Brain Tumor Segmentation [Текст] / D.-T. Nguyen, T.-T. Tran, V.-T. Pham // Soft Computing: Biomedical and Related Applications. — Springer, 2021. — С. 35—45.
46. Learning Active Contour Models for Medical Image Segmentation [Текст] / X. Chen [и др.] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. — 2019. — С. 11632—11640.
47. *Rudin, L. I.* Total Variation based Image Restoration with Free Local Constraints [Текст] / L. I. Rudin, S. Osher // Proceedings of 1st international conference on image processing. Т. 1. — IEEE. 1994. — С. 31—35.
48. Improved Residual Networks for Image and Video Recognition [Текст] / I. C. Duta [и др.] // 2020 25th International Conference on Pattern Recognition (ICPR). — IEEE. 2021. — С. 9415—9422.
49. Image Quality Assessment: From Error Visibility to Structural Similarity [Текст] / Z. Wang [и др.] // IEEE Transactions on Image Processing. — 2004. — Т. 13, № 4. — С. 600—612.
50. *Amari, S.-i.* Backpropagation and Stochastic Gradient Descent Method [Текст] / S.-i. Amari // Neurocomputing. — 1993. — Т. 5, № 4/5. — С. 185—196.
51. *Kingma, D. P.* Adam: A Method for Stochastic Optimization [Текст] / D. P. Kingma, J. Ba // CoRR. — 2014. — Т. abs/1412.6980. — URL: <https://api.semanticscholar.org/CorpusID:6628106>.
52. *Hore, A.* Image Quality Metrics: PSNR vs. SSIM [Текст] / A. Hore, D. Ziou // 2010 20th international conference on pattern recognition. — IEEE. 2010. — С. 2366—2369.
53. *Wang, Z.* Mean Squared Error: Love It or Leave It? A New Look at Signal Fidelity Measures [Текст] / Z. Wang, A. C. Bovik // IEEE signal processing magazine. — 2009. — Т. 26, № 1. — С. 98—117.
54. *Ioffe, S.* Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift [Текст] / S. Ioffe, C. Szegedy // International conference on machine learning. — pmlr. 2015. — С. 448—456.
55. *Han, D.* Deep Pyramidal Residual Networks [Текст] / D. Han, J. Kim, J. Kim // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2017. — С. 5927—5935.

56. How Does Batch Normalization Help Optimization? [Текст] / S. Santurkar [и др.] // Advances in neural information processing systems. — 2018. — Т. 31.
57. Garbin, C. Dropout vs. Batch Normalization: An Empirical Study of their Impact to Deep Learning [Текст] / C. Garbin, X. Zhu, O. Marques // Multimedia tools and applications. — 2020. — Т. 79, № 19. — С. 12777—12815.
58. Loss Functions and Metrics in Deep Learning [Текст] / J. Terven [и др.] // arXiv preprint arXiv:2307.02694. — 2023.
59. KAN: Kolmogorov-Arnold Networks [Текст] / Z. Liu [и др.] // arXiv preprint arXiv:2404.19756. — 2024.
60. KAN 2.0: Kolmogorov-Arnold Networks Meet Science [Текст] / Z. Liu [и др.] // arXiv preprint arXiv:2408.10205. — 2024.
61. Lu, L. DeepONet: Learning Nonlinear Operators for Identifying Differential Equations based on the Universal Approximation Theorem of Operators [Текст] / L. Lu, P. Jin, G. E. Karniadakis // arXiv preprint arXiv:1910.03193. — 2019.
62. Implicit Neural Representations with Periodic Activation Functions [Текст] / V. Sitzmann [и др.] // Advances in neural information processing systems. — 2020. — Т. 33. — С. 7462—7473.
63. WIRE: Wavelet Implicit Neural Representations [Текст] / V. Saragadam [и др.] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2023. — С. 18507—18516.
64. Physics-informed Neural Networks for Heat Transfer Problems [Текст] / S. Cai [и др.] // Journal of Heat Transfer. — 2021. — Т. 143, № 6. — С. 060801.
65. Implicit Neural Representation for MRI Parallel Imaging Reconstruction [Текст] / H. Li [и др.] // arXiv preprint arXiv:2309.06067. — 2023.
66. Колмогоров, А. Элементы Теории Функций и Функционального Анализа [Текст] / А. Колмогоров, С. Фомин. — Litres, 2022.
67. Rapid Prediction of Thermal Stress on Satellites via Domain Decomposition-based Hybrid Fourier Neural Operator [Текст] / K. Zhou [и др.] // Engineering Applications of Artificial Intelligence. — 2025. — Т. 153. — С. 110826.
68. Денисов, А. М. Обыкновенные Дифференциальные Уравнения [Текст] / А. М. Денисов, А. В. Разгулин. — 2009.

69. *Perona, P.* Anisotropic Diffusion [Текст] / P. Perona, T. Shiota, J. Malik // Geometry-driven diffusion in computer vision. — 1994. — С. 73—92.
70. *Kolmogorov, A. N.* On the Representations of Continuous Functions of Many Variables by Superposition of Continuous Functions of One Variable and Addition [Текст] / A. N. Kolmogorov // Dokl. Akad. Nauk USSR. T. 114. — 1957. — С. 953—956.
71. *Cybenko, G.* Approximation by Superpositions of a Sigmoidal Function [Текст] / G. Cybenko // Mathematics of control, signals and systems. — 1989. — Т. 2, № 4. — С. 303—314.
72. *Girosi, F.* Representation Properties of Networks: Kolmogorov's Theorem is Irrelevant [Текст] / F. Girosi, T. Poggio // Neural Computation. — 1989. — Т. 1, № 4. — С. 465—469.
73. *Seydi, S. T.* Exploring the Potential of Polynomial Basis Functions in Kolmogorov-Arnold Networks: A Comparative Study of Different Groups of Polynomials [Текст] / S. T. Seydi // arXiv preprint arXiv:2406.02583. — 2024.
74. Chebyshev Polynomial-based Kolmogorov-Arnold Networks: An Efficient Architecture for Nonlinear Function Approximation [Текст] / S. SS, K. AR, A. KP [и др.] // arXiv preprint arXiv:2405.07200. — 2024.
75. *Крылов, А. С.* Введение в Математические Методы Обработки Изображений [Текст] / А. С. Крылов. — 2022.
76. U-KAN Makes Strong Backbone for Medical Image Segmentation and Generation [Текст] / C. Li [и др.] // Proceedings of the AAAI Conference on Artificial Intelligence. — 2025. — Т. 39, № 5. — С. 4652—4660.
77. Convolutional Kolmogorov-Arnold Networks [Текст] / A. D. Bodner [и др.] // arXiv preprint arXiv:2406.13155. — 2024.
78. *Drokin, I.* Kolmogorov-Arnold Convolutions: Design Principles and Empirical Studies [Текст] / I. Drokin // arXiv preprint arXiv:2407.01092. — 2024.
79. UnO: Unsupervised Occupancy Fields for Perception and Forecasting [Текст] / B. Agro [и др.] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2024. — С. 14487—14496.

80. *Grünbaum, F. A.* The Eigenvectors of the Discrete Fourier Transform: A Version of the Hermite Functions [Текст] / F. A. Grünbaum // Journal of Mathematical Analysis and Applications. — 1982. — Т. 88, № 2. — С. 355—363.
81. Dataset of Breast Ultrasound Images [Текст] / W. Al-Dhabyani [и др.] // Data in brief. — 2020. — Т. 28. — С. 104863.
82. Medical Transformer: Gated Axial-attention for Medical Image Segmentation [Текст] / J. M. J. Valanarasu [и др.] // International conference on medical image computing and computer-assisted intervention. — Springer. 2021. — С. 36—46.
83. WM-DOVA Maps for Accurate Polyp Highlighting in Colonoscopy: Validation vs. Saliency Maps from Physicians [Текст] / J. Bernal [и др.] // Computerized medical imaging and graphics. — 2015. — Т. 43. — С. 99—111.
84. *Valanarasu, J. M. J.* UNeXt: MLP-based Rapid Medical Image Segmentation Network [Текст] / J. M. J. Valanarasu, V. M. Patel // International conference on medical image computing and computer-assisted intervention. — Springer. 2022. — С. 23—33.
85. Rolling-UNet: Revitalizing MLP's Ability to Efficiently Extract Long-distance Dependencies for Medical Image Segmentation [Текст] / Y. Liu [и др.] // Proceedings of the AAAI conference on artificial intelligence. Т. 38. — 2024. — С. 3819—3827.
86. *Ma, J.* U-Mamba: Enhancing Long-range Dependency for Biomedical Image Segmentation [Текст] / J. Ma, F. Li, B. Wang // arXiv preprint arXiv:2401.04722. — 2024.
87. SpectralDefense: Detecting Adversarial Attacks on CNNs in the Fourier Domain [Текст] / P. Harder [и др.] // 2021 International Joint Conference on Neural Networks (IJCNN). — IEEE. 2021. — С. 1—8.
88. *Park, S.* Digital Imaging in Pathology [Текст] / S. Park, L. Pantanowitz, A. V. Parwani // Clinics in laboratory medicine. — 2012. — Т. 32, № 4. — С. 557—584.
89. Integrating Digital Pathology into Clinical Practice [Текст] / M. G. Hanna [и др.] // Modern Pathology. — 2022. — Т. 35, № 2. — С. 152—164.
90. Digital Imaging in Pathology: Whole-Slide Imaging and Beyond [Текст] / F. Ghaznavi [и др.] // Annual Review of Pathology: Mechanisms of Disease. — 2013. — Т. 8, № 1. — С. 331—359.

91. Whole Slide Imaging, Artificial Intelligence, and Machine Learning in Pediatric and Perinatal Pathology: Current Status and Future Directions [Текст] / J. C. Hutchinson [и др.] // *Pediatric and Developmental Pathology*. — 2025. — Т. 28, № 2. — С. 91—98.
92. Whole Slide Image to DICOM Conversion as Event-Driven Cloud Infrastructure [Текст] / D. Brundage [и др.] // *arXiv preprint arXiv:2203.13888*. — 2022.
93. From Traditional to Deep Learning Approaches in Whole Slide Image Registration: A Methodological Review [Текст] / B. Elhaminia [и др.] // *arXiv preprint arXiv:2502.19123*. — 2025.
94. Whole Slide Imaging Technology and Its Applications: Current and Emerging Perspectives [Текст] / E. Jain [и др.] // *International Journal of Surgical Pathology*. — 2024. — Т. 32, № 3. — С. 433—448.
95. A State-of-The-Art Survey of Artificial Neural Networks for Whole-Slide Image Analysis: From Popular Convolutional Neural Networks to Potential Visual Transformers [Текст] / W. Hu [и др.] // *Computers in Biology and Medicine*. — 2023. — Т. 161. — С. 107034.
96. Artificial Intelligence–Assisted Classification of Gliomas using Whole Slide Images [Текст] / L. Jose [и др.] // *Archives of Pathology & Laboratory Medicine*. — 2023. — Т. 147, № 8. — С. 916—924.
97. Deep Residual Learning for Image Recognition [Текст] / K. He [и др.] // *Proceedings of the IEEE conference on computer vision and pattern recognition*. — 2016. — С. 770—778.
98. *Simonyan, K. Very Deep Convolutional Networks for Large-Scale Image Recognition [Текст] / K. Simonyan, A. Zisserman // arXiv preprint arXiv:1409.1556. — 2014.*
99. Visualization of Whole Slide Histological Images with Automatic Tissue Type Recognition [Текст] / A. Khvostikov [и др.] // *Pattern Recognition and Image Analysis*. — 2022. — Т. 32, № 3. — С. 483—488.
100. Going Deeper with Convolutions [Текст] / C. Szegedy [и др.] // *Proceedings of the IEEE conference on computer vision and pattern recognition*. — 2015. — С. 1—9.

101. Densely Connected Convolutional Networks [Текст] / G. Huang [и др.] // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2017. — С. 4700—4708.
102. Goar, V. Foundations of Machine Learning [Текст] / V. Goar, N. S. Yadav // Intelligent Optimization Techniques for Business Analytics. — IGI Global, 2024. — С. 25—48.
103. Robert, C. M. Agile Software Development [Текст] / C. M. Robert // Martin: Prentice Hall. — 2003.
104. When and Why Your Code Starts to Smell Bad [Текст] / M. Tufano [и др.] // 2015 IEEE/ACM 37th IEEE International Conference on Software Engineering. T. 1. — IEEE. 2015. — С. 403—414.
105. CHANGE, C. T. SRP: The Single Responsibility Principle [Текст] / C. T. CHANGE. — 2005.
106. Martin, R. C. The Open-Closed Principle [Текст] / R. C. Martin // More C++ gems. — 1996. — Т. 19, № 96. — С. 9.
107. Noback, M. The Liskov Substitution Principle [Текст] / M. Noback // Principles of Package Design: Creating Reusable Software Components. — Springer, 2018. — С. 31—53.
108. Noback, M. The Dependency Inversion Principle [Текст] / M. Noback // Principles of Package Design: Creating Reusable Software Components. — Springer, 2018. — С. 67—104.
109. Noback, M. The Interface Segregation Principle [Текст] / M. Noback // Principles of Package Design: Creating Reusable Software Components. — Springer, 2018. — С. 55—66.
110. Martin, R. C. Design Principles and Design Patterns [Текст] / R. C. Martin // Object Mentor. — 2000. — Т. 1, № 34. — С. 597.
111. Fowler, M. Inversion of Control Containers and The Dependency Injection Pattern [Текст] / M. Fowler // <http://www.martinfowler.com/articles/injection.html>. — 2004.

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 (SJР) [0.35 / 0.46] п.л.
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 (SJР) [0.2 / 0.35] п.л.
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 (SJР) [0.3 / 0.4] п.л.

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.

Список рисунков

- 1.1 Проявление артефактов Гиббса на изображениях магнитно-резонансной томографии. (а) – T1-взвешенное изображение МРТ мозга в низком разрешении (64×64) с осцилляциями Гиббса в виде ярких и тёмных линий, параллельных краям резких изменений интенсивности, (б) – T1-взвешенное изображение МРТ мозга в высоком разрешении (128×128), (в) – пример осцилляций Гиббса на ступенчатой функции, моделирующей резкий контур на изображении. 11
- 1.2 Осцилляции Гиббса на примере конечного разрыва. (а) – ступенчатая функция $\xi(t)$, (б) – результат восстановления сигнала $\xi(t)$ по усечённому ряду Фурье, демонстрирующий осцилляции в окрестностях точек разрыва. 13
- 1.3 Пример улучшения качества изображения МРТ головного мозга алгоритмом Кельнера. (а) – референсное изображение без артефактов, (б) – изображение с осцилляциями Гиббса, (в) – результат алгоритма Кельнера. 17
- 1.4 Примеры улучшения качества изображений МРТ головного мозга нейронной сетью GAS-CNN. (а) – референсные изображения без артефактов, (б) – изображения с осцилляциями Гиббса, (в) – результаты нейронной сети GAS-CNN. 20
- 1.5 Архитектура разработанного гибридного метода DGAS9-CNN подавления осцилляций Гиббса на изображениях МРТ головного мозга. 21
- 1.6 Сравнение разработанного гибридного метода DGAS9-CNN улучшения качества изображений МРТ головного мозга с нейронной сетью GAS-CNN. (а) – референсные изображения без артефактов, (б) – изображения с осцилляциями Гиббса, (в) – результаты нейронной сети GAS-CNN, (г) – результаты гибридного метода DGAS9-CNN. . . . 24
- 1.7 Динамика средних значений PSNR на валидационном наборе данных из 2016 изображений в зависимости от эпохи. Оценка влияния вспомогательных признаков описаний алгоритма Кельнера на качество предлагаемого гибридного метода DGAS9-CNN. 26

1.8	Средние значения PSNR по тестовому набору данных из 2617 изображений. Оценка влияния вспомогательных признаков описаний алгоритма Кельнера на качество предлагаемого гибридного метода DGAS9-CNN.	26
2.1	Мониторинг качества реконструкции изображения рентгеновского снимка грудной клетки размера 224×224 тремя сравниваемыми нейронными операторами, отличающимися видом нелинейности: ReLU, функции Габора (WIRE [63]), синусы (SIREN [62]).	30
2.2	Примеры восстановления изображения рентгеновского снимка грудной клетки размера 224×224 тремя нейронными операторами, отличающимися видом нелинейности: ReLU, функции Габора (WIRE [63]), синусы (SIREN [62]). I – референсное изображение из набора данных PneumoniaMNIST [https://medmnist.com/].	31
2.3	Архитектура нейронного оператора Фурье. I^0 – входное изображение, требующее улучшения качества, \hat{I} – изображение, обработанное нейронным оператором Фурье, \mathcal{F} – прямое преобразование Фурье, \mathcal{F}^{-1} – обратное преобразование Фурье, \mathcal{R}_θ – параметризованный ядром оператор обрезки высоких частот, W – линейное преобразование, встроенное в пространственную связь (англ. skip connection), σ – нелинейность (например, ReLU).	33
2.4	Генерация изображения магнитно-резонансной томографии головного мозга с осцилляциями Гиббса и аддитивным гауссовским шумом. (а) – референсное изображение I^1 высокого разрешения 255×255 , (б) – Фурье-спектр изображения I^1 , (в) – результат обрезки высоких частот оператором \mathcal{R} , (г) – изображение I^0 низкого разрешения 145×145 , требующее улучшения качества, (д) – остаточный кадр $ I^0 - I^1 $, демонстрирующий артефакты на снимке. .	37
2.5	Обработка изображения МРТ головного мозга нейронным оператором Фурье и классическим методом Кельнера. (а) – входное изображение с осцилляциями Гиббса (отмечены жёлтыми стрелками) и шумом, (б) – результат обработки алгоритмом Кельнера, (в) – результат обработки нейронным оператором Фурье, аппроксимирующим решение уравнения анизотропной диффузии, использованном в модели Перона-Малика.	39

- 2.6 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 регуляризирующей способности нейронного оператора Фурье, обученного аппроксимации решения уравнения анизотропной диффузии, использованном в модели Перона-Малика на масштабе $\times 1.0$. Масштаб $\times 1.0$ соответствует равномерной пиксельной сетке 145×145 . Остальные масштабы получены линейной интерполяцией. 39
- 2.7 Схема вычисления очередного нейрона $x_{l,j}$ слоем Φ_l сети Колмогорова-Арнольда. 42
- 2.8 Схема вычисления очередного нейрона $x_{l,j}$ слоем Φ_l сети Колмогорова-Арнольда через представление функций Колмогорова-Арнольда в виде линейных комбинаций дифференцируемых базисных функций $\{\psi_k\}_{k=1}^r$ с коэффициентами $\theta_{l,jik}$, формирующими матрицу $\Theta_l \in \mathbb{R}^{m \times n \times r}$ весов l -го слоя, где m – число нейронов в l -ом слое, n – число нейронов в $(l - 1)$ -ом слое, r – число базисных функций. 43
- 2.9 Признаковое описание изображения МРТ головного мозга, состоящее из n функций, значения каждой из которых известны на сетке $h \times w$. . . 44
- 2.10 Схема слоя функциональной сети Колмогорова-Арнольда. 46
- 2.11 Архитектура модуля деформации сетки, иллюстрирующая поле смещений $\{\Delta q_{l,x}, \Delta q_{l,y}\}$, сгенерированное обученным остаточным свёрточным блоком. Предсказанные смещения аддитивно комбинируются с равномерной сеткой $\{q_x, q_y\}$, чтобы создать деформированную сетку для вычисления базисных функций Эрмита. . . 46
- 2.12 FunKAN как многоцелевой метод обработки и анализа медицинских изображений. Обучение обработке входных изображений I^0 контролируется функцией потерь $\mathcal{L}(I^*, I^1)$ и референсными изображениями I^1 47

2.13	Обработка изображения МРТ головного мозга гибридным и классическими методами. (а) – входное изображение с осцилляциями Гиббса (отмечены жёлтыми стрелками) и шумом, (б) – результат алгоритма Кельнера, (в) – результат модели Перона-Малика, (г) – результат гибридного метода, осуществляющего фильтрацию глубоких признаков описаний функциональной сетью Колмогорова-Арнольда по первым шести функциям Эрмита.	48
2.14	Качественное сравнение KAN и FunKAN в рамках одной свёрточной нейронной сети. Выделены тонкие анатомические структуры.	52
2.15	Исследование влияния количества базисных функций Эрмита в U-FunKAN на качество сегментации рака молочной железы на ультразвуковом наборе данных BUSI. Для каждого эксперимента приведены усреднённые значения метрик IoU и F1 за последние 50 эпох обучения.	56
2.16	Исследование влияния количества базисных функций Эрмита в U-FunKAN на скорость вывода (эффективность). Для каждого эксперимента приведено медианное время вывода на процессоре Intel Core i7-14700HX, усреднённое по 100 запускам для входного изображения размером 256×256 пикселей.	56
2.17	Качественные результаты работы U-FunKAN для трёх разнородных медицинских сценариев. (а) – входные изображения, (б) – референсные маски, (в) – результаты предлагаемого метода.	57
2.18	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.	58
3.1	Полнослайдовое гистологическое изображение стенки желудка из набора данных PATH-DT-MSU [https://imaging.cs.msu.ru/en/research/histology/path-dt-msu]. Снимок сделан на оптическом увеличении $\times 40$ с разрешением 65809×99600 пикселей.	60

3.2	Визуализация многомасштабной структуры полнослайдового изображения, включающего несколько уровней оптического увеличения.	61
3.3	Участок полнослайдового гистологического изображения в различных оптических увеличениях (масштабах).	62
3.4	Результат классификации полнослайдового гистологического изображения стенки желудка свёрточной нейронной сетью DenseNet121. Увеличенной яркостью выделены области, размеченные специалистами.	64
3.5	Результаты применения метода к различным многомасштабным фрагментам полнослайдового изображения. На графиках показаны зависимости расстояний (ρ) до разделяющих гиперплоскостей в зависимости от масштаба (s) обрабатываемого фрагмента. (а) – фрагмент класса АТ, (б) – фрагмент класса ВG, (в) – фрагмент класса LP, (г) – фрагмент класса ММ, (д) – фрагмент класса TUM.	69
3.6	Гистограмма по множеству предложенных методом масштабов \hat{S} для 100 случайных фрагментов LP.	70

Список таблиц

1	Средние значения PSNR, полной вариации и времени выполнения методов улучшения качества изображений МРТ головного мозга на тестовом наборе данных из 2617 изображений.	23
2	Сравнительный анализ модулей нелинейной фильтрации в рамках одной свёрточной нейронной сети для повышения качества изображений МРТ. Результаты отражают среднее пиковое отношение сигнал/шум и полную вариацию для 2617 тестовых изображений (145×145) из набора данных IXI. I^0 , I^1 обозначают изображения с артефактами и без артефактов соответственно.	51
3	Сравнение качества методов сегментации в трёх клинически различных медицинских сценариях. Результаты включают средние значения IoU и F1 со стандартным отклонением для трёх случайных запусков обучений.	53
4	Сравнение вычислительной эффективности методов сегментации. Результаты получены с помощью инструмента THOR Python для профилирования моделей PyTorch.	53
5	Исследование влияния масштабирования каналов в U-FunKAN на качество сегментации и эффективность. Приведены наилучшие значения метрик IoU и F1.	55
6	Точность (%) классификации фрагментов (224×224) полнослайдового изображения стенки желудка в зависимости от выбранного масштаба входных данных. Асс – общая точность, Асс@АТ – точность распознавания участков подслизистой основы, Асс@BG – точность распознавания фона, Асс@LP – точность распознавания неизменённых участков собственной пластинки слизистой оболочки, Асс@ММ – точность распознавания неизменённых участков мышечной пластинки слизистой оболочки, Асс@TUM – точность распознавания участков аденокарциномы. . . .	65