

Московский государственный университет имени М.В. Ломоносова
Факультет вычислительной математики и кибернетики
Кафедра математической физики

Пчелинцев Яков Антонович

**Математические методы адаптивного повышения качества
биомедицинских изображений**

Специальность 1.2.2 —
«Математическое моделирование, численные методы и комплексы программ»

Научно-квалификационная работа

Научный руководитель:
доктор физико-математических наук
Крылов Андрей Серджевич

Москва, 2023

Оглавление

Стр.

Введение	4
---------------------------	----------

Глава 1. Регуляризирующий метод повышения разрешения изображений мигающей флуоресцентной микроскопии	8
---	----------

1.1 Модель искажения изображения	9
1.1.1 Модель размытия изображения	10
1.1.2 Модель ядра размытия	10
1.2 Описание данных	11
1.2.1 Источники данных	11
1.2.2 Детектор Airyscan	11
1.3 Рассматриваемые методы	13
1.3.1 Визуализация оптических флуктуаций сверхвысокого разрешения (SOFI)	13
1.3.2 Алгоритм классификации множественных сигналов для флуоресцентной микроскопии сверхвысокого разрешения MUSICAL	15
1.3.3 Байесовский анализ мигания и выгорания 3В	17
1.4 Итерационный регуляризирующий алгоритм	18
1.5 Эксперименты и результаты	21
1.6 Выводы	23

Глава 2. Повышение резкости медицинских изображений методом деформации пиксельной сетки	25
--	-----------

2.1 Реалистичные ядра размытия изображений	25
2.2 Построение тестового набора изображений	28
2.3 Деформационный метод повышения резкости изображений	28
2.4 Модели функции смещения	30
2.5 Методика оптимизации параметров моделей функций смещения	31
2.6 Результаты экспериментов	32
2.7 Выводы	35

Глава 3. Нейросетевой алгоритм контроля качества рентгеновских снимков грудной клетки в задаче диагностики туберкулёза лёгких	36
--	-----------

3.1 Метод оценки жёсткости рентгеновских снимков грудной клетки	37
3.1.1 Использованные данные	38
3.1.2 Адаптивная предобработка входных изображений	41
3.1.3 Метод определения жёсткости рентгеновских снимков грудной клетки	43
3.1.4 Эксперименты и результаты	44

3.2 Использование результатов оценки качества рентгеновских снимков при нейросетевой диагностике туберкулёза лёгких	49
3.2.1 Формирование набора данных	49
3.2.2 Метод компьютерной диагностики туберкулёза лёгких	51
3.2.3 Эксперименты и результаты	51
3.3 Выводы	54
Глава 4. Программный комплекс реализации алгоритмов анализа биомедицинских изображений	55
4.1 Программный модуль повышения разрешения изображений мигающей флуоресцентной микроскопии	55
4.2 Программный повышения резкости медицинских изображений методом деформации пиксельной сетки	57
4.3 Программный модуль анализа и обработки рентгенограмм грудной клетки	58
4.4 Набор рентгеновских изображений грудной клетки для компьютерной диагностики туберкулёза лёгких	59
4.4.1 Общедоступные данные	59
4.4.2 Набор рентгенограмм грудной клетки Sakha-TB	61
4.4.3 Эксперименты	63
4.4.4 Результаты	64
Заключение	67
Список литературы	68

Введение

Актуальность темы.

В настоящее время в разных областях биологии и медицины повсеместно используются цифровые изображения и развиваются алгоритмы их обработки, в том числе и на основе глубокого обучения. Несмотря на быстрый научно-технический прогресс, в силу различных причин оборудование, используемое для получения таких изображений, не всегда способно решить возникающие задачи, однако применение математических методов позволяет приблизиться к их решению и повысить качество уже полученных данных. Поэтому разработка алгоритмов повышения качества изображений с учётом современных условий остаётся актуальной.

Изображения клеточных структур, возникающие в области флуоресцентной микроскопии, являются одним из классов изображений, увеличение разрешения которых необходимо на практике. Одними из примеров таких задач являются ситуации, при которых разрешающая способность оптической системы микроскопа упирается в теоретический предел из-за дифракции света или когда увеличение оптической системы микроскопа не позволяет реализовать весь потенциал теоретической разрешающей способности.

Хотя появились методики преодоления дифракционного предела с использованием специального оборудования, на практике более доступным подходом является использование специальных стохастический мигающих красителей. Серия снимков, на которых молекулы красителя имеют разную яркость, преобразуется в изображение повышенного разрешения. Благодаря повышенному по сравнению с одним снимком количеству информации и слабой корреляции миграции молекул флуорофора, вычислительные методы позволяют существенно повысить разрешающие способности микроскопов без дорогостоящей модификации оборудования. В то же время из-за необходимости получать большое число изображений за короткий срок, повышается значимость таких параметров, как уровень шума и скорость выгорания красителей.

Для макроскопических же изображений в медицине по причине строгих ограничений со стороны закона и специфики области, связанной с повышенными рисками ошибок, важен вопрос аккуратной обработки изображений для избежания потери или привнесения информации в имеющиеся данные и контроля качества получаемых изображений. Очень актуальна задача повышения резкости изображений. А для методов глубокого обучения в медицинской диагностике, которые внедряются в практику во всём мире, является анализ и предобработка входных данных. Необходим контроль соответствия входной информации и применяемого обученного или обучаемого алгоритма глубокого обучения.

При изучении рентгеновских снимков и, в частности, при диагностике туберкулёза лёгких важным фактором является жёсткость снимка, так как она напрямую влияет на его информативность: на изображении должны в достаточном количестве

присутствовать важные для вынесения верного решения детали. Кроме того, необходима правильная настройка контрастности снимков, чтобы различия изображений из обучающего набора и тестовых снимков, которые получаются на различном оборудовании в разных условиях, не были велики настолько, чтобы влиять на качество работы алгоритмов диагностики.

Основное внимание в данной работе уделено повышению качества изображений клеточных структур и анализу качества медицинских изображений. Под качеством понимается как наличие у изображения определённых характеристик (резкость, контрастность, разрешение), так и их степень сходства с эталонными изображениями, для работы с которыми оптимизирован алгоритм машинного или глубокого обучения.

Цели и задачи работы

Цель данной работы состоит в разработке адаптивных методов анализа и обработки биомедицинских изображений различных модальностей на основе методов математического моделирования, их алгоритмическая и программная реализация для решения задач повышения разрешения изображений флуоресцентной микроскопии, повышения резкости и контроля качества медицинских изображений.

Научная новизна

В данной работе были получены новые методы и алгоритмы:

1. регуляризующий метод повышения разрешения и резкости изображений флуоресцентной мигающей микроскопии;
2. малопараметрический деформационный метод повышения резкости изображений;
3. нейросетевой метод контроля качества рентгеновских снимков грудной клетки, основанный на автоматическом анализе жёсткости рентгенограммы;
4. алгоритм компьютерной диагностики туберкулёза лёгких по рентгеновскому снимку грудной клетки.

Теоретическая и практическая ценность

Создан программный комплекс повышения разрешения изображений флуоресцентной мигающей микроскопии, повышения резкости медицинских изображений и определения качества рентгенограмм грудной клетки для задачи диагностики туберкулёза лёгких.

Предложенные методы могут применяться как отдельно, так и в составе других систем анализа и повышения качества биомедицинских изображений, в качестве вспомогательных инструментов в работе учёных-биологов и врачей, а также могут быть модифицированы для применения в иных областях медицины и биологии, отличных от рассмотренных в работе.

Степень разработанности темы.

Исследование, проведённое в данной работе, затрагивает три различных области анализа и повышения качества биомедицинских изображений.

В области флуоресцентной микроскопии, получившей распространение благодаря таким своим достоинствам, как высокий контраст получаемых изображений,

возможность избирательного окрашивания структур разной природы в разные цвета и наблюдения за живыми образцами, к настоящему времени был разработан ряд методов повышения разрешающей способности микроскопов. Некоторые из них (конфокальная микроскопия, STED, SIM) требуют особого оборудования, другие же (STORM, PALM, SOFI, SRRF, MUSICAL) появились уже в 21 веке и полагаются на использование специальных мигающих красителей и вычислительные алгоритмы.

В прошлом и в текущем веке было разработано множество эффективных методов повышения резкости и восстановления смазанных изображений. Многие из них требуют указания типа размытия для работы, и хотя за последние 20 лет появились алгоритмы автоматического определения размытия по входному изображению, чувствительность методов повышения резкости к шуму и ошибкам в определении и указании типа и силы размытия остаётся значимым фактором. Метод повышения резкости изображений с помощью деформации пиксельной сетки не предназначен для решения задачи полноценного восстановления смазанных изображений, но может служить в качестве этапа постобработки восстановленного другим алгоритмом изображения, повышая резкость без порождения артефактов на изображении. Выбор оптимальной функции смещения пикселей позволяет регулировать эффективность обработки изображения.

С новым витком развития области науки в середине 2010-х годов и повсеместном распространении алгоритмов на основе искусственного интеллекта и глубокого обучения вопрос о влиянии качества входных данных на результаты обработки и о контроле качества этих данных снова получил высокую актуальность. За последнее десятилетие были проведены исследования на тему целенаправленных атак на алгоритмы анализа данных с целью повлиять на результаты их работы путём незаметного для человека изменения входных данных и защиты от таких атак. В то же время в области медицинской диагностики заболеваний по рентгенограммам грудной клетки за последние 10 лет были разработаны методы контроля некоторых условий съёмки, изучено влияние качества снимков на точность диагностики некоторых заболеваний.

Методология и методы исследования.

В основе методологии исследования лежат методы математического моделирования в обработке и анализе изображений, ряд вычислительных экспериментов реализован в рамках задач машинного обучения и анализа изображений с помощью искусственных и реальных данных.

Степень достоверности результатов

Достоверность результатов проведённых исследований обеспечивается опорой на теоретическую базу, математической обоснованностью разработанных методов, воспроизводимыми вычислительными экспериментами и тестированием алгоритмов на искусственных и реальных данных. Значительная часть данных, использованных для создания и тестирования разработанных методов, находится в открытом доступе.

Апробация работы

Основные результаты работы докладывались на:

1. 7-й Международной конференции по теории обработки изображений, методам и применению «IPTA 2017» (Монреаль, Канада, 2017);
2. 7-м Европейском семинаре по обработке визуальной информации «EUVIP 2018» (Тампере, Финляндия, 2018);
3. 4-й Международной конференции по обработке биомедицинских изображений и сигналов «ICBSP 2019» (Нагоя, Япония, 2019);
4. Всероссийской конференции «Ломоносовские чтения-2020» (Москва, 2020);
5. 6-й Международной конференции по обработке биомедицинских изображений и сигналов «ICBSP 2021» (Сямынь, Китай, 2021);

Публикации

По теме исследования опубликовано 8 работ, из них 6 работ в изданиях, индексируемых системами Web of Science и Scopus, рекомендованных для защиты в диссертационном совете МГУ имени М.В. Ломоносова по специальности 1.2.2. «Математическое моделирование, численные методы и комплексы программ», и 2 работы, опубликованные в иных изданиях.

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 работы в полученные результаты математического моделирования, численные методы и разработку комплекса программ является определяющим.

Основные положения, выносимые на защиту

1. Метод повышения разрешения изображений мигающей флуоресцентной микроскопии, основанный на итерационном процессе максимизации регуляризующего функционала.
2. Малопараметрический метод повышения резкости медицинских изображений на основе деформации пиксельной сетки.
3. Метод автоматического контроля качества рентгенограмм грудной клетки для задач машинного обучения, основанный на анализе и классификации уровня жёсткости рентгеновского снимка грудной клетки.
4. Программный комплекс повышения разрешения изображений мигающей флуоресцентной микроскопии, повышения резкости медицинских изображений, анализа качества рентгенограмм грудной клетки.

Объем и структура работы.

Работа состоит из введения, 4 глав, заключения.

Глава 1. Регуляризующий метод повышения разрешения изображений мигающей флуоресцентной микроскопии

Флуоресцентная микроскопия обладает рядом достоинств, среди которых высокая контрастность изображений и возможность маркировки разных структур разными красителями. Благодаря этому данный вид микроскопии получил широкое распространение. Однако разрешающая способность микроскопа ограничена дифракцией, а использование видимого света делает разрешающую способность изображений флуоресцентной микроскопии недостаточной для ряда современных задач. В связи с этим развиваются различные подходы решения данной проблемы.

Двумя основными направлениями методов повышения разрешающей способности микроскопов являются восстановление размытых изображений и повышение разрешения изображений. Получение изображения с более высоким разрешением необходимо, если сам микроскоп не обеспечивает должного увеличения для различения мелких объектов. Если же микроскоп имеет сильное увеличение, то основным препятствием становится дифракция света, поэтому в таком случае задача устранения размытия выходит на передний план.

Использование стохастически мигающих флуорофоров даёт дополнительную информацию для анализа изображения и позволяет выделять сигналы отдельных участков объекта или отдельных флуорофоров на фоне друг друга при наличии временной серии изображений. Такой подход не требует специального оборудования, поэтому в настоящее время получил широкое распространение и продолжает активно развиваться [1]. В этой главе рассматриваются именно такие методы повышения качества изображений флуоресцентной микроскопии на основе анализа временной серии снимков образца, подсвеченного стохастически мигающим красителем. Пример последовательных кадров из рассматриваемых серий изображений представлен на Рис. 1.1.

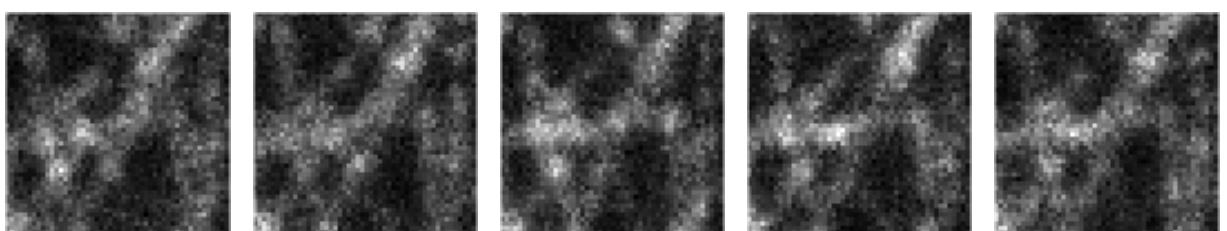


Рисунок 1.1: Пример серии изображений, полученных с использованием мигающих флуорофоров

В последние годы для этой области микроскопии было разработано множество алгоритмов, основанных на различных подходах, для решения задачи построения резкого изображения с высоким разрешением на основе серии снимков. В частности, алгоритмы PALM [2] и STORM [3] использует редко мигающие красители, чтобы находить центры отдельных пятен в каждом кадре и строить карту молекул флуорофора. Ограничением

является требование низкой частоты мигания молекул и длинных серий снимков для визуализации всех участков изучаемого объекта.

Метод SOFI [4; 5] использует корреляцию и полуинварианты [6; 7] пикселей для уменьшения размытия и увеличения разрешения изображения. Однако он чувствителен к шуму, поэтому требует длинных последовательностей данных, а также этапа устранения размытия.

В то время как полуинварианты, используемые в SOFI, оценивают только часть ковариационной матрицы серии изображений, SPARCOM [8] полностью использует эту матрицу и использует предположение о разреженности в пространстве корреляций сигналов флуорофоров для достижения более высокого пространственного разрешения и уменьшения требований к длине последовательности изображений.

Метод MUSICAL [9] вдохновлён алгоритмом обработки радиолокационных сигналов MUSIC [10]. Он использует собственные вектора серии изображений для построения карты флуорофоров с высоким разрешением.

Другой подход предложен авторами ЗВ [11]. Их метод байесовского анализа мигания и выгорания моделирует изменение состояний флуорофоров и вычисляет расположение молекул. Однако его вычислительная сложность очень высока, и для практического использования этого алгоритма требуется кластерный компьютер.

Для этой задачи также разработаны методы, основанные на глубоком обучении. Например, свёрточная нейтральная сеть Deep-STORM [12] основана на архитектуре кодировщик-декодировщик и напрямую создаёт изображения с высоким разрешением, тогда как метод DLBI [13] сочетает в себе как глубокое обучение, так и байесовский вывод, который уточняет результат нейронной сети.

Однако проблема всё ещё остаётся открытой, и требуется более высокая производительность алгоритмов улучшения изображения. Поэтому в этой главе будет представлен регуляризующий метод реконструкции изображения повышенного разрешения и сравнение его эффективности с некоторыми другими современными методами.

1.1 Модель искажения изображения

Используется следующая модель искажения изображения:

$$y_{ij} = (Kx + n)_{ij}, \quad K = DB,$$

где x — исходное резкое изображение высокого разрешения, y — наблюдаемое изображение, K — оператор размытия и понижения разрешения, действие которого эквивалентно последовательному действию оператора размытия B и оператора понижения разрешения D (производящего, например, усреднение групп соседних пикселей или простое

прореживание пиксельной сетки изображения), а n — аддитивный шум с нормальным распределением.

1.1.1 Модель размытия изображения

Рассматриваемая модель размытия изображения имеет следующий вид:

$$x'(p, q) = (B * x)(p, q) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B(s, t) x(p - s, q - t) ds dt,$$

где x — исходное резкое изображение, x' — искажённое изображение, B — ядро размытия (англ. point spread function, PSF), что в дискретном случае принимает вид:

$$x'_{ij} = (B * x)_{ij} = \sum_{k=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} B_{kp} x_{i-k, j-p}.$$

1.1.2 Модель ядра размытия

В лазерном сканирующем микроскопе ядро размытия $PSF_{eff}(r)$ симметрично и определяется произведением ядра размытия лазерного луча подсветки образца $PSF_{exc}(r, r_{0,exc})$ и ядра размытия объектива $PSF_{em}(r, r_{0,em})$. Каждое из этих ядер задаётся формулами:

$$\begin{aligned} PSF(r) &= I_0 \cdot \left(\frac{2 \cdot J_1(v)}{v} \right)^2, \\ v &= \frac{2 \cdot \pi \cdot d_a \cdot |r - r_0|}{\lambda}, \\ J_1(v) &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2^{2n+1} n! (n+1)!}, \end{aligned}$$

где I_0 — нормирующий множитель, J_1 — функция Бесселя первого рода первого порядка, λ — длина волны лазера подсветки или флуоресцентного излучения, r — точка в плоскости изображения, r_0 — центр ядра в плоскости изображения.

Характер отдельного ядра размытия $PSF(r)$ определён дифракцией света, а само такое ядро называется диском Эйри. Расстояние d_a между ближайшими к центру этого ядра нулями называется диаметром диска Эйри (англ. airy unit, AU) и равно $\frac{\mu \cdot \lambda}{\pi A_n}$, где μ — наименьший положительный нуль функции Бесселя первого рода первого порядка, а A_n — числовая апертура фокусирующей линзы.

Итоговое ядро размытия получаемого микроскопом изображения имеет вид $PSF_{eff}(r) = PSF_{exc}(r, r_{0,exc}) \cdot PSF_{em}(r, r_{0,em})$, где r — точка на изображении, $r_{0,exc}$,

$r_{0,em}$ — центры ядер размытия лазера подсветки и флуоресцентного излучения соответственно (см. Рис. 1.2).

1.2 Описание данных

1.2.1 Источники данных

Искусственные данные были подготовлены с помощью инструмента для симуляции процесса получения серий снимков мигающей флуоресцентной микроскопии SOFI Simulation Tool [14], модифицированном для получения структуры в виде сходящихся отрезков.

Экспериментальные данные были получены в сотрудничестве с биологами Токийского университета с помощью микроскопа Zeiss LSM 880 с детектором Airyscan [15], состоящем из 32 сенсоров, и мигающего флуоресцентного красителя HMSiR [16]. Серия снимков представляла собой 32 синхронных серии с разных сенсоров микроскопа, которые были предварительно совмещены.

1.2.2 Детектор Airyscan

Так как итоговое ядро размытия лазерного сканирующего микроскопа представляет собой произведение ядра размытия луча подсветки и ядра размытия излучения объекта (см. п. 1.1.2), то при смещении детектора излучённого света в сторону от оптической оси микроскопа результирующее ядро размытия будет меняться, становясь более узким в центральной части, но приобретая более весомый «хвост», как показано на Рис. 1.3. На этом принципе основана работа детектора Airyscan.

Изображение получается синхронно 32 детекторами: одним центральным и расположеннымми вокруг него тремя кольцами детекторов (Рис. 1.4). Суммарный размер детектора Airyscan равен 1.25 диаметра диска Эйри, а размер одного элемента равен $\frac{1.25}{6} d_a$. Затем изображения с детекторов смещаются в сторону оптической оси с тем, чтобы положения максимумов ядер размытия детекторов совпали. Этот процесс называется переназначением пикселей (англ. pixel reassignment) [17]. Величина смещения v изображения с одного детектора к центру определяется по формулам:

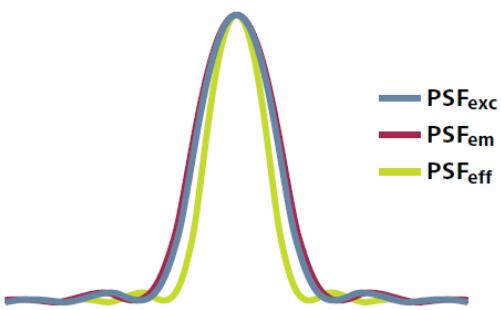


Рисунок 1.2: Нормированные профили ядер размытия лазера подсветки (PSF_{exc}), объектива (PSF_{em}) и микроскопа (PSF_{eff})

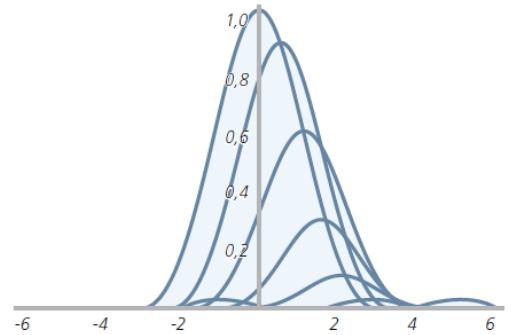


Рисунок 1.3: Изменение профиля ядра размытия при смещении детектора излучаемого образцов света от оптической оси

$$v = -\frac{1}{1+\beta} \cdot \frac{1.25 d_a}{6} \cdot v_d,$$

$$\beta = \frac{\lambda_{em}}{\lambda_{exc}},$$

где v_d — относительное смещение детектора d от оптической оси в размерах детектора, λ_{em} и λ_{exc} — длины волн излучённого света и подсветки. После этого изображения суммируются, образуя изображение с более узким ядром размытия (Рис. 1.5) и повышенным отношением уровня сигнала к шуму.

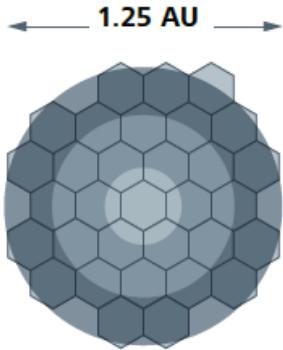


Рисунок 1.4: Схема детектора Airyscan. Сенсоры расположены в центрах правильных шестиугольников

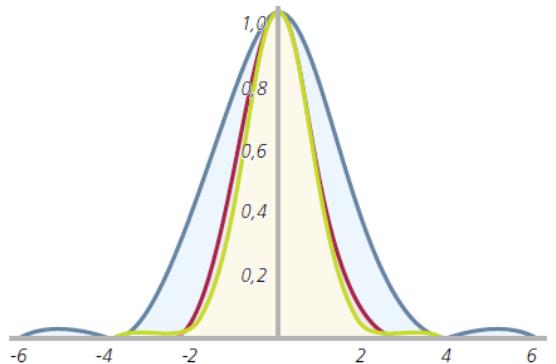


Рисунок 1.5: Нормированные профили ядер размытия обычного широкопольного (синий), конфокального сканирующего (красный) микроскопов и микроскопа с детектором Airyscan (зелёный)

За счёт такого подхода достигается резкость конфокальных сканирующих микроскопов одновременно с высоким уровнем сигнала обычных широкопольных микроскопов, что позволяет делать длинные серии снимков с высокой частотой кадров. Примеры ядер размытия для центрального элемента и элементов каждого из колец Airyscan представлены на Рис. 1.6.

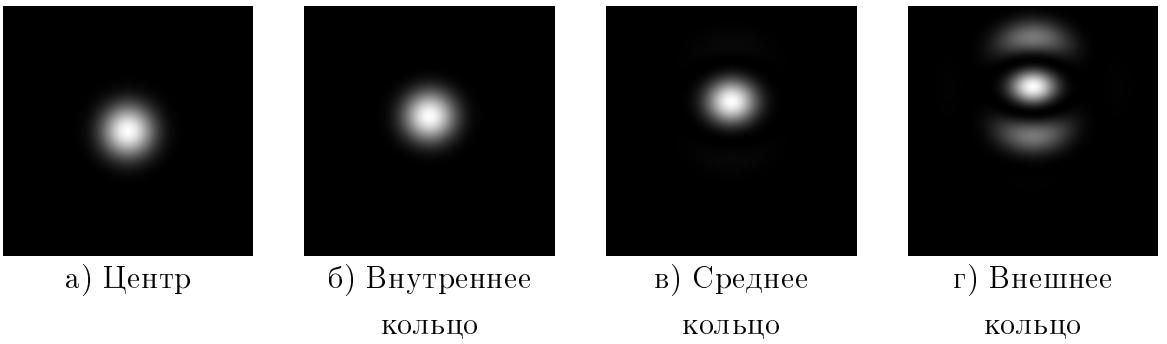


Рисунок 1.6: Нормированные по яркости ядра размытия для элементов из разных слоёв детектора Airyscan

1.3 Рассматриваемые методы

В данном параграфе рассматривается ряд современных методов повышения разрешения и резкости флуоресцентных изображений, полученных с использованием мигающих красителей. Далее следует описание их базовых принципов.

1.3.1 Визуализация оптических флюктуаций сверхвысокого разрешения (SOFI)

Метод SOFI (super-resolution optical fluctuation imaging) основан на рассмотрении интенсивностей излучения молекул флуорофора как независимых случайных величин и использовании полуинвариантов (англ. cumulants) — коэффициентов разложения в степенной ряд логарифма характеристической функции случайной величины, также могут быть выражены через моменты случайной величины [6; 7] — для получения резкого изображения повышенного разрешения [4; 5].

В рамках этого подхода рассматриваются N независимо мигающих флуорофоров, находящихся в позициях \mathbf{r}_k . Их яркости меняются со временем по закону $\varepsilon_k \cdot s_k(t)$, а итоговая интенсивность в позиции r задаётся следующим образом:

$$\sum_{k=1}^N \delta(\mathbf{r} - \mathbf{r}_k) \cdot \varepsilon_k \cdot s_k(t), \quad \delta(\mathbf{r} - \mathbf{r}_k) = \begin{cases} 1, & \mathbf{r} = \mathbf{r}_k, \\ 0, & \mathbf{r} \neq \mathbf{r}_k, \end{cases}$$

где ε_k — постоянная яркость молекулы красителя, а $s_k(t)$ — зависящая от времени функция.

Считая, что позиции молекул флуорофора не меняются со временем, а ядро размытия изображения $U(\mathbf{r})$ одинаково для всех участков изображения, интенсивность $F(\mathbf{r}, t)$ в позиции r в момент времени t можно записать в следующем виде:

$$F(\mathbf{r},t) = \sum_{k=1}^N U(\mathbf{r} - \mathbf{r}_k) \cdot \varepsilon_k \cdot s_k(t).$$

Изменение яркости также может быть выражено как флуктуации с нулевым математическим ожиданием:

$$\begin{aligned}\delta F(\mathbf{r},t) &= F(\mathbf{r},t) - \langle F(\mathbf{r},t) \rangle_t = \\ &= \sum_k U(\mathbf{r} - \mathbf{r}_k) \cdot \varepsilon_k \cdot [s_k(t) - \langle s_k(t) \rangle_t] = \\ &= \sum_k U(\mathbf{r} - \mathbf{r}_k) \cdot \varepsilon_k \cdot \delta s_k(t),\end{aligned}$$

где $\langle \dots \rangle_t$ означает усреднение по времени. Тогда автокорреляционная функция второго порядка задаётся формулой

$$\begin{aligned}G_2(\mathbf{r},\tau) &= \langle \delta F(\mathbf{r},t + \tau) \cdot \delta F(\mathbf{r},t) \rangle_t = \\ &= \sum_{j,k} U(\mathbf{r} - \mathbf{r}_j) U(\mathbf{r} - \mathbf{r}_k) \cdot \varepsilon_j \cdot \varepsilon_k \cdot \langle \delta s_j(t + \tau) \delta s_k(t) \rangle_t = \\ &= \sum_k U^2(\mathbf{r} - \mathbf{r}_k) \cdot \varepsilon_k^2 \cdot \langle \delta s_k(t + \tau) \delta s_k(t) \rangle_t.\end{aligned}$$

Последний переход обусловлен независимостью сигналов от разных молекул крашителя, перекрёстные члены суммы принимают нулевые значения для всех \mathbf{r} и t .

Как видно из последнего уравнения, автокорреляционная функция второго порядка имеет вид свёртки некоторого резкого изображения, задаваемого формулой $\sum_{k=1}^N \delta(\mathbf{r} - \mathbf{r}_k) \cdot \varepsilon_k^2 \cdot \langle \delta s_k(t + \tau) \delta s_k(t) \rangle_t$, с квадратом исходного ядра размытия. Благодаря тому, что полная ширина на полувысоте (англ. full width at half maximum, FWHM) возведённого в степень $\alpha > 1$ ядра меньше, чем у исходного ядра, результат работы алгоритма SOFI обладает более высокой резкостью, чем необработанное изображение, а также позволяет достичь лучшего результата при дальнейшем восстановлении размытого изображения из-за усиленных высоких частот в спектре изображения. В частности, в случае гауссовского размытия квадрат ядра размытия в $\sqrt{2}$ раз уже самого ядра.

Для порядков выше второго используется не автокорреляционная функция, а полуинварианты соответствующих порядков, так как в противном случае перекрёстные члены в сумме остаются.

Хотя теоретического ограничения на степень повышения резкости нет, на практике используются полуинварианты порядков не выше четвёртого из-за роста влияния шума на результат с ростом порядка полуинварианта. А поскольку результат преобразования содержит не яркость молекул, а значения, зависящие от колебания яркости, которые могут быть в том числе равными нулю или меньше него, то только полуинвариант второго порядка (дисперсия яркости) даёт наиболее устойчивый результат с минимальной потерей данных и деталей изображения. Это и обусловило выбор второго порядка SOFI для проведения сравнения методов.

При замене полуинвариантов на смешанные полуинварианты (англ. joint cumulants) — коэффициенты разложения в степенной ряд логарифма характеристической функции случайного вектора, составленного из нескольких случайных величин [7] — получившийся алгоритм XC-SOFI n -го порядка, помимо повышения резкости, обеспечивает и повышение разрешения в n раз. В этом случае в преобразовании участвует несколько соседних пикселей изображения. В частности, для второго порядка получаем следующее выражение:

$$\begin{aligned} XC_2(\mathbf{r}_1, \mathbf{r}_2, \tau_1, \tau_2) &= \\ &= \sum_{i=1}^N U(\mathbf{r}_1 - \mathbf{r}_i) U(\mathbf{r}_2 - \mathbf{r}_i) \cdot \varepsilon_k^2 \cdot \langle \delta s_i(t + \tau_1) \delta s_i(t + \tau_2) \rangle_t = \\ &= U\left(\frac{\mathbf{r}_1 - \mathbf{r}_2}{\sqrt{2}}\right) \sum_{i=1}^N U^2\left(\frac{\mathbf{r}_1 + \mathbf{r}_2}{2} - \mathbf{r}_i\right) \cdot \varepsilon_i^2 \cdot \langle \delta s_i(t + \tau_1) \delta s_i(t + \tau_2) \rangle_t. \end{aligned}$$

Отсюда видно, что:

1. Позиция результирующего пикселя лежит между двумя пикселями, участвующими в преобразовании.
2. Результирующий пиксель имеет дополнительный множитель $U\left(\frac{\mathbf{r}_1 - \mathbf{r}_2}{\sqrt{2}}\right)$, поэтому необходимо произвести перевзвешивание результирующего изображения для приведения уровня яркости разных групп пикселей к одному. При неизвестном ядре размытия это можно сделать, например, нормировав яркости пикселей каждой группы на среднее отклонение или математическое ожидание их яркостей.

Для более высоких порядков формула приобретает вид:

$$\begin{aligned} XC_n(\mathbf{r}_1, \dots, \mathbf{r}_n, \tau_1, \dots, \tau_n) &= \\ &= \prod_{j < l}^n U\left(\frac{\mathbf{r}_j - \mathbf{r}_l}{\sqrt{n}}\right) \cdot \sum_{i=1}^N U^n\left(\frac{\sum_k^n \mathbf{r}_k}{n} - \mathbf{r}_i\right) \cdot \varepsilon_i^n \cdot w_i(\tau_1, \dots, \tau_n), \end{aligned}$$

где w_i — смешанный полуинвариант функций $s_i(t)$ порядка n .

1.3.2 Алгоритм классификации множественных сигналов для флуоресцентной микроскопии сверхвысокого разрешения MUSICAL

Метод повышения разрешения и резкости изображений MUSICAL (Multiple signal classification algorithm for super-resolution fluorescence microscopy) [9] основан на методе разделения нескольких сигналов MUSIC [10].

В рамках этого метода интенсивность n -го пикселя k -го кадра ($1 \leq k \leq K$, где K — общее число изображений в обрабатываемой серии) задаётся формулой

$$I_k(\vec{r}_n) = \int_{y_n - \frac{w}{2}}^{y_n + \frac{w}{2}} \int_{x_n - \frac{w}{2}}^{x_n + \frac{w}{2}} \sum_{m=1}^M G(\vec{r}_n, \vec{r}'_m) s_m(k) dx dy \approx \sum_{m=1}^M w^2 G(\vec{r}_n, \vec{r}'_m) s_m(k),$$

где $\vec{r}_n = (x_n, y_n)$ — позиция n -го пикселя в плоскости изображения, w — длина его стороны, \vec{r}'_m — позиция m -го флуорофора в плоскости образца, $s_m(k) \geq 0$ — суммарное излучение от m -го флуорофора за время выдержки k -го кадра (при этом $\exists k : s_m(k) > 0$), а $G(\vec{r}, \vec{r}')$ — функция, определяющая ядро размытия, отображающая интенсивность излучения из точки \vec{r}' в плоскости образца в точку \vec{r} в плоскости изображения.

Таким образом, интенсивности всех пикселей k -го изображения могут быть записаны в виде матричного уравнения: $\bar{I}_k = \mathbf{G} \bar{s}_k$, где

$$\begin{aligned} \bar{I}_k &= \left[I_k(\vec{r}_1) \quad I_k(\vec{r}_2) \quad \cdots \quad I_k(\vec{r}_N) \right]^T, \\ \mathbf{G} &= w^2 \left[\bar{G}(\vec{r}'_1) \quad \bar{G}(\vec{r}'_2) \quad \cdots \quad \bar{G}(\vec{r}'_M) \right], \\ \bar{G}(\vec{r}') &= \left[G(\vec{r}_1, \vec{r}') \quad G(\vec{r}_2, \vec{r}') \quad \cdots \quad G(\vec{r}_N, \vec{r}') \right]^T, \\ \bar{s}_k &= \left[s_1(k) \quad s_2(k) \quad \cdots \quad s_M(k) \right]^T. \end{aligned}$$

Полный набор изображений может быть записан матрицей

$$\mathbf{I} = \left[\bar{I}_1 \quad \bar{I}_2 \quad \cdots \quad \bar{I}_K \right] = \mathbf{G} \cdot \left[\bar{s}_1 \quad \bar{s}_2 \quad \cdots \quad \bar{s}_K \right].$$

Отсюда видно, что каждый столбец матрицы \mathbf{I} — это линейная комбинация столбцов матрицы \mathbf{G} , а пространство столбцов матрицы \mathbf{I} , являющееся линейной оболочкой полученных изображений, определяется пространством столбцов матрицы \mathbf{G} . Одновременно с этим пространство столбцов матрицы \mathbf{I} задаётся левыми сингулярными векторами из сингулярного разложения этой матрицы, у которых соответствующие им сингулярные числа не равны нулю: $\{\bar{u}_{\sigma_i \neq 0}\}$.

Если \mathbf{I} — матрица неполного ранга, то существует линейное пространство векторов $\{\bar{u}_{\sigma_i=0}\}$, ортогональное пространству столбцов матрицы \mathbf{I} , и выполняется $\bar{G}(\vec{r}'_m) \cdot \bar{u}_{\sigma_i=0} = 0$ для всех m . При условии, что флуорофоры мигают независимо, одному элементу из $\{\bar{G}(\vec{r}'_m); m = \overline{1; M}\}$ соответствует один и только один элемент из $\{\bar{u}_{\sigma_i \neq 0}\}$. Поэтому вектора $\bar{G}(\vec{r}')$ для позиций $\vec{r}' \notin \{\vec{r}'_m, m = \overline{1; M}\}$ имеют ненулевую проекцию на линейное пространство векторов $\{\bar{u}_{\sigma_i=0}\}$.

В случае полного ранга выбирается некоторый порог σ_0 и за базис пространства столбцов матрицы берётся множество $\{\bar{u}_{\sigma_i \geq \sigma_0}\}$, а за базис ортогонального сигнала пространства — $\{\bar{u}_{\sigma_i < \sigma_0}\}$. Обусловлен такой подход тем, что разные вектора \bar{u}_{σ_i} соответствуют разным структурным деталям изображений, соответствующие им σ_i^2 определяют энергию (силу) этих векторов, а вектора с низкой энергией могут быть

сильно искажены шумом и не могут считаться надёжными при восстановлении изображения высокого разрешения.

Для проверки условий принадлежности вектора $\bar{G}(\vec{r}')$ к $\{\bar{G}(\vec{r}'_m); m = \overline{1; M}\}$ в MUSICAL используется индикаторная функция с параметром α , отвечающим за контраст:

$$f(\vec{r}'_{test}) = \left(\frac{d_{PR}(\vec{r}'_{test})}{d_{PN}(\vec{r}'_{test})} \right)^\alpha,$$

$$d_{PR}(\vec{r}'_{test}) = \sqrt{\sum_{\sigma_i \geq \sigma_0} \|\bar{G}(\vec{r}'_{test}) \cdot \bar{u}_i\|^2},$$

$$d_{PN}(\vec{r}'_{test}) = \sqrt{\sum_{\sigma_i < \sigma_0} \|\bar{G}(\vec{r}'_{test}) \cdot \bar{u}_i\|^2}.$$

Значения такой индикаторной функции будут тем больше, чем ближе точка \vec{r}' к позиции одного из m флуорофоров. Посчитав значения этой функции для всех пикселей изображения высокого разрешения, можно построить своего рода карту расположения молекул флуорофоров.

Для ускорения работы алгоритма и устойчивости определения базиса пространства сигнала алгоритм MUSICAL обрабатывает изображение не целиком, а с использованием скользящего окна, сопоставимого по размеру с диаметром ядра размытия, с дальнейшим усреднением значений индикаторной функции в областях перекрытия окон.

1.3.3 Байесовский анализ миграции и выгорания ЗВ

Ещё одним рассмотренным методом является метод локализации молекул флуорофора с помощью байесовского алгоритма локационной микроскопии [11]. Метод использует скрытую марковскую модель для представления переходов флуорофора между тремя состояниями (излучает, не излучает, разрушен) и делает вывод о присутствии флуорофора в точке на основе отношения вероятности его присутствия (событие \mathcal{F}) в этой точке при учёте наблюдаемых данных D к вероятности его отсутствия (событие \mathcal{N}) при тех же данных: $\frac{p(\mathcal{F}|D)}{p(\mathcal{N}|D)} = \frac{p(D|\mathcal{F})p(\mathcal{F})}{p(D|\mathcal{N})p(\mathcal{N})}$. Так как $p(\mathcal{F})$ и $p(\mathcal{N})$ постоянные, то оценить нужно только вероятности $p(D|\mathcal{F})$ и $p(D|\mathcal{N})$. Последнее — это вероятность наблюдения данных при учёте наличия в них только шума известной модели (например, шума с нормальным распределением). Первая величина вычисляется следующим образом:

$$p(D|\mathcal{F}) = \int_{a \in \mathbb{R}^4} \int_{b \in \mathbb{Z}_3^N} p(D, a, b | \mathcal{F}) \, db \, da,$$

где a — 4 параметра каждого флуорофора (координаты в плоскости изображения, яркость и размер), а b — набор состояний флуорофора в каждом кадре. Вероятность $p(D, a, b | \mathcal{F})$ равна произведению вероятности получения наблюдаемых данных при условии наличия флуорофора в рассматриваемой точке на априорные вероятности параметров a и b . Априорные вероятности координат флуорофоров заданы равномерным распределением на рассматриваемой области изображения и нулём вне этой области. Размер изображения флуорофора имеет логнормальное априорное распределение с параметрами $\sigma = 0.1$ и μ таким, чтобы мода распределения имела значение полной ширины на полувысоте заданного ядра размытия. Априорное распределение яркости флуорофора тоже логнормальное, с параметрами $\sigma = 3$ и $\mu = 1$. Априорное распределение состояния флуорофора в первом кадре задано так, чтобы вероятности состояний «излучает» и «не излучает» были равновероятными, а событие «флуорофор разрушен» — невозможным. Дальнейшие априорные вероятности вычисляются с использованием скрытой марковской модели в зависимости от предыдущих состояний.

С учётом всех M молекул флуорофора уравнение приобретает вид:

$$p(D|\mathcal{F}_M) = \int_{a \in \mathbb{R}^{4 \cdot M}} \int_{b \in \mathbb{Z}_3^{N \cdot M}} p(D, a, b | \mathcal{F}_M) db da.$$

Из-за большого объёма вычислений в случае многих молекул интегрирование по a проводится приближённо методом Лапласа, а интеграл по b оценивается с помощью генерации выборки $\{b_i\}_{i=1}^n$ методом Монте-Карло по схеме марковской цепи и вычисления суммы

$$p(D|\mathcal{F}_M) = \int_{a \in \mathbb{R}^{4 \cdot M}} \int_{b \in \mathbb{Z}_3^{N \cdot M}} p(D, a, b | \mathcal{F}_M) db da \approx \frac{1}{n} \int_{a \in \mathbb{R}^{4 \cdot M}} \sum_{i=1}^n p(D, a, b_i | \mathcal{F}_M).$$

1.4 Итерационный регуляризующий алгоритм

В ходе выполнения исследования был разработан и реализован алгоритм повышения резкости и разрешения изображений на основе байесовского подхода с использованием регуляризации.

В основе метода лежит предположение о том, что интенсивности пикселей резкого изображения можно рассмотреть как независимые случайные величины, а серию резких изображений — как результат серии наблюдений этих случайных величин. Влияние дифракции представлено свёрткой резкого изображения с одним и тем же ядром для всех пикселей и всех кадров, а полученные размытые снимки содержат высокочастотный шум.

В рамках предлагаемого подхода размытые зашумлённые изображения серии рассматриваются как наблюдаемые переменные, соответствующие им исходные резкие

изображения — как скрытые, а среднее значение и дисперсия интенсивности каждого пикселя исходной серии изображений — как неизвестные параметры распределения интенсивности этого пикселя. Затем эти параметры распределений интенсивностей пикселей вычисляются с помощью байесовского алгоритма. При этом полученные дисперсии интенсивностей можно трактовать как высококонтрастную карту объектов с изменяющейся яркостью, на которой подсвеченный мигающим флуорофором образец будет присутствовать, а постоянный фон — нет.

На вход алгоритма поступает набор размытых и зашумлённых изображений $Y = \{y^t\}_{t=1}^T$, $y^t \in \mathbb{R}^l$, ему соответствует неизвестный набор резких изображений высокого разрешения $X = \{x^t\}_{t=1}^T$, $x^t \in \mathbb{R}^n$. Благодаря природе шума, его можно моделировать случайнм процессом с нормальным распределением [18]. Поэтому в данном пункте считается, что $y^t \sim N(Kx^t, \Theta^{-1})$, где $K \in \mathbb{R}^{n \times n}$ — оператор свёртки и понижения разрешения (имеет вид матрицы Тёплица с элементами ядра размытия на диагоналях, умноженной на матрицу понижения разрешения слева), а $\Theta \in \mathbb{R}^{n \times n}$ — обратная ковариационная матрица шума, диагональная из-за попиксельной независимости шума. Вследствие малости размера молекулы флуорофора предполагается, что в каждом пикселе изображения может находиться достаточно много независимо мигающих молекул, чтобы можно было приблизить распределение суммарной яркости точки нормальным распределением: $x^t \sim N(\mu, \Lambda^{-1})$, где $\mu \in \mathbb{R}^n$ — средняя по всем кадрам яркость объекта, а $\Lambda \in \mathbb{R}^{n \times n}$ — обратная ковариационная матрица мерцания объекта.

Финальным результатом работы алгоритма являются математическое ожидание μ и диагональ ковариационной матрицы Λ^{-1} , поэлементный квадратный корень которой является стандартными отклонениями яркостей точек резких изображений.

Рассматривая X , μ и Λ как скрытые переменные, можно записать их совместное апостериорное распределение:

$$p(X, \mu, \Lambda | Y, \Theta) = \frac{p(Y|X, \Theta) p(X|\mu, \Lambda) p(\mu) p(\Lambda)}{p(Y|\Theta)}.$$

Для нахождения распределений μ и Λ и матрицы Θ максимизируется правдоподобие набора кадров с помощью аппроксимации среднего поля и ЕМ-алгоритма [19]:

$$p(Y|\Theta) \rightarrow \max_{p(X, \mu, \Lambda | Y, \Theta), \Theta} .$$

Апостериорное совместное распределение аппроксимируется факторизуемой плотностью вероятности:

$$p(X, \mu, \Lambda | Y, \Theta) \approx q(X, \mu, \Lambda) = \prod_{t=1}^T q_{1,t}(x^t) \cdot q_2(\mu) \cdot q_3(\Lambda),$$

где в качестве множителей $q_{1,t}(x^t)$ и $q_2(\mu)$ используются нормальные распределения, а для $q_3(\Lambda)$ — распределение Уишарта.

Априорные распределения $p(\mu)$ и $p(\Lambda)$, определяющие вид стабилизатора в алгоритме максимизации, в рамках исследования задаются следующим образом:

$$\begin{aligned}\mu &\sim N(m_0, \eta_0^{-1} I_n), \\ \Lambda &\sim W(W_0, v_0) = \frac{1}{C} |\Lambda|^{\frac{v_0-n-1}{2}} \exp\left(-\frac{1}{2} \operatorname{tr}[W_0^{-1} \Lambda]\right),\end{aligned}$$

где $W(\cdot \cdot \cdot)$ — распределение Уишарта, I_n — единичная матрица, $m_0 \in \mathbb{R}^n$, $\eta_0, v_0 \in \mathbb{R}$ и $W_0 \in \mathbb{R}^{n \times n}$ — параметры регуляризации, а C — нормировочная константа. В частности для μ после преобразований это эквивалентно добавлению стабилизатора вида $-\frac{\eta_0}{2} \|x - m_0\|$. Вид распределений Λ выбран в том числе с тем, чтобы получить аналитическое решение одного из шагов алгоритма оптимизации.

При таких условиях совместная плотность вероятности наблюдаемых размытых и зашумлённых изображений и соответствующих им скрытых резких изображений с учётом регуляризации имеет следующий вид:

$$\begin{aligned}p(Y, X, \mu, \Lambda | \Theta) &= \prod_{t=1}^T \left(\frac{|\Theta| \cdot |\Lambda|}{2\pi^n} \right)^{\frac{1}{2}} \exp \left(-\frac{1}{2} (y^t - Kx^t)^T \Theta (y^t - Kx^t) - \right. \\ &\quad \left. - \frac{1}{2} (x^t - \mu)^T \Lambda (x^t - \mu) \right) \cdot \\ &\quad \cdot \left(\frac{\eta_0}{2\pi} \right)^{\frac{n}{2}} \exp \left(-\frac{1}{2} (\mu - m_0)^T \eta_0 I_n (\mu - m_0) \right) \cdot \\ &\quad \cdot \frac{1}{C} |\Lambda|^{\frac{v_0-n-1}{2}} \exp \left(-\frac{1}{2} \operatorname{tr}[W_0^{-1} \Lambda] \right).\end{aligned}$$

Тогда параметры распределений X , μ и Λ и матрица Θ находятся итерационным алгоритмом, который на каждом шаге пересчитывает их по очереди в соответствии с уравнениями:

$$\begin{aligned}\log q_{1,k}(x^k) &= \mathbb{E}_{\{x^{t \setminus k}\}, \mu, \Lambda} \log p(Y, X, \mu, \Lambda | \Theta) + C_{1,k}, \\ \log q_2(\mu) &= \mathbb{E}_{\{x^t\}, \Lambda} \log p(Y, X, \mu, \Lambda | \Theta) + C_2, \\ \log q_3(\Lambda) &= \mathbb{E}_{\{x^t\}, \mu} \log p(Y, X, \mu, \Lambda | \Theta) + C_3, \\ \Theta &= \arg \max_{\Theta} \mathbb{E}_{\{x^t\}, \mu, \Lambda} \log p(Y, X, \mu, \Lambda | \Theta),\end{aligned}$$

где $\{C_{1,k}\}_{k=1}^T$, C_2 и C_3 — нормировочные константы.

В итоге обновление параметров происходит следующим образом:

$$\begin{aligned}
q_{1,t}(x^t) : \Sigma_{x^t} &= (K^T \Theta K + \mathbb{E}\Lambda)^{-1}, \mathbb{E}x^t = \Sigma_{x^t} (K^T \Theta y^t + \mathbb{E}\Lambda \cdot \mathbb{E}\mu); \\
q_2(\mu) : \Sigma_\mu &= (T \cdot \mathbb{E}\Lambda + \eta_0 I_n)^{-1}, \mathbb{E}\mu = \Sigma_\mu \left(\mathbb{E}\Lambda \cdot \sum_{t=1}^T \mathbb{E}x^t + \eta_0 I_n m_0 \right); \\
q_3(\Lambda) : W(W', v') , v' &= v_0 + T \\
W' &= \left(W_0^{-1} + \sum_{t=1}^T \begin{bmatrix} \Sigma_{x^t} + \mathbb{E}x^t (\mathbb{E}x^t)^T - \mathbb{E}x^t (\mathbb{E}\mu)^T - \\ - \mathbb{E}\mu (\mathbb{E}x^t)^T + \Sigma_\mu + \mathbb{E}\mu (\mathbb{E}\mu)^T \end{bmatrix} \right)^{-1}, \\
\mathbb{E}\Lambda &= v' W'; \\
\Theta_{ii} &= \left(\frac{1}{T} \sum_{t=1}^T \left[((K \mathbb{E}x^t - y^t)_i)^2 + (K \Sigma_{x^t} K^T)_{ii} \right] \right)^{-1}.
\end{aligned}$$

Для ускорения работы алгоритма он может быть запущен без учёта повышения разрешения, а затем полученный результат после простой интерполяции может быть использован в качестве начального приближения при нахождении изображения повышенного разрешения.

1.5 Эксперименты и результаты

В ходе исследования методы SOFI, MUSICAL и разработанный итерационный алгоритм были реализованы на языке программирования Python. Для получения результатов обработки изображений алгоритмом 3B было использовано программное обеспечение от авторов этого метода (доступно по ссылке <http://www.coxphysics.com/3b>).

Была проведена оптимизация алгоритма SOFI для уменьшения влияния пристранственно некоррелированного шума на интенсивность пикселей изображения высокого разрешения в позициях исходных пикселей путём замены вычисления автокорреляции исходных пикселей на вычисление усреднённых смешанных полуинвариантов сигналов с пикселей вокруг исходных.

Было проведено тестирования рассматриваемых алгоритмов на искусственных и реальных данных, примеры результатов обработки представлены на Рис. 1.7-и-1.8. Разработанный итерационный метод был ускорен с использованием технологии Nvidia CUDA, тестирование остальных методов проводилось на ЭВМ с четырёхядерным процессором Intel Core i5 6500. Среднее время работы на представленных данных составило для XC-SOFI порядка 2 вместе с дальнейшим устранением размытия — 2 секунды, для 3B — 45 минут, для разработанного итерационного метода — 2 минуты, для MUSICAL — 1 минута.

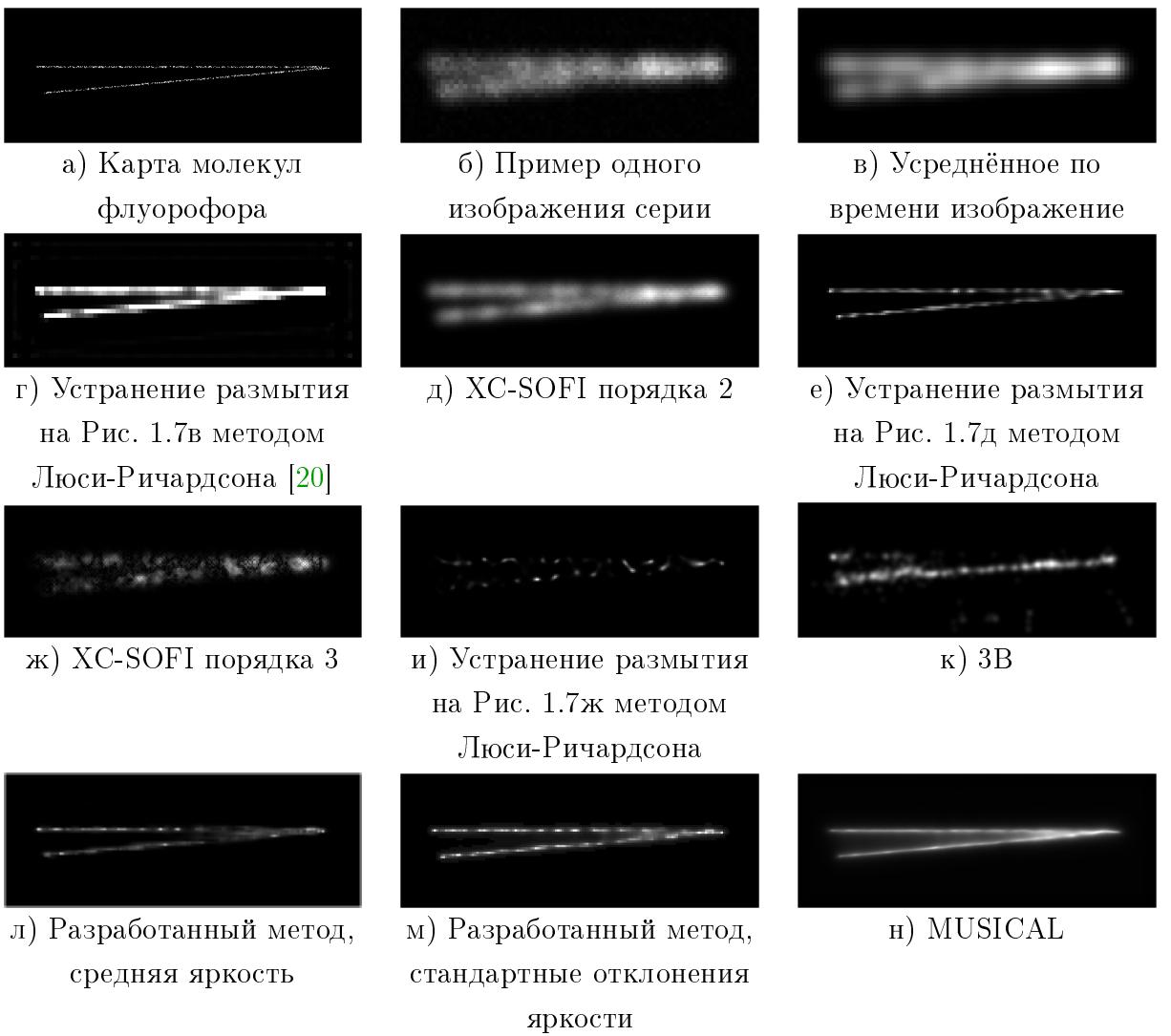


Рисунок 1.7: Примеры изображений из искусственного набора данных и результаты работы рассмотренных алгоритмов

Сравнение показало следующее:

1. результаты работы алгоритма 3В страдают от потери данных, получаемый результат не окупает долгое время работы алгоритма;
2. XC-SOFI является самым быстрым из рассмотренных, однако варианты метода порядков выше второго страдают от потери информации, а также сильно зависят от используемого алгоритма устранения размытия;
3. разработанный метод повышения качества изображений хорошо восстанавливает детали, однако с ростом степени повышения разрешения объём вычислений растёт крайне быстро, что ограничивает возможности по повышению разрешения;
4. алгоритм MUSICAL, несмотря на меньшую резкость результатов, хорошо восстанавливает детали и может быть использован при жёстких ограничениях по времени.

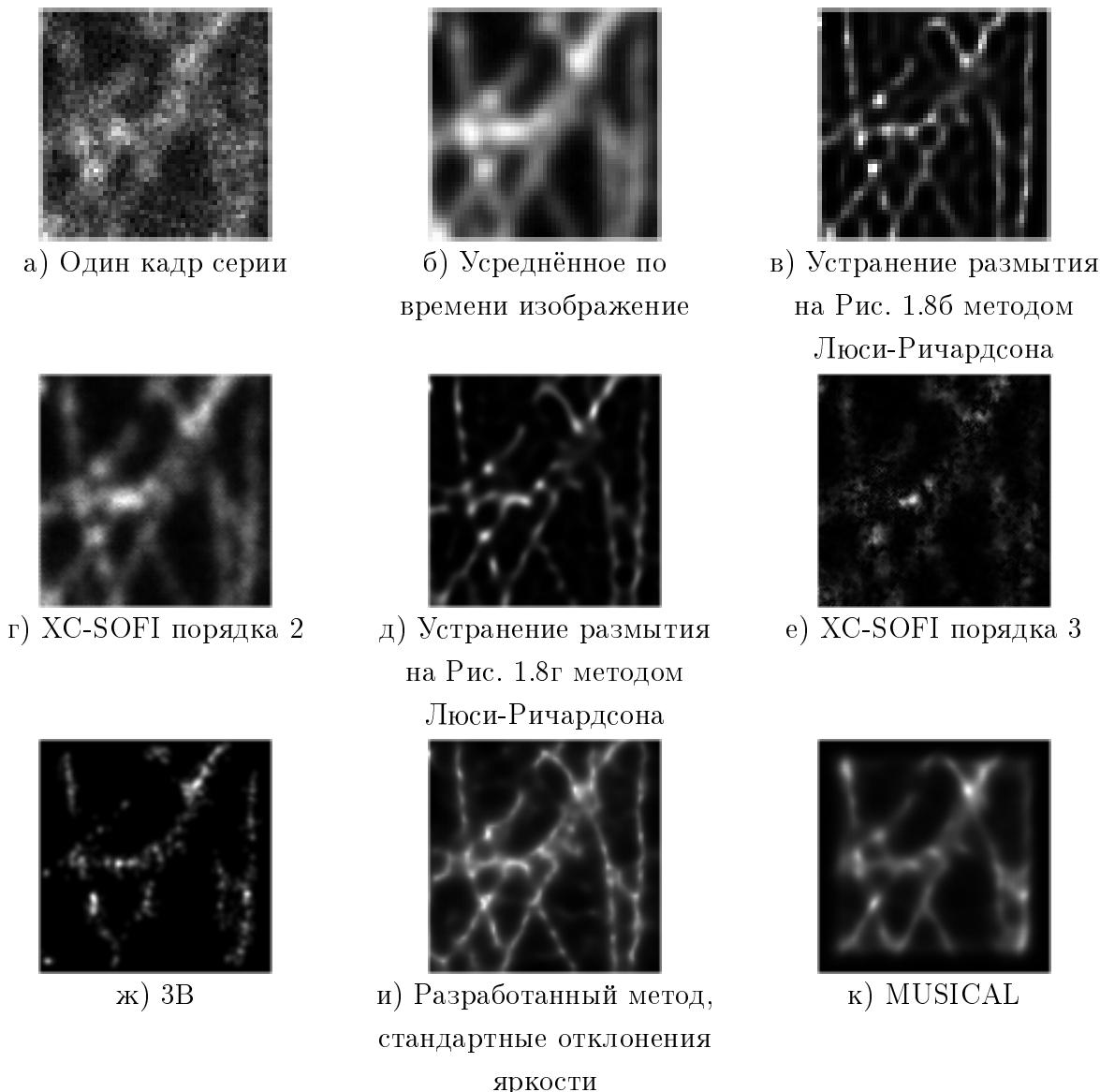


Рисунок 1.8: Примеры изображений из реального набора данных и результаты работы рассмотренных алгоритмов

1.6 Выводы

Разработанный в рамках исследования алгоритм показал качество работы, сопоставимое с результатами алгоритма MUSICAL. Его преимуществом является большая резкость полученных клеточных структур, а также имеющие физический смысл восстановленные яркость и дисперсия яркости точек. Недостатком является высокая вычислительная сложность, не позволяющая использовать алгоритм в текущем виде для увеличения разрешения в 4 и более раз за приемлемое время.

Алгоритм MUSICAL предоставляет лучшие возможности по увеличению разрешения, однако результат содержит лишь некий показатель наличия молекул флуоресцентного поглощения, поэтому область применения ограничивается локализационной микроскопией.

SOFI имеет наивысшую скорость работы, однако недостаточное повышение разрешения и резкости делают нецелесообразным его использование при низком увеличении микроскопа.

Глава 2. Повышение резкости медицинских изображений методом деформации пиксельной сетки

На практике часто возникает необходимость повышения резкости изображений по разным причинам: из-за физических ограничений оптических систем и сенсоров камер или из-за сложных условий получения изображений. В то же время в медицине при обработке изображений важно не терять присутствующие на них детали или добавлять не существовавшие элементы или артефакты.

В отличие от наиболее распространённых методов повышения резкости (обращение свёртки, нерезкое маскирование), деформационный метод, представленный в [21] и затем доработанный в [22; 23], не меняет значения интенсивности пикселей, а сдвигает сами пиксели в окрестности контуров изображений (см. Рис 2.1). Этот подход изменяет изображение в окрестности контуров и не имеет таких недостатков, свойственных некоторым другим методам, как усиление шума и возникновение эффекта ложного оконтурирования.

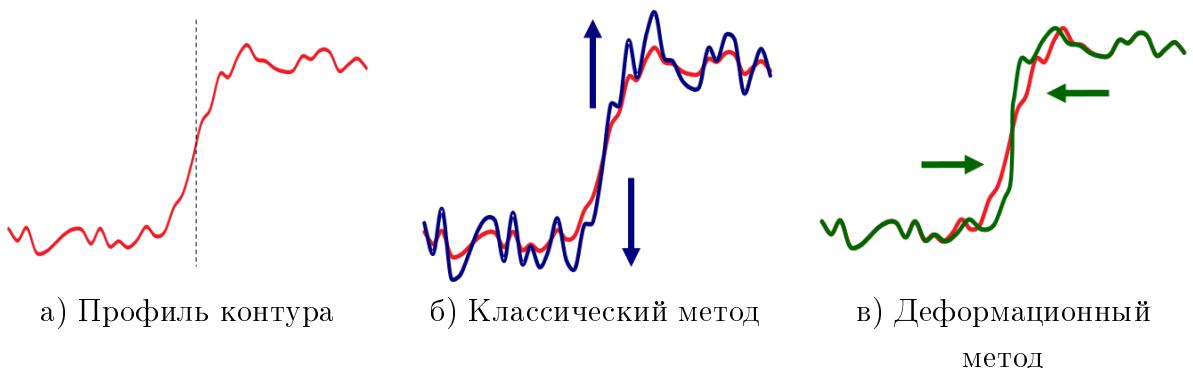


Рисунок 2.1: Принцип работы классических методов и деформационного метода повышения резкости изображений

В данной главе рассматривается деформационный алгоритм повышения резкости изображения [22; 23] и набор из 24 тестовых изображений из базы TID2013 [24], размытых с разной интенсивностью с использованием трёх типов реалистичных ядер размытия. Целью является нахождение оптимальной функции смещения для каждого типа ядра размытия отдельно и для всех типов сразу.

2.1 Реалистичные ядра размытия изображений

Модель размытия изображения, рассматриваемая в этой главе, совпадает с моделью, описанной в п. 1.1.1.

В процессе исследования рассматривались три типа ядер размытия, возникающих на практике: ядро, определяемое функцией Гаусса; и два ядра, которые часто возникают в реальных изображениях, полученных с помощью фотокамер.

Ядро Гаусса выбрано по той причине, что оно соответствует одной из наиболее распространённых моделей размытия изображения, а также потому, что такое ядро (или очень близкое к нему) можно встретить в изображениях, полученных с помощью некоторых оптических систем, где дифракция света играет заметную роль (например, с помощью микроскопов).

Также, согласно модели Зейделя, существует пять видов оптических aberrаций [25]:

1. сферическая aberrация;
2. кома;
3. астигматизм;
4. кривизна поля изображения;
5. дисторсия.

Ядра, отвечающие сферическим aberrациям и астигматизму, также схожи с ядром Гаусса [25; 26], что послужило дополнительным аргументом в пользу рассмотрения этого ядра. Примеры этих искажений представлены на Рис. 2.2.

Помимо этих пяти aberrаций также встречаются просто расфокусированные изображения, либо же расфокусированные участки изображений, полученных с использованием камер с малой глубиной резкости изображаемого пространства. В таком случае полученное изображение точки имеет вид круга с приблизительно равномерной яркостью, однако также может иметь увеличение яркости в центре или на краях (что выглядит как круг с кольцом). Таким образом, ядро размытия часто имеет вид круга или круга с кольцом. Конкретный вид зависит в том числе от положения объекта на изображении относительно области, находящейся в фокусе. Примеры этого эффекта приведены на Рис. 2.3.

Ниже приведено описание использованных в ходе исследования моделей ядер размытия:

1. Ядро Гаусса. Задаётся формулой

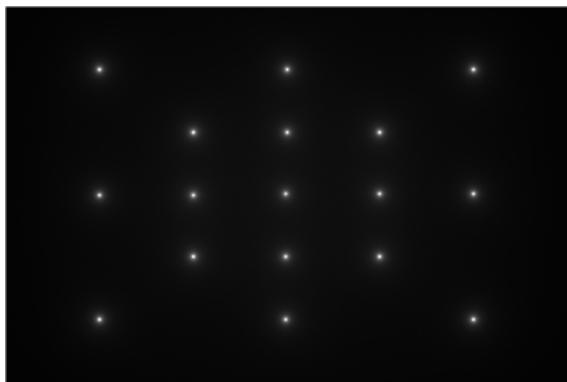
$$f_\sigma(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right).$$

Радиусом этого ядра будет считаться величина 3σ .

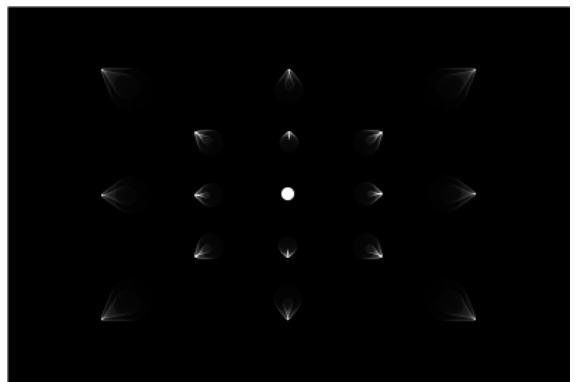
2. Круг. Задаётся формулой

$$f_r(x,y) = \begin{cases} 1, & x^2 + y^2 \leq r^2, \\ 0, & \text{иначе.} \end{cases}$$

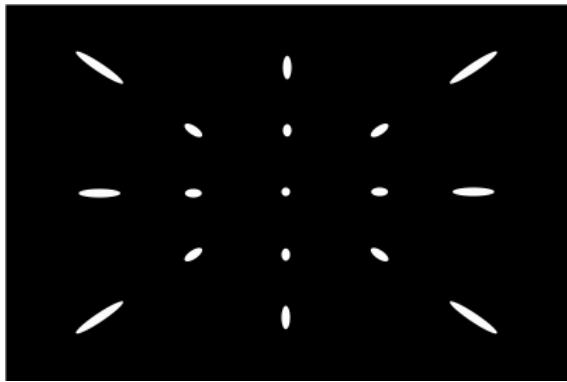
Радиус ядра — радиус круга r .



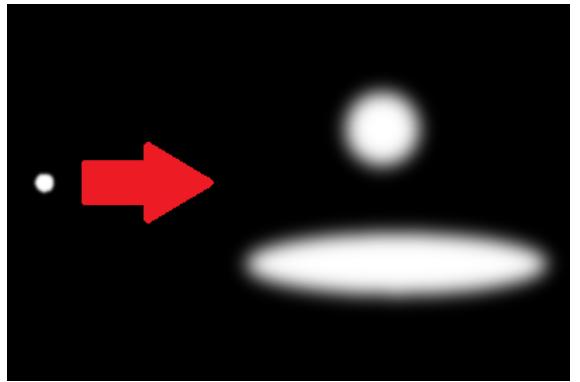
а) Сферическая aberrация



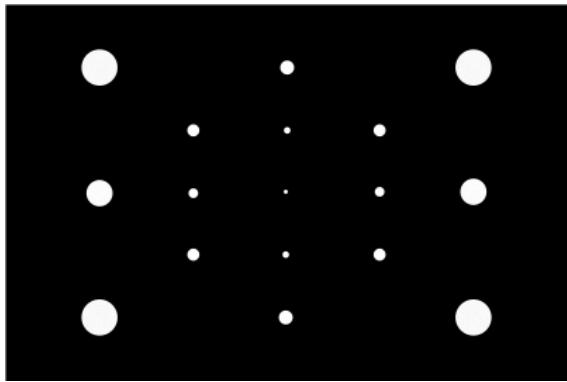
б) Кома



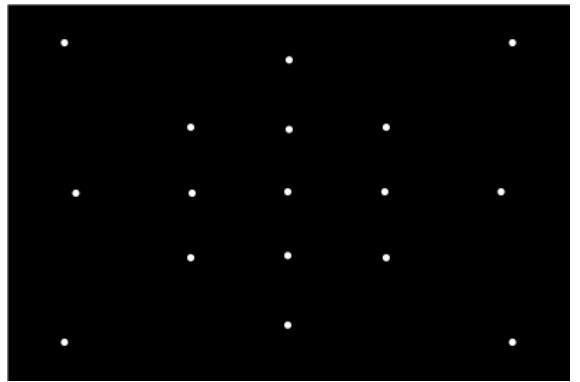
в) Астигматизм



г) Разные виды астигматизма



д) Кривизна поля изображения



е) Дисторсия

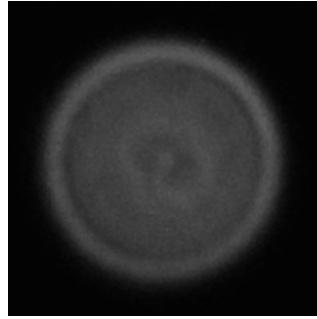
Рисунок 2.2: Виды оптических аберраций [26]

3. Круг с кольцом. Задаётся формулой

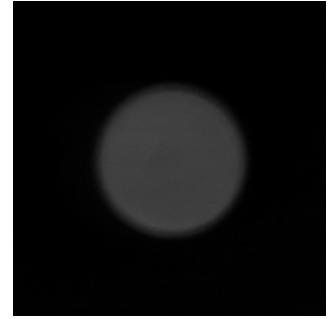
$$f_r(x, y) = \begin{cases} 0.25, & x^2 + y^2 \leqslant 0.75r^2, \\ 1, & 0.75r^2 < x^2 + y^2 \leqslant r^2, \\ 0, & \text{иначе.} \end{cases}$$

Радиус ядра — радиус круга r .

Примеры этих ядер представлены на Рис. 2.4.

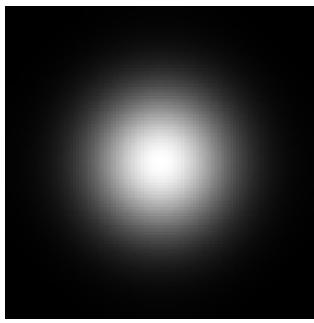


а) Объект за
областью резкости

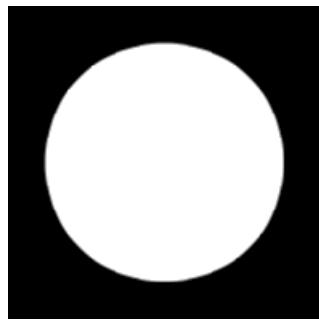


б) Объект перед
областью резкости

Рисунок 2.3: Примеры эффекта расфокусировки



а) Гауссово ядро (Γ)



б) Круг (K)



в) Круг с кольцом
(КЦ)

Рисунок 2.4: Рассматриваемые ядра размытия

2.2 Построение тестового набора изображений

Для разработки метода повышения резкости использовался набор из 24 естественных изображений из базы TID2008 [27], которая является одним из стандартных наборов данных для тестирования методов обработки изображений.

При построении тестового набора к каждому изображению применялся оператор свёртки с каждым из трёх ядер размытия с разным радиусом ядра. Для круга и круга с кольцом радиус ядра размытия изменялся в пределах от 1.5 до 5 пикселей с шагом 0.5 пикселя. Для гауссова ядра — в пределах от 2.25 до 7.5 с шагом 0.75, что соответствует изменениям параметра σ в пределах от 0.75 до 2.5 с шагом 0.25.

В результате для каждого типа размытия был сформирован набор из 192 изображений.

2.3 Деформационный метод повышения резкости изображений

В процессе работы алгоритма на изображении:

1. обнаруживаются контуры алгоритмом Кэнни [28],

2. производится расчёт векторов смещения пикселей в окрестности контуров,
3. в соответствии с вычисленным векторным полем смещений деформируется пиксельная сетка,
4. полученный результат проецируется на регулярную сетку выходного изображения с помощью интерполяции, описанной в [21].

Направление и величина смещения пикселей задаётся функцией смещения $d(x)$, $x_{new} \leftarrow x_{old} + d(x_{old})$. В одномерном случае, если центр контура находится в точке $x = 0$, вектора смещения определяются функцией близости $p(x) = 1 + d'(x)$.

Она может быть выражена через $p(x)$ с помощью уравнения

$$d(x) = \int_{-\infty}^x (p(y) - 1) dy.$$

Функция $p(x)$ выражает расстояние между соседними пикселями изображения: если её значение меньше 1, в точке x уплотнение; если больше 1 — разрежение. Для недеформированных изображений $p(x) \equiv 1$.

Чтобы взаимное расположение точек при смещении не менялось, для функции $d(x)$ должно быть выполнено $x_1 < x_2 \Rightarrow x_1 + d(x_1) \leq x_2 + d(x_2)$. Отсюда получаем ограничение $d'(x) \geq -1$. Так же деформироваться должна только сетка вблизи контуров, поэтому $d(x) \rightarrow 0$ при $|x| \rightarrow \infty$. Пример функции близости приведён на Рис. 2.5.

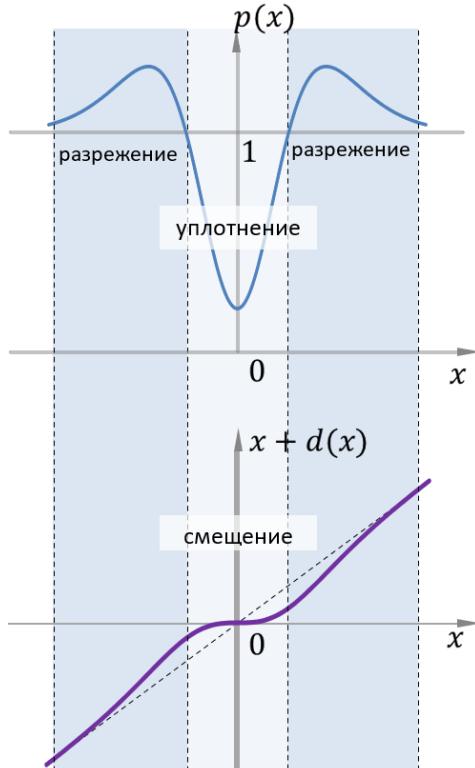


Рисунок 2.5: Пример функции близости и результата смещения пикселей с её помощью

В двумерном случае смещение задаётся векторным полем $\vec{d}(x,y)$, которое удовлетворяет уравнению $p(x,y) = 1 + \operatorname{div} \vec{d}(x,y)$. Помимо этого, во избежание возникновения завихрений требуется ограничение $\operatorname{rot} \vec{d} = 0$, а так как $\operatorname{rot} \nabla u \equiv 0$, то, если функция

$p(x, y)$ известна, векторное поле смещений можно представить как $\vec{d}(x, y) = \nabla u(x, y)$, где функция $u(x, y)$ является решением уравнения

$$\begin{cases} \Delta u = p(x, y) - 1, \\ u|_{\Gamma} = 0, \end{cases}$$

где Γ — граница изображения.

В работе [23] предлагается для двумерного случая использовать функцию близости одной переменной для построения функции близости на плоскости:

$$p(x, y) = \frac{\sum_{(x_e, y_e) \in N(x, y)} w(x_e, y_e) p(x_n)}{\sum_{(x_e, y_e) \in N(x, y)} w(x_e, y_e)},$$

$$w(x_e, y_e) = G_{\sigma}(x_t) |\vec{g}(x_e, y_e)|,$$

где $N(x, y)$ — это множество соседних точек вокруг т. (x, y) ; x_n и x_t — это проекции вектора $(x - x_e, y - y_e)$ на вектор градиента изображения $\vec{g}(x_e, y_e)$ и на нормаль к нему соответственно; а $G_{\sigma}(x_t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-x^2}{2\sigma^2}\right)$.

В этом случае итоговая функция смещения имеет следующий вид [23]:

$$\vec{d}(x, y) = \frac{\sum_{(x_e, y_e) \in N(x, y)} d(x_n) G_{\sigma}(x_t) \vec{g}(x_e, y_e)}{\sum_{(x_e, y_e) \in N(x, y)} G_{\sigma}(x_t) |\vec{g}(x_e, y_e)|}.$$

Таким образом, можно перейти к работе с функцией смещения напрямую и проанализировать её влияние на результаты работы деформационного алгоритма повышения резкости.

2.4 Модели функции смещения

Были рассмотрены три модели функции смещения, первая из которых была предложена в оригинальной работе авторов метода и используется здесь для оценки повышения эффективности деформационного метода повышения резкости изображений:

$$1. d_0(x; \sigma) = \sigma\sqrt{\pi} \left[\operatorname{erf}\left(\frac{x}{2\sigma}\right) - \operatorname{erf}\left(\frac{x}{\sigma}\right) \right], \text{ где } \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt;$$

$$2. d_2(x; a, b, c) = \begin{cases} \frac{c}{a}x, & |x| < a, \\ c \frac{b - |x|}{b - a} \cdot \operatorname{sign}(x), & a \leq |x| < b, \\ 0, & |x| \geq b; \end{cases}$$

$$3. d_1(x; a, c), \text{ имеющая тот же вид, что и } d_2(x), \text{ но } b = 1.5a.$$

При этом координата x — это расстояние точки от контура, делённое на радиус ядра размытия, что позволяет масштабировать функцию для разных уровней размытия без изменения самих параметров a , b , c и σ .

Пример функции $d_2(x; a, b, c)$ приведёт на Рис. 2.6.

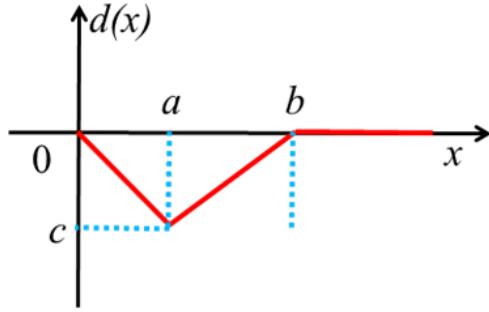


Рисунок 2.6: Вид функции смещения $d_2(x; a, b, c)$

2.5 Методика оптимизации параметров моделей функций смещения

Проводился поиск оптимальных параметров рассматриваемых малопараметрических (одно- и двухпараметрических) моделей функций смещения для разных типов размытия, а затем — сравнительный анализ результатов обработки размытых изображений деформационным методом повышения резкости с использованием полученных результатов.

Для нахождения оптимальных параметров проводилась минимизация среднего значения показателя MSE по всем изображениям для одного типа размытия:

$$\frac{1}{|V|} \sum_V MSE(u, v) \rightarrow \min_z, \quad MSE(u, v) = \frac{1}{WH} \sum_{i,j} (u_{ij} - v_{ij})^2,$$

а в качестве показателя качества — значение $PSNR(u, v) = 10 \cdot \log_{10} \left(\frac{255^2}{MSE(u, v)} \right)$.

Здесь v — размытое изображение, обработанное методом повышения резкости, u — соответствующее ему истинное резкое изображение, V — множество всех размытых изображений для одного ядра размытия, W и H — ширина и высота изображения в пикселях соответственно, z — соответствующие рассматриваемой функции смещения параметры из множества $\{a, b, c, \sigma\}$.

Для решения задачи минимизации применялся метод Нелдера-Мида [29]. Число итераций было ограничено числом 100, минимальный шаг был установлен в 10^{-3} . С начальным приближением $a = 1, b = 1.5, c = 0.5, \sigma = 1$ процесс оптимизации завершился примерно на 50-й итерации. Было сделано несколько запусков процесса оптимизации с разными начальными приближениями для достижения глобального оптимума.

2.6 Результаты экспериментов

В ходе экспериментов было установлено, что наилучшие результаты достигаются при $a = -c$. Такое соотношение параметров обеспечивает наибольшее смещение пикселей в районе контура без нарушения ограничения на производную функции смещения. Поэтому в дальнейшем это соотношение всегда выполняется, что позволяет перейти к модели с двумя (функция $d_2(x; a, b, c = -a)$) и одним (функция $d_1(x; a, c = -a)$) параметрами.

После проведения исследования и поиска оптимальных параметров функций смещения для каждого типа размытия были получены следующие результаты. Условные обозначения типов ядер размытия: Г — гауссово ядро, К — круг, КЦ — круг с кольцом.

Усреднённые по набору изображений значения PSNR входных размытых и обработанных изображений для функций смещения $d_0(x)$ и $d_2(x)$ с оптимальными параметрами приведены в Табл. 1. Из неё видно, что предложенная функция $d_2(x)$ позволяет достичь более качественного результата.

Таблица 1: Средние значения PSNR размытых и обработанных изображений для функций смещения $d_0(x)$ и $d_2(x)$ (сила размытия обозначена номером радиуса размытия в порядке возрастания)

Сила размытия	Размытые изображения			$d_0(x)$			$d_2(x)$		
	Г	К	КЦ	Г	К	КЦ	Г	К	КЦ
1	25.543	28.351	27.300	25.773	28.483	27.468	25.842	28.561	27.608
2	24.335	26.657	25.508	24.587	26.853	25.735	24.660	26.923	25.870
3	23.571	25.290	24.355	23.842	25.535	24.621	23.908	25.612	24.755
4	23.008	24.623	23.843	23.276	24.869	24.104	23.335	24.930	24.207
5	22.578	23.998	23.276	22.841	24.256	23.546	22.894	24.310	23.634
6	22.220	23.577	22.903	22.478	23.827	23.165	22.527	23.873	23.239
7	21.922	23.208	22.568	22.182	23.461	22.832	22.228	23.499	22.895
8	21.658	22.896	22.286	21.912	23.141	22.543	21.959	23.175	22.599
Среднее	23.104	24.825	24.005	23.361	25.053	24.252	23.419	25.110	24.351

В Табл. 2 представлены показатели среднего прироста PSNR (ISNR) относительно размытых изображений для вариантов алгоритма на основе функции смещения с 2 и с 1 параметром. Модуль относительной разницы не превышает 4.2%, а в среднем равен 1.1%, что позволяет говорить о сравнимом качестве работы обоих вариантов метода и возможности использовать однопараметрический вариант без значительных потерь

в качестве. Оптимальные значения параметра a однопараметрического метода с функцией $d_1(x)$ для гауссовского ядра размытия, ядра типа "круг" и ядра типа "круг с кольцом" составили 1.28, 1.12 и 1.16 соответственно.

Таблица 2: Средние значения прироста PSNR для изображений, обработанных с использованием функции смещения $d_2(x)$ и $d_1(x)$, по сравнению с размытыми изображениями

Сила размытия	$d_2(x)$			$d_1(x)$		
	Г	К	КЦ	Г	К	КЦ
1	0.299	0.210	0.308	0.299	0.204	0.321
2	0.325	0.266	0.362	0.325	0.272	0.361
3	0.337	0.322	0.400	0.337	0.324	0.386
4	0.327	0.307	0.364	0.327	0.306	0.354
5	0.316	0.312	0.358	0.316	0.309	0.351
6	0.307	0.296	0.336	0.307	0.293	0.331
7	0.306	0.291	0.327	0.306	0.290	0.321
8	0.301	0.279	0.313	0.301	0.278	0.305
Среднее	0.315	0.285	0.346	0.315	0.285	0.341

Дополнительный анализ изменения среднего прироста показателя PSNR показал, что при несильном отклонении параметра от оптимального значения качество результирующего изображения снижается слабо (см. Рис. 2.7).

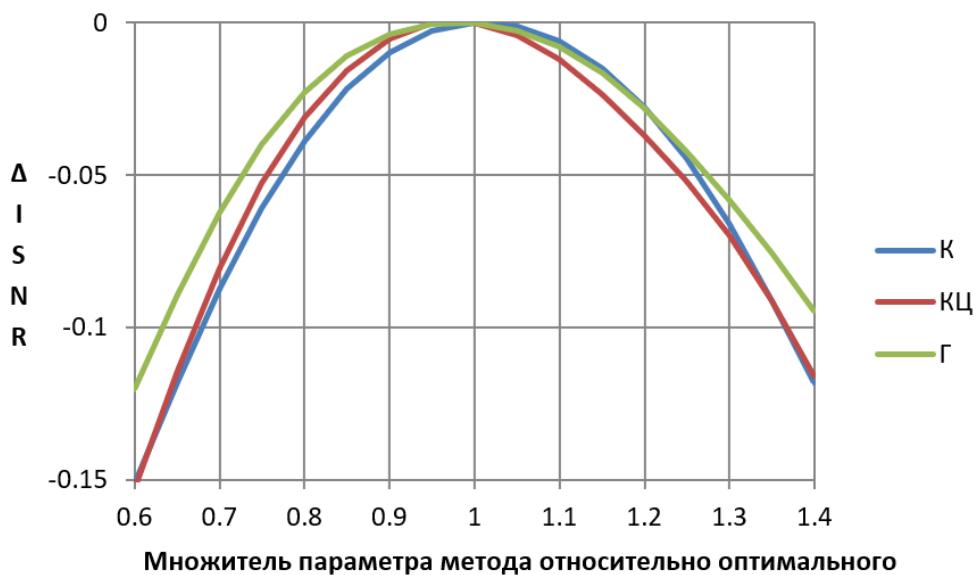


Рисунок 2.7: Изменение прироста показателя PSNR при относительном изменении параметра a

На Рис. 2.8 и 2.9 приведены примеры результата обработки медицинского изображения полученным методом. Видно повышение резкости контуров сосудов и диска зрительного нерва.

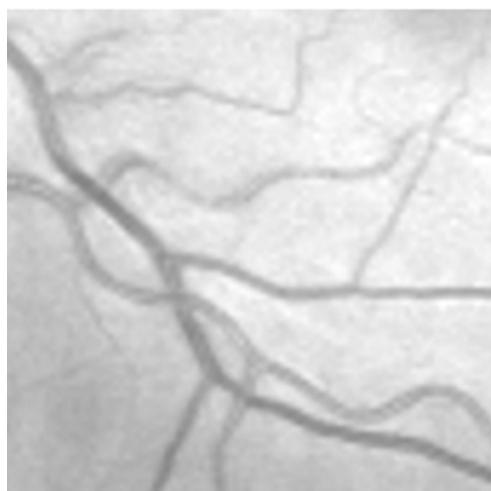


а) Входное изображение

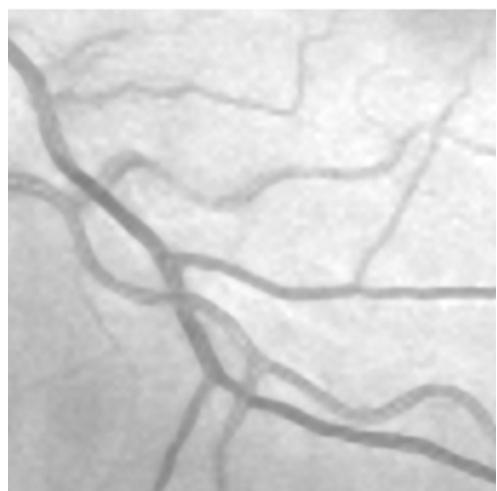


б) Обработанное изображение

Рисунок 2.8: Применение деформационного алгоритма с использованием функции смещения $d_1(x)$ как шага постобработки в задаче повышения разрешения медицинских изображений



а) Входное изображение



б) Обработанное изображение

Рисунок 2.9: Применение деформационного алгоритма с использованием функции смещения $d_1(x)$ как шага постобработки в задаче повышения разрешения медицинских изображений

2.7 Выводы

Предложенные функции смещения $d_2(x)$ и $d_1(x)$ демонстрируют улучшение качества обработки изображений по сравнению с оригинальной функцией $d_0(x)$, выражаемое в приросте показателя PNSR, в среднем на 27%. Вариант алгоритма на основе однопараметрической функции смещения $d_1(x)$ почти не отличается от двухпараметрического варианта качеством результатов, благодаря этому можно использовать его для всех трёх рассмотренных ядер размытия.

Глава 3. Нейросетевой алгоритм контроля качества рентгеновских снимков грудной клетки в задаче диагностики туберкулёза лёгких

Согласно клиническим рекомендациям Минздрава России, одним из первых этапов выявления и диагностики туберкулёза лёгких является рентгенофлюорографическое исследование [30]. При этом в последние годы в практику внедряются компьютерные системы автоматической обработки медицинских данных и диагностики заболеваний на основе искусственного интеллекта, в том числе для в категории флюорографии [31–34].

Одной из актуальных задач при применении методов глубокого обучения в медицинской диагностике является анализ и предобработка входных данных. Необходим контроль соответствия входной информации и применяемого обученного (обучаемого) метода глубокого обучения. Это, например, требуется при контроле наличия адвокативных атак на данные [35].

Контроль качества рентгеновских снимков грудной клетки на основании разных параметров важен для полноценного анализа снимка и постановки правильного диагноза. Автоматическое определение контроля пространственных условий снимка (позы пациента, положения грудной клетки в кадре и т.п.) рассматривается в работах [36–38]. В работе [39] изучено влияние качества снимка на результаты автоматической диагностики COVID-19.

В данной главе контроль входных рентгеновских снимков используется для проверки соответствия уровня облучения снимка уровню, требуемому для качественной диагностики туберкулёза лёгких. Рассматриваются:

1. задача автоматического определения уровня жёсткости рентгеновского снимка грудной клетки с помощью нейросетевого алгоритма;
2. влияние предварительной фильтрации обучающей и валидационной выборок на качество работы алгоритма классификации в задаче диагностики туберкулёза лёгких по рентгеновским снимкам грудной клетки.

Несмотря на то, что использование при анализе рентгенограммы грудной клетки, помимо фронтальной, также и боковой проекции повышает качество анализа, в том числе компьютерного, прирост качества для разных задач разный и далеко не всегда значительный [40; 41]. Кроме того, на практике в подавляющем числе случаев имеются только снимки, сделанные во фронтальной проекции [42], поэтому в рамках данной главы рассматриваются фронтальные изображения грудной клетки.

3.1 Метод оценки жёсткости рентгеновских снимков грудной клетки

При изучении рентгеновских снимков грудной клетки, и в частности при диагностике туберкулёза лёгких, важным фактором является жёсткость снимка, так как она напрямую влияет на его информативность [43; 44]. Уровень жёсткости обусловлен дозой радиации, длиной волны излучения и особенностями тела пациента. При условии правильного контрастирования снимка уровень жёсткости можно определить визуально, подсчитав число отчётливо видимых на снимке верхних грудных позвонков: оптимальному уровню в задаче диагностики туберкулёза лёгких соответствуют 3-4 видимых позвонка, а меньшее или большее число свидетельствует о том, что снимок слишком мягкий или жёсткий [44; 45]. Пример рентгеновского изображения позвоночного столба с указанием номеров позвонков приведён на Рис. 3.1, а примеры снимков разного уровня жёсткости представлены на Рис. 3.2.



Рисунок 3.1: Пример рентгеновского изображения позвоночника с номерами шейных (С) и грудных (Т) позвонков

Настоящее исследование развивает идеи работы [46], где было показано, что оптимизация алгоритма диагностики заболеваний по рентгеновскому снимку грудной клетки для работы с изображениями близкой жёсткости вместе с автоматическим контролем качества рентгеновских изображений позволяет достичь точности классификации выше, чем у алгоритма, созданного в расчёте на обработку разнородных по уровню жёсткости изображений. В ходе выполнения данного исследования был применён полностью нейросетевой подход.

Предлагаемый метод оценки жёсткости включает 2 этапа: обучение и применение. В процессе обучения нейросетевая модель, являющаяся основой метода, однократно настраивается с целью оптимального решения задачи и сохраняется для дальнейшего



Рисунок 3.2: Примеры рентгеновских снимков грудной клетки разной жёсткости

использования. На этапе применения происходит непосредственный анализ изображений.

Процесс обучения выглядит следующим образом:

1. предобработка входных данных;
2. итеративный процесс минимизации функции потерь (функционала, отображающего выходные значения нейросетевой модели в показатель их близости к оптимальным значениям для решения задачи) повторением шагов:
 - а) подача предобработанных выходных данных на вход нейросетевой модели и получение её выходных значений;
 - б) получение значения функции потерь для данных выходных значений модели и входных данных;
 - в) шаг оптимизации;
 - г) замер качества работы алгоритма;
 - д) проверка условия ранней остановки процесса минимизации;
3. сохранение модели.

Процесс применения метода состоит из следующих шагов:

1. предобработка входных данных;
2. подача предобработанных выходных данных на вход нейросетевой модели и получение её выходных значений;

3.1.1 Использованные данные

При решении рассматриваемой задачи были использованы 2 набора рентгенограмм грудной клетки, снятых во фронтальной проекции.

Первый набор изображений был собран в сотрудничестве с медиками из НПЦ «Фтизиатрия» им. Е.Н. Андреева в г. Якутск и применялся для обучения нейросетевой модели для метода определения жёсткости рентгенограмм.

Второй набор был сформирован из двух наиболее часто используемых при разработке методов компьютерной диагностики туберкулёза лёгких общедоступных наборов рентгенограмм, ставших, можно сказать, эталонными [47–50]: Montgomery County [51] и Shenzhen [52]. Так как в большинстве работ наборы Montgomery County и Shenzhen используются вместе [47; 48], то было принято решение их объединить и обозначать «Montgomery-Shenzhen» (MC-SZ). Полученный набор снимков использовался для оценки качества работы разработанного алгоритма на снимках, сделанных в других медицинских учреждениях, на другом оборудовании и при других условиях.

Набор рентгенограмм грудной клетки SakhaTB

Набор состоит из 1298 рентгеновских снимков грудной клетки больных туберкулёзом лёгких пациентов, собранных в нескольких медицинских учреждениях Республики Саха (Якутия) при помощи стационарных и переносных комплексов оборудования. Разрешение изображений разнится и находится в диапазоне примерно от 2000x2000 до 3000x3000 пикселей. Большинство рентгенограмм имеет глубину цвета 16 бит, у остальных снимков она равна 8 битам.

Для дальнейшей работы глубина цвета снимков предварительно была приведена к 8 битам, а разрешение путём усреднения соседних пикселей было уменьшено в целое число раз до достижения значений около 1000x1000. Одновременно была проведена нормализация диапазона интенсивности пикселей с отсечением (англ. clipping) по 0.5% пикселей с каждого конца гистограммы изображения и обрезкой по 1.5% пикселей с каждой стороны изображений, чтобы увеличить долю диапазона допустимых значений 8 бит информации, приходящуюся на внутренние органы грудной клетки. Примеры снимков приведены на Рис. 3.3.

Подмножество снимков этого набора было выложено вместе с рентгенограммами грудной клетки здоровых пациентов для использования в открытый доступ под названием Sakha-TB (см. п. 4.4).

Набор рентгенограмм грудной клетки Montgomery-Shenzhen

Набор изображений Montgomery County (далее МС) предоставляется Национальной библиотекой медицины Национальных институтов здравоохранения США и собран Министерством здравоохранения и социальных служб США в округе Монтгомери штата Мэриленд. Он содержит рентгеновские снимки грудной клетки в оттенках серого с глубиной цвета 8 бит в формате PNG с разрешением 4020x4892 пикселей и их мет-

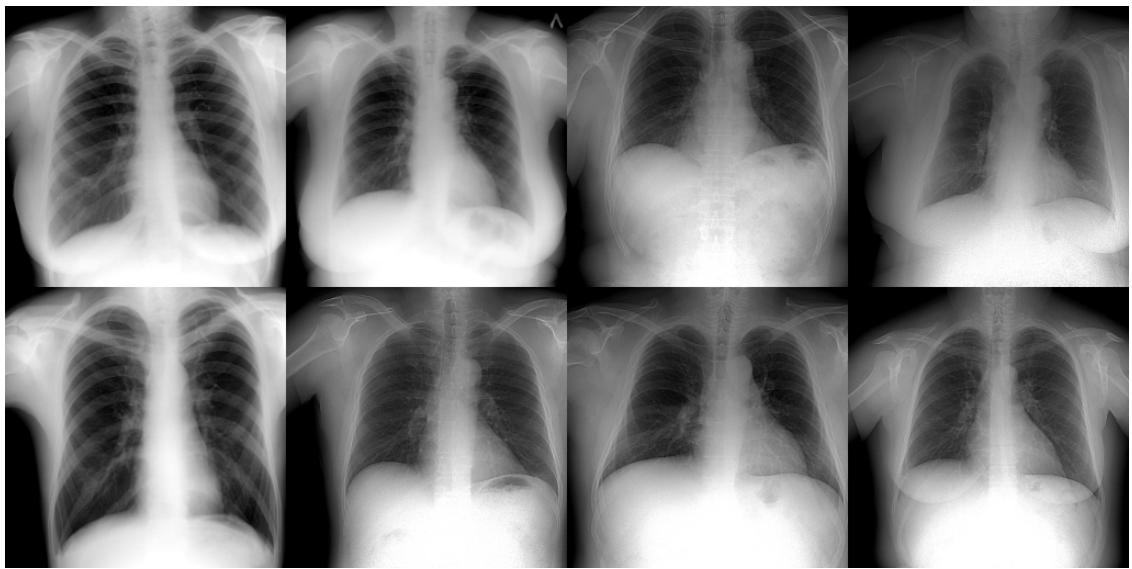


Рисунок 3.3: Примеры изображений из набора SakhaTB

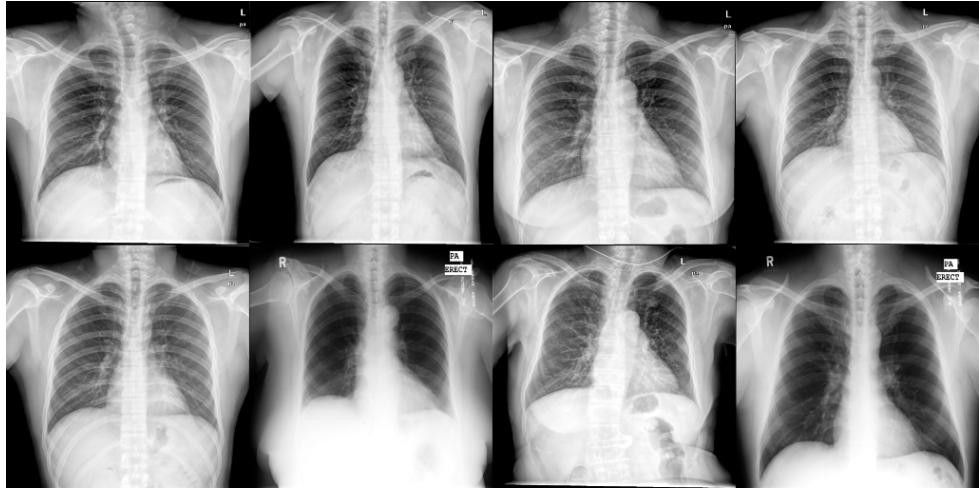
ки соответствия 2 классам: 80 рентгенограмм здоровых и 58 рентгенограмм больных туберкулём лёгких пациентов.

Набор изображений Shenzhen (далее SZ) предоставляется Национальной библиотекой медицины Национальных институтов здравоохранения США и собран Медицинским колледжем провинции Гуандун в Народном госпитале № 3 г. Шэньчжэнь в КНР. Он содержит рентгеновские снимки грудной клетки в оттенках серого с глубиной цвета 8 бит в формате PNG с различным разрешением (примерно 3000x3000 пикселей) и их метки соответствия 2 классам: 326 рентгенограмм здоровых и 336 рентгенограмм больных туберкулём лёгких пациентов.

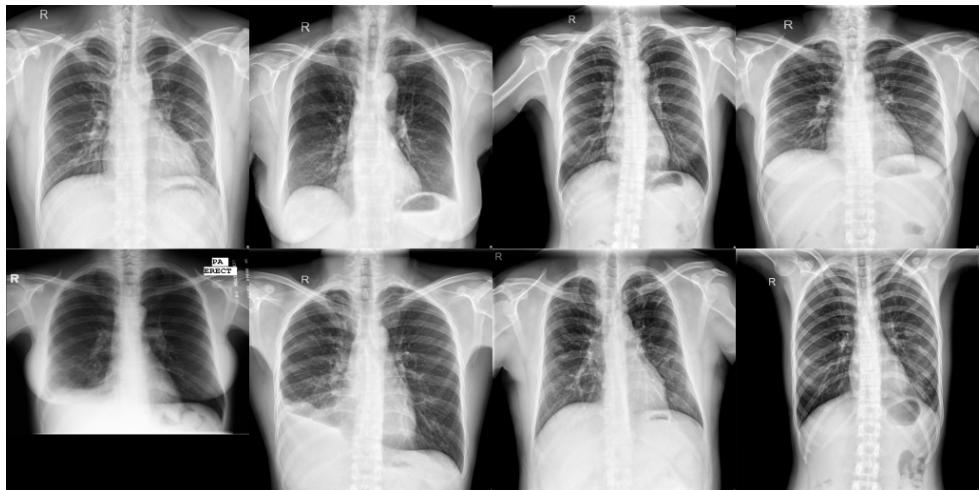
Примеры снимков объединённого набора приведены на Рис. 3.4.

Аннотирование данных

Для обоих наборов (SakhaTB, собранный в рамках исследования, и MC-SZ) была проведена процедура аннотации врачом-рентгенологом из НПЦ «Фтизиатрия» им. Е.Н. Андреева в г. Якутск. В рамках процедуры каждому изображению ставился в соответствие его уровень жёсткости, выраженный числом отчётливо видимых на этом снимке верхних грудных позвонков. Распределение изображений по такому показателю жёсткости представлено на Рис. 3.5-3.6.



а) Здоровые пациенты



б) Пациенты, больные туберкулёзом лёгких

Рисунок 3.4: Примеры изображений из объединённого набора Montgomery-Shenzhen (MC-SZ)

3.1.2 Адаптивная предобработка входных изображений

Хотя финальным критерием при определении уровня жёсткости рентгеновского снимка грудной клетки является число чётко контурируемых верхних грудных позвонков [44; 45], для определения уровня контраста перед исследованием требуется обращать внимание на видимость других областей грудной клетки, (например, на элементы лёгочного рисунка) и органов [45]. На основании этого было принято решение не ограничивать область исследования на снимках пределами грудной клетки, а рассматривать изображения целиком.

Так как условия получения снимка, характеристики тела пациента и принципы считывания сигнала и хранения информации каждого конкретного устройства могут сильно влиять на характер рентгеновских изображений, перед подачей их на вход алго-

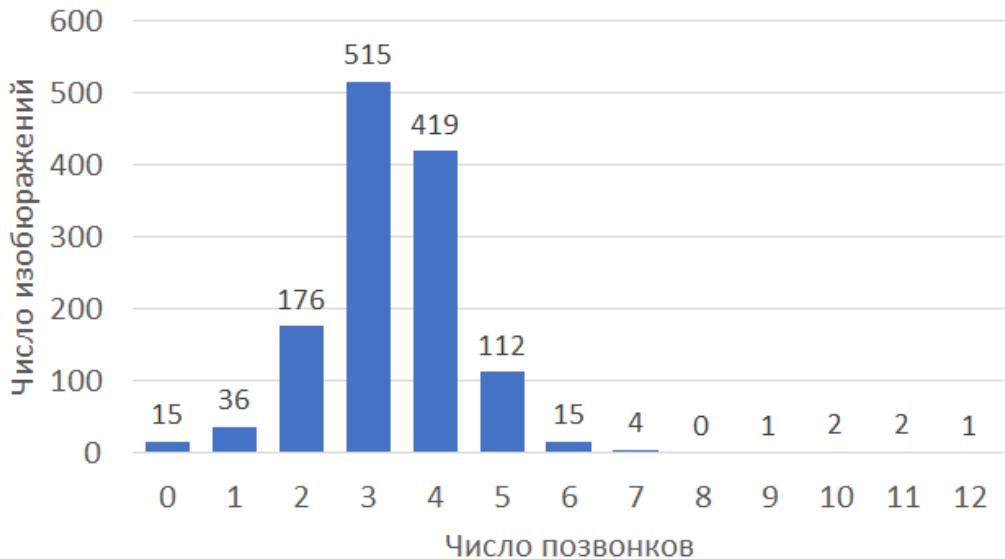


Рисунок 3.5: Гистограмма распределения снимков набора SakhaTB по числу отчётливо видимых позвонков



Рисунок 3.6: Гистограмма распределения снимков набора MC-SZ по числу отчётливо видимых позвонков

ритму определения жёсткости выполнялся этап предобработки, который заключается в последовательности следующих шагов:

- автоматическое контрастирование изображения:

$$h(x) = 255 \cdot \frac{x - p_{0.5}}{p_{99.5} - p_{0.5}},$$

где x — значение интенсивности пикселя входного изображения, $p_{0.5}$ и $p_{99.5}$ — это 0.5%- и 99.5%-процентили значений интенсивности всех пикселей изображения;

- автоматическая гамма-коррекция интенсивности пикселей изображения:

$$g(x) = 255 \cdot \left(\frac{x}{255} \right)^\gamma, \quad \gamma = \log_{\mu/255} 0.5,$$

где μ — средняя интенсивность всего изображения;

3. уменьшение изображения до входного разрешения, используемого нейронной сетью (512x512 пикселей для модели ResNet-18 и 384x384 пикселей для EfficientNetV2-S).
4. опциональный шаг глобальной или локальной (CLAHE [53]) эквализации гистограммы. Размер стороны квадратного окна в пикселях, используемого в методе локальной эквализации гистограммы, прямо зависит от размера стороны изображения и составляет $\frac{1}{2^n}$ от него, где $n \in \mathbb{N}$ — параметр метода. Влияние наличия этого шага и размера окна на качество работы алгоритма будет показано ниже.

3.1.3 Метод определения жёсткости рентгеновских снимков грудной клетки

Так как уровень жёсткости рентгеновского снимка — это упорядоченная величина, для сохранения отношений упорядоченности между классами задача его определения рассматривалась как задача порядковой регрессии (также иногда называется порядковой классификацией) [54; 55].

Для решения задачи автоматического определения жёсткости рентгеновского снимка грудной клетки был разработан нейросетевой метод. На вход алгоритма поступает рентгеновский снимок грудной клетки и подвергается предобработке, затем при помощи нейронной сети ему ставится в соответствие вещественное число на отрезке $[0, 1]$, которое является внутренним безразмерным показателем жёсткости снимка, на основании которого после сравнения с настраиваемыми порогами изображение относится к одному из рассматриваемых классов жёсткости. Пороги являются частью модели и настраиваются вместе с весами слоёв нейронной сети в процессе обучения. Преимуществом такого подхода является то, что с помощью внутреннего показателя жёсткости можно ранжировать изображения относительно друг друга даже в том случае, когда обрабатываемое изображение значительно отличается от обучающей выборки и для такого снимка пороги разделения классов жёсткости могут быть настроены неверно.

Для решения задач обработки медицинских изображений и, в частности, диагностики заболеваний широко используются [47] свёрточные нейронные сети семейств ResNet [56], DenseNet [57] и др.

В ходе исследования для решения задачи определения жёсткости снимков, вследствие малого размера выборки, была использована компактная сеть ResNet-18 с меньшим числом параметров и, следовательно, меньшей склонностью к переобучению по сравнению с более крупными сетями этой же архитектуры или представителями других вышеупомянутых архитектур.

Также было проведено сравнение качества решения этой задачи с таким у компактной сети более новой архитектуры свёрточных нейронных сетей EfficientNetV2-S [58], так как в задачах классификации изображений она проявляет себя лучше перечисленных выше. Её основное отличие от них заключается в оптимизации работы свёрточных слоёв с помощью их пропорционального масштабирования, смены порядка операций разной размерности, уменьшения размера ядер свёртки и отказа от «тяжёлых» слоёв, что позволяет уменьшить потребление памяти и задействовать освободившиеся ресурсы на увеличение глубины и обобщающей способности нейронной сети.

3.1.4 Эксперименты и результаты

Как видно из Рис. 3.5, набор изображений значительно несбалансирован и для ряда подуровней жёсткости содержит крайне мало примеров. Исходя из этого было принято решение, используя определённое врачом-рентгенологом число отчётливо видимых на снимках верхних грудных позвонков, разделить все снимки по уровню жёсткости на 3 группы, ориентируясь на медицинские критерии: мягкие (видно менее 3 позвонков), нормальные (видно 3-4 позвонка) и жёсткие (видно более 4 позвонков) [44; 45]. Количество снимков в сформированных группах представлено на Рис. 3.7. Именно эти три класса рассматривались в качестве допустимых значений целевой переменной для задачи порядковой регрессии.



Рисунок 3.7: Гистограмма распределения снимков набора SakhaTB по уровню жёсткости

Для ускорения процесса обучения за счёт наличия готовых низкоуровневых фильтров при решении обеих задач в качестве начального состояния нейронной сети использовались веса соответствующей модели, обученной на решение задачи классификации изображений реального мира ImageNet-1K [59]. Последний слой был заменён на

полносвязный слой с 1 выходом и 2 настраиваемыми в процессе обучения параметрами порогами пороговой модели порядковой регрессии [60], во время дальнейшего обучения были задействованы все слои.

Все базовые наборы изображений и их рассмотренные комбинации делились на обучающую, валидационную и тестовую выборки в соотношении 64:16:20 с предварительным случайным перемешиванием изображений и стратификацией по классу жёсткости для сохранения пропорций между классами. Изображения обучающей выборки в процессе обучения последовательно подвергались четырём случайным преобразованиям, а именно:

1. поворотам (в пределах 15 градусов в каждую сторону),
2. масштабированию (коэффициент выбирался случайно из отрезка [0.8, 1.2]),
3. сдвигам (до 30% размера изображения по каждой оси),
4. изменению яркости и контраста (до 20% в каждую сторону).

В качестве функции потерь в процессе обучения и валидации была выбрана функция потерь пороговой порядковой регрессии в виде суммы слагаемых, число которых зависит от числа классов (all-threshold) [60]:

$$\begin{cases} L(z, y) = \sum_{k=1}^{K-1} f(s(k, y) \cdot (\theta_k - z)), \\ s(k, y) = \begin{cases} -1, k < y, \\ +1, k \geq y, \end{cases} \end{cases}$$

где z — выход нейронной сети (безразмерный показатель жёсткости), принимающий значения из отрезка $[0, 1]$ (чем ближе к 1, тем выше жёсткость снимка), y — истинный класс соответствующего этому выходу изображения, K — общее число классов, $\theta_1 < \theta_2 < \dots < \theta_{K-1}$ — пороги, делящие действительную прямую на K частей, а $f(x)$ — базовая функция потерь бинарной классификации, в качестве которой использовалась логистическая функция потерь:

$$f(x) = \ln \frac{1}{1 + e^{-x}}.$$

Так как даже после сведения количества классов в задаче порядковой регрессии к трём выборка всё равно осталась сильно несбалансированной, то применялось взвешивание значений функции потерь для каждого примера с весами, обратно пропорциональными количеству изображений соответствующего класса. В роли показателя качества порядковой регрессии на этапе валидации выступала сбалансированная по классам средняя абсолютная ошибка (macro-averaged MAE, далее mMAE) [61].

Также в целях получения базовой оценки качества решения задачи была обучена модель для обычновенной классификации изображений на 3 класса. Последний слой модели-основы был заменён на полносвязный с 3 выходами. Роль функции потерь выполняла кросс-энтропия:

$$\begin{cases} CE(z, y) = \sum_{k=1}^K \mathbb{I}[y = k] \cdot \ln(softmax(z)_k), \\ softmax(z)_k = \frac{e^{z_k}}{\sum_{i=1}^K e^{z_i}}, \quad \mathbb{I}[y = k] = \begin{cases} 1, & y = k, \\ 0, & y \neq k, \end{cases} \end{cases}$$

а роль показателя качества на этапе валидации — сбалансированная точность (далее BalAcc) [62].

Для оптимизации функции потерь применялся алгоритм градиентного спуска AdamW [63] с параметрами $lr = 5 \cdot 10^{-6}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\lambda = 0.01$. Размер пакета (англ. batch) изображений был равен 64 для модели на основе ResNet-18 и 16 для модели на основе EfficientNetV2-S; в обоих случаях применялось накопление градиента на протяжении 8 и 2 итераций соответственно (до достижения размера «виртуального пакета» в 128 объектов). В конце каждой эпохи качество модели замерялось на валидационной выборке; при отсутствии уменьшения значения функции потерь на ней в течение 10 эпох шаг градиентного спуска уменьшался в 5 раз, а при отсутствии улучшения в течение 31 эпохи обучение прекращалось. Переобучение контролировалось замерами функции потерь и вышеуказанных показателей качества, однако значительного ухудшения качества работы модели на валидационной выборке с течением времени не наблюдалось (см. Рис. 3.8; можно заметить небольшой рост значений функции потерь у модели на основе EfficientNetV2-S, что свидетельствует о некотором переобучении модели), поэтому в качестве финального состояния принималось значение весов модели в конце последней эпохи.

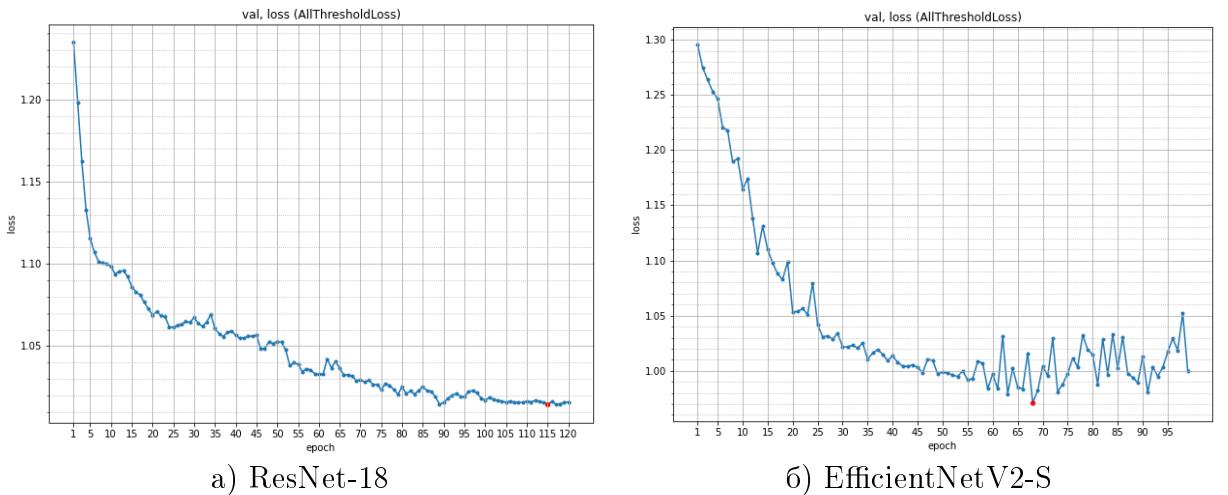


Рисунок 3.8: Примеры графиков зависимости функции потерь от количества эпох на валидационной выборке в задаче анализа жёсткости снимков грудной клетки

Итоговые значения показателей качества, полученные на тестовой выборке, представлены в Табл. 3. Модели для решения задачи порядковой регрессии содержат «ord» в своём названии, модель для решения задачи классификации — «clf». Модель «ord-eff», была создана на основе EfficientNetV2-S, а остальные модели — на основе ResNet-18. В столбце «Эквалайзация гистограммы» стоит прочерк, если эквалайзация не применялась; «глобальная» в случае использования глобальной эквалайзации; размер окна локальной

эквализации гистограммы в виде доли от размера всего изображения в случае использования локальной эквализации.

На основании сбалансированной точности и сбалансированной МАЕ лучшей моделью оказалась модель порядковой регрессии с локальной эквализацией гистограммы на этапе предобработки снимков с окном, длина стороны которого была равна $\frac{1}{2}$ длины стороны изображения. Далее для краткости она будет обозначена «ord-clahe2». Дальнейший анализ изображений проводится с использованием этой модели.

Таблица 3: Значения показателей качества работы алгоритмов определения жёсткости на тестовой выборке набора SakhaTB

Модель	Эквализация гистограммы	BalAcc	mMAE
clf	-	0.563	0.452
ord-eff	-	0.623	0.399
ord	-	0.609	0.452
ord-glob	глобальная	0.609	0.452
ord-clahe2	$\frac{1}{2}$ стороны изображения	0.636	0.398
ord-clahe4	$\frac{1}{4}$ стороны изображения	0.610	0.449
ord-clahe8	$\frac{1}{8}$ стороны изображения	0.593	0.468
ord-clahe16	$\frac{1}{16}$ стороны изображения	0.600	0.468

На Рис. 3.9 представлена зависимость вероятностей классов «Жёсткий» (Hard) и «Мягкий» (Soft), предсказанных моделью «clf» для решения задачи простой классификации, а также безразмерной величины жёсткости, предсказанной моделью «ord-clahe2» для задачи порядковой регрессии, и истинного значения уровня жёсткости объектов тестовой выборки. Можно заметить, что разделение классов далеко от идеального.

Как выяснилось после проверки ошибочно классифицированных снимков, причиной тому послужила шумность использованного набора данных из-за неоднозначности и недостаточной определённости критериев разметки. Подтверждением этого предположения также служит близость значений меры качества на обучающей выборке к таковым на тестовой выборке: сбалансированная точность около 0.70 и 0.67 у моделей «clf» и «ord-clahe2» соответственно.

Рассмотрение задачи определения жёсткости рентгеновского снимка как задачи порядковой регрессии позволяет, используя полученную для её решения модель, с некоторой точностью ранжировать изображения по жёсткости на основе внутреннего показателя жёсткости нейронной сети. В качестве меры качества ранжирования был выбран коэффициент ранговой корреляции Спирмена [64], так как он позволяет обнаруживать в том числе нелинейные зависимости рассматриваемых величин. Замеры

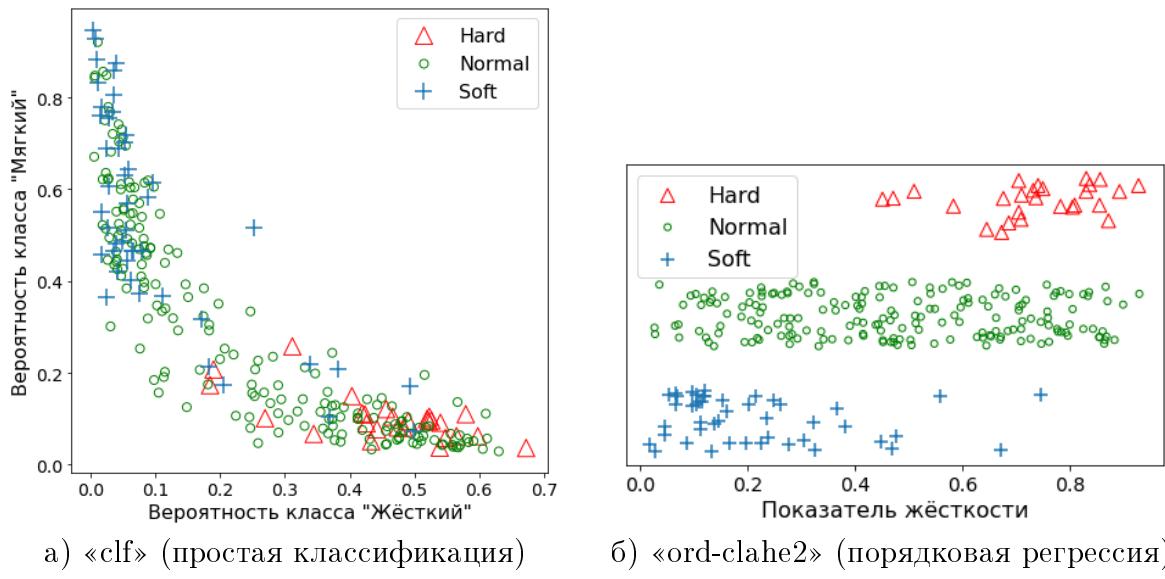


Рисунок 3.9: Зависимость предсказаний моделей в зависимости от истинного класса объекта для набора SakhaTB

делались относительно указанных врачом количества чётко видимых на снимке позвонков и относительно классов жёсткости снимков. Полученные значения приведены в Табл. 4.

Таблица 4: Значения показателя качества ранжирования тестовой выборки набора SakhaTB алгоритмом определения жёсткости

Модель	Spearman (позвонки)	Spearman (жёсткость)
ord-eff	0.564	0.457
ord	0.576	0.497
ord-glob	0.599	0.514
ord-clahe2	0.606	0.534
ord-clahe4	0.602	0.519
ord-clahe8	0.588	0.498
ord-clahe16	0.596	0.507

Была проведена оценка качества работы модели «ord-clahe2» на наборе MC-SZ. Результаты сравнения качества работы алгоритма на наборе MC-SZ и на тестовой выборке набора SakhaTB приведены в Табл. 5.

Таблица 5: Значения показателя качества ранжирования тестовой выборки набора SakhaTB алгоритмом определения жёсткости

Набор данных	BalAcc	mMAE	Spearman (позвонки)	Spearman (жёсткость)
Набор SakhaTB	0.636	0.398	0.606	0.534
MC-SZ	0.546	0.565	0.325	0.203

3.2 Использование результатов оценки качества рентгеновских снимков при нейросетевой диагностике туберкулёза лёгких

Результаты оценки жёсткости рентгенограмм могут быть использованы, например, для предварительной фильтрации обучающей выборки и входных данных нейросетевого алгоритма компьютерной диагностики туберкулёза лёгких. Как показано в работе [46], такая фильтрация, а также создание отдельных моделей для разных уровней жёсткости снимков позволяют повысить рабочесть и точность метода компьютерной диагностики. В этом разделе демонстрируется влияние предварительной фильтрации обучающей и валидационной выборок на качество работы алгоритма классификации в задаче диагностики туберкулёза лёгких по рентгеновским снимкам грудной клетки.

3.2.1 Формирование набора данных

Так как в экспериментах этапу обучения модели в задаче диагностики туберкулёза лёгких предшествует этап фильтрации снимков на основании уровня жёсткости, то в целях поддержания размера обучающей и валидационной выборок после прореживания на достаточном уровне, а также для создания более разнородных выборок, как и в работе [46], для создания нейросетевого метода диагностики и замера влияния фильтрации на качество диагностики использовался набор рентгенограмм, состоящий из двух частей: уже упомянутый в п. 3.1.1 составной набор MC-SZ и ещё один из часто применяемых при разработке методов компьютерной диагностики туберкулёза лёгких общедоступных наборов рентгенограмм [49; 50] – TBX11K [65].

Набор TBX11K подготовлен в Наньцзинском университете г. Тяньцзин в КНР. Он состоит из 11200 рентгеновских снимков грудной клетки в оттенках серого с глубиной цвета 8 бит в формате PNG с разрешением 512x512 пикселей. Он изначально поделён на обучающую, валидационную и тестовую выборки, при этом последняя не имеет разметки, поэтому из всего набора только для 8400 изображений доступны метки при-

надлежности к одному из 3 классов (3800 здоровых и 800 больных туберкулёмом лёгких пациентов, а также больные иными заболеваниями пациенты) и границы поражённых областей лёгких. В рамках исследования использовались только снимки здоровых и больных туберкулёмом лёгких пациентов, а остальные были исключены из рассмотрения. Примеры изображений из набора представлены на Рис. 3.10.

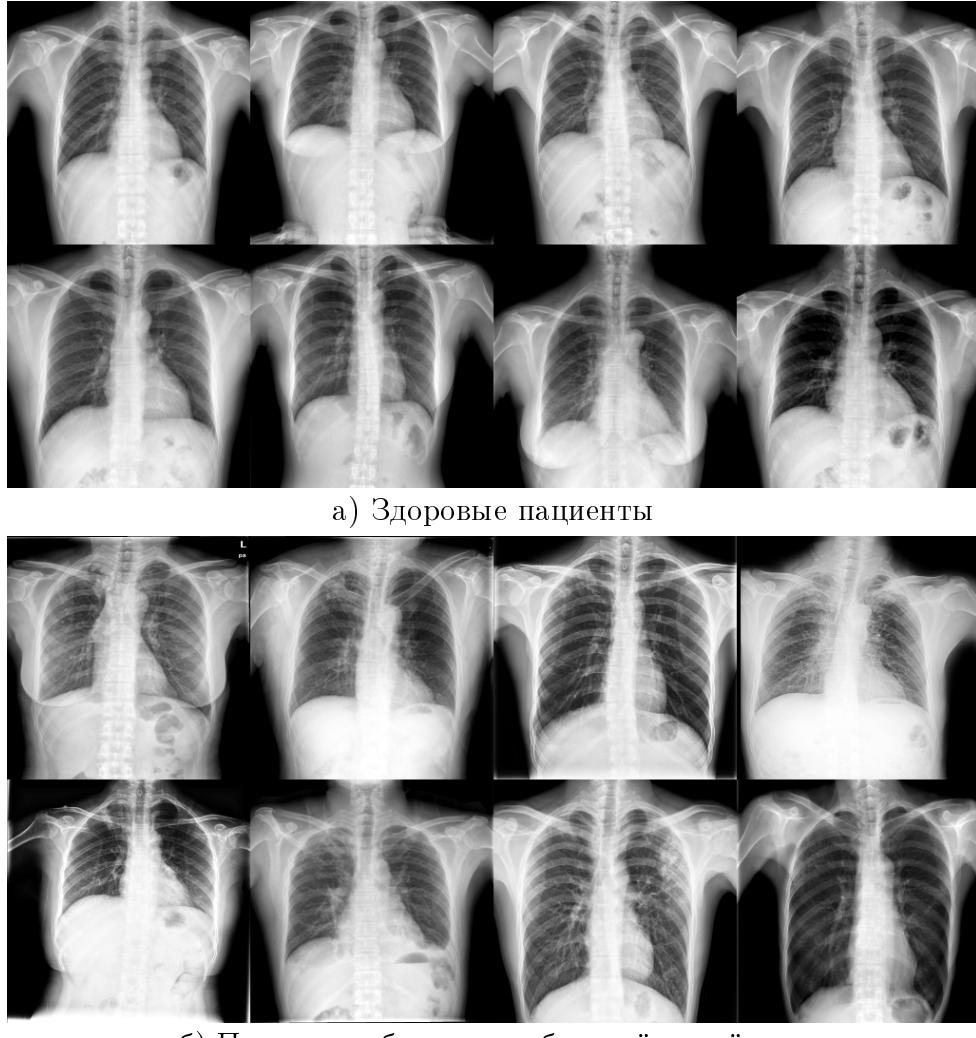


Рисунок 3.10: Примеры изображений из набора TBX11K

Рентгенограммам здоровых пациентов из получившегося объединённого набора были поставлены в соответствие метки класса NORMAL, а больных туберкулёмом лёгких пациентов — класса TB. Итоговое число снимков каждого класса в объединённом наборе, а также в его отдельных частях приведено в Табл. 6.

Таблица 6: Размеры использованных наборов данных

Название набора	Число снимков класса NORMAL	Число снимков класса ТВ	Общее число снимков
Montgomery	80	58	138
Shenzhen	326	336	662
TBX11K	3800	800	4600
Всего	4206	1194	5400

3.2.2 Метод компьютерной диагностики туберкулёза лёгких

Использованный в этом разделе для экспериментов алгоритм диагностики туберкулёза лёгких выглядел следующим образом:

1. поступивший на вход рентгеновский снимок грудной клетки подвергается предобработке, совпадающей с описанной в п. 3.1.2 за исключением того, что эквализация гистограммы не проводилась;
2. при помощи нейросетевой модели ему ставятся в соответствие 2 вещественных числа из отрезка $[0, 1]$, которые выражают предсказанные веса классов NORMAL и ТВ (сумма весов для каждого изображения равна 1);
3. класс, чей вес больше, принимается за выходное значение алгоритма.

В качестве основы для нейросетевой модели использовалась сеть EfficientNetV2-S, последний слой которой был заменён на полно связанный слой с 2 выходами. Процедуры деления набора данных на подвыборки и обучения модели со взвешиванием классов для балансировки, а также начальные состояния весов сети совпадали с описанными в п. 3.1.4. В роли функции потерь выступала кросс-энтропия, а в роли меры качества — сбалансированная точность.

3.2.3 Эксперименты и результаты

Рентгенограммы из наборов MC-SZ и TBX11K были обработаны разработанным алгоритмом оценки жёсткости. Распределение снимков из тестовой выборки сформированного в п. 3.1.1 набора и из этих двух наборов по безразмерному показателю жёсткости, предсказанному моделью «ord-clahe2», представлен на Рис. 3.11. Однако следует осторожно относиться к разделению снимков из наборов MS-SZ и TBX11K на классы жёсткости, так как истинные пороги разделения классов для этих данных мо-

гут сильно отличаться от полученных моделью на этапе обучения в п. 3.1.4 из-за их визуального отличия от изображений её обучающей выборки.

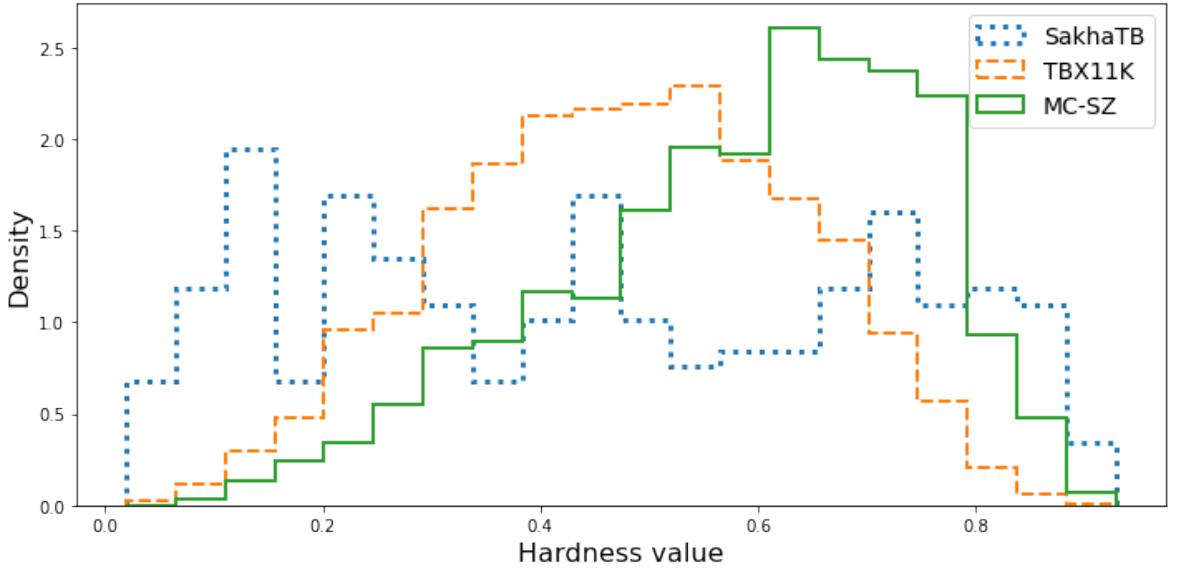


Рисунок 3.11: Гистограммы распределения снимков трёх использованных наборов изображений по предсказанному моделью «ord-clahe2» показателю жёсткости

Отдельные гистограммы для классов NORMAL и TB для наборов MC-SZ и TBX11K представлены на Рис. 3.12. Небольшие различия гистограмм классов совпадают с визуальной разницей между снимками этих классов: класс TB в обоих из них содержит больше мягких снимков, а классы NORMAL — больше жёстких (см. Рис. 3.4 и 3.10).

На основании предсказанных величин жёсткости было проведено удаление из объединённого набора MC-SZ-TBX11K одинаковой доли самых жёстких и самых мягких снимков (то есть с обеих сторон гистограммы распределения величин жёсткости). После этого качество обученного на прореженной обучающей выборке метода диагностики замерялось на прореженной тестовой выборке и сравнивалось с качеством работы на такой же тестовой выборке модели диагностики, обученной на непрореженной обучающей выборке.

Исходя из визуального различия составных частей используемого набора данных и относительно малого размера набора MC-SZ, было принято решение прореживать не объединённый набор изображений, а каждую из двух его частей отдельно, чтобы изменение их пропорций в итоговой выборке не повлияло на качество работы. Такой выборочный подход обусловлен необходимостью более точного контрастирования снимков, полученных на разных аппаратах и в разных условиях, и приведения их к как можно более близкому внешнему виду перед определением жёсткости.

Помимо прореживания выборок с обеих сторон гистограммы уровня жёсткости был рассмотрен и случай отбрасывания только самых жёстких снимков, так как на мягких снимках ещё могут сохраняться некоторые детали лёгочной ткани, в том время как на жёстких они могут быть полностью утеряны.

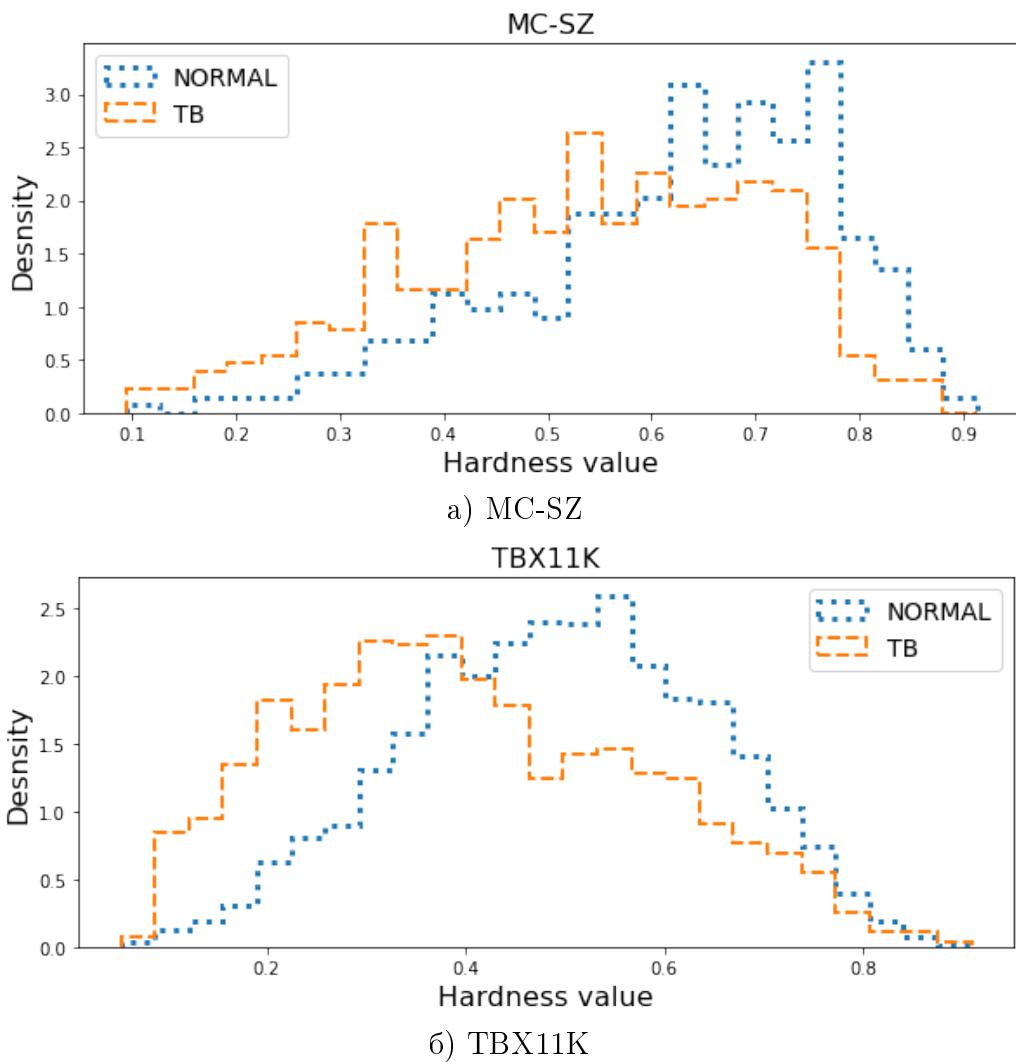


Рисунок 3.12: Гистограммы распределения снимков каждого класса для использованных в задаче диагностики туберкулёза лёгких наборов изображений MC-SZ и TBX11K по предсказанному моделью «ord-clahe2» показателю жёсткости

Значения сбалансированной точности, полученные в результате экспериментов представлены в Табл. 7. Из неё видно, что изменение качества зависит от степени прореживания изображений, однако остаётся стабильно положительным. Сравнение более распространённых в медицине показателей чувствительности и специфичности для класса TB приведено в Табл. 8.

Таблица 7: Сравнение качества классификации моделей, обученных на полном и прореженном наборе (сбалансированная точность)

Доля удалённых	Удаление жёстких и мягких изображений			Удаление только жёстких изображений		
	5%	10%	15%	5%	10%	15%
До	0.958	0.951	0.951	0.962	0.961	0.965
После	0.961	0.962	0.953	0.968	0.966	0.975

Таблица 8: Сравнение качества классификации моделей, обученных на полном и прореженном наборе (чувствительность / специфичность)

Доля удалённых	Удаление жёстких и мягких изображений			Удаление только жёстких изображений		
	5%	10%	15%	5%	10%	15%
До	0.923 / 0.994	0.909 / 0.994	0.908 / 0.995	0.930 / 0.994	0.927 / 0.995	0.934 / 0.996
	0.933 / 0.990	0.933 / 0.991	0.915 / 0.990	0.943 / 0.994	0.941 / 0.992	0.958 / 0.993
После						

3.3 Выводы

В данной главе была продемонстрирована возможность использования нейросетевых методов для определения уровня жёсткости рентгеновских снимков грудной клетки. Хотя из-за особенности имеющихся данных не удалось достичь высоких показателей качества, полученные результаты заметно превышают возможности случайного выбора ответа: сбалансированная точность составила 0.636, а коэффициенты корреляции Спирмена — 0.606 и 0.534. Однако качество работы полученного алгоритма ощутимо снижается в случае применения его к данным из других источников: для набора MC-SZ сбалансированная точность упала до 0.546, а значения показателя качества ранжирования объектов снизились примерно вдвое.

Однако даже такая несовершенная модель определения жёсткости позволяет увеличить качество работы алгоритма диагностики туберкулёза лёгких при условии предварительной фильтрации изображений перед обучением классификатора и получением предсказаний. Снижение разброса жёсткости данных и увеличение их однородности заметно повысило точность обнаружения больных пациентов при сохранении или малом снижении специфичности: наибольший абсолютный и относительный прирост чувствительности алгоритма для класса ТВ наблюдался при удалении 10% самых жёстких и 10% самых мягких снимков (с 0.909 до 0.933) и при удалении 15% самых жёстких снимков (с 0.934 до 0.958). Во втором случае также было достигнуто наибольшее значение чувствительности: 0.958.

Глава 4. Программный комплекс реализации алгоритмов анализа биомедицинских изображений

Данная глава посвящена программной реализации разработанных ранее алгоритмов и методов. Так как рассмотренные в работе задачи относятся к разным областям медицины и биологии и решают задачи разного типа, то разработанные методы и алгоритмы были реализованы в виде независимых программных модулей.

Программный комплекс, разработанный в рамках данной главы, содержит в себе три отдельных модуля:

1. модуль повышения разрешения изображений мигающей флуоресцентной микроскопии;
2. модуль повышения резкости медицинских изображений методом деформации пиксельной сетки;
3. модуль анализа и обработки рентгенограмм грудной клетки.

Кроме того, в рамках исследования был сформирован набор рентгеновских изображений грудной клетки для компьютерной диагностики туберкулёза лёгких.

Ниже представлено детальное описание каждого из модулей и сформированного набора рентгенограмм.

4.1 Программный модуль повышения разрешения изображений мигающей флуоресцентной микроскопии

Данный программный модуль реализует разработанный в Главе 1 метод повышения разрешения изображений мигающей флуоресцентной микроскопии.

С целью упрощения разработки и отладки, для написания программного модуля использовался язык программирования Python 3. При разработке применялись программные библиотеки для работы с изображениями, математическими функциями и векторными вычислениями: imageio, scikit-image, NumPy, SciPy, PyTorch.

Для удобства потенциального внедрения в программные продукты в качестве расширения и в автоматизированные системы модуль представляет собой консольное приложение. Формат аргументов командной строки имеет следующий вид:

```
warping.exe -i FILE -o FILE -a FLOAT -x INT -m INT -p 0|1 -s FLOAT  
-z INT -n INT -e FLOAT -v FLOAT
```

и определяет следующие параметры:

-i FILE

задаёт имя файла входной серии изображений;

-o FILE

задаёт имя файла выходного изображения;

-x INT
задаёт длину волны лазера подсветки в нанометрах;

-m INT
задаёт длину волны флуоресцентного излучения в нанометрах;

-p 0|1
задаёт тип детектора микроскопа: 0 — обычновенный, 1 — детектор Airyscan;

-s FLOAT
задаёт длину стороны пикселя в нанометрах;

-z INT
задаёт степень увеличения изображения;

-n INT
задаёт число итераций алгоритма оптимизации;

-e FLOAT
задаёт значение параметра η_0 ;

-v FLOAT
задаёт значение параметра v_0 .

Поддерживаемые форматы входных и выходных изображений:

- Windows bitmap (*.bmp),
- JPEG (*.jpg, *.jpeg),
- Portable network graphics (*.png),
- Tagged Image File Format (*.tif, *.tiff).

Исходный код распределён по файлам в соответствии с функциями, которые он выполняет:

- run.py — содержит код обработки параметров командной строки, запуска обработки входных данных и сохранения результата;
- process.py — содержит функции и классы, реализующие алгоритм;
- psf_functions.py — содержит функции для построения матриц размытия и понижения разрешения для указанных параметров оптической системы;
- pixel_reassignment.py — содержит функции для выполнения переназначения пикселей (англ. pixel reassignment).

Сборка интерпретатора языка Python, необходимых библиотек и файлов с исходным кодом программного модуля в консольное приложение осуществлялась с помощью инструмента PyInstaller.

4.2 Программный повышения резкости медицинских изображений методом деформации пиксельной сетки

Данный программный модуль реализует разработанный в Главе 2 однопараметрический алгоритм повышения резкости медицинских изображений деформационным методом.

Для написания программного модуля использовался язык программирования Python 3 и программные библиотеки для работы с изображениями, математическими функциями и векторными вычислениями: imageio, scikit-image, NumPy, SciPy.

Программный модуль представляет собой консольное приложение. Формат аргументов командной строки имеет следующий вид:

```
warping.exe -i FILE -o FILE [-l FLOAT] [-h FLOAT] -t g|c|r -r FLOAT
```

и определяет следующие параметры:

-i FILE

задаёт имя файла входного изображения;

-o FILE

задаёт имя файла выходного изображения;

-l FLOAT

задаёт нижний порог детектора контуров Кэнни в виде вещественного числа от 0 до 1;

-h FLOAT

задаёт верхний порог детектора контуров Кэнни в виде вещественного числа от 0 до 1;

-t g|c|r

задаёт тип ядра размытия: «g» — гауссово ядро, «c» — круг, «r» — круг с кольцом;

-r FLOAT

задаёт радиус ядра размытия в виде вещественного числа, большего 0.

Поддерживаемые форматы входных и выходных изображений:

– Windows bitmap (*.bmp),

– JPEG (*.jpg, *.jpeg),

– Portable network graphics (*.png).

Поддерживается работа как с чёрно-белыми, так и с цветными изображениями: во втором случае для обнаружения контуров используется чёрно-белая версия изображения, для получения которой применяется формула:

$$Y = 0.2125R + 0.7154G + 0.0721B,$$

где Y — значение основного канала для изображения в градациях серого, а R , G и B — интенсивности красного, зелёного и синего каналов соответственно.

Сборка интерпретатора языка Python, необходимых библиотек и файлов с исходным кодом программного модуля в консольное приложение осуществлялась с помощью инструмента PyInstaller.

4.3 Программный модуль анализа и обработки рентгенограмм грудной клетки

Данный программный модуль реализует разработанный в Главе 3 однопараметрический алгоритм повышения резкости медицинских изображений деформационным методом.

Для написания программного модуля использовался язык программирования Python 3 и программные библиотеки для работы с изображениями, математическими функциями, векторными вычислениями и машинного обучения: imageio, scikit-image, NumPy, SciPy, PyTorch, scikit-learn.

Программный модуль представляет собой консольное приложение. Формат аргументов командной строки имеет следующий вид:

```
chest-xray.exe -i FILE -h|-d
```

и определяет следующие параметры:

-i FILE

задаёт имя файла входного изображения;

-h

включает режим анализа жёсткости рентгенограммы;

-d

включает режим диагностики туберкулёза.

В качестве результата программа выводит значение программной жёсткости и степень уверенности в диагнозе «туберкулём».

Поддерживаемые форматы входных и выходных изображений:

- Windows bitmap (*.bmp),
- JPEG (*.jpg, *.jpeg),
- Portable network graphics (*.png),
- DICOM (*.dcm).

Исходный код распределён по файлам в соответствии с функциями, которые он выполняет:

- base.py — содержит код функций загрузки, сохранения, обучения нейросетевой модели, получения предсказания и предобработки входных изображений;
- predict.py — содержит функции запуска обработки и реализации алгоритма.

Сборка интерпретатора языка Python, необходимых библиотек и файлов с исходным кодом программного модуля в консольное приложение осуществлялась с помощью инструмента PyInstaller.

4.4 Набор рентгеновских изображений грудной клетки для компьютерной диагностики туберкулёза лёгких

В отличие от таких заболеваний лёгких, как бактериальная или вирусная пневмония, данных по диагностике туберкулёза лёгких в открытом доступе не так много. Кроме того, из-за частичной несовместимости различных наборов рентгеновских изображений грудной клетки качество работы алгоритмов диагностики на рентгенограммах из наборов, не попавших в обучающую выборку, может сильно падать. Причины подобной низкой робастности (то есть низкой устойчивости качества работы алгоритма с различными наборами входных данных) включают в себя различия в условиях получения снимков, в диапазонах и методах преобразования интенсивности пикселей, в сложности визуального разделения классов.

Анализ доменного сдвига (явления, при котором наблюдается снижение качества работы алгоритмов при его замере на наборах данных, относящихся к той же области, но отличных от использованных при разработке этих алгоритмов) для рентгенограмм органов грудной клетки был проведён: в [66] — для некоторых открытых наборов (наборов, находящихся в открытом доступе) применительно к задаче обнаружения изображения лёгких; и в [67] — для крупнейших открытых наборов применительно к задаче диагностики патологии лёгких. Кроме того, влияние различия в распределении пациентов по возрасту на качество сегментации грудной клетки на рентгеновских снимках упоминается в [68], а влияние гендерного дисбаланса на точность классификации рассмотрено в [69].

В рамках исследования было принято решение подготовить набор рентгенограмм грудной клетки, который позволил бы в некоторой степени закрыть пробелы в доступных в свободном доступе наборах изображений и повысить устойчивость качества работы алгоритмов автоматической диагностики туберкулёза лёгких для входных изображений, полученных в разных учреждениях и условиях.

4.4.1 Общедоступные данные

Согласно недавним обзорам методов диагностики туберкулёза лёгких на основе глубокого обучения, таким как [47—49], основными открытыми источниками данных

для задачи диагностики туберкулёза лёгких по рентгеновским изображениям грудной клетки всё ещё являются наборы Montgomery County [51] и Shenzhen [52]. Другие репрезентативные наборы снимков в свободном доступе [50] включают DA и DB [70] и TBX11K [65].

Перечисленные наборы были использованы для проведения экспериментов и проверки влияния добавления предлагаемого набора рентгенограмм на качество диагностики туберкулёза.

Описание наборов Montgomery County, Shenzhen и TBX11K было приведено ранее в п. 3.1.1 и 3.2.1, использованные в данной главе варианты этих наборов совпадают с описанными там.

Наборы DA и DB [70] были собраны в Национальном институте туберкулёза и респираторных заболеваний в Нью-Дели в Индии. Набор DA содержит по 78 изображений здоровых и больных туберкулёзом лёгких пациентов в оттенках серого в формате JPEG с разрешением 1024x1024 или 2320x2828 пикселей и глубиной цвета 8 бит. Набор DB содержит изображения в оттенках серого в формате DICOM с разрешением 3008x3008 пикселей и глубиной цвета 16 бит, и в рамках данного исследования был преобразован в формат PNG с глубиной цвета 8 бит. Некоторые снимки больных туберкулёзом лёгких из набора DB отсутствуют в его репозитории на сайте SourceForge, поэтому итоговое число снимков в нём составило 75 рентгенограмм здоровых и 47 рентгенограмм больных туберкулёзом лёгких пациентов. Эти два набора в силу их родства и малого размера по отдельности в рамках данной работы были объединены вместе. Примеры изображений из набора (обрезанные до соотношения сторон 1:1) представлены на Рис. 4.1.

Фактическое количество изображений лёгких здоровых пациентов (далее — класс «Healthy») и изображений лёгких пациентов с проявлениями туберкулёза (класс «TB»), для которых имеются метки классов, в использованных открытых наборах представлены в Табл. 9.

Таблица 9: Размеры рассматриваемых открытых наборов рентгенограмм грудной клетки

Набор данных	Число снимков по классам		
	TB	Healthy	Общее число снимков
Montgomery County (MC)	58	80	138
Shenzhen (SZ)	336	326	662
DA	78	78	156
DB	47	75	122
TBX11K	800	3800	4600

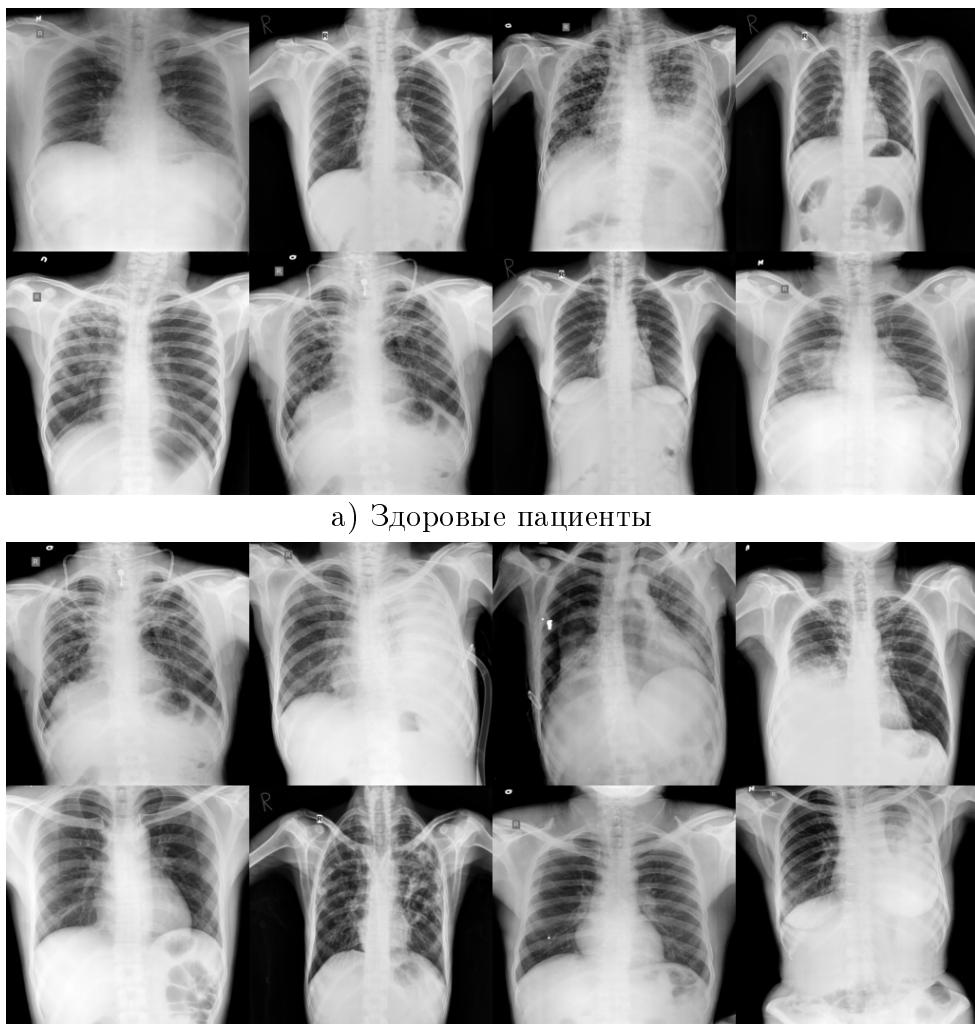


Рисунок 4.1: Примеры изображений из объединённого набора DA + DB

4.4.2 Набор рентгенограмм грудной клетки Sakha-TV

Сформированный в рамках исследования набор рентгенограмм грудной клетки, сделанных во фронтальной проекции, собран в результате сотрудничества с несколькими медицинскими учреждениями Республики Саха (Якутия); ниже он обозначен «Sakha-TV». Процедура формирования набора изображений из имеющихся коллекций рентгенограмм включала в себя:

- удаление снимков, сделанных в боковой проекции;
- контроль правильности установленных значений «Photometric Interpretation» в файлах DICOM;
- удаление повторяющихся сессий;
- удаление пациентов с несколькими сессиями;
- удаление несовершеннолетних пациентов;
- а также некоторую балансировку соотношения полов и распределений возрастов и диагнозов.

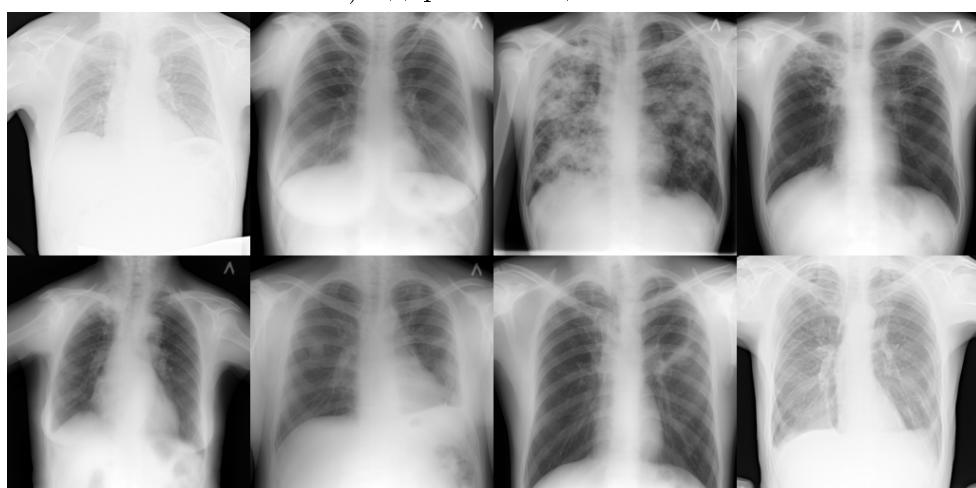
В результате было отобрано 400 изображений здоровых и 400 изображений больных туберкулём лёгких пациентов, где каждому пациенту соответствует только 1 снимок. Для получения рентгенограмм использовались стационарные и переносные комплексы оборудования. В основном разрешение изображений примерно равно 3000x3000 пикселей, но часть снимков имеет меньший размер вплоть до около 2000x2000 пикселей. Большинство изображений имеет глубину цвета 16 бит, у остальных снимков она равна 8 битам.

Врачебная диагностика осуществлялась посредством независимого двойного чтения снимков с подтверждением диагноза «туберкулёт» специалистами НПЦ «Фтизиатрия» им. Е.Н. Андреева на основании клинико-лабораторных и микробиологических данных.

Примеры рентгенограмм из сформированного набора изображений представлены на Рис. 4.2. Сравнение распределений снимков по диагнозу, полу и возрасту полученного набора и набора MC + SZ (для наборов TBX11K, DA и DB такие данные недоступны) изображены на Рис. 4.3-4.5. Сам набор изображений Sakha-TB доступен для загрузки на странице по ссылке: <https://imaging.cs.msu.ru/en/research/TB/Sakha-TB>.



а) Здоровые пациенты



б) Пациенты, больные туберкулём лёгких

Рисунок 4.2: Примеры изображений из набора Sakha-TB

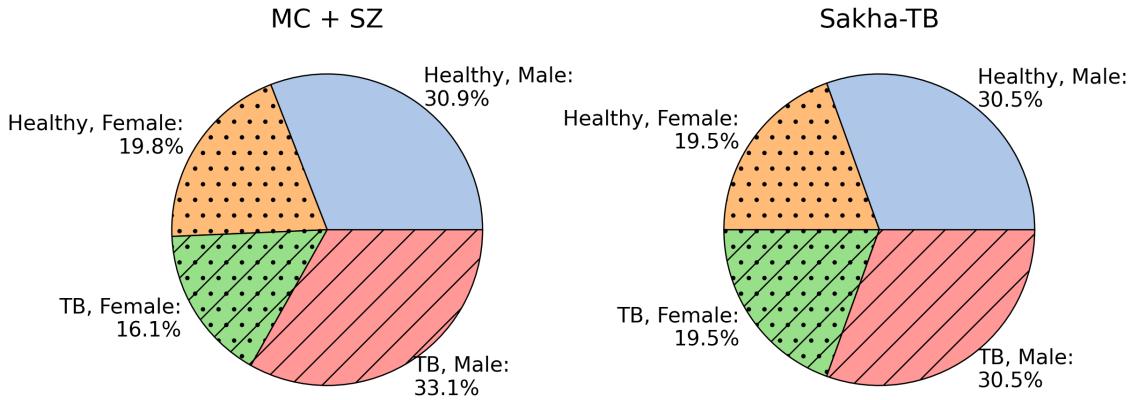


Рисунок 4.3: Распределение изображений в наборах MC + SZ и Sakha-TB по диагнозу и полу

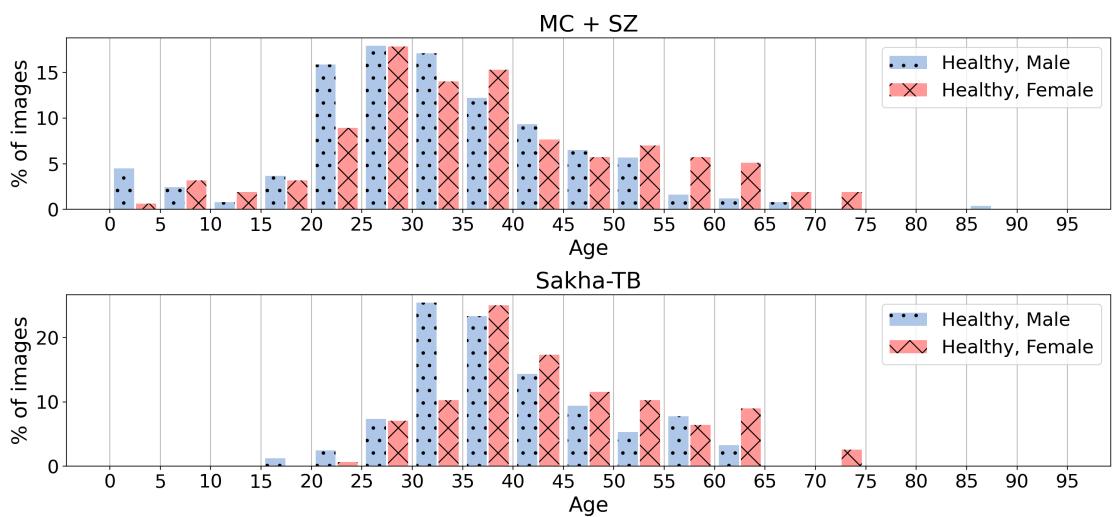


Рисунок 4.4: Распределение изображений здоровых пациентов в наборах MC + SZ и Sakha-TB по возрасту в зависимости от пола

4.4.3 Эксперименты

Для оценки возможностей применения предлагаемого набора изображений в задаче компьютерной диагностики туберкулёза были сделаны замеры качества работы методов глубокого обучения при обучении и тестировании на разных комбинациях рассматриваемых открытых наборов снимков:

1. MC + SZ,
2. DA + DB,
3. TBX11K,
4. Sakha-TB,
5. (MC + SZ) + Sakh-TB,
6. (DA + DB) + Sakha-TB,
7. TBX11K + Sakha-TB,

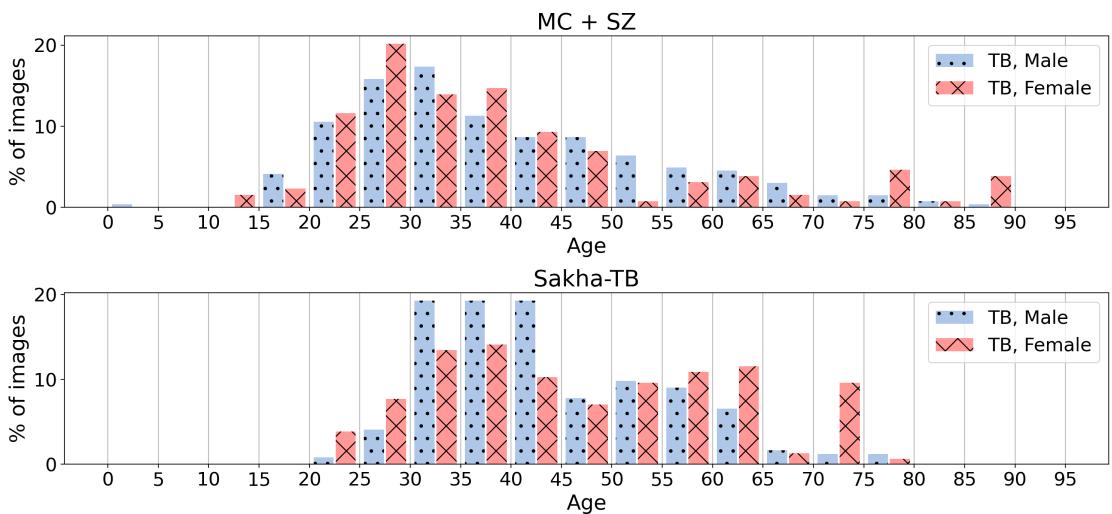


Рисунок 4.5: Распределение изображений больных туберкулёзом пациентов в наборах MC + SZ и Sakha-TB по возрасту в зависимости от пола

8. TBX11K + (MC + SZ),
9. TBX11K + (MC + SZ) + (DA + DB),
10. TBX11K + (MC + SZ) + (DA + DB) + Sakha-TB.

Снимки из набора Sakha-TB предварительно подверглись обработке, совпадающей с описанной в п. 3.1.1.

Использованный в этом разделе для экспериментов алгоритм диагностики туберкулёза лёгких совпадал с описанным в п. 3.2.2, за исключением следующих деталей.

В качестве основы нейросетевой модели были рассмотрены две современные модели EfficientNetV2-M [58] и VisionTransformer-Base/16 [71], которые демонстрируют высокую точность в задачах классификации изображений. Далее варианты алгоритма на их основе будут обозначаться «EffNetV2-M» и «ViT-B16» соответственно.

Размер пакета изображений составлял 64 для модели EffNetV2-M и 32 для ViT-B16. В конце каждой эпохи качество модели замерялось на валидационной выборке, при отсутствии уменьшения значения функции потерь на ней в течение 5 эпох градиентного спуска уменьшался в 5 раз. Для предотвращения переобучения при отсутствии уменьшения значения функции потерь на валидационной выборке в течение 16 эпох обучение прерывалось. В качестве финального состояния весов модели принимались значение весов, соответствующие эпохе с наивысшим значением показателя качества на валидационной выборке.

4.4.4 Результаты

Итоговые замеры качества работы обученных алгоритмов на рассматриваемых наборах производились с использованием широко используемых в медицине таких ста-

тистических показателей, как чувствительность (англ. sensitivity) и специфичность (англ. specificity). Были построены ROC-кривые, которые позволяют оценить общие возможности алгоритма классификации с учётом возможности варьирования порога уверенности для каждого класса. Интегральным показателем возможностей алгоритма в таком случае может считаться величина ROC AUC, представляющая собой площадь под ROC-кривой рассматриваемого алгоритма, которая принимает значения от 0 до 1, где 1 соответствует абсолютно точному алгоритму. ROC-кривые с показателями ROC AUC для рассмотренных моделей, обученных на некоторых из приведённых выше комбинаций рассматриваемых наборов рентгеновских снимков грудной клетки, представлены на Рис. 4.6–4.9. Для построения кривых для наборов, вошедших в обучающую выборку алгоритма, использовались их тестовые выборки, такие наборы выделены звёздочкой в описании графика; в противном случае использовались все наборы целиком.

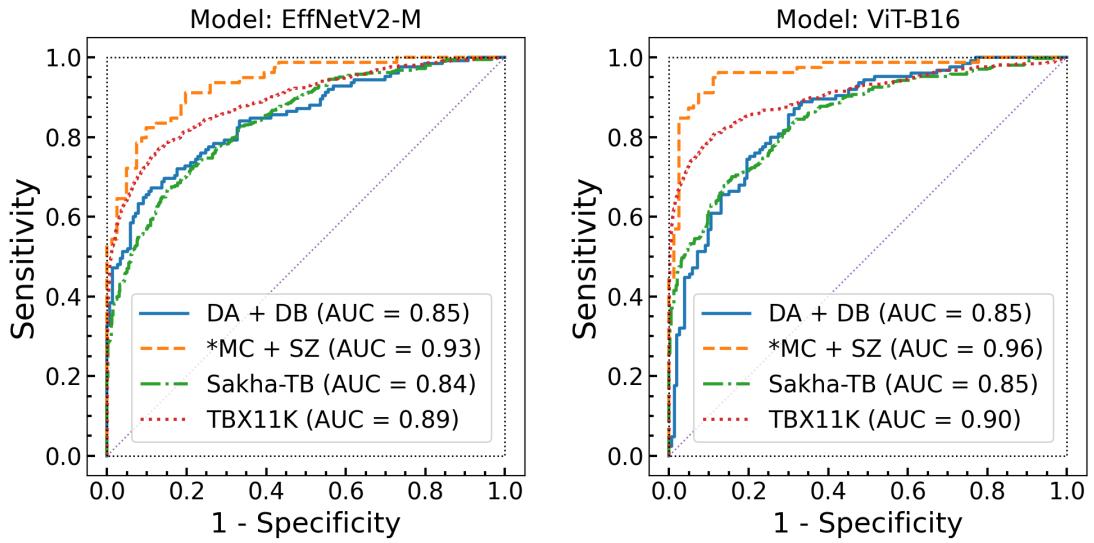


Рисунок 4.6: Графики ROC-кривых для моделей, обученных на наборе MC + SZ

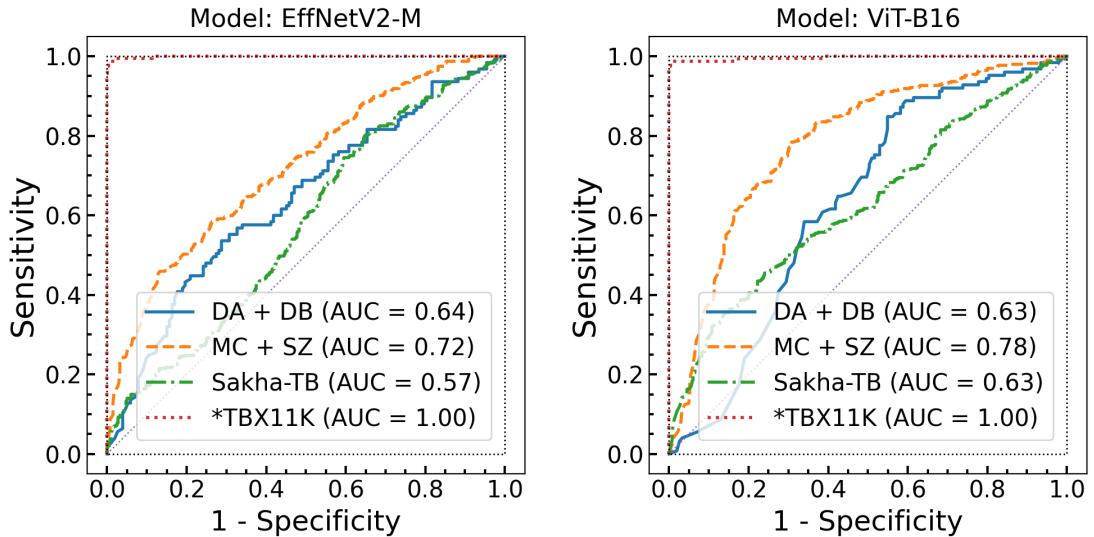


Рисунок 4.7: Графики ROC-кривых для моделей, обученных на наборе TBX11K

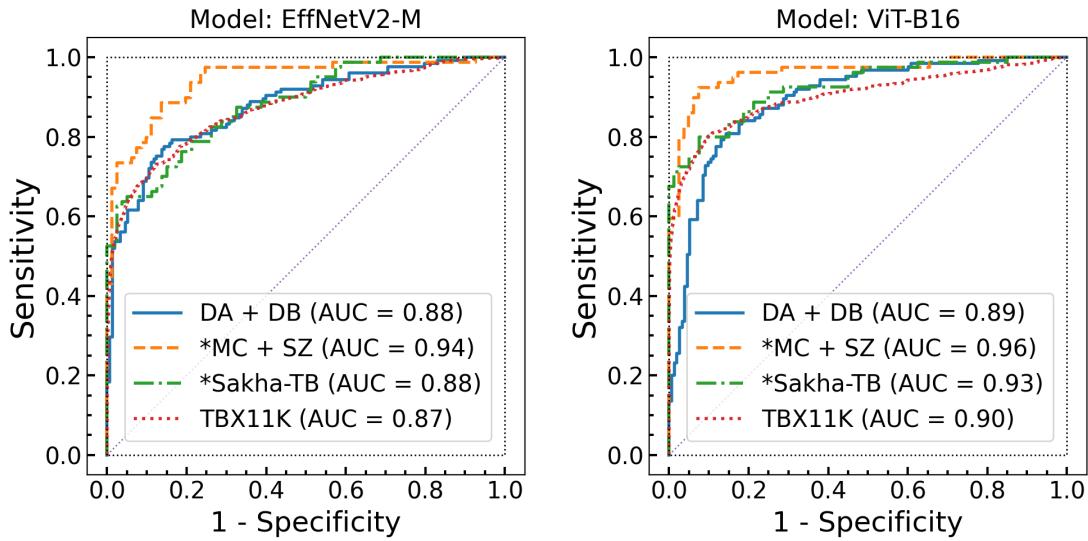


Рисунок 4.8: Графики ROC-кривых для моделей, обученных на наборе $(MC + SZ) + Sakha-TB$

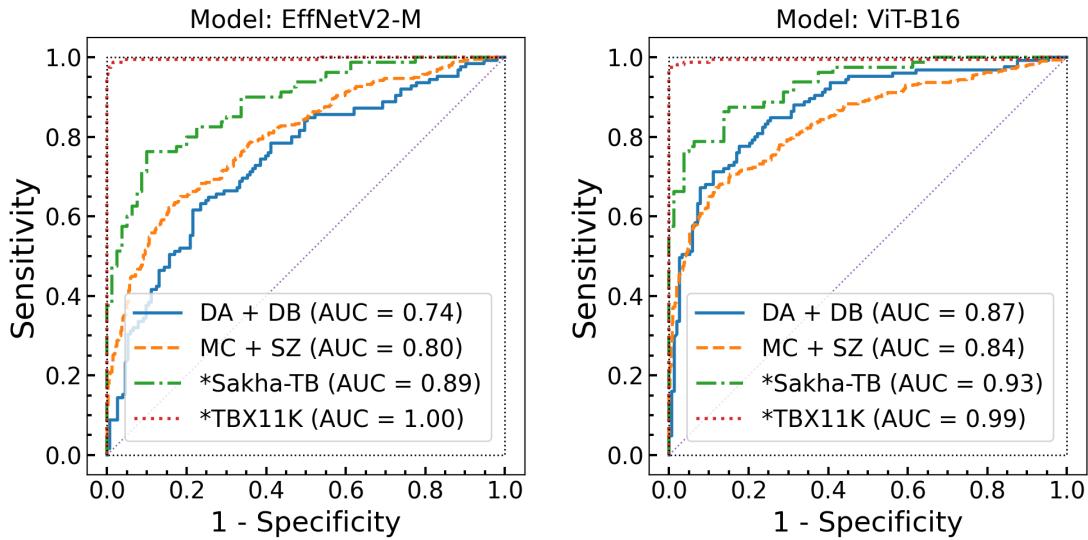


Рисунок 4.9: Графики ROC-кривых для моделей, обученных на наборе $TBX11K + Sakha-TB$

Из графиков видно, что добавление набора *Sakha-TB* в обучающую выборку повышает качество работы алгоритма диагностики в том числе на других наборах. При этом в случае использования представленного набора в качестве основы обучающей выборки не обязательно требуется большой набор изображений: комбинация его с набором *MC + SZ* позволила достичь высокого качества диагностики. Таким образом, представляется полезным включение набора *Sakha-TB* в обучающую выборку нейросетевых моделей компьютерной диагностики туберкулёза для повышения стабильности их работы на практике.

Заключение

1. Разработан итерационный регуляризирующий алгоритм повышения разрешения и резкости изображений флуоресцентной мигающей микроскопии.
2. Найдены оптимальные функции смещения для деформационного метода повышения резкости изображений для трёх видов ядер размытия, возникающих на практике. Предложен малопараметрический вариант алгоритма.
3. Разработан и применён для повышения качества диагностики туберкулёза лёгких нейросетевой метод контроля качества рентгеновских снимков грудной клетки, основанный на автоматическом анализе жёсткости рентгенограммы.
4. Реализован программный комплекс, состоящий из модулей повышения разрешения изображений флуоресцентной мигающей микроскопии, повышения резкости медицинских изображений, определения качества рентгенограмм грудной клетки для задачи диагностики туберкулёза лёгких и компьютерной диагностики туберкулёза лёгких.

Список литературы

1. *Мишин А., Лукъянов К.* Флуоресцентная микроскопия сверхвысокого разрешения живых клеток // Успехи биологической химии. — 2019. — Т. 59. — С. 39—66.
2. Imaging intracellular fluorescent proteins at nanometer resolution / E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, H. F. Hess // Science. — 2006. — Т. 313, № 5793. — С. 1642—1645.
3. *Rust M. J., Bates M., Zhuang X.* Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM) // Nature methods. — 2006. — Т. 3, № 10. — С. 793—796.
4. Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI) / T. Dertinger, R. Colyer, G. Iyer, S. Weiss, J. Enderlein // Proceedings of the National Academy of Sciences. — 2009. — Т. 106, № 52. — С. 22287—22292.
5. Achieving increased resolution and more pixels with Superresolution Optical Fluctuation Imaging (SOFI) / T. Dertinger, R. Colyer, R. Vogel, J. Enderlein, S. Weiss // Optics express. — 2010. — Т. 18, № 18. — С. 18875—18885.
6. *Mendel J.* Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications // Proceedings of the IEEE. — 1991. — Т. 79, № 3. — С. 278—305.
7. *Малахов А. Н.* Кумулянтный анализ случайных негауссовых процессов и их преобразований. — М.: Советское радио, 1978. — 376 с.
8. Sparsity-based super-resolution microscopy from correlation information / O. Solomon, M. Mutzafi, M. Segev, Y. C. Eldar // Opt. Express. — 2018. — Т. 26, № 14. — С. 18238—18269.
9. *Agarwal K., Macháň R.* Multiple signal classification algorithm for super-resolution fluorescence microscopy // Nature communications. — 2016. — Т. 7, № 1. — С. 13752.
10. *Schmidt R.* Multiple emitter location and signal parameter estimation // IEEE transactions on antennas and propagation. — 1986. — Т. 34, № 3. — С. 276—280.
11. Bayesian localization microscopy reveals nanoscale podosome dynamics / S. Cox, E. Rosten, J. Monypenny, T. Jovanovic-Talisman, D. T. Burnette, J. Lippincott-Schwartz, G. E. Jones, R. Heintzmann // Nature methods. — 2012. — Т. 9, № 2. — С. 195—200.
12. Deep-STORM: super-resolution single-molecule microscopy by deep learning / E. Nehme, L. E. Weiss, T. Michaeli, Y. Shechtman // Optica. — 2018. — Т. 5, № 4. — С. 458—464.

13. DLBI: deep learning guided Bayesian inference for structure reconstruction of super-resolution fluorescence microscopy / Y. Li, F. Xu, F. Zhang, P. Xu, M. Zhang, M. Fan, L. Li, X. Gao, R. Han // Bioinformatics. — 2018. — T. 34, № 13. — C. i284—i294.
14. *Girsault A., Lukes T., Sharipov A., Geissbuehler S., Leutenegger M., Vandenberg W., Dedecker P., Hofkens J., Lasser T.* SOFI Simulation Tool: A Software Package for Simulating and Testing Super-Resolution Optical Fluctuation Imaging // PLOS ONE. — 2016. — T. 11, № 9. — C. 1—13.
15. *Weisshart K.* The basic principle of Airyscanning // Zeiss Technology Note. — 2014. — T. 22.
16. A spontaneously blinking fluorophore based on intramolecular spirocyclization for live-cell super-resolution imaging / S.-n. Uno, M. Kamiya, T. Yoshihara, K. Sugawara, K. Okabe, M. C. Tarhan, H. Fujita, T. Funatsu, Y. Okada, S. Tobita [и др.] // Nature chemistry. — 2014. — T. 6, № 8. — C. 681—689.
17. *Sheppard C. J., Mehta S. B., Heintzmann R.* Superresolution by image scanning microscopy using pixel reassignment // Optics letters. — 2013. — T. 38, № 15. — C. 2889—2892.
18. *Gadsden M.* Some statistical properties of pulses from photomultipliers // Applied Optics. — 1965. — T. 4, № 11. — C. 1446—1452.
19. *Bishop C. M.* Pattern recognition and machine learning. — New York: Springer, 2006. — 738 c.
20. *Richardson W. H.* Bayesian-based iterative method of image restoration // JoSA. — 1972. — T. 62, № 1. — C. 55—59.
21. *Krylov A., Nasonova A., Nasonov A.* Gridwarping for image sharpening using one-dimensional approach // 2014 12th International Conference on Signal Processing (ICSP). — IEEE. 2014. — C. 672—677.
22. *Nasonova A., Krylov A.* Deblurred images post-processing by Poisson warping // IEEE Signal Processing Letters. — 2014. — T. 22, № 4. — C. 417—420.
23. *Gusev A., Nasonov A., Krylov A.* Parallel implementation of image sharpening method using grid warping // Proceedings of the 26th International Conference on Computer Graphics and Vision, GraphiCon 2016. — 2016. — C. 294—297.
24. Image database TID2013: Peculiarities, results and perspectives / N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti [и др.] // Signal processing: Image communication. — 2015. — T. 30. — C. 57—77.
25. *Simpkins J., Stevenson R. L.* Parameterized modeling of spatially varying optical blur // Journal of Electronic Imaging. — 2014. — T. 23, № 1. — C. 013005—013005.
26. *Simpkins J.* Modeling and Estimation of Spatially-Varying Point-Spread Functions due to Lens Aberrations and Defocus. — 2011. — 131 c.

27. TID2008 - a database for evaluation of full-reference visual quality assessment metrics / N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti // Advances of modern radioelectronics. — 2009. — Т. 10, № 4. — С. 30—45.
28. Canny J. A computational approach to edge detection // IEEE Transactions on pattern analysis and machine intelligence. — 1986. — № 6. — С. 679—698.
29. Nelder J. A., Mead R. A Simplex Method for Function Minimization // The Computer Journal. — 1965. — Т. 7, № 4. — С. 308—313.
30. Клинические рекомендации «Туберкулез у взрослых» / И. А. Васильева, Г. С. Баласанянц, С. Е. Борисов, И. Бурмистрова, Р. Валиев, Е. Кульчавеня [и др.]. — 2022. — 151 с. — URL: https://cr.minzdrav.gov.ru/schema/16_2 (дата обр. 08.08.2023).
31. Применение искусственного интеллекта в медицине: достижения и перспективы. Обзор литературы. Часть 1 / В. А. Бердутин, О. П. Абаева, Т. Е. Романова, С. В. Романов // Социология медицины. — 2023. — Т. 21, № 1. — С. 83—96.
32. Невинная И. Михаил Мурашко: Началось широкое внедрение в практику программ с искусственным интеллектом // Российская газета. — 2023. — 16 июля. — URL: <https://rg.ru/2023/07/16/mihail-murashko-nachalos-shirokoe-vnedrenie-v-praktiku-programm-s-iskusstvennym-intellektom.html> (дата обр. 08.08.2023).
33. Живым в помощь: как внедряются нейросети в российской медицине // РБК. — 2023. — 7 авг. — URL: <https://www.rbc.ru/economics/07/08/2023/64cb69909a79477320ae5ebe> (дата обр. 08.08.2023).
34. Программы с искусственным интеллектом внедряются в систему здравоохранения, они облегчают работу медиков - Мурашко // Интерфакс-Россия. — 2023. — 8 авг. — URL: <https://www.interfax-russia.ru/volga/news/programmy-s-iskusstvennym-intellektom-vnedryayutsya-v-sistemu-zdravoohraneniya-oni-oblegchayut-rabotu-medikov-murashko> (дата обр. 08.08.2023).
35. Adversarial attacks on medical machine learning / S. G. Finlayson, J. D. Bowers, J. Ito, J. L. Zittrain, A. L. Beam, I. S. Kohane // Science. — 2019. — Т. 363, № 6433. — С. 1287—1289.
36. Automating chest radiograph imaging quality control / K. Nousiainen, T. Mäkelä, A. Piilonen, J. I. Peltonen // Physica Medica. — 2021. — Т. 83. — С. 138—145.
37. Robust chest X-ray quality assessment using convolutional neural networks and atlas regularization / J. von Berg, S. Krönke, A. Gooßen, D. Bystrov, M. Brück, T. Harder, N. Wieberneit, S. Young // Medical Imaging 2020: Image Processing. Т. 11313. — SPIE. 2020. — С. 391—398.

38. Application value of convolutional neural network in quality control of direct digital chest X-ray images / X.-Q. JIA, X.-l. ZHANG, Z. LIU, Q. ZENG, Z.-f. ZHAO, Y.-s. LI, H. HUANG, A.-m. DONG, J. YANG, J.-x. GUO // Journal of Xi'an Jiaotong University (Medical Sciences). — 2019. — № 5. — C. 784—788.
39. Validating deep learning inference during chest X-ray classification for COVID-19 screening / R. Sadre, B. Sundaram, S. Majumdar, D. Ushizima // Scientific reports. — 2021. — T. 11, № 1. — C. 16075.
40. The additional value of the lateral chest radiograph for the detection of small pulmonary nodules – a ROC analysis / R. A. Kluthke, R. Kickuth, P. M. Bansmann, C. Tüshaus, S. Adams, D. Liermann, J. Kirchner // The British journal of radiology. — 2016. — T. 89, № 1067. — C. 20160394.
41. Hashir M., Bertrand H., Cohen J. P. Quantifying the value of lateral views in deep learning for chest X-rays // Medical Imaging with Deep Learning. — PMLR. 2020. — C. 288—303.
42. Вишнякова Н. Частота и уровни облучения пациентов и населения России за счет лучевой диагностики с применением источников ионизирующего излучения // Радиационная гигиена. — 2016. — Т. 3, № 3. — С. 17—22.
43. Chuiko G., Tsvetkov V. Effects of X-ray hardness on fluorogram informativeness // Biomedical Engineering. — 1982. — Т. 16, № 4. — С. 117—119.
44. Тимофеева Л. А., Алешина Т. Н., Быкова А. В. Основные рентгенологические синдромы патологии легочной ткани: учеб. пособие. — Чебоксары: Изд-во Чуваш. ун-та, 2013. — 104 с.
45. Сидоров А. И., Шербатых А. А., Покровская Л. Н. Методика анализа рентгенограмм: учебное пособие. — Иркутск: ИГМУ, 2012. — 22 с.
46. Automatic out-of-distribution detection methods for improving the deep learning classification of pulmonary X-ray images / A. Dovganich, A. Khvostikov, Y. Pchelintsev, A. Krylov, Y. Ding, M. Farias // Journal of Image and Graphics. — 2022. — Т. 10, № 2. — С. 56—63.
47. Oloko-Oba M., Viriri S. A systematic review of deep learning techniques for tuberculosis detection from chest radiograph // Frontiers in Medicine. — 2022. — Т. 9. — С. 830515.
48. Zeyu D., Yaakob R., Azman A. A Review of Deep Learning-Based Detection Methods for Tuberculosis // 2022 IEEE International Conference on Computing (ICOCO). — IEEE. 2022. — С. 68—73.
49. Evolution of machine learning in tuberculosis diagnosis: a review of deep learning-based medical applications / M. Singh, G. V. Pujar, S. A. Kumar, M. Bhagyalalitha, H. S. Akshatha, B. Abuhaija, A. R. Alsoud, L. Abualigah, N. M. Beeraka, A. H. Gandomi // Electronics. — 2022. — Т. 11, № 17. — С. 2634.

50. Advances in Deep Learning for Tuberculosis Screening Using Chest X-Rays: The Last 5 Years Review / K. Santosh, S. Allu, S. Rajaraman, S. Antani // Journal of Medical Systems. — 2022. — Т. 46, № 11. — С. 82.
51. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration / S. Candemir, S. Jaeger, K. Palaniappan, J. P. Musco, R. K. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, C. J. McDonald // IEEE transactions on medical imaging. — 2013. — Т. 33, № 2. — С. 577—590.
52. Automatic tuberculosis screening using chest radiographs / S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. K. Singh, S. Antani [и др.] // IEEE transactions on medical imaging. — 2013. — Т. 33, № 2. — С. 233—245.
53. Adaptive histogram equalization and its variations / S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, K. Zuiderweld // Computer vision, graphics, and image processing. — 1987. — Т. 39, № 3. — С. 355—368.
54. Ordinal Regression Methods: Survey and Experimental Study / P. A. Gutiérrez, M. Pérez-Ortiz, J. Sánchez-Monedero, F. Fernández-Navarro, C. Hervás-Martínez // IEEE Transactions on Knowledge and Data Engineering. — 2016. — Т. 28, № 1. — С. 127—146.
55. Winship C., Mare R. D. Regression Models with Ordinal Variables // American Sociological Review. — 1984. — Т. 49, № 4. — С. 512—525.
56. Deep residual learning for image recognition / K. He, X. Zhang, S. Ren, J. Sun // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2016. — С. 770—778.
57. Densely connected convolutional networks / G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2017. — С. 4700—4708.
58. Tan M., Le Q. EfficientNetV2: Smaller models and faster training // International conference on machine learning. — PMLR. 2021. — С. 10096—10106.
59. ImageNet large scale visual recognition challenge / O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein [и др.] // International journal of computer vision. — 2015. — Т. 115. — С. 211—252.
60. Rennie J. D., Srebro N. Loss functions for preference levels: Regression with discrete ordered labels // Proceedings of the IJCAI multidisciplinary workshop on advances in preference handling. Т. 1. — AAAI Press, Menlo Park, CA, 2005.
61. Baccianella S., Esuli A., Sebastiani F. Evaluation measures for ordinal regression // 2009 Ninth international conference on intelligent systems design and applications. — IEEE. 2009. — С. 283—287.

62. The balanced accuracy and its posterior distribution / K. H. Brodersen, C. S. Ong, K. E. Stephan, J. M. Buhmann // 2010 20th international conference on pattern recognition. — IEEE. 2010. — C. 3121—3124.
63. *Loshchilov I., Hutter F.* Decoupled Weight Decay Regularization // International Conference on Learning Representations. — 2019.
64. *Zwillinger D., Kokoska S.* CRC standard probability and statistics tables and formulae. — New York: Chapman & Hall, 2000. — 234 c.
65. Rethinking computer-aided tuberculosis diagnosis / Y. Liu, Y.-H. Wu, Y. Ban, H. Wang, M.-M. Cheng // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. — 2020. — C. 2646—2655.
66. Cross Dataset Analysis of Domain Shift in CXR Lung Region Detection / Z. Xue, F. Yang, S. Rajaraman, G. Zamzmi, S. Antani // Diagnostics. — 2023. — T. 13, № 6. — C. 1068.
67. *Pooch E. H., Ballester P., Barros R. C.* Can we trust deep learning based diagnosis? The impact of domain shift in chest radiograph classification // Thoracic Image Analysis: Second International Workshop, TIA 2020. — Springer International Publishing. 2020. — C. 74—83.
68. Improving anatomical plausibility in medical image segmentation via hybrid graph neural networks: applications to chest X-ray analysis / N. Gaggion, L. Mansilla, C. Mosquera, D. H. Milone, E. Ferrante // IEEE Transactions on Medical Imaging. — 2022.
69. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis / A. J. Larrazabal, N. Nieto, V. Peterson, D. H. Milone, E. Ferrante // Proceedings of the National Academy of Sciences. — 2020. — T. 117, № 23. — C. 12592—12594.
70. *Chauhan A., Chauhan D., Rout C.* Role of Gist and PHOG features in computer-aided diagnosis of tuberculosis without segmentation // PloS one. — 2014. — T. 9, № 11. — e112980.
71. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale / A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby // International Conference on Learning Representations. — 2021.