Казахский национальный университет имени аль-Фараби

УДК 004.8:61(043)

САРСЕМБАЕВА ТАЛШЫН САҒДАТБЕККЫЗЫ

Машинная диагностика медицинских изображений и её адаптация к потребностям популяционного скрининга на примере хронической обструктивной болезни лёгких

«8D06114 –Искусственный интеллект в медицине»

Диссертация на соискание степени
доктора философии (PhD)

Научные консультанты:
Заведующая кафедрой искусственного интеллекта и Big Data,
к.ф.-м.н., профессор
(КазНУ)
Мансурова Мадина Есимхановна,
к.мед.н., ассоциированный профессор,
Ошибаева Айнаш Есимбековна
(МКТУ)
Dr. habil., профессор
Даулетбаев Нурлан,
(Марбургский университет имени Филиппа, Германия)

Республика Казахстан
Алматы, 2026

СОДЕРЖАНИЕ

ОПРЕДЕЛЕНИЯ 4
ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ 6
НОРМАТИВНЫЕ ССЫЛКИ 8
ВВЕДЕНИЕ 9
1 ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИССЛЕДОВАНИЯ 14
1.1 Анализ научной литературы по диагностике ХОБЛ 14
1.1.1 КТ в диагностике ХОБЛ 14
1.1.2 Роль эмфиземы 18
1.1.3 LAA% и порог HU < −950 19
1.1.4 Ограничения традиционных методов и переход к ИИ 22
1.2 Современные методы анализа КТ-изображений 23
1.3 Теоретическая модель исследования 32
2 МЕТОДИКА И ОРГАНИЗАЦИЯ ИССЛЕДОВАНИЯ 39
2.1 Цель и гипотеза исследования 39
2.2 Данные и их подготовка 41
2.3 Подготовка данных и извлечение признаков 44
2.4 Методы машинного обучения 51
2.5 Архитектура модифицированной модели 56
2.6 Этические аспекты исследования 58
3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ 61
3.1 Представление полученных данных 61
3.1.1 Отбор КТ-серий и оценка полноты исходных данных 61
3.1.2 Контроль HU-диапазона и исключение технически проблемных случаев 62
3.1.3 Формирование итоговой аналитической выборки 64
3.1.4 Распределение LAA%-950 и формирование исследовательских классов 64
3.1.5 Связь LAA%-950 со средней плотностью лёгочной ткани 65
3.1.6 Связь LAA%-950 с общим объёмом лёгких 66
3.1.7 Визуальная проверка LAA-карт 67
3.1.8 Итоговая характеристика подготовленного массива данных 68
3.2 Анализ результатов 69
3.2.1 Баланс классов и стратегия формирования обучающих подвыборок 69
3.2.2 Классификация на основе полного пространства эмбеддингов ResNet-152 70
3.2.3 Отбор информативных признаков и снижение размерности 72
3.2.4 Весовые профили отобранных признаков 73
3.2.5 Проверка устойчивости модели на кросс-валидации 75
3.2.6 Предложенная модифицированная модель с объединением признаков ResNet-152 и QCT 75
3.2.7 Анализ ошибок классификации 79
3.2.8 Визуальная интерпретация результатов с использованием Grad-CAM 80
3.2.9 Внешняя валидация на данных открытого доступа NLST 83
3.3 Реализация прототипа системы автоматического анализа КТ-изображений 84
ЗАКЛЮЧЕНИЕ 90

СПИСОК ЛИТЕРАТУРЫ 95
ПРИЛОЖЕНИЕ А- Ссылка программного кода 102
ПРИЛОЖЕНИЕ Б – Список публикаций 103

ОПРЕДЕЛЕНИЯ

В настоящей диссертации применяют следующие термины с соответствующими определениями:

Хроническая обструктивная болезнь лёгких (ХОБЛ) - хроническое заболевание дыхательной системы, характеризующееся персистирующим ограничением воздушного потока, связанным с патологическими изменениями дыхательных путей и/или лёгочной паренхимы.

Эмфизема лёгких - морфологическое проявление поражения лёгочной ткани, связанное с расширением воздушных пространств дистальнее терминальных бронхиол и снижением плотности лёгочной паренхимы на КТ.

КТ-изображение - цифровое медицинское изображение, полученное методом компьютерной томографии и представленное в виде набора срезов, отражающих рентгеновскую плотность тканей.

Единица Хаунсфилда (HU) - количественная шкала плотности тканей на КТ, где воздух имеет значения около −1000 HU, вода - около 0 HU, а костная ткань - положительные значения.

Денситометрический анализ - количественная оценка распределения плотности тканей на КТ-изображении, основанная на анализе значений HU внутри заданной области.

Количественная компьютерная томография (QCT) - метод объективного анализа КТ-изображений, при котором рассчитываются числовые показатели плотности, объёма и структуры лёгочной ткани.

Низкоплотные участки лёгких (LAA) - области лёгочной паренхимы, значения плотности которых ниже заданного порога HU и которые могут отражать эмфизематозно-подобные изменения.

LAA%-950 - процент вокселей лёгочной ткани с плотностью ниже −950 HU от общего числа вокселей внутри маски лёгких; используется как количественный показатель выраженности низкоплотных зон.

Perc15 - 15-й перцентиль распределения плотности лёгочной ткани, то есть значение HU, ниже которого располагается 15% всех вокселей лёгочной маски.

Mean HU - среднее значение плотности лёгочной ткани в единицах Хаунсфилда, рассчитанное внутри сегментированной области лёгких.

TLV - общий объём лёгких, вычисляемый по количеству вокселей внутри лёгочной маски с учётом физического размера вокселя.

Сегментация лёгких - процесс автоматического или полуавтоматического выделения лёгочной ткани на КТ-изображении с исключением костей, мягких тканей, средостения и внешнего воздуха.

Маска лёгких - бинарное изображение, в котором воксели лёгочной ткани отделены от остальных анатомических структур и фона.

Ресемплинг - преобразование КТ-объёма к единому пространственному разрешению, необходимое для сопоставимого анализа данных, полученных с разными параметрами сканирования.

Изотропный воксель - воксель, имеющий одинаковый физический размер по трём пространственным осям.

Автоматическая разметка (auto-labeling) - процедура присвоения меток данным на основе заранее заданных алгоритмических или количественных критериев без ручной экспертной аннотации.

Silver Standard-разметка - суррогатная разметка, сформированная на основе воспроизводимого количественного критерия; в отличие от экспертной клинической разметки, она не является окончательной диагностической истиной.

Эмбеддинг - числовой вектор признаков, формируемый нейронной сетью и отражающий латентное представление входного изображения или его фрагмента.

Экстрактор признаков - модель или её часть, используемая для получения информативных признаков из входных данных без обязательного выполнения финальной классификации внутри этой же модели.

ResNet-152 - глубокая сверточная нейронная сеть с остаточными связями, используемая в исследовании для извлечения латентных визуально-текстурных признаков из КТ-изображений.

Остаточные связи - элементы архитектуры ResNet, обеспечивающие передачу сигнала через несколько слоёв сети и улучшающие обучение глубоких моделей.

Визуально-текстурные признаки - признаки КТ-изображения, отражающие неоднородность лёгочной паренхимы, распределение плотности, локальные контрастные переходы и особенности лёгочного рисунка.

Fusion - объединение признаков различной природы в едином классификационном контуре; в настоящей работе применяется объединение эмбеддингов ResNet-152 и QCT-показателей.

Late Fusion - объединение признаков после их независимого извлечения, например конкатенация нейросетевого вектора признаков и QCT-вектора перед классификацией.

Классификатор - алгоритм машинного обучения, присваивающий объекту один из заданных классов на основе входных признаков.

Grad-CAM - метод визуальной интерпретации нейронной сети, позволяющий определить области изображения, наиболее значимые для формирования классификационного решения.

Популяционный скрининг - массовое обследование целевой группы населения с целью раннего выявления признаков заболевания или группы риска.

Программный pipeline - последовательность автоматизированных этапов обработки данных, включающая загрузку, препроцессинг, извлечение признаков, классификацию и формирование результата.

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

AI - Artificial Intelligence, искусственный интеллект.

AP - Average Precision, средняя точность по precision–recall-кривой.

AUC - Area Under the Curve, площадь под кривой.

bACC - Balanced Accuracy, сбалансированная точность классификации.

CNN - Convolutional Neural Network, сверточная нейронная сеть.

CT / КТ - Computed Tomography, компьютерная томография.

DICOM - Digital Imaging and Communications in Medicine, стандарт хранения и передачи медицинских изображений.

DL - Deep Learning, глубокое обучение.

Edge AI - выполнение алгоритмов искусственного интеллекта на локальных вычислительных устройствах или вблизи места получения данных.

Embedded AI - встроенный искусственный интеллект, интегрированный в программно-аппаратные системы.

FN - False Negative, ложноотрицательный результат.

FP - False Positive, ложноположительный результат.

F1-score - гармоническое среднее между precision и recall.

GOLD - Global Initiative for Chronic Obstructive Lung Disease, международная инициатива по диагностике и лечению ХОБЛ.

Grad-CAM - Gradient-weighted Class Activation Mapping, метод визуализации областей изображения, значимых для решения нейронной сети.

HU - Hounsfield Unit, единица Хаунсфилда, используемая для измерения рентгеновской плотности тканей на КТ.

LAA - Low Attenuation Area, участок лёгочной ткани с низкой плотностью.

LAA%-950 - процент вокселей лёгочной ткани с плотностью ниже −950 HU.

LDCT - Low-Dose Computed Tomography, низкодозовая компьютерная томография.

LIDC-IDRI - Lung Image Database Consortium and Image Database Resource Initiative, открытый датасет КТ-изображений лёгких.

Mean HU - среднее значение плотности ткани в единицах Хаунсфилда.

MIL - Multiple Instance Learning, обучение по множественным экземплярам.

ML - Machine Learning, машинное обучение.

PACS - Picture Archiving and Communication System, система хранения и передачи медицинских изображений.

Perc15 - 15-й перцентиль распределения плотности лёгочной ткани на КТ.

PR-AUC - площадь под precision–recall-кривой.

QCT - Quantitative Computed Tomography, количественный анализ компьютерной томографии.

ROC-AUC - площадь под ROC-кривой, интегральная метрика качества бинарной классификации.

ResNet-152 - глубокая остаточная сверточная нейронная сеть, включающая 152 слоя.

ROI - Region of Interest, область интереса на медицинском изображении.

Silver Standard - суррогатный стандарт разметки, сформированный на основе воспроизводимого алгоритмического или количественного критерия, но не равный экспертной клинической разметке.

SMOTE - Synthetic Minority Over-sampling Technique, метод синтетического увеличения объектов меньшинства.

TTA - Test-Time Augmentation, усреднение предсказаний модели при нескольких преобразованиях изображения на этапе инференса.

TLV - Total Lung Volume, общий объём лёгких.

TN - True Negative, истинно отрицательный результат.

TP - True Positive, истинно положительный результат.

U-Net - архитектура сверточной нейронной сети, широко применяемая для сегментации медицинских изображений.

Voxel-based quantification - количественный анализ на уровне вокселей.

ХОБЛ - хроническая обструктивная болезнь лёгких.

Эмбеддинг - числовой вектор признаков, извлечённый нейронной сетью из изображения или его фрагмента.

Auto-labeling - автоматическое присвоение меток данным на основе заданных алгоритмических критериев.

Feature extractor - модуль или модель, используемые для извлечения признаков из входных данных.

Fusion - объединение признаков различной природы в едином классификационном контуре. В данной работе используется объединение эмбеддингов ResNet-152 и QCT-показателей.

Late Fusion - объединение признаков после их независимого извлечения, например конкатенация нейросетевого embedding-вектора и QCT-вектора перед классификацией.

НОРМАТИВНЫЕ ССЫЛКИ

Хельсинкская декларация Всемирной медицинской ассоциации: Этические принципы медицинских исследований с участием человека в качестве субъекта исследования. - Принята 18-ой Генеральной Ассамблеей ВМА, Хельсинки, Финляндия, июнь 1964. - (Последняя редакция: Форталеза, Бразилия, октябрь 2013).

Регламент (ЕС) 2024/1689 Европейского Парламента и Совета от 13 июня 2024 года, устанавливающий гармонизированные правила в области искусственного интеллекта

Закон Республики Казахстан «О персональных данных и их защите» от 21 мая 2013 года № 94-V.

Закон РК «Об искусственном интеллекте» (от 17 ноября 2025 года № 230-VIII) - создает правовую основу для систем ИИ в Казахстане.

Кодекс РК «О здоровье народа и системе здравоохранения» (2020 г.) - глава 7 «Цифровое здравоохранение» регулирует использование IT-технологий в медицине.

Закон РК «Об информатизации» (2015 г.) - устанавливает правила работы с информационными системами в стране.

ВВЕДЕНИЕ

Хроническая обструктивная болезнь лёгких является одной из наиболее значимых медико-социальных проблем современной пульмонологии, поскольку характеризуется высокой распространённостью, длительным прогрессирующим течением, значительным влиянием на трудоспособность и качество жизни пациентов, а также существенным вкладом в структуру хронической заболеваемости и смертности. Особенность ХОБЛ состоит в том, что заболевание длительное время может протекать с малоспецифическими симптомами, тогда как структурные изменения лёгочной ткани и дыхательных путей формируются задолго до выраженного клинического ухудшения. В связи с этим раннее выявление признаков заболевания имеет принципиальное значение для своевременного наблюдения, коррекции факторов риска и предотвращения дальнейшего прогрессирования патологического процесса.

Важно подчеркнуть, что ХОБЛ не должна рассматриваться исключительно как заболевание пожилых и курильщиков. Хотя табакокурение остаётся одним из ведущих факторов риска, развитие хронической бронхолёгочной патологии возможно и у лиц без активного курительного анамнеза. Существенную роль могут играть пассивное курение, загрязнение атмосферного воздуха, профессиональный контакт с пылью и химическими аэрозолями, воздействие продуктов сгорания биомассы, перенесённые респираторные инфекции, сопутствующая бронхиальная астма, генетическая предрасположенность и неблагоприятные факторы раннего периода жизни. Кроме того, морфологические изменения лёгочной ткани могут формироваться до появления выраженной клинической симптоматики и до достижения пожилого возраста. В связи с этим актуальность автоматизированного анализа КТ-изображений выходит за рамки диагностики уже сформированной тяжёлой ХОБЛ и связана с необходимостью раннего выявления структурных признаков поражения лёгочной паренхимы у более широкой группы населения, включая лиц с недооценённым или нетипичным профилем риска.

Республика Казахстан несёт значительное бремя ХОБЛ: расчётная численность поражённых лиц достигает 1,4 миллиона человек, а исследование, проведённое в Жамбылской области, выявило высокую распространённость респираторных симптомов и подтвердило табакокурение как ведущий модифицируемый фактор риска. Эпидемиологический анализ смертности от болезней органов дыхания за период 2011–2021 гг. показал, что предотвратимая летальность в Казахстане более чем удвоилась, достигнув 156,2 на 100 000 населения, при среднегодовом темпе прироста 7,5%, причём для женщин этот показатель оказался выше, чем для мужчин (9,2% против 6,6%) [1]. В этом контексте особую значимость приобретает проект Казахстанской академии превентивной медицины — кросс-секционное исследование распространённости ХОБЛ среди курильщиков, экс-курильщиков и никогда не куривших жителей Алматы, в дизайне которого предусмотрено сочетанное применение спирометрии и высокоразрешающей компьютерной томографии [2].

Данный проект подтверждает востребованность методов количественной оценки КТ-изображений для популяционного анализа ХОБЛ в Казахстане и служит дополнительным обоснованием целесообразности разработки автоматизированных систем выявления эмфизематозно-подобных изменений по данным компьютерной томографии.

Классическим функциональным методом диагностики ХОБЛ остаётся спирометрия, позволяющая оценить наличие стойкого ограничения воздушного потока. Однако в условиях популяционного скрининга данный подход имеет ряд ограничений. Во-первых, выполнение спирометрии требует стандартизированной техники, участия обученного персонала и достаточной кооперации пациента. Во-вторых, функциональные нарушения не всегда отражают ранние морфологические изменения лёгочной ткани. В-третьих, спирометрия не позволяет детально оценить структурный фенотип поражения, включая выраженность и распределение эмфиземы. Поэтому компьютерная томография лёгких рассматривается как важный источник морфологической информации, позволяющий выявлять низкоплотные участки лёгочной паренхимы, оценивать признаки эмфизематозной деструкции и количественно характеризовать плотностные изменения ткани.

Особое значение в автоматизированном анализе КТ имеет количественная компьютерная томография. QCT-показатели, такие как LAA%-950, Perc15, Mean HU и общий объём лёгких, позволяют формализовать изменения плотности лёгочной ткани и перейти от субъективной визуальной оценки к воспроизводимым числовым признакам. Метрика LAA%-950, отражающая долю вокселей лёгочной паренхимы с плотностью ниже −950 HU, широко используется как денситометрический показатель низкоплотных зон, ассоциированных с эмфиземой. Вместе с тем изолированная пороговая денситометрия имеет ограничения: она чувствительна к параметрам сканирования, уровню вдоха, реконструкционному ядру, шуму изображения и качеству сегментации лёгких. Кроме того, QCT-показатели преимущественно описывают плотностное распределение, но не полностью отражают пространственную организацию и визуально-текстурный паттерн поражения.

Развитие методов машинного обучения и глубоких нейронных сетей создало новые возможности для анализа медицинских изображений. Сверточные нейронные сети способны извлекать из КТ-изображений латентные признаки, отражающие сложные визуальные закономерности, которые не сводятся к одному HU-порогу. В задачах анализа лёгких применяются различные архитектуры: U-Net и её модификации используются преимущественно для сегментации анатомических структур, тогда как ResNet-подобные сети применяются для классификации и извлечения высокоуровневых признаков. Однако полностью нейросетевые модели нередко характеризуются недостаточной интерпретируемостью, что ограничивает их применение в медицинской практике. Для задач скрининга особенно важно, чтобы модель не только показывала высокие метрики качества, но и сохраняла связь результата с физически понятными параметрами КТ.

В связи с этим актуальным является разработка модифицированного метода машинной диагностики, объединяющего нейросетевые признаки и количественные денситометрические показатели. В такой постановке ResNet-152 может использоваться как экстрактор латентных визуально-текстурных признаков, а QCT-показатели - как физически интерпретируемые дескрипторы плотности лёгочной ткани. Объединение этих двух уровней информации позволяет повысить устойчивость классификации и одновременно сохранить объяснимость результата. Для популяционного скрининга это имеет особое значение, поскольку автоматизированная система должна быть пригодна для обработки большого числа КТ-исследований, работать воспроизводимо и выдавать результат, понятный для последующей клинической интерпретации.

Степень разработанности проблемы показывает, что существующие исследования развиваются в нескольких направлениях. Первое направление связано с использованием классических QCT-метрик, прежде всего LAA%-950 и перцентильных показателей плотности. Второе направление представлено методами глубокого обучения, включая CNN, ResNet, attention-based MIL и другие архитектуры, ориентированные на выявление ХОБЛ и эмфиземы по КТ-изображениям. Третье направление связано с радиомикой и комбинированными моделями, включающими изображения, количественные признаки и клинические данные. Несмотря на достигнутые результаты, остаются нерешёнными вопросы интерпретируемости нейросетевых признаков, работы с датасетами без специализированной COPD-разметки, автоматической генерации меток и адаптации моделей к масштабным скрининговым сценариям. Эти обстоятельства определили выбор темы настоящего исследования.

Объектом исследования является процесс автоматического выявления признаков эмфиземы по данным КТ-изображений.

Предметом исследования является модификация метода автоматического выявления признаков хронической обструктивной болезни лёгких (ХОБЛ) по данным КТ-изображений.

Цель исследования - разработка системы автоматического выявления признаков хронической обструктивной болезни лёгких по данным КТ-изображений с помощью модифицированного метода машинной диагностики, ориентированного на массовый популяционный скрининг.

Для достижения поставленной цели были сформулированы следующие задачи:

Провести анализ латентного слоя предобученной ResNet-152 для получения сокращённого подмножества признаков при сохранении диагностической эффективности.

Реализовать стратегию автоматической разметки данных на основе количественного анализа КТ с использованием QCT-показателей.

Разработать модифицированный метод, объединяющий эмбеддинги ResNet-152 и автоматическую аннотацию по QCT, для выявления признаков эмфиземы лёгких по КТ-изображениям в условиях масштабируемого скринингового сценария.

Разработать прототип системы автоматического выявления признаков эмфиземы по данным КТ-изображений на основе предложенного модифицированного метода.

Научная новизна исследования- Предложен метод диагностики эмфиземы при ХОБЛ по объединенным данным компьютерной томографии, основанный на совместном использовании эмбеддингов, извлекаемых предобученной моделью ResNet-152, и количественных показателей QCT (LAA%) для снижения доли диагностических ошибок..

Положения, выносимые на защиту. Использование представленного метода, объединяющего признаки QCT и эмбеддинги ResNet-152, по сравнению с использованием только количественных QCT (AUC 0.732) или только признаков из эмбеддинга (AUC 0.766), позволило существенно повысить качество классификации (до ROC-AUC ≈ 0.996 и PR-AUC ≈ 0.962).

Практическая значимость исследования состоит в возможности использования предложенного подхода как основы для автоматизированного предварительного анализа КТ-изображений в задачах выявления эмфизематозно-подобных изменений лёгочной паренхимы. Разработанный pipeline позволяет обрабатывать КТ-исследования без ручной разметки, рассчитывать количественные денситометрические показатели, извлекать нейросетевые признаки и формировать структурированный результат. В перспективе такой подход может быть использован в системах поддержки принятия решений, скрининговых программах, локальных медицинских информационных системах и Edge AI-сценариях, где требуется быстрая и воспроизводимая обработка большого числа исследований.

Методы исследования включали методы обработки медицинских изображений, количественного анализа КТ, машинного обучения и статистической оценки качества классификации. На этапе препроцессинга применялись перевод DICOM-изображений в шкалу Хаунсфилда, ресемплинг, сегментация лёгочной ткани, маскирование области анализа и расчёт QCT-показателей. Для формирования признакового пространства использовалась предобученная ResNet-152 как экстрактор эмбеддингов. Для отбора информативных признаков применялись статистические и регуляризационные методы, включая t-test/FDR, L1-регуляризацию и RFE. Качество классификации оценивалось по ROC-AUC, PR-AUC, F1-score, balanced accuracy, чувствительности, специфичности и матрицам ошибок. Для визуальной интерпретации нейросетевого компонента использовался Grad-CAM.

Структура диссертации соответствует логике поставленных задач. В первой главе рассмотрены теоретические основы диагностики ХОБЛ, роль КТ в выявлении эмфиземы, QCT-показатели и современные подходы к машинному анализу медицинских изображений. Во второй главе описаны данные, процедура их подготовки, метод автоматической разметки, препроцессинг КТ-изображений, извлечение признаков, архитектура предложенной модели и этические аспекты исследования. В третьей главе представлены результаты экспериментов: характеристика подготовленного массива данных, анализ QCT-

разметки, оценка эмбеддингов ResNet-152, сравнение вариантов модели, интерпретация Grad-CAM, реализация программного прототипа, сопоставление с предыдущими исследованиями, ограничения и перспективы дальнейшего развития.

Настоящее исследование направлено на решение актуальной задачи автоматизированного выявления КТ-признаков эмфизематозных изменений, ассоциированных с ХОБЛ. Предложенный подход сочетает преимущества глубокого обучения и количественной компьютерной томографии, обеспечивая как повышение качества классификации, так и сохранение физической интерпретируемости результата.

Таким образом, **предложенная модель ResNet-152 + QCT** может рассматриваться как промежуточный этап между экспериментальной нейросетевой диагностикой и практическим Edge AI-решением. Она объединяет способность CNN выявлять сложные визуально-текстурные паттерны КТ с объективностью денситометрических показателей, что повышает как диагностическую ценность, так и потенциальную пригодность модели для масштабируемого популяционного скрининга.

Нормативно-правовую и этическую основу исследования составили международные акты и законодательство Республики Казахстан в области медицины, цифровизации и искусственного интеллекта (ИИ).

В части обеспечения этических стандартов и защиты прав субъектов данных исследование опирается на положения Хельсинкской декларации Всемирной медицинской ассоциации (1964 г., с изменениями от 2013 г.), определяющей приоритет интересов человека над научными интересами.

Правовой режим разработки и внедрения алгоритмов ИИ в работе рассматривается через призму Закона РК «Об искусственном интеллекте» (2025 г.) и Регламента (ЕС) 2024/1689 (EU AI Act), задающих гармонизированные требования к безопасности и прозрачности высокорисковых систем.

Вопросы обработки биомедицинской информации и обеспечения конфиденциальности регулируются Кодексом РК «О здоровье народа и системе здравоохранения» (в частности, нормами главы 7 о цифровом здравоохранении), а также профильными законами: Законом РК «О персональных данных и их защите» и Законом РК «Об информатизации», устанавливающими порядок работы с государственными и частными информационными системами.

1 ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИССЛЕДОВАНИЯ

1.1 Анализ научной литературы по диагностике ХОБЛ

1.1.1 КТ в диагностике ХОБЛ

Компьютерная томография при ХОБЛ занимает особое место среди методов диагностики, поскольку позволяет перейти от функционального подтверждения бронхиальной обструкции к морфологической и фенотипической характеристике заболевания. Спирометрия остаётся базовым клиническим стандартом диагностики ХОБЛ, так как фиксирует снижение отношения объёма форсированного выдоха за 1 секунду к форсированной жизненной ёмкости лёгких (FEV1/FVC) менее 0,70 и позволяет классифицировать тяжесть обструкции по критериям Глобальной инициативы по хронической обструктивной болезни лёгких (GOLD) [3, 4, 6, 12, 15]. Однако её диагностический потенциал ограничен тем, что она отражает преимущественно суммарную функцию лёгких и не позволяет дифференцировать морфологические подтипы заболевания, включая эмфизему, поражение малых дыхательных путей, бронхиальный компонент, air trapping и сосудистые изменения [6, 14, 15, 17, 24, 27].

В этом контексте КТ-диагностика ХОБЛ рассматривается как ключевой инструмент фенотипирования. В отличие от спирометрии, КТ визуализирует структурные изменения лёгочной паренхимы: эмфизематозную деструкцию, утолщение стенок бронхов, воздушные ловушки, снижение плотности лёгочной ткани, региональную неоднородность поражения и изменения сосудистого русла [3, 5, 6, 8, 24, 27]. Bhaskar и соавторы подчёркивают, что FEV1 недостаточен для описания гетерогенности ХОБЛ, тогда как высокоразрешающая компьютерная томография (HRCT) помогает определить, какой компонент преобладает - поражение дыхательных путей или паренхимы [5]. Kang и соавторы показывают, что визуальная КТ-классификация позволяет выделять нормальный тип, парасептальную эмфизему, бронхиальный тип и разные степени центролобулярной эмфиземы, что отражает патологическую неоднородность ХОБЛ [6].

В соответствии с рисунком 1.1.1-1 представлены карты параметрического ответа (PRM) трёх пациентов с ХОБЛ различных степеней тяжести и компьютерно-томографических подтипов. Нормальная лёгочная ткань отображается зелёным цветом (PRMnormal), функциональное заболевание мелких дыхательных путей - жёлтым (PRMfSAD), эмфизематозная ткань - красным (PRMemphy). Случаи иллюстрируют спектр структурно-функциональных нарушений: от изолированного бронхиального фенотипа с минимальной эмфиземой, но значительной дисфункцией мелких дыхательных путей (A), до комбинированного поражения с выраженной парасептальной эмфиземой (B) и тяжёлого деструктивного эмфизематозного фенотипа (C). Данные изображения наглядно демонстрируют гетерогенность морфофункциональных изменений лёгочной ткани при ХОБЛ и подчёркивают

важность количественного анализа с помощью PRM для точного фенотипирования заболевания.

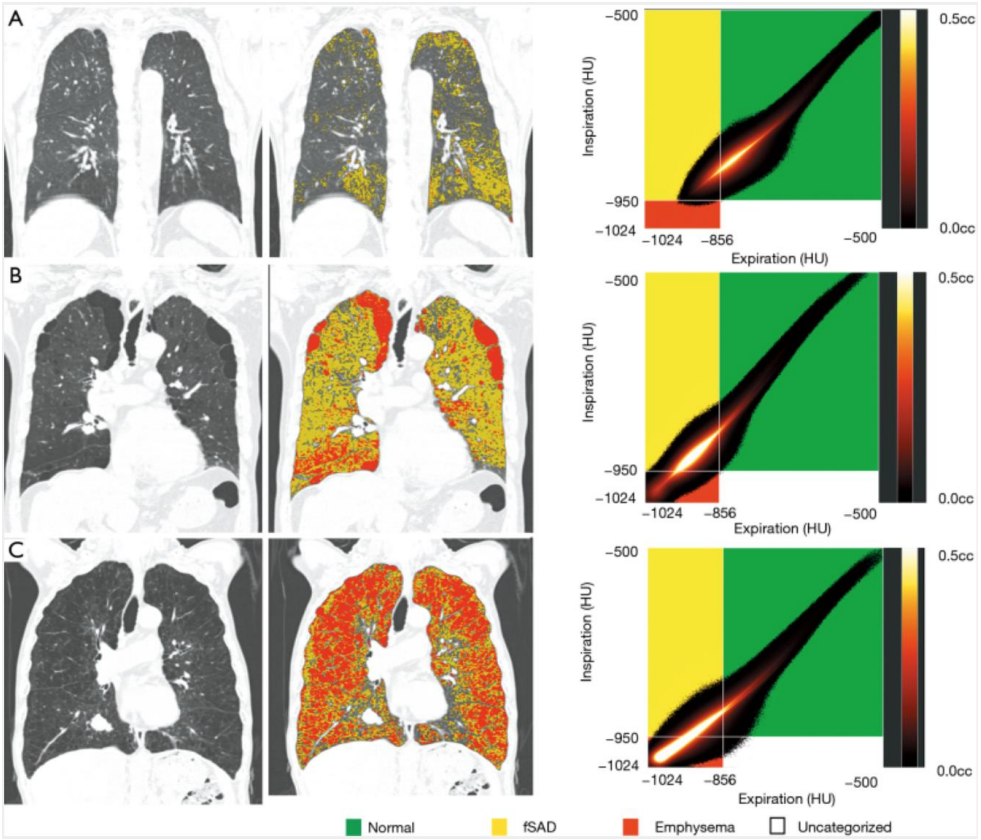


Рисунок 1.1.1-1 - Карты параметрического ответа (PRM) у трёх пациентов

Нормальная лёгочная ткань обозначена зелёным цветом (PRMnormal), функциональное заболевание мелких дыхательных путей - жёлтым (PRMfSAD), эмфизематозная ткань - красным (PRMemphy).

А - мужчина 56 лет, ХОБЛ, GOLD 2. На КТ определяется выраженное утолщение стенки бронхов без признаков эмфиземы; КТ-подтип - бронхиальный. PRMemphy = 0,7 %, PRMfSAD = 19,2 %, Pi-10 = 5,5 мм.

В - мужчина 78 лет, ХОБЛ, GOLD 4. На КТ визуализируется парасептальная эмфизема в верхних отделах обоих лёгких; КТ-подтип - парасептальная эмфизема. PRMemphy = 18,2 %, PRMfSAD = 50,4 %.

С - мужчина 78 лет, ХОБЛ, GOLD 2. На КТ определяются диффузные дольковые просветления с гиперинфляцией обоих лёгких; КТ-подтип - продвинутая деструктивная эмфизема. PRMemphy = 38,5 %, PRMfSAD = 30,0 %.

PRM - parametric response mapping; КТ - компьютерная томография; ХОБЛ - хроническая обструктивная болезнь лёгких; GOLD - Global Initiative for Chronic Obstructive Lung Disease [6].

Количественный анализ КТ при ХОБЛ усиливает диагностическую ценность метода, поскольку переводит визуальные признаки в воспроизводимые числовые параметры. Одним из основных таких параметров является доля низкоплотных участков (LAA%, порог – 950 HU), наглядно представленная на

Рисунке 1.1.1-2 в виде синих зон. К другим количественным маркерам относят Perc15, среднюю плотность лёгких, отношение плотности на выдохе и вдохе (E/I-ratio MLD), Pi10, показатели воздушной ловушки и параметрическое ответное картирование (parametric response mapping) [6, 8, 18, 24, 25].

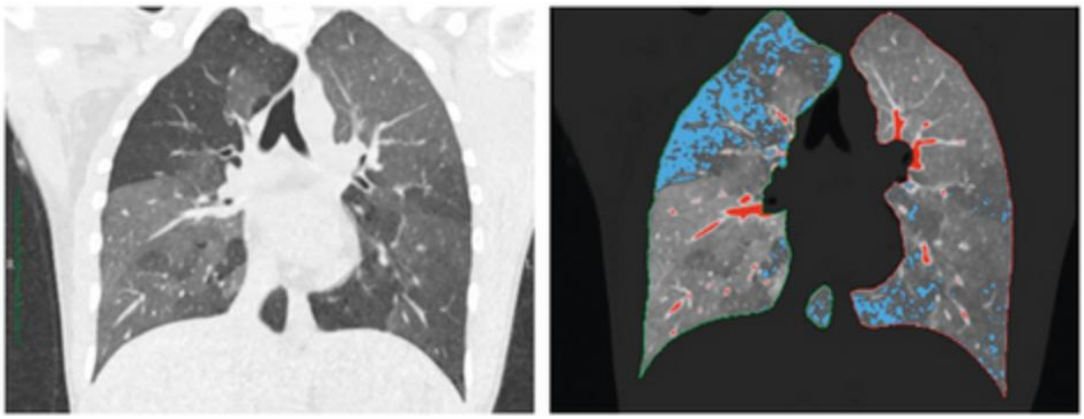


Рисунок 1.1.1-2 - Автоматическое выделение границ лёгких и количественная оценка лёгочной паренхимы

Синие зоны соответствуют объёму низкой плотности (LAV) со значениями ниже − 950 HU (эмфизематозная ткань); красные зоны - объёму высокой плотности (HAV) со значениями выше −200 HU. КТ - компьютерная томография; HU - единицы Хаунсфилда[8].

Gawlitza и соавторы показали, что экспираторная КТ и delta-параметры между вдохом и выдохом улучшают корреляцию КТ-признаков с функциональными показателями лёгких [8]. Oh и соавторы также подчёркивают, что количественная компьютерная томография (QCT) позволяет объективно оценивать эмфизему, воздушные ловушки и утолщение стенок дыхательных путей, а визуальная оценка ограничена субъективностью, вариабельностью между экспертами и зависимостью от настроек просмотра [24].

Для преодоления ограничений однофазовой денситометрии был разработан метод параметрического ответного картирования (PRM), который использует парные инспираторную и экспираторную КТ для дифференциации типов поражения лёгочной ткани. В рамках PRM каждый воксель классифицируется на основе изменения плотности между фазами дыхания: нормальная паренхима (плотность на вдохе ≥ –950 HU и на выдохе ≥ –856 HU), функциональное поражение малых дыхательных путей (fSAD; плотность на вдохе ≥ –950 HU и на выдохе < –856 HU) и эмфизема (плотность на вдохе < –950 HU и на выдохе < –856 HU). Это позволяет получить трёхмерную карту распределения патологических процессов с высокой пространственной точностью [6, 24].

Современная литература постепенно расширяет роль КТ от подтверждения эмфиземы к построению интегральной модели диагноза. В исследовании COPDGene 2019 предложено учитывать четыре компонента: экспозицию, симптомы, КТ-признаки структурного поражения и спирометрические

нарушения [18]. Такой подход важен для популяционного скрининга, поскольку часть пациентов имеет структурные изменения на КТ ещё до выраженного спирометрического нарушения. В перспективе именно сочетание КТ-фенотипов, количественных биомаркеров и алгоритмов машинного анализа может стать основой более раннего выявления ХОБЛ.

Новые работы усиливают этот вывод. Sørensen и соавторы показали, что текстурный анализ КТ может превосходить классические денситометрические показатели: их texture-based measure достиг AUC 0,713, тогда как лучшие плотностные показатели имели AUC около 0,598 [21]. Li и соавторы развивают это направление через unsupervised 3D deep learning: на данных 541 бывшего курильщика и 59 здоровых некурящих из SPIROMICS они использовали 3D convolutional autoencoder-feature constructor с двумя каналами - CT intensity и Jacobian value - и выделили 80 тканевых pattern-clusters и 7 латентных факторов [25]. Эти факторы предсказывали спирометрические показатели, а две факторные компоненты были связаны с подтипами эмфиземы, PRM-метриками, вариантами дыхательных путей и отношением airway tree к объёму лёгких [25]. Это показывает, что признаки, извлекаемые глубокими моделями из КТ, могут описывать скрытую морфологическую структуру ХОБЛ шире, чем заранее заданные показатели LAA%, MLD или Pi10.

Отдельно следует отметить работы, расширяющие КТ-анализ за пределы паренхимы и эмфиземы. Zhao и соавторы предложили подход на основе обучения с частичной разметкой (semi-supervised teacher–student framework) для сегментации лёгочных сосудов на КТ у пациентов с ХОБЛ [27]. В исследовании использованы бесконтрастные КТ-сканы 125 пациентов; модель Semi2 повысила точность (precision) сегментации сосудов на 2,3 % и достигла precision 90,3 % [27]. Это важно, поскольку сосудистые признаки - общий объём крови, BV5, BV5/TBV, число сегментов, количество ветвей и длина сосудистого дерева - могут использоваться для оценки микрососудистых и перфузионных изменений при ХОБЛ [27]. Следовательно, современные КТ-биомаркеры ХОБЛ включают не только низкоплотные зоны и бронхиальные стенки, но и сосудистую архитектонику.

В контексте популяционного скрининга также значимы работы, которые не ограничиваются КТ, но показывают необходимость мультимодальной стратификации риска. Liu и соавторы на данных UK Biobank разработали ML-модели для прогноза early-onset COPD у лиц 20–50 лет; лучшая модель достигла AUC 0,8505, specificity 0,8539 и sensitivity 0,7584 при включении персональных данных и биохимических показателей крови [22]. Lee и соавторы также использовали UK Biobank для анализа перехода из группы no-COPD в COPD, показав значимость социодемографических, клинических и генетических факторов, включая возраст, мужской пол, underweight, материальную депривацию, низкую физическую активность, курение и вариант rs2070600 [26]. Эти исследования не заменяют КТ-диагностику, но указывают, что для массового скрининга КТ целесообразно комбинировать с предварительной ML-стратификацией риска.

Таким образом, роль КТ в диагностике ХОБЛ заключается не только в выявлении эмфиземы, но и в более широком фенотипировании заболевания. КТ позволяет описывать радиологические признаки ХОБЛ, количественно оценивать степень поражения, выявлять региональные изменения, строить imaging biomarkers и формировать основу для систем поддержки принятия решений. Новые работы показывают, что будущее КТ-диагностики связано с переходом от отдельных плотностных порогов к комплексному анализу текстуры, латентных признаков, сосудистых параметров и мультимодальных моделей риска [21, 22, 25, 26, 27].

1.1.2 Роль эмфиземы

Эмфизема при ХОБЛ является одним из ключевых морфологических фенотипов заболевания. Она связана с разрушением альвеолярных стенок, расширением воздушных пространств, снижением эластичности лёгочной ткани и нарушением газообмена [3, 4, 14, 23, 24]. Jung и Vij рассматривают прогрессирование ХОБЛ как процесс, при котором воспаление дыхательных путей, повреждение альвеол, воздушная ловушка и повторные обострения приводят к нарастающему снижению функции лёгких и переходу к тяжёлой эмфиземе [3]. Поэтому выявление эмфиземы имеет не только диагностическое, но и прогностическое значение.

КТ-признаки эмфиземы позволяют визуализировать структурный субстрат заболевания раньше, чем он может быть полноценно отражён в спирометрических показателях. Kimura и соавторы подчёркивают, что HRCT может выявлять небольшие эмфизематозные поражения у пациентов без субъективных симптомов или без выраженных отклонений по спирометрии [4]. Это особенно важно для популяционного скрининга, где задача состоит не только в подтверждении уже сформированной ХОБЛ, но и в выявлении ранних морфологических изменений у лиц группы риска.

Эмфизематозный фенотип ХОБЛ связан с клинической гетерогенностью заболевания. Kang и соавторы показывают, что по мере усиления визуальной выраженности центролобулярной эмфиземы ухудшаются показатели функции лёгких, увеличиваются показатели симптомной нагрузки и возрастают количественные КТ-характеристики эмфиземы и air trapping [6]. Kellerer и соавторы подчёркивают, что КТ является наиболее точным методом выявления, количественной оценки и наблюдения эмфиземы, а дифференциация эмфизематозного и airway-dominated фенотипов имеет значение для ведения пациентов [16]. Oh и соавторы дополняют этот тезис, указывая, что %LAA-950 и Perc15 ассоциированы не только с морфологией эмфиземы, но и с клиническими исходами: снижением функции лёгких, ухудшением качества жизни, риском обострений, прогрессированием заболевания и смертностью [24].

При этом новые источники уточняют: эмфизема является центральным, но не единственным ранним маркером ХОБЛ. Fazleen и Wilkinson, Choi и Rhee, а также Yip и соавторы указывают, что поражение малых дыхательных путей может предшествовать выраженной эмфизематозной деструкции [11, 14, 17]. Li и соавторы подтверждают это на уровне латентных КТ-признаков: выделенные

их моделью факторы были связаны не только с подтипами эмфиземы, но и с PRM-показателями функционального поражения малых дыхательных путей [25]. Следовательно, роль эмфиземы в ХОБЛ должна рассматриваться в динамике: на ранних стадиях заболевание может проявляться изменениями малых дыхательных путей и air trapping, тогда как эмфизема отражает более выраженный структурный этап прогрессирования.

Отдельное значение имеет автоматическое выявление эмфиземы. Karthik, Murugan и Goel предлагают автоматизированный подход к сегментации эмфизематозных поражений на КТ с учётом трёх классов эмфиземы: centrilobular emphysema, panlobular emphysema и paraseptal emphysema [23]. Авторы подчёркивают, что CLE преимущественно поражает центральные отделы ацинусов, PLE характеризуется более распространённым разрушением в пределах долек, а PSE локализуется около плевральных поверхностей [23]. В их исследовании автоматическая сегментация сочетается с 3D-визуализацией, что позволяет анализировать не только факт наличия эмфиземы, но и её пространственное распределение, форму и выраженность.

Количественные результаты Karthik и соавторов показывают, что разные типы эмфиземы имеют разные морфометрические профили: CLE характеризовалась наибольшим числом поражений на изображение, но меньшей суммарной burden, тогда как PSE имела меньшее число очагов, но более высокую выраженность поражения [23]. Это поддерживает тезис о необходимости не только бинарного выявления эмфиземы, но и её фенотипической классификации. Для диссертации это важно, поскольку автоматическая диагностика должна учитывать не только наличие низкоплотных зон, но и их распределение, связность, размер, форму и анатомическую локализацию.

Li и соавторы дополнительно показывают, что эмфизематозный фенотип может быть описан не только экспертно заданными признаками, но и скрытыми латентными факторами, автоматически извлекаемыми из КТ [25]. Их подход важен тем, что не требует ручного выбора заранее известных признаков: модель обучается на локальных 3D областях интереса (ROI) и формирует pattern-кластеры, а затем факторный анализ выявляет скрытые компоненты, связанные с функциональными и морфологическими характеристиками ХОБЛ [25]. Это создаёт методологическую основу для перехода от классической фенотипизации эмфиземы к data-driven фенотипизации.

Таким образом, роль эмфиземы в диагностике ХОБЛ состоит в том, что она выступает морфологическим маркером альвеолярной деструкции, радиологическим фенотипом заболевания и объектом количественной оценки. Однако современный анализ эмфиземы должен выходить за пределы визуального описания и включать автоматическую сегментацию, количественные признаки, текстурные показатели, латентные КТ-факторы и методы пространственной интерпретации [21, 23, 24, 25].

1.1.3 LAA% и порог HU < −950

LAA% является одной из наиболее распространённых количественных метрик КТ-оценки эмфиземы. Под LAA% понимают долю низкоплотных

участков лёгочной ткани - Low Attenuation Areas - в общем объёме лёгких. На денситометрическом уровне эта метрика отражает процент вокселей, плотность которых находится ниже заданного порога, чаще всего HU < −950 на инспираторной КТ [4, 6, 8, 18, 24]. Такие низкоплотные участки лёгких интерпретируются как зоны эмфизематозной деструкции, поскольку разрушение альвеолярных стенок и увеличение воздушных пространств приводят к снижению рентгеновской плотности паренхимы.

Порог −950 HU занимает центральное место в количественной КТ-оценке эмфиземы. Kimura и соавторы описывают LAA как низкоабсорбционные зоны, возникающие вследствие разрушения альвеолярного региона, и связывают пороговый анализ с автоматическим выявлением эмфизематозных изменений [4]. Kang и соавторы используют порог < −950 HU на вдохе в рамках parametric response mapping для выделения эмфизематозного компонента [6]. Gawlitza и соавторы также применяют low attenuation volume с порогом −950 HU и указывают, что этот cutoff широко оценивался в предшествующих исследованиях и коррелирует с микроскопической и макроскопической эмфиземой [8].

В COPDGene 2019 структурная КТ-аномалия также определяется через наличие ≥5% эмфиземы, измеренной по LAA ≤ −950 HU на инспираторной КТ [18]. Это подтверждает, что LAA% при пороге −950 HU является не просто техническим показателем, а устойчивым imaging biomarker эмфизематозного поражения. В терминах FDA-NIH BEST Resource, биомаркер должен иметь определённую идентичность, биологическую правдоподобность, метод измерения и единицы измерения [20]. LAA% соответствует этой логике: его источник - КТ, тип - радиологический/визуализационный биомаркер, биологическая основа - эмфизематозная деструкция паренхимы, метод измерения - подсчёт доли вокселей ниже порога плотности, единица - процент объёма лёгких.

Oh и соавторы уточняют физико-морфологическое основание этой метрики: нормальная лёгочная ткань на инспираторной КТ имеет типичную плотность около −850 HU, тогда как эмфизематозная ткань характеризуется снижением плотности примерно в диапазоне от −900 до −970 HU [24]. Среди разных порогов %LAA-950 является наиболее распространённой метрикой и, по данным обзора, наиболее тесно коррелирует с выраженностью эмфиземы по гистологическим данным [24]. Вторым широко используемым подходом является percentile density, особенно Perc15, который отражает значение HU, ниже которого располагается 15% всех вокселей лёгочной ткани [24]. Следовательно, LAA%-950 и Perc15 представляют две основные ветви денситометрической оценки эмфиземы: пороговую и гистограммную.

Однако LAA% и порог HU < −950 имеют ограничения. Sørensen и соавторы отмечают, что классические денситометрические меры, такие как relative area below threshold и percentile density, рассматривают каждый воксель (voxel) независимо и игнорируют пространственные отношения между вокселями, а также паттерны более крупного масштаба [21]. Кроме того, такие меры зависят от одного порогового параметра, что делает их чувствительными к калибровке

сканера, шуму КТ-изображения и уровню вдоха [21]. Oh и соавторы также подчёркивают, что вариабельность объёма лёгких является одним из главных источников ошибок при density-based QCT, особенно при продольном наблюдении прогрессирования эмфиземы; для коррекции этого ограничения применяются подходы volume-adjusted lung density, включая VALDPerc15 [24].

Важным ограничением LAA% является то, что он отражает только уже сформированное снижение плотности и не учитывает топологическую организацию поражения. Li и соавторы прямо указывают, что традиционные переменные, извлекаемые из КТ, могут быть недостаточны для описания всех топологических особенностей лёгочной ткани при ХОБЛ [25]. Их модель, выделившая 80 pattern-clusters и 7 латентных факторов, показывает, что КТ содержит более сложную информацию, чем процент вокселей ниже фиксированного порога [25]. Следовательно, LAA% целесообразно рассматривать как базовый интерпретируемый биомаркер, но не как исчерпывающий метод описания эмфизематозной деструкции.

Oh и соавторы дополнительно показывают, что для air trapping используются другие пороги - например, −856 или −850 HU на экспираторной КТ [24]. В рамках PRM воксель с инспираторной плотностью < −950 HU и экспираторной плотностью > −850 HU классифицируется как преимущественно эмфизематозный, тогда как сочетание инспираторной плотности > −950 HU и экспираторной плотности < −850 HU указывает на функциональное поражение малых дыхательных путей [24]. Это важное уточнение: порог −950 HU применим прежде всего к эмфиземе на вдохе, но не исчерпывает весь спектр КТ-биомаркеров ХОБЛ.

Дополнительное развитие LAA% связано с автоматической сегментацией. Karthik и соавторы используют LAA(%) как процент площади лёгкого, поражённой эмфиземой, рассчитывая его через отношение пикселей маски поражения к общей площади лёгочного поля [23]. Такой подход показывает, что LAA% может использоваться не только как глобальная денситометрическая метрика, но и как результат сегментационного анализа, где предварительно выделяются эмфизематозные очаги. В перспективе такие методы позволяют объединять пороговую денситометрию с lesion-level analysis, 3D-визуализацией и морфометрией поражений.

Таким образом, LAA% при HU < −950 можно рассматривать как базовый количественный биомаркер эмфиземы на КТ. Его преимущества заключаются в автоматизируемости, воспроизводимости и клинической интерпретируемости. В то же время его ограничения связаны с пороговой природой метода, зависимостью от качества изображения, фазы дыхания, параметров сканирования, объёма вдоха и неспособностью учитывать пространственно-текстурную организацию поражения. Поэтому в рамках диссертации LAA% целесообразно использовать как исходную количественную основу и инструмент разметки, но дополнять его методами глубокого обучения, способными анализировать более сложные признаки эмфизематозной деструкции [21, 23, 24, 25].

1.1.4 Ограничения традиционных методов и переход к ИИ

Анализ литературы показывает, что традиционные методы диагностики ХОБЛ имеют несколько устойчивых ограничений. Спирометрия остаётся необходимым функциональным тестом, но она недостаточно чувствительна к ранним и региональным изменениям, зависит от правильности выполнения дыхательного манёвра и не описывает морфологические фенотипы заболевания [3, 14, 15, 17, 24]. Визуальная КТ-оценка позволяет выявлять эмфизему и другие структурные признаки, но ограничена субъективностью, межэкспертной вариабельностью и зависимостью от опыта радиолога [4, 6, 21, 24]. Количественная КТ-денситометрия повышает объективность, но зависит от параметров сканирования, шума, калибровки, объёма вдоха, фазы дыхания и выбранного порога HU [8, 21, 24].

Эти ограничения объясняют растущий интерес к методам машинного и глубокого обучения. Kimura и соавторы уже рассматривали deep learning как способ автоматической диагностики ранней ХОБЛ по КТ [4]. Sørensen и соавторы показали, что texture-based CT analysis может превосходить классические density-based measures [21]. Li и соавторы продемонстрировали, что unsupervised 3D deep learning способен выделять латентные тканевые паттерны и факторы, связанные с функцией лёгких, PRM и подтипами эмфиземы [25]. Zhao и соавторы показали перспективность semi-supervised learning для сегментации лёгочных сосудов при ХОБЛ, что расширяет спектр автоматизируемых КТ-биомаркеров [27].

Новые не-КТ исследования также подтверждают общий тренд к интеллектуальной диагностике и предварительной стратификации риска. Alve и соавторы применили CNN к анализу дыхательных звуков на базе ICBHI и достигли 96% accuracy при 10-fold cross-validation и 90% accuracy без cross-validation; модель также классифицировала тяжесть заболевания как mild, moderate или severe [28]. Dhar предложил multistage ensemble learning model с weighted voting, genetic algorithm optimization, LightGBM-RFE, KNN-imputation и isolation forest; на Exasens dataset модель достигла precision 0,9800, recall 0,9600, F1 0,9667, AUC 0,9912 и accuracy 0,9820 [29]. Эти результаты не относятся напрямую к КТ-денситометрии, но демонстрируют, что автоматизированные модели могут быть полезны в предварительном скрининге, особенно при ограниченных ресурсах.

Вместе с тем данные работы выявляют и новые методологические ограничения. Часть моделей обучается на небольших или специфических наборах данных, что ограничивает обобщаемость [27, 29].

Некоторые исследования используют ретроспективные базы или одну популяцию, что снижает переносимость результатов на другие страны и группы пациентов [22, 26]. Модели с высокой точностью могут быть трудно интерпретируемыми, вычислительно сложными или требовать качественной разметки [23, 25, 27, 29]. Поэтому дальнейшее развитие ИИ-диагностики ХОБЛ должно сочетать высокую точность, клиническую интерпретируемость, устойчивость к вариабельности КТ-протоколов и возможность применения в массовом скрининге.

Итоговый вывод для раздела 1.1 состоит в том, что современная диагностика ХОБЛ движется от изолированной спирометрии и визуальной КТ-оценки к комплексной модели, включающей QCT, LAA%-950, PRM, текстурные признаки, сосудистые параметры, клинические данные и методы искусственного интеллекта. LAA% остаётся важной отправной точкой для автоматизированной оценки эмфиземы, но новые исследования показывают необходимость более сложного анализа КТ-изображений, способного учитывать не только плотность, но и пространственную организацию, латентные тканевые паттерны и мультимодальные признаки риска.

1.2 Современные методы анализа КТ-изображений

Современный анализ КТ-изображений лёгких развивается как переход от ручной визуальной оценки и простых пороговых алгоритмов к гибридным системам, объединяющим количественную КТ-денситометрию, машинное обучение и глубокие нейронные сети. Для диагностики ХОБЛ и эмфиземы это особенно важно, поскольку заболевание проявляется не только снижением спирометрических показателей, но и сложными пространственными изменениями паренхимы: низкоплотными зонами, нарушением текстуры лёгочной ткани, изменением сосудистого рисунка, воздушными ловушками и региональной неоднородностью поражения.

На Рисунке 1.2-1 показана эволюция методов анализа КТ при ХОБЛ: от классических алгоритмов (сегментация, морфология) через машинное обучение (текстурные признаки, SVM) к глубокому обучению (CNN, ResNet, U-Net) и современным гибридным моделям (CNN + QCT). Данная схема отражает методологический переход от ручного конструирования признаков к автоматическому извлечению сложных паттернов и их интеграции с физически интерпретируемыми QCT-параметрами.

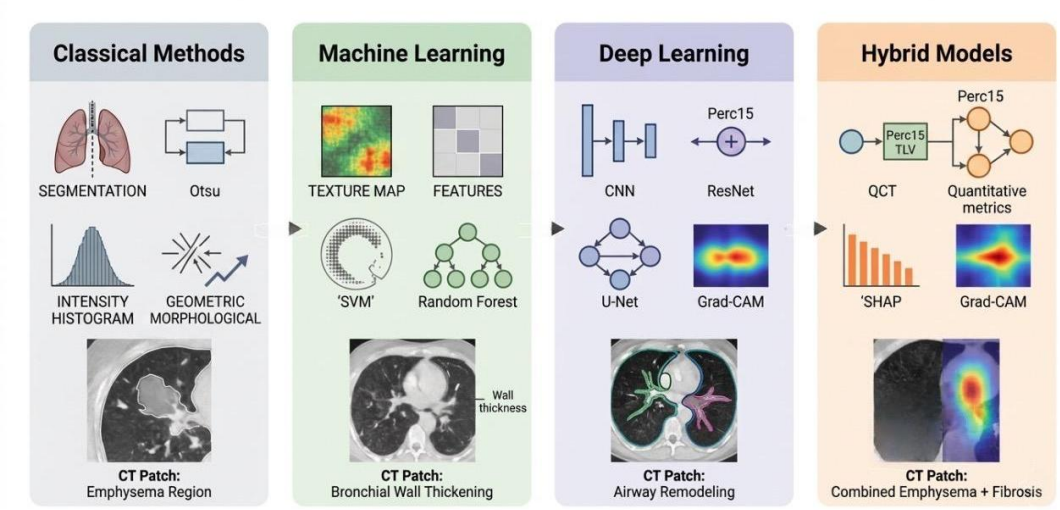


Рисунок 1.2-1 - Эволюция методов анализа КТ-изображений при диагностике ХОБЛ [24]

Четыре этапа развития на рисунке 1.2-1:

1) Классические алгоритмы (Classical Methods): сегментация (Segmentation), метод Оцу (Otsu), анализ гистограмм интенсивности (Intensity Histogram), геометрическая и морфологическая обработка (Geometric Morphological);

2) Машинное обучение (Machine Learning): текстурные карты (Texture Map), извлечение признаков (Features), метод опорных векторов (SVM), случайный лес (Random Forest);

3) Глубокое обучение (Deep Learning): свёрточные нейронные сети (CNN), ResNet, U-Net, визуализация активации (Grad-CAM);

4) Гибридные модели (Hybrid Models): интеграция QCT-метрик (Perc15, TLV) с количественными показателями, SHAP для интерпретации, Grad-CAM для локализации паттернов.

Примеры КТ-патчей иллюстрируют применение методов для выявления: области эмфиземы (Emphysema Region), утолщения бронхиальных стенок (Bronchial Wall Thickening), ремоделирования дыхательных путей (Airway Remodeling), комбинированной эмфиземы и фиброза (Combined Emphysema + Fibrosis).

Deep Learning в медицине: переход от классических алгоритмов к глубокому обучению.

Ранние методы автоматического анализа КТ-изображений лёгких в основном опирались на классическую обработку изображений: пороговую сегментацию, метод Оцу, анализ гистограмм интенсивности, локальные средние, математическую морфологию и текстурные дескрипторы. Так, Смелкина и соавторы предложили технологию автоматического распознавания эмфиземы по двумерным КТ-срезам, включающую выделение области интереса, вычисление локальных средних интенсивности и морфологическую обработку; эксперимент на реальных данных показал погрешность распознавания по объёму поражённых областей около 7,5% [30]. В другой группе классических подходов Ibrahim и Mukundan использовали каскад признаков на основе мультифрактального анализа, alpha-histograms и multi-fractal descriptors для классификации эмфиземы на КТ; такой подход отражает этап hand-crafted feature engineering, когда диагностическая модель строилась вокруг заранее заданных математических признаков изображения [31].

Ограничение классических алгоритмов состоит в том, что они требуют ручного выбора признаков и плохо учитывают сложную пространственную организацию лёгочной ткани. Sørensen и соавторы показали, что стандартные денситометрические показатели, включая relative area ниже порога и percentile density, рассматривают каждый воксель независимо и не используют пространственные связи между вокселями [21]. Их texture-based подход, основанный на фильтрах Гаусса и классификации локальных ROI, достиг AUC 0,713 для различения пациентов с ХОБЛ и без ХОБЛ, тогда как лучшие

плотностные показатели достигали AUC около 0,598 [21]. Это стало одним из аргументов в пользу перехода от простого анализа плотности к анализу текстуры.

В частности, текстурный анализ на основе банков фильтров Габора позволяет выделять локальные ориентационные и частотные характеристики изображения, однако подобные подходы остаются представителем hand-crafted features - параметры фильтров (частоты, ориентации, масштаб) должны задаваться экспертом заранее. Это принципиально ограничивает способность модели адаптироваться к многообразию патологических паттернов в лёгочной ткани. Данное ограничение стимулировало закономерный переход к методам глубокого обучения, которые способны автоматически извлекать иерархические признаки непосредственно из данных, выявляя сложные пространственные зависимости без необходимости ручного конструирования.

Дальнейшим развитием методов количественной оценки КТ стала интеграция алгоритмов глубокого обучения, позволяющая извлекать скрытые (латентные) признаки легочной ткани без какой-либо предварительной врачебной разметки. Один из таких подходов схематично изображен на Рисунке 1.2-2. Трехмерный сверточный автоэнкодер (3D CAE) анализирует случайно выбранные объемные фрагменты легкого, формируя их компактное представление (эмбеддинг), а блок «конструктора признаков» группирует эти представления в паттерн-кластеры, соответствующие различным КТ-текстурам. Применение обученной модели ко всему объему легких позволяет построить для каждого пациента гистограмму встречаемости паттернов, из которой с помощью разведочного факторного анализа (EFA) выделяются интегральные латентные факторы. Эти факторы демонстрируют высокую корреляцию с функциональными показателями и могут служить основой для более точного фенотипирования ХОБЛ, дополняя классические денситометрические параметры и PRM [25].

На Рисунке 1.2-2 наглядно показаны ключевые этапы этого процесса: от извлечения трёхмерных областей интереса (3D ROI) и обучения автоэнкодера до кластеризации эмбеддингов и последующего выделения латентных факторов с помощью разведочного факторного анализа.

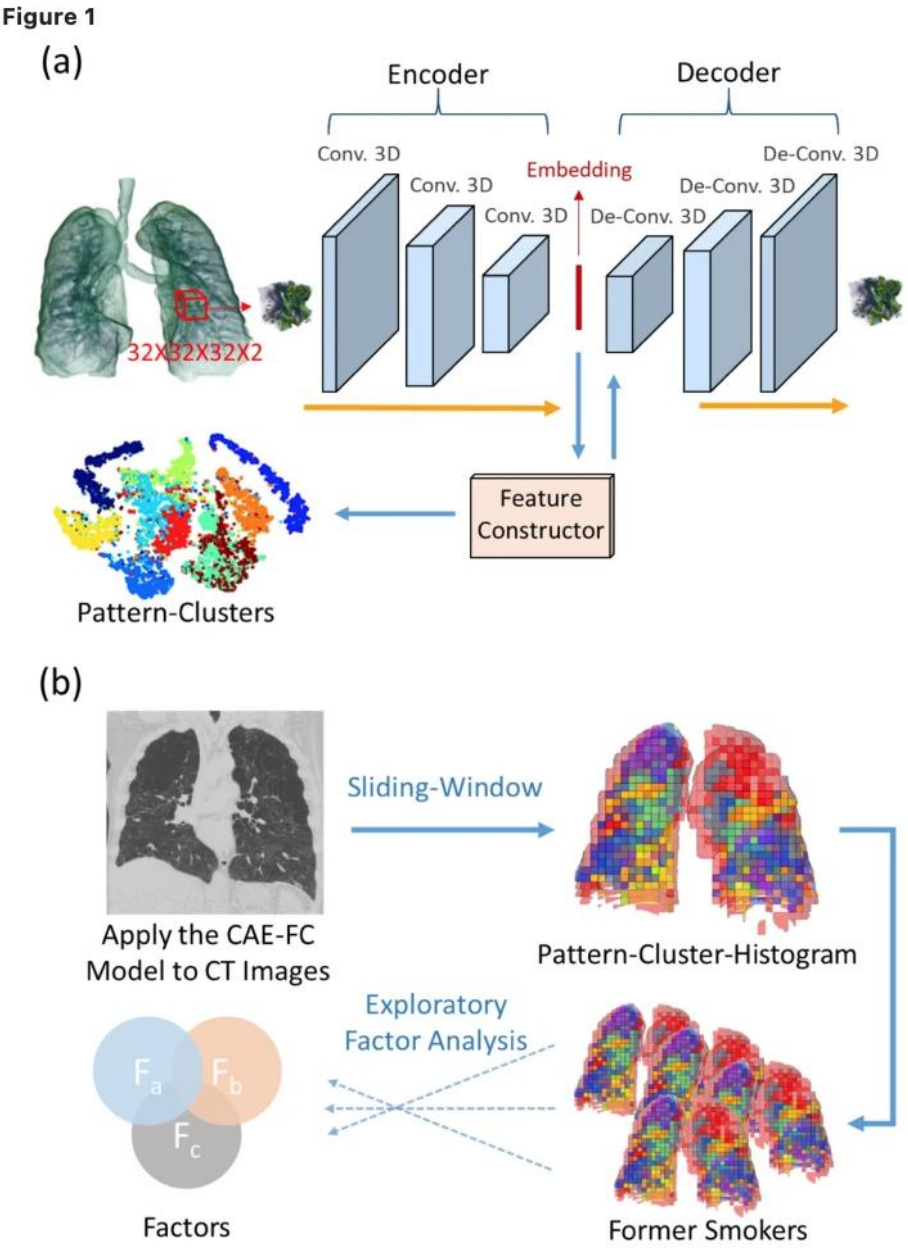


Рисунок 1.2-2 - Извлечение латентных признаков легочной ткани с помощью глубокого обучения[25]

(a) Трехмерные области интереса (3D ROI) случайным образом извлекаются из КТ-изображения легких и подаются в 3D-сверточный автоэнкодер (CAE) для обучения компактному представлению (эмбеддингу). Конструктор признаков (FC) группирует эмбеддинги в паттерн-кластеры.

(b) Обученная модель классифицирует легочную ткань методом скользящего окна, формируя для каждого обследуемого гистограмму паттерн-кластеров. Последующий разведочный факторный анализ (EFA) выделяет латентные факторы, характеризующие структурные изменения легких.

Глубокое обучение расширило эту логику: вместо ручного конструирования признаков CNN извлекают иерархические признаки

автоматически - от локальных краёв и контрастных зон до сложных паттернов эмфиземы, сосудистого рисунка и неоднородности паренхимы. Tarando в диссертационной работе описывает переход от фильтров, wavelet-преобразований, GLCM, LBP, SVM и других классических методов к CNN-архитектурам, включая convolutional, pooling, fully connected и softmax-слои, а также backpropagation как механизм обучения признаков [32]. В современных работах глубокое обучение применяется не только для бинарной классификации «эмфизема/норма», но и для стадирования ХОБЛ, оценки подтипов эмфиземы, сегментации лёгких, сегментации сосудов, построения activation maps и прогнозирования функциональных показателей [8, 24, 33, 34].

Архитектуры CNN: ResNet для классификации и U-Net для сегментации.

В задачах классификации КТ-изображений ХОБЛ и эмфиземы особое значение имеют ResNet-подобные архитектуры. Их принципиальное преимущество состоит в использовании residual connections, которые позволяют обучать глубокие сети без выраженной деградации градиента. В контексте КТ-анализа это даёт возможность извлекать признаки из сложных объёмных или патчевых (patch) структур лёгочной ткани. В работе Sarsembayeva и соавторов предложен интерпретируемый pipeline для выявления эмфиземы на низкодозовой КТ (LDCT), где используется ResNet-152 как замороженный feature extractor: из слоя global average pooling извлекается 2048-мерное представление, после чего применяются логистическая регрессия, LASSO и рекурсивное исключение признаков [34]. Такой дизайн отличается от end-to-end классификатора: CNN используется для получения глубоких признаков, а финальная классификация выполняется прозрачными статистическими моделями. Это снижает риск «чёрного ящика» и облегчает интерпретацию вклада отдельных признаков.

На Рисунке 1.2-3 наглядно представлена архитектура ResNet-152 в конфигурации извлекателя признаков. Оранжевые стрелки демонстрируют механизм skip-connections (остаточных связей), которые передают информацию напрямую через слои, предотвращая затухание градиента при обучении глубоких сетей. Выходной блок Global Average Pooling агрегирует пространственные признаки и формирует компактный 2048-мерный вектор эмбеддинга, который в дальнейшем подаётся на вход классификатора. Такая конфигурация позволяет использовать предобученную сеть без полной перестройки весов, сохраняя вычислительную эффективность и клиническую интерпретируемость.

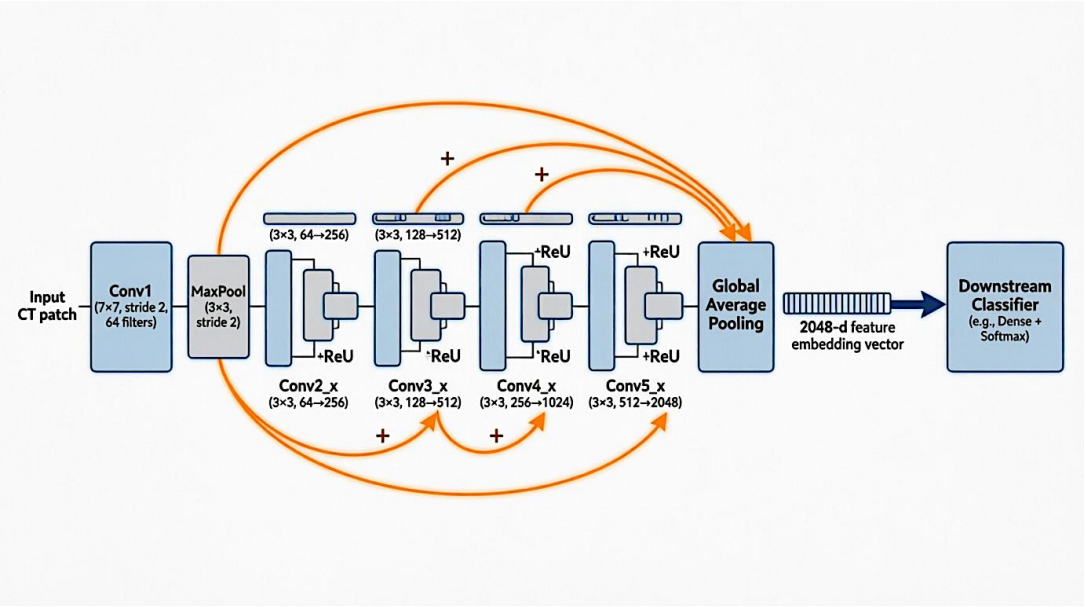


Рисунок 1.2-3 - Архитектура резидуальной свёрточной сети (ResNet-152) как извлекателя признаков

Блок-схема: входной КТ-патч → Conv1 → residual-блоки (Conv2_x–Conv5_x) с skip-connections → Global Average Pooling → 2048-мерный вектор признаков → Downstream Classifier. Визуализация по [34].

Технически pipeline Sarsembayeva et al. включает выбор наиболее полной DICOM-серии, сортировку срезов по Z-координате, перевод интенсивностей в Hounsfield Units, ресэмплинг к изотропному разрешению 1×1×1 мм³, сегментацию лёгких, автоматическую разметку по LAA% и последующее извлечение ResNet-152 embeddings (векторные представления) из патчей средней зоны лёгких [34]. В этой работе deep embeddings объединяются с QCT-маркерами LAA-950, Perc15 и общий объём лёгких (TLV); итоговая fusion-модель достигла ROC-AUC 0,996, PR-AUC 0,962 и balanced accuracy 0,931 [34]. Этот результат важен методологически: он показывает, что глубокие признаки и физические КТ-параметры не конкурируют, а дополняют друг друга.

ResNet-подход также используется в задаче подтипирования эмфиземы. Xie и соавторы обучали deep neural networks для автоматизации визуальной шкалы Fleischner Society на данных 9650 участников COPDGene [35]. Модель предсказывала тяжесть центролобулярной и парасептальной эмфиземы и использовала regression training strategy вместо прямой многоклассовой классификации. Авторы преобразовали визуальные категории Fleischner в интервалы процента поражения лёгкого: например, mild centrilobular emphysema соответствует диапазону 1–5% вовлечения лёгкого [35]. Регрессионная сеть формировала dense regression activation maps для центролобулярного и парасептального подтипов, что позволяло не только получить категориальный score, но и визуализировать локализацию эмфиземы. По сравнению с ранее опубликованным Humphries algorithm точность классификации тяжести центролобулярной эмфиземы увеличилась с 45% до 52% [35].

Для сегментации лёгких, сосудов и очагов поражения центральную роль играют U-Net-подобные архитектуры. Их логика основана на encoder–decoder структуре с skip connections: encoder сжимает изображение и извлекает контекстные признаки, decoder восстанавливает пространственное разрешение, а skip-связи передают локальные детали, необходимые для точной границы сегментации. В pipeline Sarsembayeva et al. сегментация лёгких выполнялась с помощью lungmask tool, использующего trunk U-Net model, обученный на VESSEL12 [34, 73]. Это критический этап, поскольку все последующие QCT-метрики - LAA%, Perc15, Mean HU и TLV - вычисляются только внутри корректной маски лёгких.

Zhao и соавторы демонстрируют развитие U-Net-парадигмы в более сложной задаче сегментации лёгочных сосудов при ХОБЛ [27]. В их работе сравнивались nnFormer, UNETR, Swin UNETR и nnU-Net; лучшую fully supervised performance показал nnFormer с Dice coefficient 0,855, IoU 0,747, sensitivity 0,832 и precision 0,880, тогда как nnU-Net также показал устойчивый результат с Dice 0,835 и IoU 0,739 [27]. Эти результаты показывают, что U-Net-подобные и transformer-U-Net архитектуры применимы не только для сегментации лёгочного поля, но и для более тонких структур - сосудов, малых ветвей и микрососудистых признаков ХОБЛ.

Сегментационные методы важны и для очагов эмфиземы. Karthik, Murugan и Goel предложили pipeline, включающий min–max scaling, CLAHE для усиления контраста, Gaussian blur для подавления шума, пороговую сегментацию лёгких, morphological closing и последующее выделение эмфизематозных поражений с учётом разных классов эмфиземы [23]. Хотя этот подход не является чистой U-Net-моделью, он показывает, какие задачи должна решать нейросетевая сегментация: выделять лёгочные границы, отделять поражённые участки от нормальной паренхимы, сохранять анатомические контуры и обеспечивать количественную оценку burden, распределения и формы очагов.

Однако все рассмотренные выше архитектуры - ResNet, U-Net и их модификации - предполагают наличие достаточного объёма качественных эталонных разметок (ground truth). В реальных клинических задачах анализа ХОБЛ получение детальных воксельных или пожаровых аннотаций крайне трудоёмко, требует участия экспертов и часто сопровождается высокой межэкспертной вариабельностью. Этот дефицит точных разметок стимулировал активное развитие методов обучения со слабой разметкой (weakly supervised learning), позволяющих обучать модели по менее детальным, но более доступным сигналам.

Weakly supervised learning: решение проблемы нехватки детальной разметки.

Одной из главных проблем анализа КТ при ХОБЛ является дефицит детальной воксельной или региональной разметки. Эмфизематозные изменения могут быть диффузными, слабовыраженными и неоднородными, поэтому ручная разметка требует высокой квалификации, занимает много времени и обладает межэкспертной вариабельностью. В этой ситуации обучение с частичным

привлечением учителя (weakly supervised learning) становится методологически важным направлением: модель обучается не на точных масках поражения, а на слабых метках - например, на уровне пациента, скана, визуального score или функционального теста.

Sørensen и соавторы предложили один из ранних вариантов такой логики: локальные ROI получают метку, основанную не на ручной аннотации конкретной области, а на PFT/GOLD-статусе пациента [21]. Авторы прямо отмечают, что случайно выбранные ROI у пациента с ХОБЛ могут содержать как поражённую, так и здоровую ткань, но всё равно получают метку COPD; обратная ситуация возможна и у здоровых субъектов. Такая постановка создаёт шумные, перекрывающиеся классы, но позволяет избежать ручной разметки ROI и использовать метаданные, связанные со всем сканом [21]. Это типичный пример weak labeling: метка корректна на уровне субъекта, но не обязательно точна на уровне локального участка ткани.

В работе Sarsembayeva et al. weak supervision реализована через quality-controlled auto-labeling: сканы получают бинарную метку эмфиземы на основе LAA%, то есть доли вокселей ниже −950 HU внутри сегментированной маски лёгких [34]. Такая стратегия не требует ручного указания очагов, но связывает обучение CNN-признаков с объективным QCT-критерием. Авторы дополнительно применяют patient-level splits, class weighting, bootstrap confidence intervals и feature-level attribution, чтобы снизить риск смещения и повысить воспроизводимость [34].

Xie и соавторы развивают более сложную weakly supervised схему. Вместо воксельной разметки эмфиземы они используют визуальные Fleischner scores, преобразуют их в интервалы процента поражения и обучают regression network, которая выдаёт dense regression activation maps [35]. Дополнительно вводятся три компонента функции потерь: interval regression loss, overlapping loss для предотвращения перекрытия карт разных подтипов эмфиземы и segmentation loss, где union dense regression activation maps согласуется с LAA-950 pseudo-target [35]. Такой подход позволяет получить локализационные карты без истинной voxel-wise разметки, что особенно важно для клинических данных, где подробная сегментация эмфиземы обычно отсутствует.

Semi-supervised стратегия Zhao et al. решает близкую проблему нехватки аннотаций для сосудистой сегментации [27]. Авторы используют teacher–student self-training: сначала teacher model обучается на ограниченном наборе размеченных CT, затем генерирует pseudo-labels для неразмеченных данных, после чего отбираются только надёжные pseudo-labels. На первой итерации выбирались 40 pseudo-labels с Mean Precision >0,9, на второй - pseudo-labels с Mean Precision >0,95 и Mean Dice >0,85 [27]. Такая схема особенно релевантна для анализа ХОБЛ, поскольку ручная разметка малых сосудов и терминальных структур крайне трудоёмка.

Метрики QCT: Perc15, TLV и Mean HU.

Количественная КТ остаётся физически интерпретируемой основой анализа ХОБЛ. В отличие от CNN-признаков, QCT-метрики имеют прямой

смысл в шкале Хаунсфилда, объёме лёгких и распределении плотности ткани. Наиболее часто используются LAA%, Perc15, TLV и Mean HU/Mean Lung Density.

Perc15 - это 15-й перцентиль распределения плотности лёгочной ткани: значение HU, ниже которого располагается 15% всех вокселей лёгкого. В обзоре Oh и соавторов указано, что percentile density является альтернативой LAA-порогам, а Perc15 является наиболее распространённой перцентильной метрикой, поскольку коррелирует с гистопатологической выраженностью эмфиземы [24]. В отличие от LAA%-950, Perc15 не фиксирует один жёсткий порог, а описывает нижнюю часть гистограммы плотности. Это делает его полезным для отслеживания прогрессирования эмфиземы, особенно в продольных исследованиях, хотя он также зависит от объёма вдоха и требует коррекции при различиях в lung volume.

TLV - total lung volume, общий объём лёгких, вычисляемый по сегментированной маске лёгких с учётом размера вокселя. TLV важен не только как самостоятельный показатель гиперинфляции или объёмного состояния лёгких, но и как нормирующий параметр для денситометрии. Oh и соавторы подчёркивают, что вариабельность объёма лёгких является одним из главных источников ошибок density-based QCT, особенно при продольной оценке прогрессирования эмфиземы [24]. Wysoczanski также рассматривает TLV как важный КТ-показатель и предлагает multi-view CNN для оценки total lung volume по cardiac CT и lung masks, поскольку во многих крупных когортах лёгкие отображены частично, а полный TLV необходимо регрессировать из неполного поля зрения [36].

Mean HU, или mean lung density/mean lung attenuation, представляет среднюю плотность лёгочной ткани в HU внутри маски лёгких. Эта метрика отражает общее смещение плотности паренхимы: при эмфиземе средняя плотность обычно становится более отрицательной, поскольку увеличивается доля воздуха и низкоплотных зон. Gawlitza и соавторы использовали mean lung density наряду с lung volume, low-attenuated volume и full width at half maximum в QCT-моделях для предсказания функциональных параметров лёгких; среди моделей k-nearest neighbours показал наименьшую mean relative error - 16% [7]. Abd elsalam и соавторы также показали значимую корреляцию между PFT-показателями FEV1 и FEV1/FVC и параметрами количественной оценки при пороге −950 HU, подчёркивая применимость автоматической КТ-денситометрии для оценки тяжести эмфиземы [37].

Важно, что QCT-метрики могут расширяться за пределы эмфиземы. Schwartz и соавторы использовали HUI - процент вокселей <−950 HU на вдохе, и HUE - процент вокселей <−856 HU на выдохе; в их модели air trapping по HUE был независимым предиктором риска рака лёгкого наряду с FEV1/FVC <0,70 [38]. Этот пример показывает, что количественная КТ позволяет анализировать не только низкоплотную эмфизему на вдохе, но и функциональные признаки задержки воздуха на выдохе.

Синтез методов: обоснование объединенного подхода.

Критическая точка современного анализа КТ-изображений состоит в том, что ни классическая QCT, ни CNN по отдельности не решают задачу полностью. QCT обеспечивает физически интерпретируемые показатели: LAA%-950, Perc15, TLV, Mean HU, low-attenuated volume, air trapping и другие параметры. Эти метрики хорошо стандартизируются, легко объясняются клиницисту и связаны с плотностью ткани в шкале Хаунсфилда. Однако QCT-показатели часто сводят сложную трёхмерную структуру лёгкого к гистограмме или порогу и потому могут не учитывать пространственное распределение, форму очагов, текстуру, локальную неоднородность и региональные паттерны поражения [21, 24, 34].

Сверточные нейросети, напротив, эффективно распознают сложные визуальные паттерны и текстуры. Они способны извлекать признаки, которые не описываются одним HU-порогом: неоднородность паренхимы, локальные переходы плотности, структурные паттерны эмфиземы, сосудистую редукцию и форму очагов. Работы Sørensen et al. и Xie et al. показывают, что texture/deep learning approaches могут выявлять информацию, недоступную простой денситометрии [21, 35]. Однако CNN-модели часто сложнее интерпретировать, требуют контроля за смещениями, зависят от качества данных и могут иметь ограниченную переносимость между протоколами КТ.

Поэтому наиболее обоснованным направлением является гибридный подход. В нём CNN используется как инструмент извлечения сложных визуально-текстурных признаков, а QCT - как физически интерпретируемый слой объективных параметров. Sarsembayeva et al. демонстрируют именно такую логику: ResNet-152 embeddings объединяются с LAA-950, Perc15 и TLV, а классификация выполняется интерпретируемыми моделями с LASSO/RFE и bootstrap-валидацией [34]. Полученная fusion-модель показывает высокий диагностический результат, но при этом сохраняет связь с клинически понятными QCT-маркерами.

С практической точки зрения Fusion система особенно перспективна для популяционного скрининга. QCT-компонент обеспечивает воспроизводимую автоматическую разметку и контроль физического смысла результата, а CNN-компонент позволяет выявлять более тонкие признаки, включая предпороговые изменения, текстурную неоднородность и сложные паттерны эмфизематозного поражения. Таким образом, оптимальная стратегия анализа КТ при ХОБЛ должна строиться не как замена денситометрии нейросетью, а как синтез: количественный анализ фиксирует объективные физические параметры плотности ткани, тогда как CNN распознаёт сложные визуальные паттерны, которые не сводятся к одному порогу HU.

1.3 Теоретическая модель исследования

Теоретическая модель настоящего исследования строится на предпосылке, что автоматическая диагностика эмфиземы по КТ-изображениям не может быть сведена ни к одной изолированной метрике плотности, ни к исключительно нейросетевой классификации. Эмфизема при ХОБЛ представляет собой

пространственно неоднородный процесс деструкции лёгочной паренхимы, проявляющийся снижением плотности ткани, формированием низкоплотных зон, изменением текстуры, региональной неоднородностью поражения и нарушением анатомического рисунка лёгкого. Поэтому задача классификации эмфиземы должна формализоваться как многофакторная задача анализа медицинского изображения, в которой учитываются как пространственно-визуальные паттерны, так и физически измеримые денситометрические параметры.

С методологической точки зрения входными данными модели является КТ-объём лёгких, предварительно приведённый к стандартизованному виду: выделяется маска лёгких, исключаются внепульмональные структуры, нормализуются пространственные параметры и рассчитываются количественные признаки. После этого изображение представляется в двух взаимодополняющих пространствах. Первое пространство - нейросетевое: оно формируется глубокими признаками, извлекаемыми сверточной архитектурой ResNet-152. Второе пространство - физико-денситометрическое: оно задаётся QCT-показателями, прежде всего LAA%, Perc15 и TLV. Такая постановка соответствует современному направлению развития анализа КТ при ХОБЛ, где глубокие модели используются для распознавания сложных паттернов, а QCT сохраняет связь результата с измеряемыми физическими характеристиками ткани [24], [34], [39].

ResNet-152 в качестве извлекателя признаков.

Использование ResNet-152 в качестве feature extractor методологически оправдано тем, что эмфизематозная деструкция не всегда выражается простой бинарной границей между нормальной и патологической тканью. В ранних и умеренных формах эмфиземы патологический процесс может проявляться тонкими изменениями текстуры, локальными зонами снижения плотности, нарушением сосудистого рисунка и неоднородным распределением низкоплотных областей. Классический пороговый анализ фиксирует только воксели ниже заданного HU-порога, тогда как глубокая CNN-архитектура способна извлекать признаки более высокого порядка.

ResNet-152 относится к глубоким residual-сетям, принципиальной особенностью которых являются skip-connections. Эти связи позволяют передавать информацию через большое число слоёв, снижая риск деградации градиента и обеспечивая обучение более глубокого признакового представления. В контексте КТ лёгких это означает, что сеть может извлекать не только локальные контрастные признаки, но и более сложные семантические паттерны: распределение очагов деструкции, структурную неоднородность паренхимы, текстурные переходы между нормальной и патологической тканью, а также пространственные признаки, которые трудно задать вручную.

В исследовании Sarsembayeva и соавторов ResNet-152 используется именно как извлекатель признаков, а не как полностью переобучаемая end-to-end модель. Авторы извлекали 2048-мерные embeddings из mid-lung patches, после чего применяли логистическую регрессию, LASSO и recursive feature elimination;

далее эти embeddings объединялись с QCT-маркерами %LAA, Perc15 и TLV для повышения устойчивости и интерпретируемости модели [34]. Такой подход принципиально важен для диссертационной модели: глубокая сеть не заменяет клиническую и физическую логику диагностики, а преобразует КТ-изображение в компактное пространство высокоуровневых признаков, пригодное для последующего объяснимого анализа.

Сходную методологическую ценность residual-подходов подтверждает работа Ahmed и соавторов, где использовалась 3D CNN на основе VoxResNet для классификации ХОБЛ и эмфиземы по объёмным КТ-данным с volume-wise annotations [41]. Авторы подчёркивают, что 3D-сети позволяют использовать межсрезовый контекст, который теряется при анализе отдельных 2D-срезов. Хотя в настоящей модели ResNet-152 применяется как feature extractor для извлечения признаков из выбранных КТ-представлений, сама логика residual learning остаётся общей: глубокие признаки позволяют моделировать пространственную структуру поражения более полно, чем ручные признаки или простая денситометрия.

QCT как физические признаки.

Количественный анализ КТ вводится в модель как слой «жёстких» физических признаков, связывающих результат классификации с измеряемыми параметрами плотности ткани. В отличие от CNN embeddings, которые являются абстрактными признаками, QCT-метрики имеют прямую физическую интерпретацию в шкале Хаунсфилда и в объёмных характеристиках лёгких. Именно поэтому LAA%, Perc15 и TLV следует рассматривать не как дополнительные статистические переменные, а как физические якоря модели.

LAA% отражает долю низкоплотных участков лёгочной ткани, чаще всего определяемых по порогу −950 HU на инспираторной КТ. Этот показатель фиксирует объёмную долю паренхимы, соответствующей эмфизематозной деструкции. Perc15 представляет 15-й перцентиль распределения плотности: значение HU, ниже которого находится 15% вокселей лёгкого. В отличие от LAA%, Perc15 не зависит от одного фиксированного порога, а описывает нижнюю часть гистограммы плотности. TLV, или общий объём лёгких, необходим для учёта объёмного состояния лёгких и гиперинфляции, поскольку вариации вдоха и общего объёма могут влиять на денситометрические показатели [24], [42].

Роль QCT в теоретической модели состоит в снижении риска неконтролируемой нейросетевой интерпретации. CNN может обнаруживать статистические паттерны, связанные не только с заболеванием, но и с особенностями протокола сканирования, шумом, реконструкцией или популяционными смещениями. Введение QCT-признаков ограничивает модель физически проверяемыми параметрами: если нейросеть выявляет визуальный паттерн эмфиземы, то LAA% и Perc15 должны подтверждать наличие соответствующего снижения плотности. Поэтому QCT выполняет функцию физической регуляризации: оно не устраняет необходимость глубокого обучения, но задаёт клинически понятную и измеряемую основу результата.

Данные Gomes и соавторов дополнительно показывают, что объём эмфиземы целесообразно соотносить не только с измеренным CT lung volume, но и с predicted total lung volume, поскольку гиперинфляция может искажать оценку протяжённости поражения [42]. Это особенно важно для теоретической модели исследования: TLV не является второстепенной переменной, а выступает параметром, позволяющим корректно интерпретировать LAA% и Perc15 в условиях различного дыхательного объёма и степени гиперинфляции.

Концепция Fusion: объединение CNN + QCT

Ключевым элементом теоретической модели является fusion - объединение нейросетевых и количественных признаков в едином диагностическом пространстве. Формально модель можно представить как двухпоточный pipeline. Первый поток принимает КТ-представление лёгких и преобразует его с помощью ResNet-152 в вектор глубоких признаков:

$$z_{CNN} = ResNet - 152(X_{CN}), \quad (1.3.1)$$

где $X_{CN} -$ входное КТ-представление (патч или объём), z_{CNN}- embedding, отражающий визуально-текстурные и семантические признаки эмфизематозной деструкции. Второй поток рассчитывает физические QCT-параметры:

$$z_{QCT} = [LAA\%, Perc15, TLV, Mean\ HU] \quad (1.3.2)$$

После нормализации и отбора признаков оба представления объединяются:

$$z_{Fusion} = z_{CNN} \,||\, z_{QCT} \quad (1.3.3)$$

где символ || обозначает конкатенацию признаковых векторов. Далее объединённый вектор поступает в интерпретируемый классификатор, например логистическую регрессию с L1-регуляризацией, LASSO/RFE-отбором или другой моделью, позволяющей оценивать вклад отдельных признаков. ...Такая схема позволяет сохранить преимущества глубокого обучения, но не превращать всю систему в полностью непрозрачный классификатор.

На Рисунке 1.3 наглядно представлена архитектура гибридной модели, объединяющей два независимых потока анализа. Левый поток (Stream 1) отвечает за извлечение высокоразмерных семантических признаков с помощью ResNet-152, формируя 2048-мерный вектор эмбеддинга (z_{CNN}). Правый поток (Stream 2) вычисляет физически интерпретируемые QCT-метрики (LAA%, Perc15, TLV, Mean HU), формируя вектор z_{QCT}. В блоке Fusion Block происходит конкатенация представлений, после чего применяется отбор наиболее информативных признаков (LASSO/RFE) и финальная интерпретируемая классификация. Карта активации в правом нижнем углу подтверждает, что модель фокусирует своё внимание именно на анатомических зонах эмфизематозного поражения, а не на артефактах или фоновых структурах.

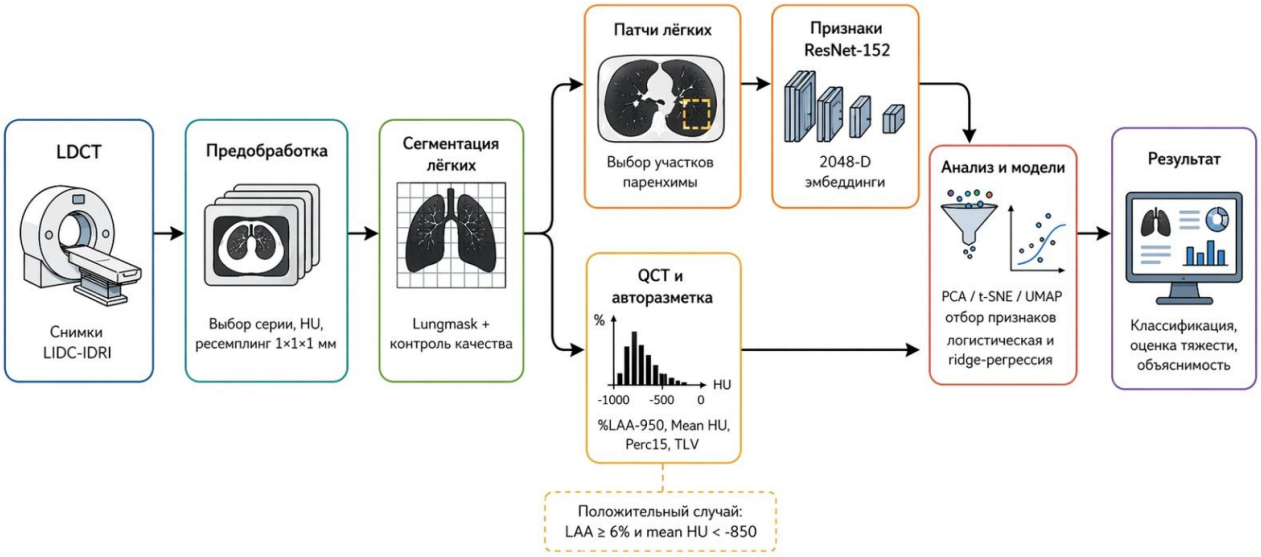


Рисунок 1.3 - Fusion мультимодальный метод диагностики эмфиземы (CNN + QCT Fusion)

Двухпоточная архитектура: поток CNN (ResNet-152 → z_{CNN} embeddings) и поток QCT (LAA%, Perc15, TLV → z_{QCT})

→ конкатенация признаков → отбор признаков (LASSO/RFE) → интерпретируемая классификация.

Именно такой принцип реализован в работе Sarsembayeva и соавторов: ResNet-152 embeddings объединяются с QCT-маркерами %LAA, Perc15 и TLV, а итоговая fusion-модель демонстрирует высокую диагностическую эффективность - ROC-AUC 0,996, PR-AUC 0,962 и balanced accuracy 0,931 [34]. Для диссертационной модели это является прямым методологическим основанием: наилучший результат достигается не при противопоставлении CNN и QCT, а при их синтезе.

Широкую обоснованность fusion-подхода подтверждают и другие исследования. Bodduluri и соавторы подчёркивают, что отдельные density- и texture-based features описывают повреждение и ремоделирование лёгочной ткани, но для понимания ХОБЛ важно объединять плотностные, текстурные и механические признаки, получаемые из КТ-регистрации вдоха и выдоха [48]. Zhang и соавторы в модели EDLNet также используют adaptive gated fusion mechanism, который селективно объединяет encoder- и decoder-признаки, подавляя нерелевантную информацию и улучшая локализацию эмфиземы [23]. Хотя их работа относится к self-supervised anomaly detection, сама идея адаптивного объединения признаков подтверждает общий принцип: диагностика эмфиземы выигрывает от контролируемого синтеза разнородных представлений.

Мультимодальная интерпретируемость

Предлагаемая теоретическая модель является мультимодальной не в смысле использования различных источников медицинских данных, а в смысле интеграции двух уровней представления одного КТ-исследования: визуально-

семантического и физико-денситометрического. Свёрточная нейронная сеть реализует анализ визуальной структуры изображения, выделяя паттерны эмфизематозной деструкции, пространственную неоднородность, текстурные особенности и морфологию поражённых участков. Блок количественной компьютерной томографии (QCT), напротив, выполняет объективное измерение денситометрических характеристик: определяет долю лёгочной ткани с патологически низкой плотностью, смещение гистограммы распределения единиц Хаунсфилда и общий объём лёгких.

Подобное разделение функций повышает обоснованность диагностического заключения для клинического применения. В ситуациях, когда модель указывает на высокую вероятность эмфиземы, врач получает возможность оценить не только итоговый класс, но и основания решения: вклад признаков, извлечённых ResNet, характеризует визуальный паттерн, тогда как показатели LAA%, Perc15, TLV и Mean HU обеспечивают физически измеряемое подтверждение диагноза. Это обстоятельство приобретает особую значимость в задачах популяционного скрининга, где автоматическая система должна сочетать высокую точность с интерпретируемостью, воспроизводимостью и устойчивостью к вариациям протоколов сканирования.

Клиническая значимость CNN-признаков подтверждается результатами Humphries и соавторов, показавших, что классификация эмфиземы по шкале Fleischner с помощью глубокого обучения коррелирует с функциональными показателями и смертностью [46]. Almeida и соавторы, применяя метод self-supervised anomaly detection, установили, что величина аномалии (anomaly score), вычисленная по КТ, согласуется со снижением лёгочной функции, выраженностью эмфиземы, воздушной ловушки и клиническими исходами, а в ряде случаев превосходит по информативности традиционные supervised-модели [39]. Эти результаты свидетельствуют о том, что нейросетевые представления способны улавливать клинически значимые патологические изменения. Однако без QCT-компонента такие признаки остаются менее прозрачными для интерпретации лечащим врачом.

Вместе с тем использование исключительно QCT-метрик также сопряжено с ограничениями. Sørensen и соавторы продемонстрировали, что классические денситометрические показатели рассматривают каждый воксель независимо, игнорируя пространственные взаимосвязи и текстурные паттерны более высокого порядка [21]. Соответственно, модель, опирающаяся только на LAA% или Perc15, хотя и является физически интерпретируемой, может обладать недостаточной чувствительностью к ранним или неоднородным формам эмфиземы. Таким образом, интерпретируемость должна сочетать физическую обоснованность (денситометрический компонент) с визуально-пространственной объяснимостью (паттерны, выявляемые CNN), что позволяет отразить как плотностный субстрат заболевания, так и структурную организацию поражения.

В данном контексте концепция интеграции CNN и QCT формирует теоретический базис машинной диагностики ХОБЛ. ResNet-152 осуществляет

извлечение высокоуровневых визуальных признаков деструкции лёгочной ткани, блок QCT обеспечивает связь с физическими параметрами сканирования, а механизм fusion объединяет оба уровня представления в единую диагностическую систему. Такая архитектура не ограничивается бинарной классификацией эмфиземы, а формирует объяснимое заключение, в котором нейросетевой паттерн верифицируется количественными денситометрическими параметрами.

Итоговая теоретическая модель

Теоретическая модель исследования может быть представлена как последовательность взаимосвязанных этапов:

Предобработка КТ: стандартизация изображения, выделение лёгочной маски, контроль качества, нормализация пространственных параметров.

Извлечение QCT-признаков: расчёт LAA%, Perc15, TLV и Mean HU как физически интерпретируемых показателей плотности и объёма лёгочной ткани.

Извлечение CNN-признаков: применение ResNet-152 в качестве экстрактора признаков (feature extractor) для получения высокоуровневых эмбеддингов z_{CNN}, отражающих визуально-текстурные паттерны эмфиземы.

Интеграция признаков: объединение векторов z_{CNN} и z_{QCT} в единое признаковое пространство с последующим отбором наиболее информативных признаков.

Интерпретируемая классификация: использование модели, позволяющей оценивать вклад отдельных признаков и формировать объяснимое диагностическое решение.

Клиническая интерпретация: сопоставление результата с визуальным паттерном, значениями LAA%, Perc15, TLV, Mean HU и пороговыми критериями эмфиземы.

Таким образом, исследование опирается на теоретическую модель, в которой глубокое обучение и количественная компьютерная томография не конкурируют, а выполняют различные, но взаимодополняющие эпистемологические функции. Свёрточная нейронная сеть выявляет сложные визуальные паттерны, тогда как QCT фиксирует объективные физические параметры плотности ткани. Интеграция этих двух компонентов позволяет создать мультимодальную интерпретируемую систему, потенциально пригодную как для масштабируемой диагностики эмфиземы, так и для адаптации КТ-анализа к задачам популяционного скрининга.

2 МЕТОДИКА И ОРГАНИЗАЦИЯ ИССЛЕДОВАНИЯ

2.1 Цель и гипотеза исследования

Целью исследования является разработка и экспериментальная апробация интеллектуальной системы автоматизированной диагностики эмфиземы по данным компьютерной томографии, основанной на совместном использовании глубоких признаков, извлекаемых свёрточной нейронной сетью ResNet-152, и объективных количественных параметров плотности лёгочной ткани, рассчитываемых методами QCT. Система ориентирована на объединение двух взаимодополняющих типов информации: латентных визуально-текстурных представлений, формируемых CNN, и физически интерпретируемых денситометрических показателей.

Актуальность такого подхода определяется ограничениями существующих методов диагностики ХОБЛ. Спирометрия, оставаясь базовым функциональным тестом, отражает преимущественно глобальное нарушение воздушного потока и не позволяет дифференцировать морфологические фенотипы заболевания. Визуальная оценка КТ способна выявлять эмфизематозные изменения, однако зависит от квалификации врача и подвержена значительной межэкспертной вариабельности [24, 46]. Количественная КТ-денситометрия, включая такие показатели, как LAA% и Perc15, обеспечивает объективное измерение низкоплотных зон, но пороговые и гистограммные метрики не в полной мере учитывают пространственную организацию поражения, текстурные характеристики паренхимы и сложные визуальные паттерны эмфизематозной деструкции [21, 24].

В связи с этим методологически обоснованным представляется переход к модели, в которой ResNet-152 выступает не в роли автономного классификатора, а в качестве экстрактора латентных признаков КТ-изображения. Эти признаки отражают особенности визуальной структуры лёгочной ткани: локальную неоднородность, характер распределения низкоплотных участков, нарушения сосудисто-паренхиматозного рисунка, текстурные переходы и пространственную организацию изменений. Сами по себе CNN-признаки не являются прямыми патологическими маркерами, поскольку нейронная сеть не измеряет альвеолярную деструкцию непосредственно. Однако после сопоставления с QCT-показателями, клиническими метками и стадиями ХОБЛ они могут интерпретироваться как вычислительные дескрипторы, ассоциированные с эмфизематозным поражением [21, 34].

Задачи исследования

Для достижения поставленной цели необходимо решить следующие задачи:

Провести анализ научных источников по современным методам диагностики ХОБЛ и эмфиземы, включая спирометрию, визуальную КТ-оценку, количественную КТ-денситометрию, машинное обучение и глубокие нейронные сети для анализа медицинских изображений.

Провести анализ латентного слоя из предобученной ResNet152 для получения сокращенного подмножества признаков, с сохранением диагностической эффективности.

Реализовать стратегию автоматической разметки (auto-labeling) данных, используя количественный анализ КТ (QCT).

Разработать модифицированный(объединенный) метод с использованием эмбеддингов и автоматической аннотации по QCT для выявления признаков эмфиземы легких по КТ-изображениям, адаптированный к масштабным скрининговым задачам.

Разработка системы автоматического выявления признаков эмфиземы по данным КТ-изображений на основе предложенного модифицированного метода.

Научная гипотеза исследования

Основная научная гипотеза исследования заключается в том, что объединение латентных визуально-текстурных признаков, извлечённых глубокой сверточной нейронной сетью ResNet-152, и объективных физических показателей плотности лёгочной ткани, рассчитанных методами количественной КТ-денситометрии, позволит статистически значимо повысить точность классификации ХОБЛ и эмфиземы по сравнению с использованием каждого из этих подходов по отдельности.

Сущность гипотезы состоит в том, что CNN-признаки и QCT-признаки описывают разные уровни информации, содержащейся в КТ-изображении. ResNet-152 формирует латентное визуально-текстурное представление изображения: сеть извлекает не диагноз как таковой, а устойчивые вычислительные признаки, связанные с распределением контрастов, неоднородностью паренхимы, пространственной конфигурацией низкоплотных зон и изменением нормального лёгочного рисунка. Эти признаки могут отражать морфологически обусловленные изменения, однако их патологическая интерпретация требует сопоставления с клиническими и денситометрическими данными.

QCT-метрики, напротив, имеют прямую физическую интерпретацию. LAA% показывает долю низкоплотных участков лёгочной ткани, Perc15 отражает нижний сегмент распределения плотности вокселей, а дополнительные объёмные показатели позволяют учитывать особенности вентиляционного состояния лёгких. Благодаря этому QCT-компонент вводит в модель «жёсткие» физические признаки, связанные со шкалой Хаунсфилда и плотностью ткани. Это снижает риск того, что нейросетевая модель будет опираться на нерелевантные визуальные закономерности, шум или особенности протокола сканирования [24], [42].

Таким образом, в рамках предлагаемой модели ResNet-152 характеризует визуальную структуру лёгочной ткани на КТ-изображении, тогда как QCT даёт объективную количественную оценку физического субстрата - степени снижения рентгеновской плотности. Интеграция обеих групп признаков в единой архитектуре позволяет сочетать высокую чувствительность CNN к сложным

визуально-текстурным паттернам с метрологической строгостью и интерпретируемостью QCT.

Обоснование гипотезы подтверждается современными исследованиями. Показано, что глубокое обучение может автоматически классифицировать эмфизематозные паттерны на КТ и связывать их с клиническими показателями и прогнозом [39], [46]. Одновременно количественная КТ позволяет объективно оценивать выраженность эмфиземы и её связь с функциональными нарушениями лёгких [24], [42]. Однако изолированные density-based показатели ограничены тем, что не учитывают пространственные отношения между вокселями и сложную текстурную структуру поражения [6]. Поэтому Fusion метод, объединяющий ResNet-embeddings и QCT-метрики, является методологически обоснованным решением для повышения точности, устойчивости и интерпретируемости автоматизированной диагностики [34].

Итак, научная гипотеза исследования формулируется следующим образом: интеграция латентных визуально-текстурных признаков, извлечённых сетью ResNet-152, и физически интерпретируемых QCT-метрик (LAA%, Perc15, TLV, Mean HU) в рамках единой диагностической модели обеспечивает статистически значимо более высокие показатели точности классификации эмфиземы и клиническую интерпретируемость результатов по сравнению с раздельным применением нейросетевого анализа и традиционной КТ-денситометрии.

2.2 Данные и их подготовка

В качестве исходного массива данных в исследовании используется открытый датасет **LIDC-IDRI** (*Lung Image Database Consortium and Image Database Resource Initiative*), размещённый в The Cancer Imaging Archive. Коллекция включает **1018 торакальных КТ-исследований** от **1010 пациентов** и была сформирована в рамках многоцентрового проекта при участии академических медицинских центров, NCI, FNIH, FDA и промышленных партнёров в области медицинской визуализации [49]. Первоначальное назначение LIDC-IDRI заключалось не в диагностике ХОБЛ, а в разработке и валидации алгоритмов компьютерного поиска и классификации лёгочных узлов. Поэтому каждый скан сопровождается разметкой узлов лёгких, выполненной четырьмя радиологами, однако специализированных меток по ХОБЛ, эмфиземе или GOLD-стадиям в датасете нет [49].

Важной особенностью LIDC-IDRI является высокая техническая разнородность данных: исследования получены на сканерах различных производителей (GE, Siemens, Philips, Toshiba) при варьирующих параметрах напряжения, силы тока, толщины среза, реконструкционного интервала и ядра реконструкции (kernel). Толщина среза варьирует приблизительно от 0,6 до 5,0 мм, что создаёт значительную неоднородность размера вокселя и уровня шума. Для задач поиска узлов такая вариабельность приемлема, однако для денситометрического анализа и расчёта доли низкоплотных зон она становится

критически значимой, поскольку LAA% непосредственно зависит от корректного пространственного разрешения и калибровки единиц Хаунсфилда.

Ключевая методологическая проблема исследования состоит в отсутствии в LIDC-IDRI клинически верифицированной разметки по ХОБЛ. Классическое обучение с учителем на основе экспертных марок эмфиземы, таким образом, невозможно. Для преодоления этого ограничения в работе применяется стратегия автоматической разметки (auto-labeling), опирающаяся на объективные физические свойства КТ-изображения. Разметка формируется на основе количественного денситометрического показателя LAA% (Low Attenuation Area percentage) и рассматривается как «серебряный стандарт» (Silver Standard) - не полноценная клинико-гистологическая истина, а воспроизводимый суррогатный стандарт, обоснованный валидированными пороговыми значениями плотности. Такой подход принципиально минимизирует субъективность, присущую визуальной оценке, и исключает межэкспертную вариабельность, поскольку решение принимается на основе строгих физических критериев, а не индивидуального опыта врача.

Расчёт LAA% выполняется после выделения маски лёгких и представляет собой объёмную долю паренхимы с плотностью ниже порога −950 HU, широко используемого для количественной оценки эмфизематозной деструкции [24, 34]. Формально показатель вычисляется по формуле:

$$LAA\% = \frac{\text{число вокселей внутри маски лёгких с } HU < -950}{\text{общее число вокселей внутри маски лёгких}} \times 100\% \qquad (2.2.1)$$

На основе порога LAA% формируется бинарная метка y_i для $i-$го КТ-скана. Формально операция авторазметки записывается как индикаторная функция:

$$y_i = \{1, \text{если } LAA\%\ i \geq \theta, 0, \text{если } LAA\%\ i < \theta\} \qquad (2.2.2)$$

где θ - пороговое значение LAA% (в базовой конфигурации θ=6%). Дополнительно для положительного класса требуется выполнение условия $Mean\ HU < -850$, что снижает вероятность ложной разметки из-за технических артефактов.

Полученная таким образом обучающая выборка обладает характерным для медицинских задач дисбалансом классов: доля сканов с эмфиземой (положительный класс) составляет ориентировочно 15–25 % от общего числа исследований (точное значение зависит от θ). Количественно дисбаланс выражается отношением числа примеров мажоритарного класса к миноритарному $\rho = Nneg/Npos$. Подобное распределение требует специальных мер при построении моделей машинного обучения, включая стратификацию разбиений, взвешивание функции потерь и выбор метрик, устойчивых к перекосу классов.

Такой подход представляет собой **Voxel-based quantification**, поскольку оценка эмфиземы выполняется не визуально, а на уровне каждого вокселя

лёгочной паренхимы. При этом LAA% отражает долю патологически низкоплотных участков, ассоциированных с разрушением альвеолярных стенок и увеличением воздушных пространств. В отличие от ручной визуальной оценки, данный показатель является автоматизируемым и воспроизводимым, однако зависит от качества сегментации, параметров реконструкции, уровня вдоха и корректности HU-нормализации.

На основе полученных значений LAA% формируется обучающая выборка. В базовой бинарной постановке сканы с низкой долей низкоплотных зон относятся к группе «норма/без выраженной эмфиземы», а сканы, превышающие заданный порог LAA%, - к группе «эмфизема». При многоклассовой постановке значения LAA% могут использоваться для формирования порядковых классов: норма, лёгкая, умеренная и тяжёлая эмфизема. Важно подчеркнуть, что такие классы являются не клиническими GOLD-стадиями, а **QCT-derived Silver Standard labels**, то есть классами, полученными из денситометрического распределения плотности ткани.

Подготовка изображений включает несколько обязательных этапов. Сначала DICOM-данные переводятся в шкалу Хаунсфилда с использованием параметров rescale slope и rescale intercept. Затем выполняется сегментация лёгких (*lung masking*), позволяющая исключить грудную стенку, средостение, кости и внепульмональные структуры. После этого объёмы приводятся к единому пространственному разрешению, например **1×1×1 мм³**, что необходимо для сопоставимости объёмных и процентных показателей между пациентами. Без ресемплинга один воксель у разных пациентов может соответствовать разному физическому объёму, что искажает расчёт LAA%.

Далее выполняется нормализация HU-значений и контроль качества. Для нейросетевого анализа интенсивности могут ограничиваться клинически значимым диапазоном, например от −1000 до 400 HU, чтобы уменьшить влияние экстремальных значений и стабилизировать входные данные. Для QCT-расчётов при этом сохраняется физический смысл HU-шкалы, поскольку именно абсолютная плотность вокселей определяет LAA% и Perc15. Сканы с некорректной сегментацией лёгких, выраженными артефактами, неполным охватом лёгочной ткани или технически неправдоподобными плотностными характеристиками должны исключаться либо направляться на дополнительную проверку.

Таким образом, адаптация набора данных LIDC-IDRI для задачи диагностики ХОБЛ и эмфиземы потребовала переориентации исходного массива с анализа лёгочных узлов на анализ лёгочной паренхимы. Отсутствие экспертных клинических меток компенсируется стратегией автоматической QCT-разметки, в которой показатель LAA% выступает в качестве объективного «серебряного стандарта» (Silver Standard). Такой подход позволяет задействовать крупный открытый КТ-датасет для обучения модели, объединяющей латентные визуально-текстурные признаки, извлекаемые сетью ResNet-152, и физически интерпретируемые денситометрические параметры.

2.3 Подготовка данных и извлечение признаков

Подготовка данных КТ-изображений в настоящем исследовании представляет собой последовательный алгоритмический конвейер, направленный на преобразование исходных DICOM-данных в стандартизированный набор количественных и нейросетевых признаков. Такая организация обработки необходима, поскольку используемый набор данных LIDC-IDRI изначально создавался для разработки алгоритмов поиска и классификации лёгочных узлов, а не для диагностики ХОБЛ или эмфиземы. Коллекция включает 1018 КТ-сканов и характеризуется выраженной технической неоднородностью: исследования получены на сканерах разных производителей, с различными параметрами напряжения, силы тока, толщины среза, реконструкционного интервала и ядер реконструкции [49]. Данная вариативность особенно значима для количественного анализа эмфиземы, поскольку денситометрические показатели зависят от пространственного разрешения, уровня шума, метода реконструкции, глубины вдоха и корректности калибровки шкалы Хаунсфилда [21, 60, 68].

Методологически конвейер включает несколько последовательных этапов: выбор полной DICOM-серии, сортировку срезов по координате Z, перевод изображений в шкалу Хаунсфилда, ресемплинг к изотропному вокселю, сегментацию лёгких, расчёт LAA%, генерацию слабых меток на основе QCT, декомпозицию трёхмерного объёма в двумерные патчи и извлечение признаков с помощью ResNet-152. Сходная логика конвейера представлена в работе Sarsembayeva и соавторов, где для каждого пациента выбиралась наиболее полная DICOM-серия, изображения сортировались по Z-координате, переводились в HU, проходили ресемплинг, сегментацию лёгких, расчёт LAA% и последующее извлечение ResNet-эмбеддингов [34].

Ресемплинг КТ-изображений

Первым ключевым этапом является приведение КТ-объёмов к единому изотропному разрешению 1×1×1 мм³. Необходимость данного шага обусловлена тем, что в LIDC-IDRI толщина среза и реконструкционный интервал существенно различаются между исследованиями. Без ресемплинга один воксель у разных пациентов соответствует разному физическому объёму, что делает некорректным прямое сравнение объёмных и процентных показателей, включая LAA%, TLV и Mean HU. В работе Sarsembayeva и соавторов все объёмы также ресемплировались к изотропному разрешению 1×1×1 мм³ для обеспечения межпациентной сопоставимости [34].

Современные исследования подтверждают ключевую роль стандартизации КТ-данных. Cheplygina и соавторы подчёркивают, что при многоцентровой классификации ХОБЛ данные различаются по центрам, сканерам и протоколам, что снижает переносимость моделей и требует доменной адаптации или устойчивых признаков [50]. Almeida и соавторы также показывают, что обучающие выборки и популяционные различия существенно влияют на обобщаемость DL-моделей ХОБЛ, а сбалансированные и разнообразные данные уменьшают смещение модели [51]. Следовательно, ресемплинг в настоящем

исследовании выполняет не только техническую, но и методологическую функцию: он уменьшает часть вариативности, связанной с различиями spatial spacing, и делает последующий voxel-based analysis более сопоставимым.

Для непрерывных HU-значений применяется интерполяция интенсивностей, а для дискретных бинарных масок - интерполяция по ближайшему соседу, чтобы избежать возникновения промежуточных неанатомических классов. Такой выбор соответствует различению между непрерывным КТ-объёмом и категориальной маской лёгких.

Сегментация лёгких

После ресемплинга выполняется сегментация лёгких, то есть выделение лёгочной паренхимы и исключение нерелевантных структур: грудной стенки, костей, средостения, мягких тканей и внешнего воздуха. Маскирование необходимо потому, что LAA%, Perc15, Mean HU и TLV должны рассчитываться строго в пределах лёгочного поля. Включение внешнего воздуха или костных структур привело бы к искажению HU-гистограммы и ложному увеличению доли низкоплотных зон.

Научная обоснованность данного шага подтверждается несколькими группами работ. В классическом текстурном анализе Sørensen использовал сегментацию лёгочных полей как предварительное условие для отбора ROI и расчёта текстурных признаков: лёгкие выделялись алгоритмом наращивания области при предположении, что лёгочная паренхима имеет плотность ниже −400 HU; затем из анализа исключались трахея, главные бронхи и структуры с плотностью выше −400 HU [21, 59]. Это демонстрирует, что даже до распространения U-Net-подходов автоматическая сегментация лёгочной области была обязательным этапом количественного анализа ХОБЛ.

Современные работы всё чаще используют нейросетевую сегментацию. Yang и соавторы в задаче классификации стадий ХОБЛ сначала сегментировали паренхиму лёгких на HRCT с помощью ResU-Net, затем извлекали радиомические признаки посредством PyRadiomics и использовали их в графовой нейронной сети для стадирования заболевания [6]. Zhou и соавторы также применяли полностью автоматическую сегментацию всего лёгкого перед извлечением whole-lung радиомических признаков и построением номограммы для классификации тяжести ХОБЛ [53]. В работе Sarsembayeva и соавторов сегментация выполнялась с помощью lungmask - инструмента, использующего trunk U-Net, обученный на VESSEL12 [34]. Таким образом, применение автоматической U-Net-подобной сегментации в настоящем исследовании согласуется с современными КТ-радиомическими и DL-подходами.

Дополнительно в литературе встречаются альтернативные стратегии сегментации. Mondal и соавторы применяли выравнивание гистограммы, медианную фильтрацию и кластеризацию Fuzzy C-Means для выделения лёгочных областей перед извлечением Adaptive Local Ternary Pattern и классификацией эмфиземы [69]. Karthik и соавторы описывают конвейер, включающий Z-score нормализацию, CLAHE, размытие по Гауссу, пороговую обработку и морфологическое закрытие для выделения лёгких и последующей

сегментации эмфизематозных поражений [23]. Эти работы не заменяют выбранный в настоящем исследовании lungmask/U-Net-подход, но подтверждают общий принцип: перед анализом эмфиземы необходимо анатомически ограничить область исследования лёгочной паренхимой и удалить шумовые и внепульмональные структуры.

После сегментации выполняется денситометрический анализ, основанный на воксельной количественной оценке. Основной метрикой является LAA% - доля вокселей лёгочной ткани с плотностью ниже заданного порога. В настоящем исследовании используется порог HU < −950, поскольку он является одним из наиболее распространённых критериев количественной оценки эмфизематозного поражения на инспираторной КТ [24, 34, 42]. Формально показатель вычисляется по формуле:

$$LAA\% = \frac{N(HU < -950)}{N_{Total}} \times 100\% \qquad (2.3.1)$$

где $N(HU<-950)$ - количество вокселей в маске лёгких с плотностью ниже –950 HU, а Ntotal - общее количество вокселей в маске лёгких.

Содержательно этот показатель отражает долю патологически низкоплотной лёгочной ткани. При эмфиземе разрушение альвеолярных стенок и увеличение воздушных пространств приводят к снижению рентгеновской плотности, поэтому увеличение LAA% интерпретируется как количественное выражение эмфизематозной деструкции. Gomes и соавторы рассчитывали LAA как долю объёма лёгких с плотностью ≤ −950 HU и показали связь регионального распространения эмфиземы на КТ с функциональными нарушениями у пациентов с ХОБЛ [42]. Moslemi указывает, что в количественной КТ используется несколько установленных порогов - −910 HU, −950 HU и −960 HU, причём измерения по этим порогам валидировались через сопоставление с гистологической оценкой эмфиземы [60]. Oh и соавторы также подчёркивают, что %LAA-950 является наиболее распространённой метрикой и коррелирует с гистологической выраженностью эмфиземы [24].

Современные исследования демонстрируют двойственную роль показателя LAA%-950: с одной стороны, подтверждают его применимость для стадирования эмфиземы, с другой - выявляют существенные ограничения. Zhang и соавторы используют %LAA-950 для стратификации степени эмфиземы: отсутствие эмфиземы при значениях <6%, лёгкая степень - 5–10%, умеренная - 10–20% и тяжёлая - ≥20% [45]. Эти результаты подтверждают пригодность LAA%-950 в качестве основы для автоматизированного стадирования. В то же время Peña и соавторы подчёркивают, что денситометрические показатели, основанные на LAA или перцентильной плотности, существенно зависят от уровня вдоха, параметров реконструкции, дозы облучения и типа сканера, вследствие чего в литературе отсутствует единый консенсус относительно оптимального порогового значения для диагностики эмфиземы [68]. Sørensen дополнительно указывает, что методы, основанные на анализе плотности, рассматривают воксели независимо, игнорируя пространственные взаимосвязи,

а также демонстрируют чувствительность к калибровке сканера и уровню шума [21]. Таким образом, в рамках настоящего исследования показатель LAA%-950 интерпретируется не как абсолютный клинический критерий, а как воспроизводимый физический дескриптор, используемый в качестве основы для формирования Silver Standard-разметки.

Генерация меток на основе Auto-labeling

Поскольку LIDC-IDRI не содержит специализированных экспертных меток ХОБЛ, GOLD-стадий или стадий эмфиземы, классическое обучение с учителем на экспертной разметке невозможно. Для решения этой проблемы используется автоматическая разметка на основе LAA%, описанного выше. Такая разметка трактуется как «серебряный стандарт» (Silver Standard), то есть как воспроизводимый суррогатный стандарт, основанный на количественной КТ-денситометрии, но не равный полноценной клинической, спирометрической или гистопатологической верификации.

Данная стратегия согласуется с направлением обучения со слабой разметкой (weakly supervised learning). Sørensen и соавторы использовали метаданные, связанные с КТ-изображением, для обучения текстурной модели ХОБЛ без ручной разметки ROI; в их работе метки основывались на PFT/GOLD-статусе, а итоговая вероятность ХОБЛ получалась через объединение вероятностей по ROI [21]. Cheplygina и соавторы формализовали аналогичную проблему как обучение с множественными экземплярами (multiple instance learning), где известна только глобальная метка скана, но неизвестно, какие именно патчи содержат признаки заболевания [54]. Braman и соавторы также анализировали трёхмерные объёмы как последовательности двумерных изображений при наличии только бинарных меток на уровне объёма, что позволило обучать модель без детальной локальной аннотации [55]. Bortsova и соавторы развивают близкую идею через обучение по пропорциям меток: модель обучается по слабым меткам, отражающим интервал доли поражённой ткани, и фактически учится сегментировать поражение так, чтобы его объём соответствовал заданной пропорции [67]. Эти работы подтверждают, что слабая разметка является не вынужденным компромиссом, а признанной методологией для задач, где детальная медицинская аннотация трудоёмка и вариабельна.

В настоящем исследовании auto-labeling выполняется следующим образом: после расчёта LAA% каждому КТ-скану присваивается метка, отражающая степень выраженности низкоплотных зон. В бинарной постановке сканы с низким значением LAA% относятся к классу «норма/без выраженной эмфиземы», а сканы с повышенным значением LAA% - к классу «эмфизема». В рабочем варианте конвейера используется правило LAA% $\geq$ 6% при дополнительном контроле средней плотности лёгких, например Mean HU < -850, что снижает риск ложной положительной разметки, связанной с техническими артефактами или неполной сегментацией [34]. При этом важно подчеркнуть, что порог 6% является операционной границей данной модели, близкой к диапазону лёгкой эмфиземы в шкалах %LAA-950, но не должен трактоваться как универсальный клинический консенсус. Категории, сформированные по LAA%,

являются QCT-производными метками, а не GOLD-стадиями, поскольку GOLD основан на спирометрии, тогда как LAA% отражает морфологически-денситометрическую выраженность низкоплотной лёгочной ткани.

Подготовка патчей: декомпозиция 3D-объёма в 2D-представления

Следующим этапом является декомпозиция трёхмерного КТ-объёма в набор информативных двумерных срезов или патчей. Этот шаг обусловлен архитектурой ResNet-152, которая изначально оптимизирована для двумерных изображений. Поэтому исходный трёхмерный объём преобразуется в набор двумерных представлений, пригодных для подачи на вход сети.

Научная обоснованность patch-based подхода подтверждается рядом близких работ. Sørensen в текстурном анализе ХОБЛ использовал случайно выбранные, возможно перекрывающиеся кубические ROI внутри сегментированных лёгочных полей; такие ROI представляли весь КТ-скан и использовались для построения текстурных гистограмм и последующего объединения вероятностей [21, 59]. Cheplygina и соавторы рассматривали КТ-скан как набор экземпляров, где отдельные ROI/патчи не имеют точной локальной метки, а диагноз известен только на уровне изображения [54]. Braman и соавторы предложили анализировать трёхмерные медицинские объёмы как последовательности двумерных изображений, что позволило обучаться на слабых бинарных метках и при этом учитывать распределение заболевания по объёму [55]. Ørting и соавторы показали, что CNN может извлекать релевантные заболеванию признаковые представления из отдельных КТ-срезов для оценки эмфиземы, используя визуальное сходство и triplet learning [56].

В настоящем конвейере сначала определяется Z-диапазон лёгочной ткани по маске, затем выбираются информативные уровни лёгких. Из этих областей формируются двумерные патчи, содержащие преимущественно лёгочную паренхиму. Согласно описанию применяемой методики, выбираются три региональные зоны - верхняя, средняя и нижняя части лёгкого; для каждой зоны формируется трёхканальное представление из соседних аксиальных срезов, которое затем обрезается по bounding box лёгких и масштабируется до размера 224×224×3 [34]. Такой подход одновременно решает две задачи: адаптирует объёмные КТ-данные к 2D-CNN-архитектуре и увеличивает число обучающих примеров без искусственного изменения класса пациента.

Современные исследования предлагают различные стратегии обработки объёмных КТ-данных, отражающие компромисс между полнотой трёхмерного контекста и вычислительной эффективностью. Nagaraj и соавторы применяют специфичную для заболевания аугментацию на основе проекции минимальной интенсивности (minIP) с толщиной слэба от 1 до 11 мм; наилучшие результаты достигнуты при толщине 11 мм, что позволило повысить чувствительность выявления эмфиземы и снизить число ложноотрицательных предсказаний [61]. В противоположность этому Argus и соавторы предлагают сквозной подход, основанный на обработке полного трёхмерного объёма без явного задания признаков, осуществляя прогноз как визуальных оценок эмфиземы, так и функциональных показателей по полным торакальным КТ-сканам [65]. Таким

образом, оба подхода подтверждают значимость трёхмерного контекста, однако различаются по вычислительным требованиям и архитектурной сложности. В рамках настоящего исследования выбран компромиссный подход: трёхмерный объём декомпозируется на двумерные патчи, что позволяет использовать проверенную архитектуру ResNet-152, снизить вычислительную нагрузку и одновременно сохранить региональную структуру лёгочной ткани.

При этом разбиение на обучающую, валидационную и тестовую выборки должно выполняться на уровне пациента, а не на уровне патча. Это принципиально важно, поскольку патчи одного пациента могут быть статистически похожи; их попадание в разные выборки привело бы к утечке данных и необоснованному завышению качества модели. В работе Sarsembayeva и соавторов прямо используется стратификация на уровне пациента, а также взвешивание классов для учёта дисбаланса положительных и отрицательных случаев [34]. Guan и соавторы также подчёркивают проблему классового дисбаланса при LDCT-скрининге ХОБЛ и используют веса классов или взвешенный случайный сэмплер для уменьшения смещения модели в сторону большинства [66].

После фильтрации сканов с дефектами сегментации, артефактами или неполным охватом лёгочной ткани итоговый объём выборки, использованной для обучения и тестирования моделей, составил N исследований (точное значение определяется в экспериментальной части). Полученные патчи и соответствующие им QCT-метки формируют основу для обучения ResNet-152 как экстрактора признаков и построения интегральной диагностической модели в рамках данного исследования.

Извлечение признаков ResNet-152

После подготовки патчей выполняется извлечение признаков с помощью ResNet-152. В настоящем исследовании ResNet-152 используется как экстрактор признаков (feature extractor), а не как полностью автономный end-to-end диагностический классификатор. Из слоя глобального усреднения (global average pooling) извлекается 2048-мерный эмбеддинг, представляющий латентные визуально-текстурные признаки КТ-патча. Эти признаки отражают не диагноз как таковой, а вычислительное представление изображения: локальную неоднородность, текстуру паренхимы, контрастные переходы, распределение низкоплотных зон и пространственную организацию лёгочного рисунка.

Такой подход обоснован несколькими направлениями литературы. Sørensen показал, что текстурный анализ КТ лучше различает ХОБЛ, чем простые денситометрические показатели, поскольку использует информацию о пространственной организации ткани [21]. Yang и соавторы продемонстрировали, что неконтролируемые текстурные прототипы могут быть визуально однородными, воспроизводимыми и способными предсказывать радиологические подтипы эмфиземы, что подтверждает ценность текстурных представлений лёгочной ткани [57]. Ørting и соавторы также показали, что CNN может обучать представления, релевантные тяжести эмфиземы, даже на основе производных триплетов визуального сходства [56].

Наиболее близкой к настоящей методике является работа Sarsembayeva и соавторов: ResNet-152 используется как замороженный экстрактор, из патчей средней зоны лёгких извлекаются 2048-мерные эмбеддинги, после чего применяются логистическая регрессия, LASSO и рекурсивное исключение признаков; далее эмбеддинги объединяются с QCT-маркерами %LAA, Perc15 и TLV [34]. Такая организация соответствует целевой логике диссертации: нейросеть извлекает латентные визуально-текстурные признаки, а QCT-компонент сохраняет физическую интерпретируемость результата. Humphries и соавторы дополнительно показали, что глубокое обучение может автоматически классифицировать эмфизематозные паттерны на КТ по шкале Fleischner и связывать полученные оценки с функциональными показателями и прогнозом [46].

Ряд работ уточняют, почему extraction-based стратегия предпочтительна для решаемой задачи. Guan и соавторы сравнили модели на основе радиомики и машинного обучения (radiomics-based ML) с моделями глубокого обучения, работающими непосредственно с изображениями (image-based DL), на низкодозовых КТ и показали, что модель, объединяющая радиомические признаки, параметры плотности лёгких и клинические характеристики, превосходила DL-модель, обученную только на изображениях. Авторы связывают это с ограниченностью размеченных данных, характером «чёрного ящика» глубоких сетей и большей сфокусированностью радиомики на лёгочной области [66]. Таким образом, использование ResNet-152 не в качестве закрытого end-to-end классификатора, а как генератора эмбеддингов с последующим интерпретируемым анализом является методологически оправданным.

Итоговая логика pipeline

В результате препроцессинг и извлечение признаков формируют два взаимодополняющих признаковых пространства. Первое - физически интерпретируемое QCT-пространство, включающее LAA%, Perc15, TLV и Mean HU. Второе - латентное нейросетевое пространство, представленное эмбеддингами ResNet-152. Такая организация методологически обоснована: QCT фиксирует объективные денситометрические параметры плотности ткани, тогда как CNN извлекает визуально-текстурные паттерны, которые не сводятся к одному HU-порогу.

Следовательно, применяемый конвейер - ресемплинг, сегментация, расчёт LAA%, auto-labeling, декомпозиция 3D→2D и извлечение ResNet-признаков - не является произвольной технической последовательностью. Каждый этап подтверждается близкими работами по количественной КТ, обучению со слабой разметкой, множественному обучению с экземплярами, patch-based анализу, радиомике и глубокому обучению применительно к ХОБЛ и эмфиземе [21, 34, 52, 54, 55, 60, 66, 67, 68]. При этом литература также указывает на ограничения выбранного подхода: LAA%-950 зависит от параметров сканирования и уровня вдоха, слабые метки не заменяют экспертную клиническую верификацию, а двумерные патчи частично теряют трёхмерный контекст. Именно поэтому в дальнейшем модель должна оцениваться не только по точности и AUC, но и через

абляционные исследования, проверку устойчивости к протоколам КТ, анализ интерпретируемости и внешнюю валидацию.

2.4 Методы машинного обучения

В настоящем исследовании методы машинного обучения применяются для построения интегративной модели автоматизированной диагностики эмфиземы по данным КТ, объединяющей два типа признаков: латентные визуально-текстурные представления, формируемые глубокой свёрточной нейронной сетью ResNet-152 в виде компактных числовых векторов (эмбеддингов), и количественные КТ-показатели (LAA%, Perc15, TLV, Mean HU), обладающие прямой физической интерпретацией в шкале Хаунсфилда. Такой подход позволяет совместить «зрительное» описание паттернов эмфизематозной деструкции, извлекаемое сетью, с объективными денситометрическими характеристиками ткани. Выбор архитектуры ResNet-152 обусловлен её значительной глубиной, обеспечивающей способность улавливать тонкие текстурные изменения лёгочной паренхимы, характерные для ранних стадий эмфиземы, которые могут быть пропущены более мелкими сетями (например, ResNet-18 или ResNet-50). Данный подход соответствует актуальной тенденции анализа КТ при ХОБЛ, в рамках которой изолированная денситометрия дополняется текстурным анализом, радиомикой, глубоким обучением и комбинированными моделями [21, 24, 34, 70].

ResNet-152 как глубокий экстрактор признаков

В качестве нейросетевого компонента используется архитектура ResNet-152, рассматриваемая не как самостоятельный end-to-end диагностический классификатор, а как глубокий экстрактор признаков (feature extractor). В данной постановке сеть преобразует входные КТ-патчи в высокоуровневые числовые представления - эмбеддинги, которые затем совместно с количественными QCT-признаками формируют итоговое признаковое пространство для классификации.

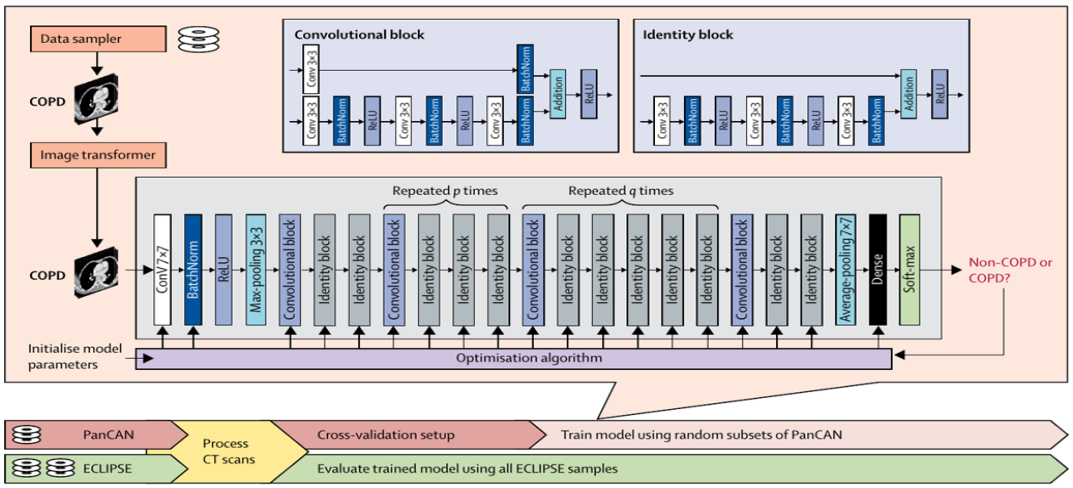


Рисунок 2.4 - Предварительно обученная архитектура ResNet-152, используемая в качестве экстрактора признаков

Показаны входной свёрточный слой, каскад остаточных блоков (свёртка, пакетная нормализация, ReLU), слой глобального усреднения и выходной 2048-мерный эмбеддинг. Источник: [34, 70]

Выбор ResNet-архитектуры обоснован принципом остаточного обучения (residual learning). Остаточные связи (skip connections) позволяют передавать сигнал через глубокие блоки сети, уменьшая риск деградации градиента при увеличении числа слоёв. Это особенно важно для анализа КТ лёгких, поскольку признаки эмфиземы могут проявляться не только как крупные низкоплотные зоны, но и как тонкие визуально-текстурные изменения: локальная неоднородность паренхимы, нарушение сосудисто-паренхиматозного рисунка, региональное разрежение ткани и сложная пространственная организация низкоплотных участков [21, 70].

Дополнительную стабилизацию обучения обеспечивает пакетная нормализация (Batch Normalization), применяемая в остаточных блоках для нормализации промежуточных активаций. В условиях КТ-данных это имеет практическое значение, поскольку изображения могут различаться по уровню шума, реконструкционным параметрам и плотностному распределению. Использование предобученной сети соответствует стратегии трансферного обучения (Transfer Learning): модель, предварительно обученная на крупном наборе естественных изображений, применяется для извлечения универсальных визуальных признаков, а не обучается с нуля на ограниченной медицинской выборке. Такая стратегия снижает риск переобучения и согласуется с работами, в которых перенос обучения и адаптация моделей рассматриваются как важные условия работы с многоцентровыми КТ-данными [34, 50, 51].

Наиболее близкой к настоящей методологии является работа Sarsembayeva и соавторов, где ResNet-152 использовалась как фиксированный экстрактор признаков: из КТ-патчей извлекались 2048-мерные эмбеддинги, которые затем объединялись с QCT-показателями для интерпретируемой диагностики эмфиземы [34]. Tang и соавторы продемонстрировали эффективность остаточных сетей для выявления ХОБЛ по низкодозовой КТ, причём модель на основе ResNet превосходила изолированное использование %LAA-950, что подтверждает наличие в глубоких признаках диагностической информации, не сводимой к единственному денситометрическому порогу [70].

Формирование embeddings

После предварительной обработки КТ-объёма и формирования 2D-патчей каждый патч подаётся на вход ResNet-152. Из финальных слоёв сети, предшествующих классификационному, извлекается вектор высокоуровневых признаков. В соответствии с применяемой методикой [34] используется выход слоя глобального усреднения (global average pooling), формирующий 2048-мерный эмбеддинг.

Данный эмбеддинг представляет собой компактное латентное описание визуальной структуры КТ-патча. Он не является прямым патологическим маркером и не должен интерпретироваться как непосредственное измерение альвеолярной деструкции. Корректнее рассматривать его как латентный

визуально-текстурный радиологический признак, отражающий распределение контрастов, локальную неоднородность, конфигурацию низкоплотных зон и особенности лёгочного рисунка. Клиническая значимость таких признаков проявляется после сопоставления с QCT-метриками, автоматически сформированными метками и результатами классификации.

Ценность извлечения подобных признаков подтверждается исследованиями текстурного и глубокого анализа КТ. Sørensen и соавторы показали, что текстурные признаки лучше различают ХОБЛ, чем простые денситометрические показатели, поскольку учитывают пространственную организацию лёгочной ткани [21]. Xie и соавторы применяли глубокие нейронные сети для подтипирования эмфиземы на КТ, что подтверждает применимость CNN-признаков к анализу сложных эмфизематозных паттернов [35]. Fuhrman и соавторы использовали attention-based multiple instance learning для оценки эмфиземы на низкодозовой КТ, дополнительно обосновывая патчевый и срезовый анализ при отсутствии детальной локальной разметки [71].

Интеграция QCT-признаков

Параллельно с извлечением CNN-эмбеддингов рассчитываются количественные КТ-признаки. В качестве физических предикторов используются показатели, подробно охарактеризованные в разделах 1.1.3 и 1.2: доля низкоплотных участков (LAA%, порог –950 HU), 15-й перцентиль распределения плотности (Perc15), общий объём лёгких (TLV) и средняя плотность лёгочной ткани (Mean HU) [24, 42]. Далее эти метрики обозначаются как вектор *zQCT*.

Перед включением в модель QCT-признаки должны быть нормализованы. Это необходимо потому, что LAA%, Perc15, TLV и Mean HU имеют различные шкалы измерения: LAA% выражается в процентах, TLV - в миллилитрах, а Perc15 и Mean HU - в единицах Хаунсфилда. Без нормализации признаки с большей числовой амплитудой могут непропорционально повлиять на обучение классификатора. Поэтому QCT-вектор стандартизируется, например посредством z-score нормализации или минимакс-масштабирования, и только после этого объединяется с CNN-эмбеддингами.

Интеграция QCT-показателей имеет принципиальное методологическое значение: они вводят в модель физически интерпретируемый слой признаков. CNN-эмбеддинги описывают визуальный паттерн поражения, тогда как QCT фиксирует измеряемые параметры плотности ткани. Это снижает риск того, что модель будет опираться исключительно на неявные визуальные закономерности, шум или особенности протокола сканирования [24, 42].

Учёт дисбаланса классов в функции потерь

Ввиду установленного на этапе авторазметки дисбаланса классов (раздел 2.2) применение стандартной бинарной кросс-энтропии привело бы к смещению модели в сторону мажоритарного (отрицательного) класса. Для его коррекции используется взвешенная бинарная кросс-энтропия (Weighted Binary Cross-Entropy, WBCE):

$$LWBCE = -\frac{1}{N}\sum i \sum_{i=1}^{N} [w_1 y_i log(\widehat{y_i}) + w_0(1 - y_i i) log(1 - \widehat{y_i})] \quad (2.4.1)$$

где $y_i \in \{0,1\}$ - истинная метка, $\widehat{y_i}$ - предсказанная вероятность наличия эмфиземы, N - объём батча, а веса определяются как $w_1 = \frac{N}{2N_{pos}}, w_0 = \frac{N}{2N_{neg}}$. При таком взвешивании ошибки на миноритарном (положительном) классе штрафуются сильнее, что способствует повышению чувствительности модели к случаям эмфиземы.

В альтернативной конфигурации, где финальным классификатором выступает L1-регуляризованная логистическая регрессия (LASSO), минимизируется регуляризованная функция потерь, также учитывающая веса классов:

$$pi = \sigma(\beta_0 + z_{Fusion}^T, \beta_i \quad (2.4.2)$$

где {β_i|}- модельная вероятность, а λ - параметр регуляризации, подбираемый перекрёстной проверкой.

Концепция Late Fusion

Ключевым элементом предлагаемой модели является Late Fusion - объединение признаков после их независимого извлечения. На первом этапе ResNet-152 формирует вектор нейросетевых признаков:

$$z_{CNN} = f_{ResNet-152}(X_{patch}) \quad (2.4.3)$$

На втором этапе из сегментированного КТ-объёма рассчитывается вектор количественных признаков. Где x_1 =LAA%, $x_2 = \text{Perc15}$, x_3 = Mean HU

$$z_{QCT} \in R^3 \qquad z_{QCT} = [x_1, x_2, x_3] \quad (2.4.4)$$

После нормализации QCT-признаков оба представления объединяются через конкатенацию:

$$z_{Fusion} = [z_{CNN}; z_{QCT}] \quad (2.4.5)$$

Полученный объединённый вектор поступает в финальный классификационный блок. Таким образом, модель является мультимодальной на уровне признаков: она объединяет визуально-текстурный паттерн, извлекаемый CNN, и физические характеристики ткани, представленные денситометрическими QCT-показателями. Такая схема отличается от раннего слияния (early fusion), где данные объединяются на уровне входного изображения; в late fusion каждое представление сначала формируется в собственной оптимальной форме, а затем интегрируется в общем признаковом пространстве.

Работа Sarsembayeva и соавторов прямо подтверждает применимость такой логики: ResNet-152 эмбеддинги объединялись с QCT-маркерами %LAA, Perc15 и

TLV, после чего использовались интерпретируемые классификационные модели [34]. Guan и соавторы показали, что комбинированные модели, включающие КТ-изображения, радиомические признаки, параметры плотности лёгких и клинико-демографические характеристики, могут превосходить модели, основанные только на одном типе данных [66]. Yang и соавторы также продемонстрировали эффективность сочетания сегментации лёгочной паренхимы, радиомических признаков, LASSO-отбора и графовой нейронной сети для классификации стадий ХОБЛ [52]. Zhou и соавторы показали, что радиомическая номограмма, объединяющая КТ-признаки и клинические факторы, применима для стратификации тяжести ХОБЛ [53]. Эти работы подтверждают методологическую ценность подхода, основанного на слиянии разнородных признаков.

Регуляризация и оптимизация

Для защиты модели от переобучения в классификационном блоке используется Dropout. Его задача состоит в случайном отключении части нейронов во время обучения, что препятствует чрезмерной зависимости модели от отдельных признаков и повышает устойчивость классификации. Это особенно важно при работе с медицинскими изображениями, где объём выборок обычно ограничен, а размерность CNN-эмбеддингов велика.

Для настройки весов классификационного блока применяется оптимизатор Adam, сочетающий адаптивную скорость обучения с оценкой первого и второго моментов градиента. Adam удобен для интегративных моделей, поскольку объединённое признаковое пространство содержит признаки разной природы: высокоразмерные CNN-эмбеддинги и низкоразмерные физические QCT-показатели. В предлагаемой архитектуре Adam используется для обучения полносвязных слоёв после объединения признаков, тогда как ResNet-152 применяется в замороженном режиме в качестве экстрактора признаков.

Данная стратегия обучения соответствует общей логике трансферного обучения: глубокая свёрточная сеть используется для генерации устойчивых визуально-текстурных представлений, а обучаемая часть модели ограничивается классификационным блоком, принимающим объединённый fusion-вектор. Это снижает вычислительную нагрузку и уменьшает риск переобучения на ограниченной медицинской выборке [34, 50, 51].

Итоговое описание модели

Таким образом, машинное обучение в настоящем исследовании реализуется в виде архитектуры, объединяющей ResNet-152, QCT и Late Fusion. ResNet-152 извлекает латентные визуально-текстурные эмбеддинги из КТ-патчей. QCT-блок формирует физически интерпретируемые признаки LAA%, Perc15, TLV и Mean HU. Затем оба признаковых пространства нормализуются и объединяются через конкатенацию в единый fusion-вектор, который поступает в классификатор.

Предлагаемый подход обладает несколькими преимуществами. Во-первых, он использует способность CNN распознавать сложные визуальные паттерны лёгочной ткани. Во-вторых, он сохраняет связь с физическими параметрами КТ

за счёт QCT-признаков. В-третьих, late fusion повышает интерпретируемость модели, поскольку итоговое решение может быть сопоставлено как с латентными признаками изображения, так и с объективными денситометрическими показателями. В-четвёртых, применение Batch Normalization, трансферного обучения, Dropout и Adam делает архитектуру более устойчивой к ограниченному объёму медицинских данных и технической вариабельности КТ.

2.5 Архитектура модифицированной модели

Программная архитектура разработанной диагностической системы реализована в виде многоэтапного конвейера, объединяющего предварительную обработку КТ-изображений, параллельное извлечение нейросетевых и денситометрических признаков, а также два классификационных контура, ориентированных на решение задач интерпретируемости и итоговой диагностики соответственно. Структурная схема системы представлена на Рисунке 2.5 и включает следующие основные блоки: предобработку и сегментацию лёгких, нейросетевой модуль на основе ResNet-152, модуль количественной денситометрии (QCT), блок автоматической разметки, два классификационных маршрута и модуль формирования заключения.

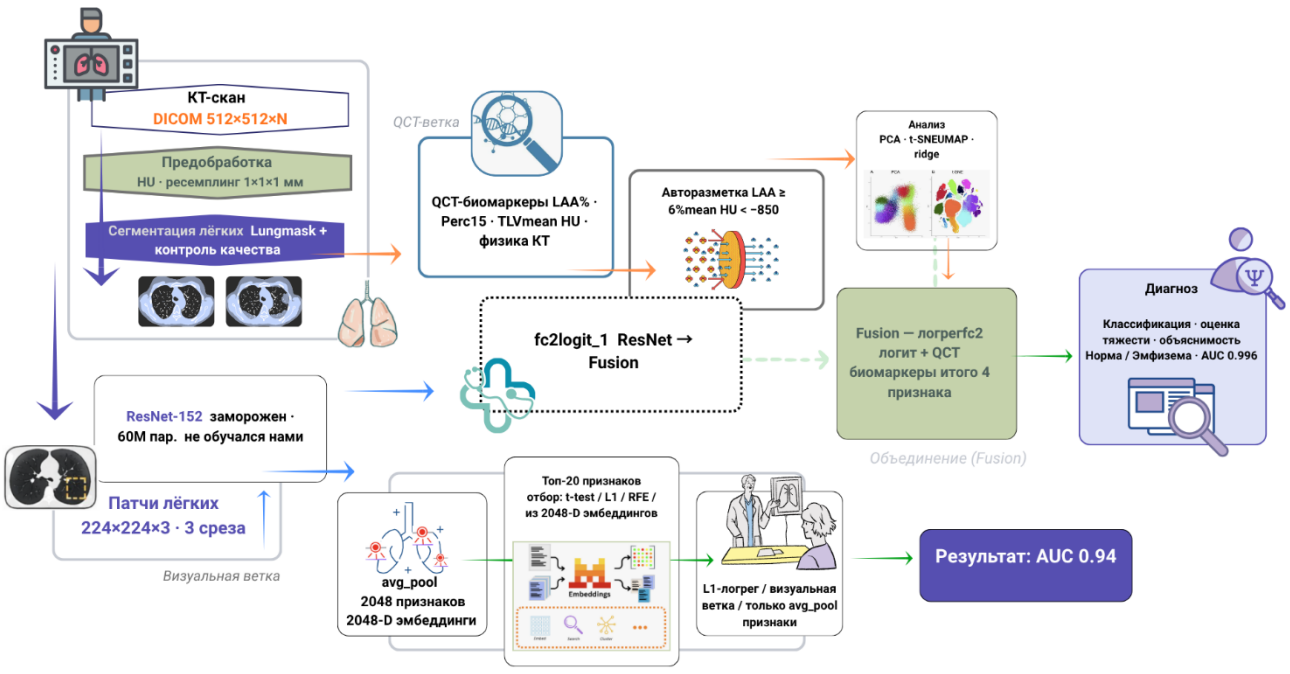


Рисунок 2.5- Архитектура программной системы: от входного КТ-скана до диагностического заключения

Показаны два классификационных пути: исследовательский (L1-логистическая регрессия на 20 отобранных компонентах эмбеддинга) и итоговый (Fusion-модель, объединяющая скалярный логит ResNet и три QCT-биомаркера).

На вход системы подаётся торакальный КТ-скан в формате DICOM (матрица 512×512×N срезов). На этапе предобработки значения переводятся в

единицы Хаунсфилда, после чего выполняется ресемплинг к изотропному разрешению 1×1×1 мм³. Сегментация лёгочной паренхимы осуществляется автоматически с помощью инструмента Lungmask, основанного на архитектуре U-Net; сканы с артефактами сегментации или неполным охватом исключаются по результатам контроля качества.

Нейросетевой модуль функционирует следующим образом. Из сегментированного объёма формируются двумерные патчи размером 224×224×3 пикселя, представляющие собой три смежных аксиальных среза из верхней, средней и нижней зон лёгких. Каждый патч подаётся на вход предварительно обученной сети ResNet-152, веса которой заморожены (около 60×10^6 параметров) и в ходе настоящего исследования не обновлялись. С выхода слоя глобального усреднения (avg_pool) извлекается 2048-мерный вектор эмбеддинга, компактно кодирующий визуально-текстурные особенности патча.

Параллельно модуль количественной денситометрии (QCT) вычисляет на основе полной маски лёгких набор физически интерпретируемых показателей: долю низкоплотных зон LAA% (порог –950 HU), 15-й перцентиль распределения плотности Perc15, общий объём лёгких TLV и среднюю плотность Mean HU. Первые три из них впоследствии включаются в итоговый классификатор.

Для формирования обучающих меток используется процедура авторазметки, опирающаяся исключительно на объективные денситометрические критерии: скану присваивается положительный класс «эмфизема» при одновременном выполнении условий LAA% ≥ 6 % и Mean HU < –850 HU. Такой подход позволяет генерировать стандартизованные метки («серебряный стандарт») без привлечения экспертной разметки.

Ключевой особенностью архитектуры является наличие двух независимых классификационных маршрутов, обрабатывающих один и тот же 2048-мерный эмбеддинг, но преследующих различные цели.

Интерпретируемый исследовательский классификатор. Из исходного 2048-мерного вектора с помощью комбинации статистических критериев (t-критерий, L1-регуляризация, рекурсивное исключение признаков) отбираются 20 наиболее информативных компонент. На этом сокращённом признаковом подпространстве обучается самостоятельная модель логистической регрессии с L1-регуляризацией. Данный классификатор, достигающий площади под ROC-кривой около 0,94, не предназначен для финальной диагностической задачи; его роль состоит в идентификации конкретных детекторов (нейронов) ResNet-152, наиболее тесно связанных с эмфизематозными изменениями. Это открывает возможность содержательной интерпретации нейросетевых признаков: исследователь получает информацию о том, какие именно визуальные паттерны, закодированные в отобранных компонентах эмбеддинга, вносят основной вклад в распознавание патологии.

Итоговый Fusion-классификатор. Второй маршрут предусматривает агрегирование информации, содержащейся в полном эмбеддинге, до единственного скалярного показателя. Для этого к выходу avg_pool подключается дополнительный полносвязный слой (fc2) с одним выходным

нейроном, формирующий логит *zCNNlogit* . Данный логит калибруется, после чего объединяется с тремя нормализованными QCT-биомаркерами - LAA%, Perc15 и TLV - в единый вектор признаков zFusion=[zCNNlogit,LAA%,Perc15,TLV],LAA%,Perc15,TLV]. На этом компактном четырёхмерном пространстве обучается логистическая регрессия (Fusion-логгер), выполняющая итоговую классификацию. Предварительная оценка эффективности данной модели показывает AUC на уровне 0,996, что существенно превосходит результаты изолированного использования нейросетевых или денситометрических признаков. Ограниченное число входных переменных обеспечивает прозрачность решения: каждый признак имеет однозначную клиническую либо вычислительную интерпретацию, что критически важно для практического применения в скрининговых задачах.

Таким образом, предлагаемая архитектура разделяет задачи достижения максимальной точности и обеспечения интерпретируемости. Итоговый Fusion-классификатор формирует диагностическое заключение (норма / эмфизема), а параллельный исследовательский контур предоставляет инструмент для анализа значимости отдельных компонент нейросетевого представления. Сочетание этих двух контуров создаёт основу для создания одновременно высокоточной и объяснимой системы поддержки принятия врачебных решений.

2.6 Этические аспекты исследования

Настоящее исследование относится к категории вторичного анализа медицинских изображений, поскольку работа проводится не с пациентами напрямую, а с уже существующим открытым набором КТ-данных LIDC-IDRI. Данный датасет был сформирован как многоцентровая коллекция торакальных КТ-исследований для разработки и валидации алгоритмов компьютерного анализа лёгочных узлов; изображения были предварительно деидентифицированы и опубликованы для научных целей [49]. Следовательно, исследование не предполагает дополнительного лучевого воздействия, вмешательства в лечебный процесс, назначения диагностических процедур или контакта с пациентами.

Ключевым этическим условием работы с LIDC-IDRI является соблюдение принципов конфиденциальности и защиты персональных данных. Несмотря на то что изображения в открытых медицинских коллекциях проходят процедуру удаления персональных идентификаторов, КТ-данные остаются чувствительной медицинской информацией. Поэтому при обработке данных не должны предприниматься попытки повторной идентификации пациентов, сопоставления изображений с внешними базами или восстановления персональных сведений. Все результаты анализа должны представляться только в агрегированном виде: через метрики качества модели, статистические показатели и обобщённые визуализации, исключающие возможность идентификации конкретного человека.

Отдельного внимания требует этический статус автоматически сформированных меток. Поскольку LIDC-IDRI не содержит специализированной экспертной разметки по ХОБЛ, эмфиземе или GOLD-стадиям, в исследовании используется auto-labeling на основе QCT-показателя LAA%-950. Такая разметка должна трактоваться как **Silver Standard**, а не как клинический диагноз. Это принципиально важно, поскольку LAA%-950 является количественным суррогатным показателем низкоплотных зон лёгочной ткани, но не заменяет экспертную радиологическую оценку, спирометрию или клиническое заключение. В литературе подчёркивается, что weak labels, multiple instance learning и learning from label proportions позволяют работать с неполной разметкой, но сохраняют риск шумных или неточных меток [55, 67, 68]. Поэтому результаты модели не должны интерпретироваться как окончательная медицинская диагностика.

Этическая значимость интерпретируемости особенно высока в исследованиях медицинского искусственного интеллекта. Модель, основанная только на скрытых нейросетевых признаках, может демонстрировать высокую точность, но оставаться трудной для клинического объяснения. В настоящем исследовании этот риск частично снижается за счёт гибридной архитектуры: ResNet-152 извлекает латентные визуально-текстурные признаки, а QCT-компонент вводит физически интерпретируемые показатели - LAA%, Perc15 и Mean HU. Такой подход повышает прозрачность результата, поскольку решение модели может быть сопоставлено не только с нейросетевым embedding, но и с объективными денситометрическими параметрами плотности лёгочной ткани [24, 66].

Важным этическим аспектом является риск алгоритмического смещения. КТ-исследования в LIDC-IDRI были получены в разных центрах, на разных сканерах и с различными техническими параметрами. Такая неоднородность полезна для обучения более устойчивых моделей, однако она также может приводить к смещению, если алгоритм начинает учитывать не признаки заболевания, а особенности протокола сканирования, реконструкции или качества изображения. Работы по multi-center classification и generalization показывают, что переносимость моделей между центрами и популяциями остаётся одной из ключевых проблем медицинского ИИ [50, 51]. Поэтому при разработке модели необходимо использовать разбиение данных на уровне пациента, контроль утечки данных, анализ устойчивости к технической вариативности и, по возможности, внешнюю валидацию.

Также необходимо учитывать риск чрезмерной автоматизации диагностического решения. Разрабатываемая система должна рассматриваться как инструмент поддержки анализа КТ-изображений, а не как замена врачу-рентгенологу или пульмонологу. Алгоритм может помочь выявить подозрительные признаки эмфиземы, количественно оценить низкоплотные зоны и повысить воспроизводимость анализа, однако клиническое решение должно приниматься специалистом с учётом анамнеза, симптомов, спирометрии, факторов риска и других данных пациента.

Таким образом, этическая рамка исследования включает несколько принципов: использование только обезличенных открытых данных, отказ от повторной идентификации пациентов, представление результатов в агрегированном виде, корректную трактовку auto-labeling как Silver Standard, обеспечение интерпретируемости модели, контроль алгоритмического смещения и позиционирование системы как вспомогательного инструмента, а не самостоятельного диагностического субъекта. Соблюдение этих условий позволяет проводить исследование в соответствии с принципами конфиденциальности, научной добросовестности и ответственного применения искусственного интеллекта в медицине.

3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В рамках настоящего эксперимента под автоматическим выявлением признаков ХОБЛ понимается не постановка клинического диагноза ХОБЛ, а выявление КТ-признаков эмфизематозно-подобных изменений лёгочной паренхимы, являющихся одним из морфологических компонентов ХОБЛ.

3.1 Представление полученных данных

Исходной эмпирической базой исследования послужил открытый набор торакальных компьютерных томограмм LIDC-IDRI, представленный в формате DICOM. Первоначальное назначение данного набора связано с задачами анализа лёгочных узлов, поэтому он не содержит специализированной клинической разметки по хронической обструктивной болезни лёгких, стадиям эмфиземы или спирометрическим показателям. В связи с этим для целей настоящего исследования был сформирован самостоятельный аналитический массив, включающий результаты извлечения КТ-объёмов, контроля HU-диапазона, ресемплинга, сегментации лёгких, расчёта денситометрических показателей и автоматического присвоения меток на основе LAA%-950.

Обработка данных была организована как последовательный вычислительный цикл: выбор DICOM-серии, перевод изображений в шкалу Хаунсфилда, приведение к единому пространственному разрешению, выделение лёгочной маски, расчёт доли низкоплотных вокселей, формирование исследовательских классов, подготовка двумерных патчей и последующее извлечение признаков ResNet-152. Такая структура позволила перейти от исходных неоднородных КТ-исследований к стандартизированному набору количественных и нейросетевых признаков.

3.1.1 Отбор КТ-серий и оценка полноты исходных данных

На первом этапе для каждого пациента анализировались доступные DICOM-серии. Поскольку один и тот же случай в LIDC-IDRI мог включать несколько серий с различным числом срезов, для дальнейшей обработки выбиралась наиболее полная серия, обеспечивающая максимальное анатомическое покрытие грудной клетки. Этот шаг был принципиален для корректного расчёта объёмных показателей, включая общий объём лёгких и долю низкоплотных участков, поскольку неполное представление лёгочных полей могло бы привести к смещению денситометрических характеристик.

Распределение числа срезов в отобранных сериях представлено на рисунке 3.1.1 всего на данном этапе было проанализировано 1018 КТ-серий. Медианное число срезов составило 201, межквартильный диапазон - от 133 до 288 срезов, минимальное значение - 48, максимальное - 764. Распределение характеризовалось выраженной неоднородностью: значительная часть серий была сосредоточена в диапазоне 100–300 срезов, однако присутствовали и более протяжённые исследования, превышающие 400–500 срезов. Это подтверждает техническую вариативность исходного набора и обосновывает необходимость предварительного отбора серии перед расчётом количественных признаков.

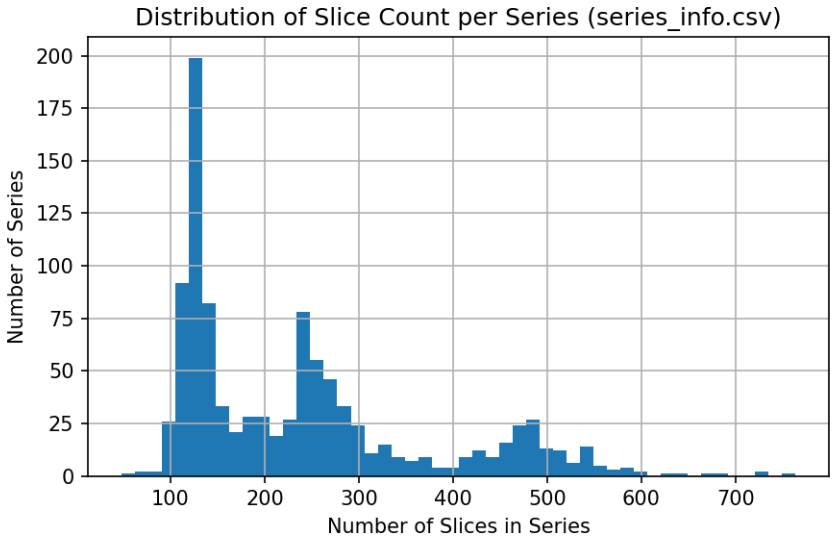


Рисунок 3.1.1 - Распределение числа срезов в отобранных КТ-сериях.

3.1.2 Контроль HU-диапазона и исключение технически проблемных случаев

После выбора DICOM-серий изображения переводились в шкалу Хаунсфилда. Для контроля корректности извлечения объёмов рассчитывались статистические характеристики интенсивности: среднее значение HU, стандартное отклонение HU и предельные значения плотности. Такой контроль был необходим, поскольку ошибки извлечения, неполное поле зрения или преобладание внелёгочных структур могли существенно исказить расчёт LAA%-950.

На уровне пациентских случаев после этапа извлечения было получено 1010 наблюдений. Из них 1006 случаев прошли контроль качества и были отнесены к технически пригодным, тогда как 4 случая были классифицированы как имеющие аномально высокое среднее значение HU. Для этих наблюдений средняя плотность находилась в диапазоне от −232,7 до −189,4 HU, что резко отличалось от основной части выборки и указывало на некорректность данных для последующего денситометрического анализа.

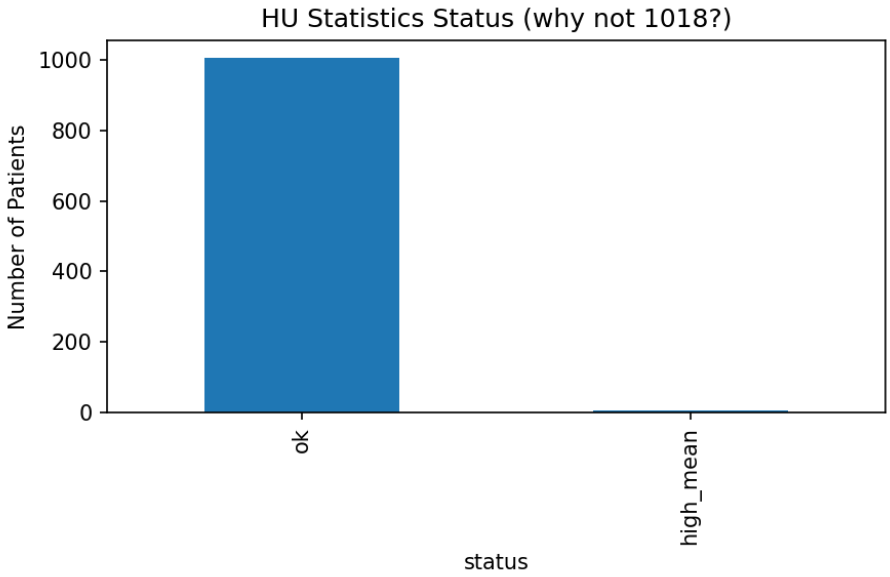


Рисунок 3.1.2 -1- Распределение статусов технической пригодности КТ-исследований по результатам HU-контроля

Распределение среднего значения HU представлено на рисунке 3.1.2-2 Среднее значение по выборке составило −519,6 HU, медиана - 521,6 HU. Основная масса наблюдений располагалась приблизительно в интервале от −570 до −470 HU, что соответствует ожидаемой плотностной характеристике всего КТ-объёма грудной клетки после включения лёгочной ткани и окружающих анатомических структур. Наличие отдельных наблюдений с менее отрицательными значениями HU подтверждало необходимость дополнительного технического контроля перед расчётом LAA%.

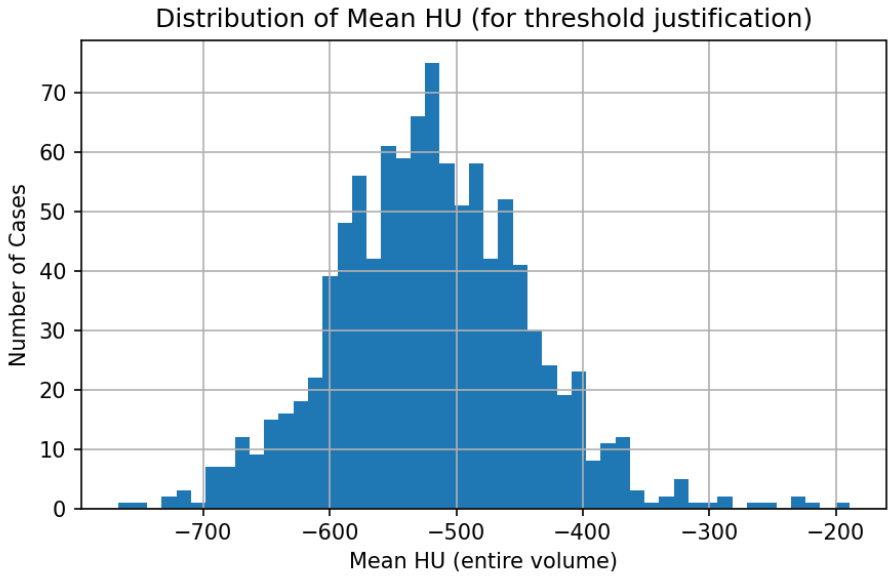


Рисунок 3.1.2-2 - Распределение среднего значения HU по извлечённым КТ-объёмам

Дополнительно оценивалось стандартное отклонение HU. Его медиана составила 479,7 HU, межквартильный диапазон - от 471,3 до 486,5 HU. Распределение было относительно компактным, однако небольшое число наблюдений с отклоняющимися значениями указывало на возможность выявления технически нестандартных объёмов. Данный показатель использовался как дополнительный критерий контроля качества извлечения.

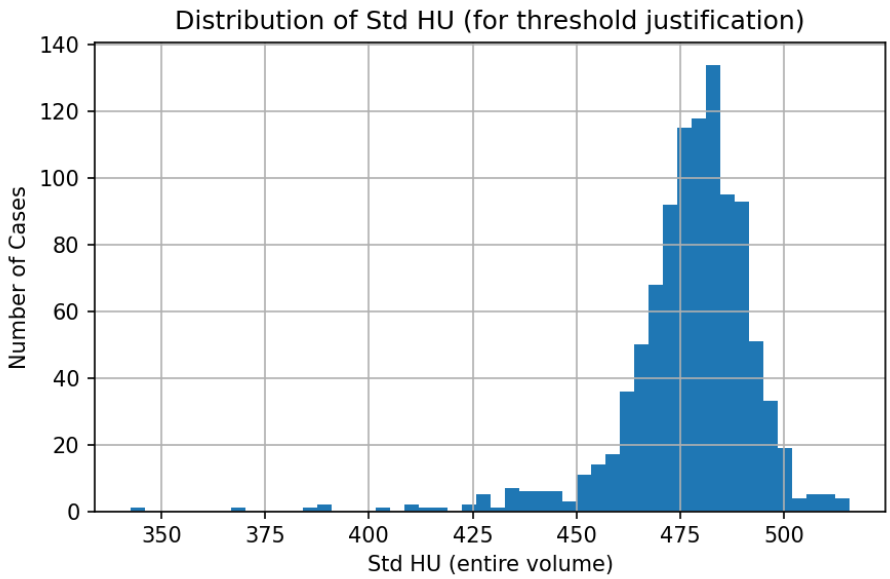


Рисунок 3.1.2-3 - Распределение стандартного отклонения HU по извлечённым КТ-объёмам

3.1.3 Формирование итоговой аналитической выборки

После завершения этапов технического контроля, сегментации лёгких и расчёта денситометрических признаков была сформирована итоговая аналитическая выборка из 679 случаев. Для каждого наблюдения были рассчитаны LAA%-950, число низкоплотных вокселей, объём лёгочной маски, средняя плотность лёгочной ткани, Perc15 HU и общий объём лёгких. Эти показатели составили количественную основу дальнейшего машинного анализа.

Итоговая выборка включала 593 случая класса 0 и 86 случаев класса 1. Доля положительного класса составила 12,7%, что указывает на выраженный дисбаланс данных. В связи с этим на последующих этапах использовалось стратифицированное разбиение, сохраняющее соотношение классов в обучающей, валидационной и тестовой частях. В обучающую выборку вошли 415 отрицательных и 60 положительных случаев, в валидационную - 89 и 13, в тестовую - 89 и 13 соответственно. Такое распределение позволило сохранить исходную структуру классов и снизить риск смещения оценки качества модели.

3.1.4 Распределение LAA%-950 и формирование исследовательских классов

Основным денситометрическим показателем в исследовании являлся LAA%-950 - доля вокселей внутри лёгочной маски с плотностью ниже −950 HU. Данный показатель использовался как физически интерпретируемый критерий

низкоплотных зон лёгочной паренхимы и как основа для формирования Silver Standard-разметки. В итоговой выборке среднее значение LAA%-950 составило 3,94%, медиана - 1,42%, межквартильный диапазон - от 0,23% до 4,79%. Максимальное значение достигало 61,59%, что свидетельствует о наличии небольшой группы случаев с резко выраженной долей низкоплотных участков. Распределение имело правостороннюю асимметрию: большинство наблюдений находилось в зоне низких значений, тогда как случаи с высоким LAA%-950 формировали длинный правый хвост.

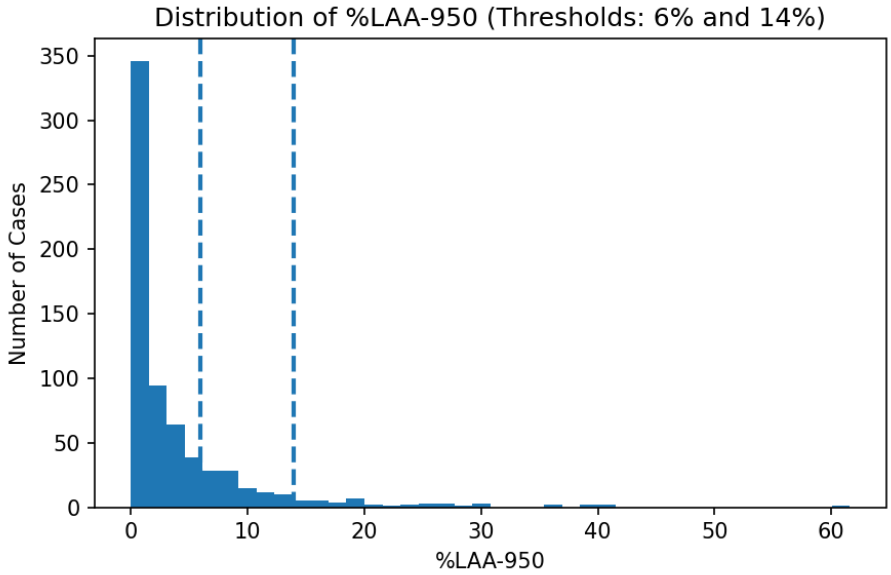


Рисунок 3.1.4 - Распределение LAA%-950 с пороговыми линиями 6% и 14%

На гистограмме отмечены две операционные пороговые линии - 6% и 14%. Порог 6% использовался для выделения группы с выраженными низкоплотными изменениями, тогда как порог 14% служил дополнительной границей для визуального обозначения случаев с более высокой долей поражения. Указанные пороги не рассматривались как клинические GOLD-стадии, поскольку GOLD-классификация основана на спирометрии. В настоящей работе они применялись как исследовательские границы для формирования QCT-derived Silver Standard-меток в условиях отсутствия исходной экспертной разметки эмфиземы в LIDC-IDRI.

По фактическому распределению 540 случаев имели LAA%-950 ниже 6%, 139 случаев - не ниже 6%, из них 41 случай превышал 14%. Окончательная положительная метка была присвоена 86 случаям, поскольку помимо LAA%-порога применялся дополнительный контроль средней плотности лёгких, направленный на исключение технически обусловленных ложноположительных случаев.

3.1.5 Связь LAA%-950 со средней плотностью лёгочной ткани

Для проверки внутренней согласованности автоматически сформированной разметки была проанализирована связь между LAA%-950 и средней плотностью лёгочной ткани. Теоретически увеличение доли вокселей ниже −950 HU должно сопровождаться смещением средней плотности лёгких в сторону более отрицательных значений. Полученные данные подтвердили данную закономерность.

Среднее значение mean lung HU в итоговой выборке составило −804,8 HU, медиана - −815,9 HU, межквартильный диапазон - от −839,5 до −779,6 HU. Корреляционный анализ выявил отрицательную связь между LAA%-950 и mean lung HU: коэффициент Пирсона составил $r = -0{,}49$ при $p = 2{,}1\times10^{-42}$, коэффициент Спирмена - $\rho = -0{,}75$ при $p = 1{,}2\times10^{-123}$. Более высокая по модулю ранговая корреляция указывает на выраженную монотонную, но нелинейную зависимость между долей низкоплотных участков и средней плотностью лёгочной паренхимы.

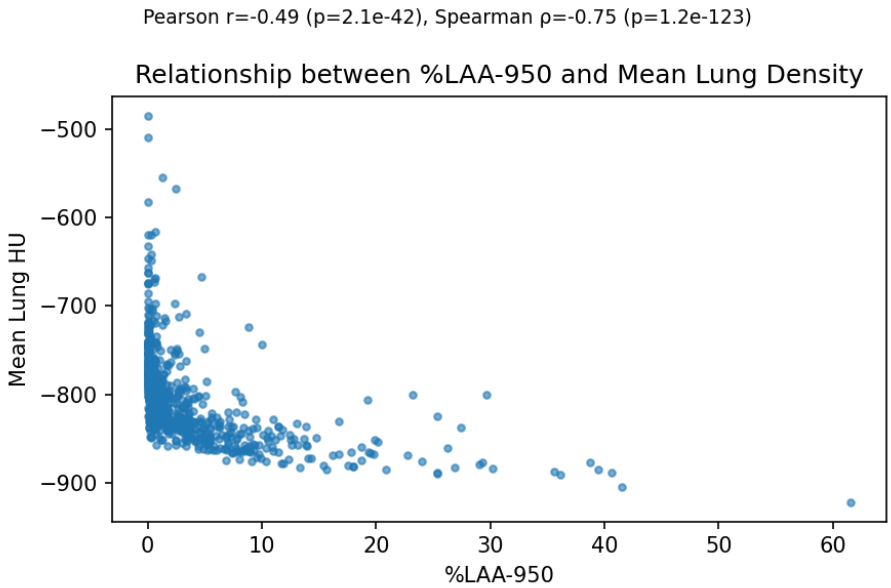


Рисунок 3.1.5 - Связь LAA%-950 со средней плотностью лёгочной ткани

Данный результат показывает, что LAA%-разметка согласуется с общей денситометрической характеристикой лёгких: чем выше доля вокселей ниже −950 HU, тем ниже средняя плотность лёгочной маски. Это подтверждает физическую непротиворечивость выбранного критерия автоматической разметки.

3.1.6 Связь LAA%-950 с общим объёмом лёгких

Дополнительно была оценена связь LAA%-950 с общим объёмом лёгких. Этот анализ необходим, поскольку степень вдоха при КТ-сканировании может влиять на плотностное распределение лёгочной ткани: при большем объёме лёгких плотность паренхимы может становиться более отрицательной, что потенциально увеличивает долю вокселей ниже −950 HU.

В исследуемом массиве средний общий объём лёгких составил 5,01 л, медиана - 4,81 л, межквартильный диапазон - от 4,06 до 5,84 л. Минимальное значение составило 0,22 л, максимальное - 10,32 л, что отражает как индивидуальные различия, так и возможную вариативность полноты сегментации и фазы вдоха. Корреляция между LAA%-950 и общим объёмом лёгких была положительной: $r = 0,37$ при $p = 1,7\times10^{-23}$; $\rho = 0,39$ при $p = 1,8\times10^{-26}$.

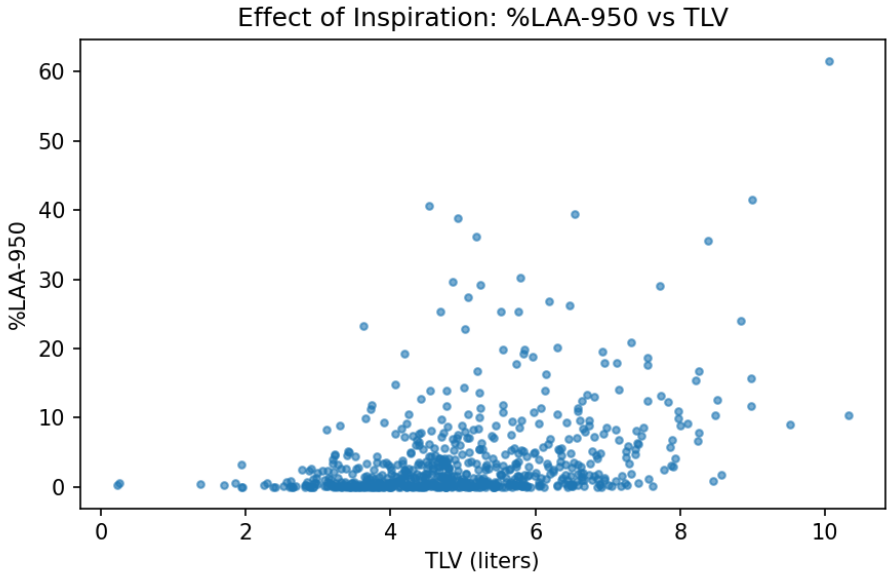


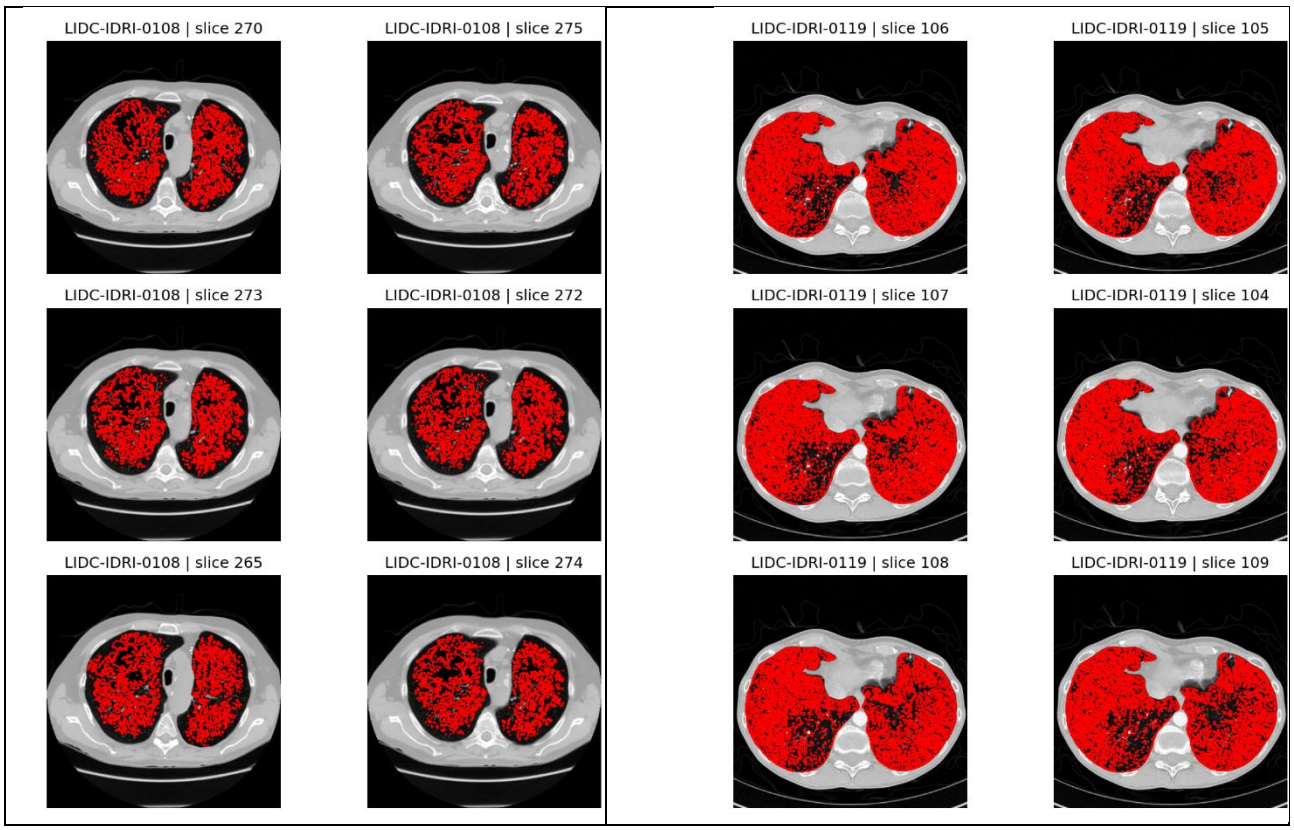
Рисунок 3.1.6 - Связь LAA%-950 с общим объёмом лёгких

Полученная зависимость указывает, что LAA%-950 частично связан с объёмным состоянием лёгких. Следовательно, общий объём лёгких должен учитываться при интерпретации денситометрической разметки, особенно в случаях с пограничными значениями LAA%-950. Рост доли низкоплотных вокселей может отражать не только эмфизематозно-подобные изменения, но и влияние различий в фазе вдоха при сканировании.

3.1.7 Визуальная проверка LAA-карт

Для качественного контроля voxel-based quantification были сформированы визуализации LAA-зон на срезах с наибольшей долей низкоплотных участков. На таких изображениях отображались области внутри лёгочного поля, соответствующие порогу $HU < -950$. Данный этап использовался для проверки того, что автоматический расчёт LAA% опирается на зоны сниженной плотности внутри лёгочной паренхимы, а не на внешний воздух, фоновые области или грубые ошибки маскирования.

Визуальная проверка не рассматривалась как экспертная радиологическая разметка. Её назначение состояло в контроле корректности автоматического порогового анализа и в повышении интерпретируемости сформированных Silver Standard-меток.



а б

Рисунок 3.1.7 - Примеры визуализации LAA-зон на КТ-срезах с максимальной долей низкоплотных участков

3.1.8 Итоговая характеристика подготовленного массива данных

В результате проведённой обработки был сформирован структурированный экспериментальный массив, включающий три взаимосвязанных уровня данных. Первый уровень представлен КТ-объёмами, прошедшими контроль качества и приведёнными к сопоставимому формату. Второй уровень включает количественные QCT-показатели: LAA%-950, mean lung HU, Perc15 HU, общий объём лёгких, объём лёгочной маски и число низкоплотных вокселей. Третий уровень представлен автоматически сформированными метками, использованными для последующего обучения и оценки предложенной модели.

Принципиальной особенностью подготовленного массива является статус его разметки. Она основана на воспроизводимых денситометрических критериях и потому может рассматриваться как Silver Standard, однако не заменяет клинический диагноз ХОБЛ, экспертную оценку эмфиземы или спирометрическую классификацию. Тем не менее полученный массив обеспечивает возможность контролируемого машинного эксперимента в условиях отсутствия исходной COPD-разметки в LIDC-IDRI и позволяет сопоставить физически интерпретируемые показатели плотности с латентными визуально-текстурными признаками, извлекаемыми ResNet-152.

Таким образом, в результате QCT-анализа была сформирована автоматическая разметка, позволившая преобразовать исходный LIDC-IDRI, не

содержащий COPD-меток, в аналитическую выборку для машинного обучения. Полученные метки имеют статус Silver Standard, поскольку основаны на воспроизводимом денситометрическом критерии, но не являются клинической экспертной разметкой.

3.2 Анализ результатов

3.2.1 Баланс классов и стратегия формирования обучающих подвыборок

Итоговая аналитическая выборка, описанная в разделе 3.1, характеризовалась выраженным дисбалансом классов: из 679 случаев 593 были отнесены к классу 0, а 86 - к классу 1. Таким образом, доля положительного класса составила 12,7%, что является принципиальным обстоятельством для интерпретации метрик качества. В условиях такого распределения простая accuracy не отражает диагностическую состоятельность модели, поскольку классификатор, склонный к предсказанию большинства, может демонстрировать внешне приемлемую общую точность при фактической неспособности выявлять положительные случаи. Поэтому основными метриками анализа были выбраны ROC-AUC, PR-AUC, F1-score и balanced accuracy.

Для оценки влияния дисбаланса были рассмотрены несколько вариантов подготовки данных: отсутствие балансировки, downsampling, upsampling, SMOTE и использование весов классов. Распределение классов при разных стратегиях представлено на рисунке 3.2.1-1 В варианте без балансировки сохранялось исходное соотношение 593:86, тогда как downsampling приводил к сокращению выборки до 172 случаев с равным числом наблюдений каждого класса. Остальные стратегии сохраняли общий объём выборки, но компенсировали дисбаланс либо через синтетическое расширение меньшинства, либо через изменение веса ошибок.



Рисунок 3.2.1-1 - Распределение классов при разных стратегиях балансировки

Общее число случаев при каждой стратегии представлено на рисунке 3.2.1-2 Наиболее заметное уменьшение объёма данных наблюдалось при downsampling: число случаев снижалось с 679 до 172. Данный вариант создавал симметричную структуру классов, однако сопровождался потерей значительной части отрицательных наблюдений. Поэтому результаты downsampling интерпретировались как вспомогательные, а не как единственная основа оценки модели.

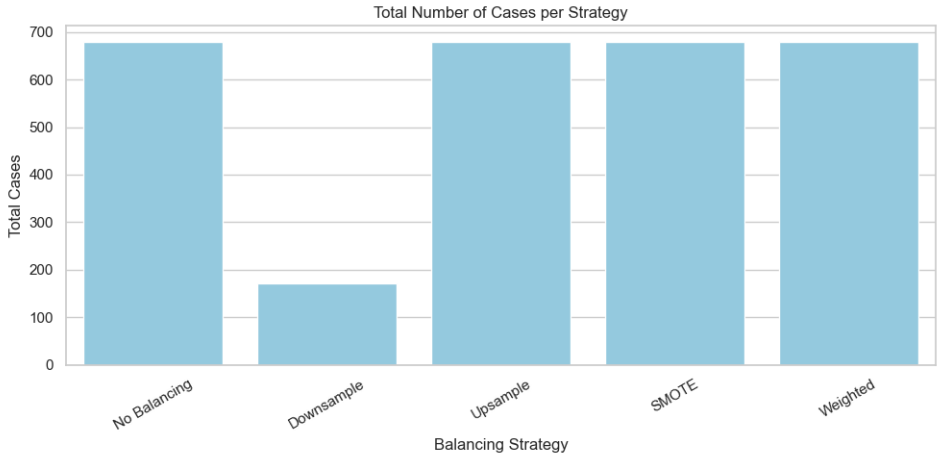


Рисунок 3.2.1-2- Общее число случаев при разных стратегиях балансировки

3.2.2 Классификация на основе полного пространства эмбеддингов ResNet-152

Первым этапом оценки была проверка того, содержит ли пространство признаков, извлечённых ResNet-152, информацию, позволяющую различать случаи с низкой и высокой долей LAA%-950 без непосредственного использования QCT-показателей. Для этого 2048-мерные эмбеддинги, полученные с финального слоя глобального усреднения, были использованы в качестве входа для логистической регрессии.

На тестовой выборке размером 59 случаев модель на полном наборе эмбеддингов достигла ROC-AUC = 0,695, PR-AUC = 0,382, F1-score = 0,267 и balanced accuracy = 0,585. Доверительный интервал для ROC-AUC составил 0,486–0,876, для PR-AUC - 0,146–0,708, для F1-score - 0,000–0,571, для balanced accuracy - 0,433–0,761. Ширина доверительных интервалов указывает на ограниченную устойчивость оценки, что объясняется малым числом положительных случаев в тестовой подвыборке.

На рисунке 3.2.2-1 представлена ROC-кривая логистической регрессии, обученной на полном пространстве 2048-мерных эмбеддингов ResNet-152. Значение ROC-AUC составило 0,695, что превышает уровень случайной классификации и указывает на наличие в эмбеддингах признаков, связанных с LAA%-разметкой. Вместе с тем полученное значение не позволяет рассматривать полное пространство эмбеддингов как достаточное для

устойчивого разделения классов. Ступенчатая форма кривой обусловлена небольшим размером тестовой выборки и ограниченным числом положительных случаев. Следовательно, данный результат следует интерпретировать как подтверждение наличия визуально-текстурного сигнала, но не как окончательную диагностическую эффективность модели без дополнительного отбора признаков или включения QCT-показателей.

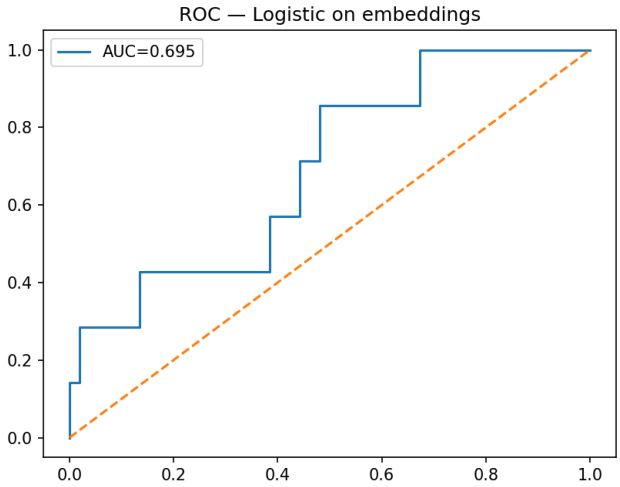


Рисунок 3.2.2-1 - ROC-кривая логистической регрессии на полном наборе эмбеддингов ResNet-152

На рисунке 3.2.2-2 представлена Precision–Recall-кривая логистической регрессии, обученной на полном наборе 2048-мерных эмбеддингов ResNet-152. Средняя точность, выраженная показателем AP, составила 0,382. С учётом выраженного дисбаланса классов данный показатель является более информативным, чем простая accuracy, поскольку отражает способность модели сохранять приемлемую точность при выявлении положительного класса. Форма кривой показывает, что при высоких значениях recall точность быстро снижается, что указывает на наличие ложноположительных срабатываний при попытке расширить охват случаев с выраженными LAA%-изменениями. Следовательно, полное пространство эмбеддингов содержит диагностически значимый сигнал, однако без отбора признаков или включения денситометрических показателей его недостаточно для устойчивого выявления положительного класса.

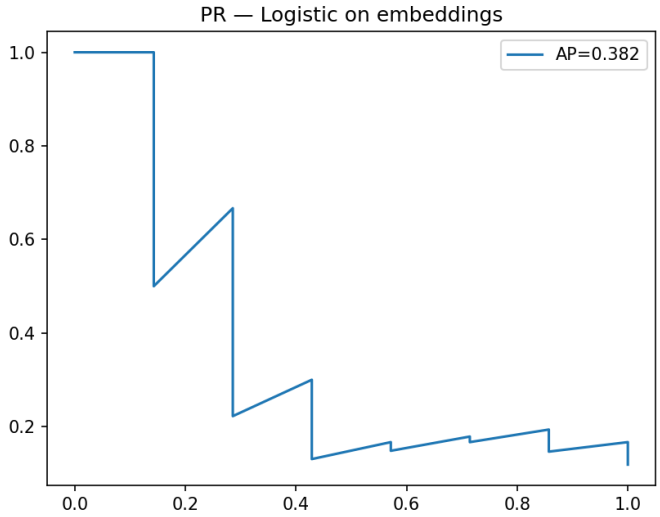


Рисунок 3.2.2-2- Precision–Recall-кривая логистической регрессии на полном наборе эмбеддингов ResNet-152

Полученные значения показывают, что латентные визуально-текстурные признаки ResNet-152 обладают умеренной дискриминативной способностью. ROC-AUC выше 0,5 свидетельствует о наличии сигнала, отличающего положительные случаи от отрицательных, однако низкие значения F1-score и balanced accuracy демонстрируют, что полного 2048-мерного пространства недостаточно для устойчивой классификации без дополнительного отбора признаков или включения физически интерпретируемых QCT-показателей. Это согласуется с общей логикой исследования: эмбеддинги отражают визуальную структуру изображения, но не являются прямыми патологическими маркерами и требуют статистической селекции.

3.2.3 Отбор информативных признаков и снижение размерности

Следующим этапом был проведён анализ информативности отдельных координат эмбеддингов. Цель состояла не только в повышении качества классификации, но и в проверке того, можно ли выделить ограниченное число признаков, статистически связанных с положительной LAA%-разметкой. Были использованы три подхода: t-test с контролем FDR, L1-регуляризованная логистическая регрессия и рекурсивное исключение признаков. Реализованный скрипт кросс-валидации использовал стратифицированное разбиение на 5 фолдов, стандартизацию признаков и логистическую регрессию с class_weight="balanced". Результаты отбора признаков показали существенные различия между стратегиями. Модель на признаках, отобранных по t-test/FDR, достигла ROC-AUC = 0,731, PR-AUC = 0,291, F1-score = 0,231 и balanced accuracy = 0,560. При использовании L1-регуляризованной логистической регрессии качество было существенно выше: ROC-AUC = 0,942, PR-AUC = 0,767, F1-score = 0,600 и balanced accuracy = 0,861. Метод RFE также показал высокое качество: ROC-AUC = 0,907, PR-AUC = 0,569, F1-score = 0,588 и balanced accuracy = 0,809.

Таблица 3.1 – Качество классификации при разных способах отбора признаков

Способ отбора признаков	**ROC-AUC**	**PR-AUC**	**F1-score**	**Balanced accuracy**	**Размер тестовой выборки**
t-test/FDR	0,731	0,291	0,231	0,560	59
L1-logistic	0,942	0,767	0,600	0,861	59
RFE	0,907	0,569	0,588	0,809	59

Различие между полным пространством эмбеддингов и отобранными признаками указывает на то, что диагностически значимый сигнал распределён неравномерно. Большая часть из 2048 координат может отражать нерелевантные вариации изображения, тогда как ограниченное подмножество признаков оказывается связанным с визуально-текстурными характеристиками низкоплотных зон. Это особенно важно для интерпретации предложенной модели: использование ResNet-152 не сводится к непрозрачной классификации, поскольку дальнейший анализ позволяет выделить конкретные координаты эмбеддинга, вносящие вклад в решение.

3.2.4 Весовые профили отобранных признаков

Для анализа вклада признаков были построены весовые диаграммы для трёх стратегий отбора. На рисунке 3.2.4-1 представлены коэффициенты L1-регуляризованной логистической модели. Наибольший положительный вклад дали признаки feat_211, feat_1577, feat_673, feat_1268, feat_253, тогда как наиболее выраженный отрицательный вклад имели feat_1152, feat_1766, feat_902, feat_820, feat_889.

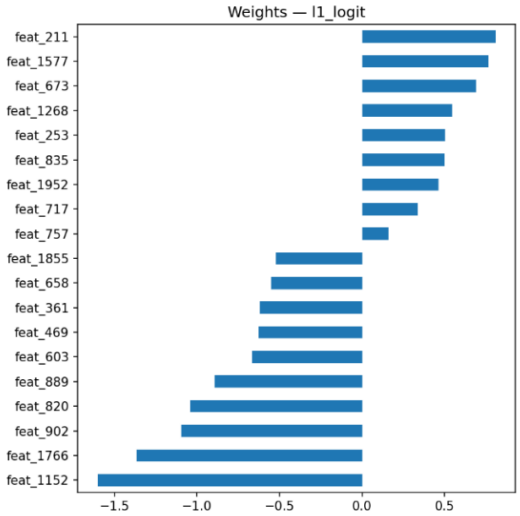


Рисунок 3.2.4-1 - Весовые коэффициенты признаков при L1-регуляризованной логистической регрессии

На рисунке 3.2.4-2 показаны признаки, отобранные методом RFE. Положительные веса были связаны с признаками feat_88, feat_965, feat_339, feat_566, feat_1041, а отрицательные - с feat_493, feat_717, feat_1146, feat_1152, feat_820.

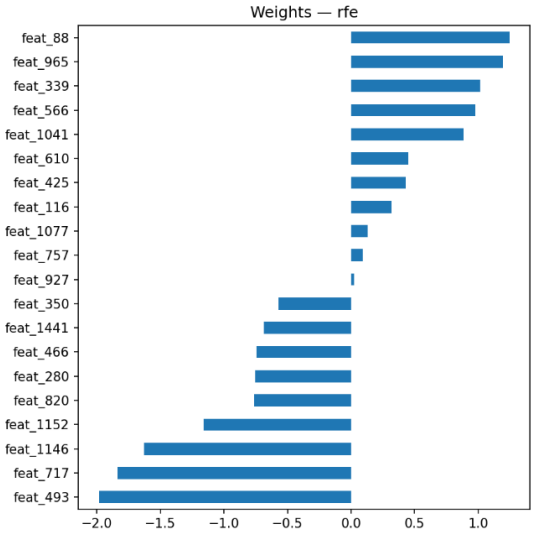


Рисунок 3.2.4-2- Весовые коэффициенты признаков, отобранных методом RFE

Рисунок 3.2.4-3 отражает весовой профиль признаков, выбранных по t-test/FDR. В этом варианте признаки имеют условную нумерацию после предварительного отбора, поэтому их интерпретация ограничивается сравнением направления и величины вклада. Наиболее выраженный положительный вес имел feat_1, а наиболее выраженный отрицательный - feat_5.

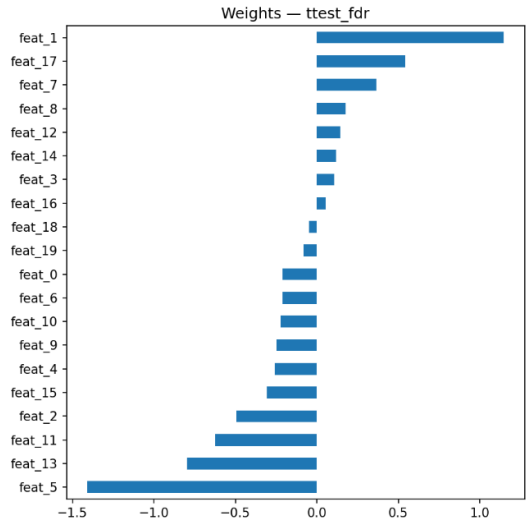


Рисунок 3.2.4-3- Весовые коэффициенты признаков, отобранных по t-test/FDR

Совпадение отдельных признаков между стратегиями отбора имеет самостоятельное значение. Например, feat_820, feat_1152, feat_717 встречались среди информативных признаков более чем в одном варианте отбора, что позволяет рассматривать их как более устойчивые координаты эмбеддингового пространства. Однако сами по себе номера признаков не имеют прямой анатомической интерпретации. Они должны пониматься как латентные визуально-текстурные дескрипторы, полученные из ResNet-152, а не как самостоятельные радиологические признаки.

3.2.5 Проверка устойчивости модели на кросс-валидации

Для оценки устойчивости результатов была выполнена 5-кратная стратифицированная кросс-валидация на top-20 признаках, выбранных по t-test. Среднее значение ROC-AUC составило 0,811 при стандартном отклонении 0,089; среднее значение PR-AUC - 0,388 ± 0,106; F1-score - 0,414 ± 0,121; balanced accuracy - 0,739 ± 0,130. Разброс между фолдами был заметным: ROC-AUC изменялся от 0,674 до 0,907, что указывает на чувствительность результата к конкретному составу обучающей и тестовой частей.

Таблица 3.2 - Результаты 5-кратной кросс-валидации на top-20 признаках

Метрика	**Среднее значение**	**Стандартное отклонение**
ROC-AUC	0,811	0,089
PR-AUC	0,388	0,106
F1-score	0,414	0,121
Balanced accuracy	0,739	0,130

Данные кросс-валидации подтверждают, что признаки ResNet-152 содержат воспроизводимый классификационный сигнал, однако его устойчивость ограничена дисбалансом классов и сравнительно небольшим числом положительных случаев. Поэтому результаты single split, особенно при высоких значениях AUC после отбора признаков, требуют осторожной интерпретации и должны рассматриваться совместно с кросс-валидационной оценкой.

3.2.6 Предложенная модифицированная модель с объединением признаков ResNet-152 и QCT

Отдельным этапом исследования была выполнена оценка предложенной модели, основанной на совместном использовании нейросетевых признаков ResNet-152 и количественных денситометрических показателей QCT. Данный эксперимент был необходим для проверки ключевой методологической идеи работы: латентные признаки CNN отражают визуально-текстурную структуру

КТ-изображения, тогда как QCT-показатели фиксируют физически измеряемые характеристики лёгочной ткани, связанные с плотностью паренхимы.

Перед оценкой предложенной модели была сформирована финальная модельная подвыборка. На этапе денситометрического анализа было получено 679 наблюдений с рассчитанными показателями LAA%-950, Perc15, Mean HU и TLV. Однако не каждое исследование из этого массива было пригодно для полного машинного pipeline, включающего подготовку КТ-патчей, извлечение эмбеддингов ResNet-152, сохранение принадлежности к train/validation/test-разбиению и последующую классификацию. Поэтому для итогового эксперимента использовалась более строгая подвыборка из 391 исследования, включавшая 341 случай отрицательного класса и 50 случаев положительного класса.

Такое сокращение объёма данных было обусловлено не изменением критерия автоматической разметки, а дополнительными техническими требованиями к пригодности случая для нейросетевого анализа. Таким образом, в главе 3 используются два уровня данных: QCT-размеченный массив из 679 случаев, применявшийся для описания денситометрического распределения, и финальная модельная подвыборка из 391 случая, на которой оценивались варианты классификационной модели.

В предложенной модели ResNet-152 использовалась как экстрактор признаков. Из финального слоя глобального усреднения извлекался вектор эмбеддингов, отражающий высокоуровневое представление КТ-патча. Параллельно для каждого КТ-исследования рассчитывались QCT-показатели: LAA%-950, Perc15, Mean HU и TLV. Перед объединением количественные признаки стандартизировались, поскольку имели различную физическую размерность и числовой диапазон: LAA%-950 выражался в процентах, Perc15 и Mean HU - в единицах Хаунсфилда, TLV - в литрах.

Объединение признаков выполнялось на уровне признакового пространства путём конкатенации нейросетевого и денситометрического векторов:

$$Z_{proposed} = [Z_{ResNet} \,||\, Z_{QCT}],$$

где Z_{ResNet} - вектор эмбеддингов ResNet-152, Z_{QCT} - нормализованный вектор количественных КТ-признаков, а символ || обозначает операцию конкатенации. Такой подход позволял сохранить два уровня описания КТ-изображения: визуально-текстурный, формируемый CNN-компонентом, и физико-денситометрический, представленный QCT-показателями.

Для уточнения вклада нейросетевого компонента до включения QCT-показателей были сопоставлены модель на основе эмбеддингов ResNet-152 и расширенный классификационный блок без прямого использования денситометрических признаков. Результаты представлены в таблице 3.3.

Таблица 3.3 - Сравнение нейросетевых вариантов модели без прямого включения QCT-признаков

Выборка	Вариант модели	ROC-AUC	PR-AUC / AP	TP	FN	TN	FP	Чувствительность	Специфичность	bACC
Validation	Embeddings-only	0,879	0,530	11	2	65	24	0,846	0,730	0,788
Validation	Расширенный блок без QCT	0,879	0,482	11	2	71	18	0,846	0,798	0,822
Test	Embeddings-only	0,683	0,207	5	8	69	20	0,385	0,775	0,580
Test	Расширенный блок без QCT	0,848	0,459	10	3	73	16	0,769	0,820	0,795

Данные таблицы 3.3 показывают, что расширенный классификационный блок без QCT повышает чувствительность на тестовой выборке с 0,385 до 0,769 и снижает число ложноотрицательных решений с 8 до 3. Это подтверждает, что структура классификатора влияет на выявление положительного класса, однако данный вариант всё ещё не использует физически интерпретируемые QCT-показатели, поэтому в дальнейшем был оценён полный вариант предложенной модели ResNet-152 + QCT.

Для оценки вклада отдельных компонентов были сопоставлены три варианта классификационного контура: логиты ResNet-152 после test-time augmentation и калибровки, линейный классификатор по эмбеддингам ResNet-152 и предложенный вариант Fusion, объединяющий нейросетевые признаки с QCT-показателями. Результаты сравнения представлены в таблице 3.4.

Таблица 3.4 - Сравнение вариантов модели на финальной модельной подвыборке

Модель / подход	Выборка	ROC-AUC	PR-AUC	F1	bACC	TN / FP / FN / TP
Logits (TTA + calibration)	Val	0,886	0,511	0,55	0,797	43 / 8 / 2 / 6
Logits (TTA + calibration)	Test	0,732	0,332	0,40	0,732	39 / 13 / 2 / 5
Embeddings Probe	Val	0,836	0,507	0,50	0,744	44 / 7 / 3 / 5
Embeddings Probe	Test	0,766	0,459	0,38	0,690	42 / 10 / 3 / 4

Fusion	Val	0,989	0,898	0,84	0,971	48 / 3 / 0 / 8
Fusion	Test	0,996	0,962	0,86	0,919	51 / 1 / 1 / 6

Примечание. ROC-AUC - площадь под ROC-кривой; PR-AUC - площадь под precision–recall-кривой; bACC - balanced accuracy; TTA - test-time augmentation; TN - истинно отрицательные случаи; FP - ложноположительные; FN - ложноотрицательные; TP - истинно положительные. Fusion обозначает предложенный вариант модели, объединяющий нейросетевые признаки ResNet-152 с количественными QCT-дескрипторами.

На валидационной выборке вариант Logits (TTA + calibration) достиг ROC-AUC = 0,886, PR-AUC = 0,511 и balanced accuracy = 0,797. На тестовой выборке качество этого варианта снизилось до ROC-AUC = 0,732 и PR-AUC = 0,332. По матрице ошибок модель правильно классифицировала 39 отрицательных и 5 положительных случаев, допустив 13 ложноположительных и 2 ложноотрицательных решения. Это показывает, что калиброванные логиты содержали диагностический сигнал, однако их устойчивость на тестовой выборке была ограниченной.

Модель Embeddings Probe, основанная на эмбеддингах ResNet-152, на тестовой выборке показала ROC-AUC = 0,766, PR-AUC = 0,459, F1 = 0,38 и balanced accuracy = 0,690. Данный результат подтверждает, что латентное пространство ResNet-152 содержит признаки, связанные с автоматически сформированной LAA%-разметкой. Однако матрица ошибок показывает, что при использовании только эмбеддингов модель выявила 4 из 7 положительных случаев и допустила 10 ложноположительных решений. Следовательно, изолированные эмбеддинги обладают самостоятельной информативностью, но не обеспечивают достаточной устойчивости классификации в условиях дисбаланса классов.

Наиболее высокие значения метрик были получены для варианта Fusion, соответствующего предложенной модели ResNet-152 + QCT. На валидационной выборке ROC-AUC составил 0,989, PR-AUC - 0,898, F1 - 0,84, balanced accuracy - 0,971. На тестовой выборке модель сохранила высокое качество: ROC-AUC = 0,996, PR-AUC = 0,962, F1 = 0,86, balanced accuracy = 0,919. По матрице ошибок модель правильно классифицировала 51 отрицательный и 6 положительных случаев, допустив только 1 ложноположительное и 1 ложноотрицательное решение[74].

Сравнение вариантов показывает, что включение QCT-показателей существенно повышает качество классификации по сравнению с использованием только нейросетевых признаков. На тестовой выборке переход от Embeddings Probe к Fusion сопровождался увеличением ROC-AUC с 0,766 до 0,996, PR-AUC - с 0,459 до 0,962, F1 - с 0,38 до 0,86, balanced accuracy - с 0,690 до 0,919. Особенно значимым является рост PR-AUC, поскольку эта метрика более чувствительна к качеству выявления положительного класса при выраженном дисбалансе данных.

Вместе с тем интерпретация высоких значений Fusion требует методической осторожности. Поскольку автоматическая разметка была сформирована на основе QCT-критериев, включение денситометрических признаков в модель частично повышает согласованность классификатора с правилом auto-labeling. Поэтому полученные показатели следует рассматривать не как окончательное доказательство клинической диагностики ХОБЛ, а как подтверждение внутренней согласованности предложенного вычислительного pipeline и эффективности объединения визуально-текстурных признаков ResNet-152 с физически интерпретируемыми QCT-показателями.

Таким образом, результаты раздела 3.2.6 подтверждают преимущество предложенной модели ResNet-152 + QCT по сравнению с изолированным использованием логитов или эмбеддингов ResNet-152. CNN-компонент обеспечивает извлечение латентных визуально-текстурных признаков КТ-изображения, тогда как QCT-компонент связывает классификационное решение с измеряемыми параметрами плотности лёгочной ткани. Такое объединение повышает как численные показатели классификации, так и интерпретируемость модели, что является принципиальным условием её дальнейшей адаптации к задачам автоматизированного анализа КТ в условиях популяционного скрининга.

3.2.7 Анализ ошибок классификации

Полученные результаты позволяют сформулировать несколько промежуточных выводов. Во-первых, автоматическая разметка на основе LAA%-950 сформировала выраженно несбалансированную выборку, что потребовало использования balanced accuracy, PR-AUC и F1-score наряду с ROC-AUC. Во-вторых, классификация на полном пространстве эмбеддингов показала умеренное качество, что свидетельствует о наличии визуально-текстурного сигнала, но не позволяет считать полные эмбеддинги достаточными без селекции. В-третьих, отбор ограниченного числа признаков заметно повышал качество, особенно при использовании L1-регуляризованной логистической регрессии и RFE. Это указывает на существование компактного подпространства признаков ResNet-152, связанного с низкоплотными изменениями лёгочной ткани.

В-четвёртых, включение QCT-показателей резко повышало качество классификации, однако это повышение обусловлено близостью QCT-признаков к правилу формирования Silver Standard-меток. Поэтому предложенная модель должна рассматриваться в двух аналитических плоскостях: как инструмент проверки воспроизводимости денситометрической разметки и как способ выявления визуально-текстурных признаков, ассоциированных с LAA%-950. Такая интерпретация позволяет избежать переоценки результатов и одновременно сохранить научную значимость эксперимента.

В целом результаты раздела 3.2 показывают, что ResNet-152-эмбеддинги не являются случайным набором признаков: после статистического отбора они обеспечивают ROC-AUC выше 0,90 на тестовой выборке в отдельных вариантах и средний ROC-AUC 0,811 при 5-кратной кросс-валидации. При этом QCT-

компонент остаётся необходимым для физической интерпретации результата и контроля связи модели с плотностными характеристиками лёгочной ткани.

3.2.8 Визуальная интерпретация результатов с использованием Grad-CAM

Для дополнительной интерпретации работы нейросетевого компонента предложенной модели была использована визуализация Grad-CAM. Данный метод применялся не как самостоятельный диагностический критерий, а как инструмент качественной проверки того, какие зоны КТ-среза вносят наибольший вклад в формирование выходной вероятности классификатора. Поскольку ResNet-152 использовалась как экстрактор признаков, Grad-CAM-карты интерпретировались как визуализация пространственно локализованных CNN-активаций, участвующих в формировании эмбеддингового представления. На рисунке 3.2.8-1 представлены Grad-CAM-визуализации случаев с высокой выходной вероятностью положительного класса. В примерах LIDC-IDRI-0240, LIDC-IDRI-0261 и LIDC-IDRI-0304 модель формировала вероятность положительного класса, близкую к 1,000. Зоны максимальной активации преимущественно располагались в пределах лёгочных полей, что свидетельствует о вовлечении внутрилёгочных визуально-текстурных признаков в классификационное решение. При этом активации не следует трактовать как прямую карту эмфиземы: они отражают области, значимые для CNN, но не измеряют плотность ткани в единицах Хаунсфилда.

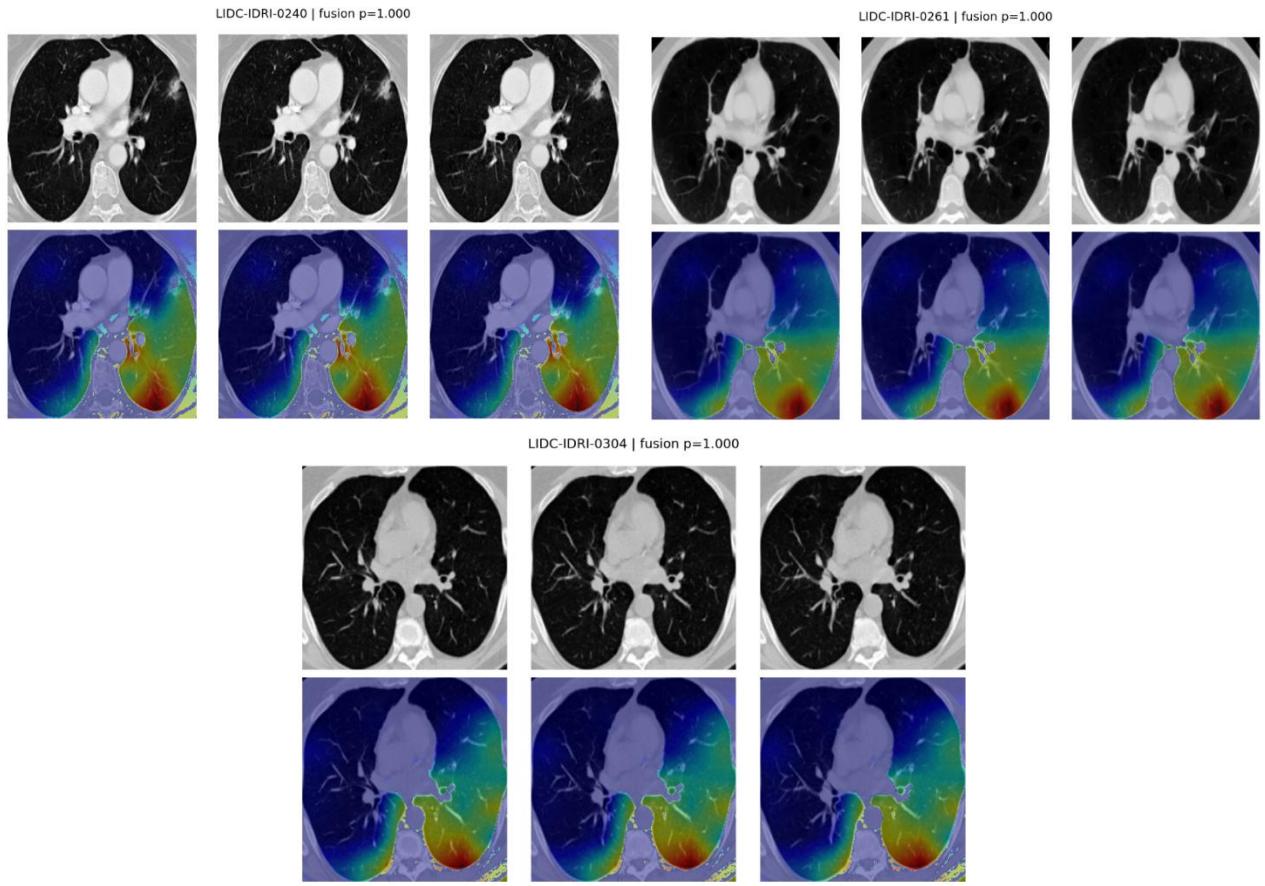


Рисунок 3.2.8-1 - Grad-CAM-визуализация случаев с высокой выходной вероятностью положительного класса

На рисунке 3.2.8-2 показаны Grad-CAM-карты для случаев со средней выходной вероятностью положительного класса. В примерах LIDC-IDRI-0324 и LIDC-IDRI-0478 вероятность составляла около 0,714, а зоны активации имели более диффузный характер. Такая картина соответствует пограничному решению модели: нейросетевой компонент фиксирует визуальные признаки неоднородности лёгочной паренхимы, однако выраженность этих признаков недостаточна для максимально уверенной классификации.

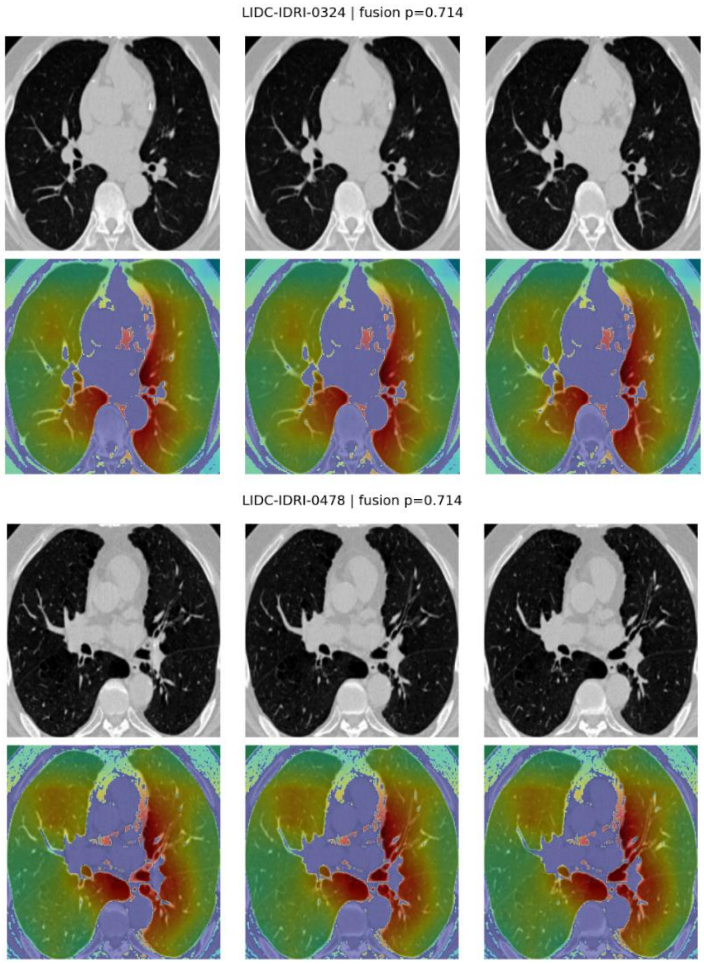


Рисунок 3.2.8-2 - Grad-CAM-визуализация случаев со средней выходной вероятностью положительного класса

На рисунке 3.2.8-3 представлены случаи с низкой выходной вероятностью положительного класса: LIDC-IDRI-0013, LIDC-IDRI-0037, LIDC-IDRI-0045, LIDC-IDRI-0061. Несмотря на наличие локальных зон повышенной активации, итоговая вероятность оставалась низкой и составляла около 0,127. Это показывает, что тепловая карта Grad-CAM не является самостоятельным аналогом LAA%-карты. Модель может реагировать на сосудисто-бронхиальные структуры, краевые участки лёгочного поля, локальные перепады плотности или

технические особенности изображения, которые сами по себе не приводят к положительной классификации.

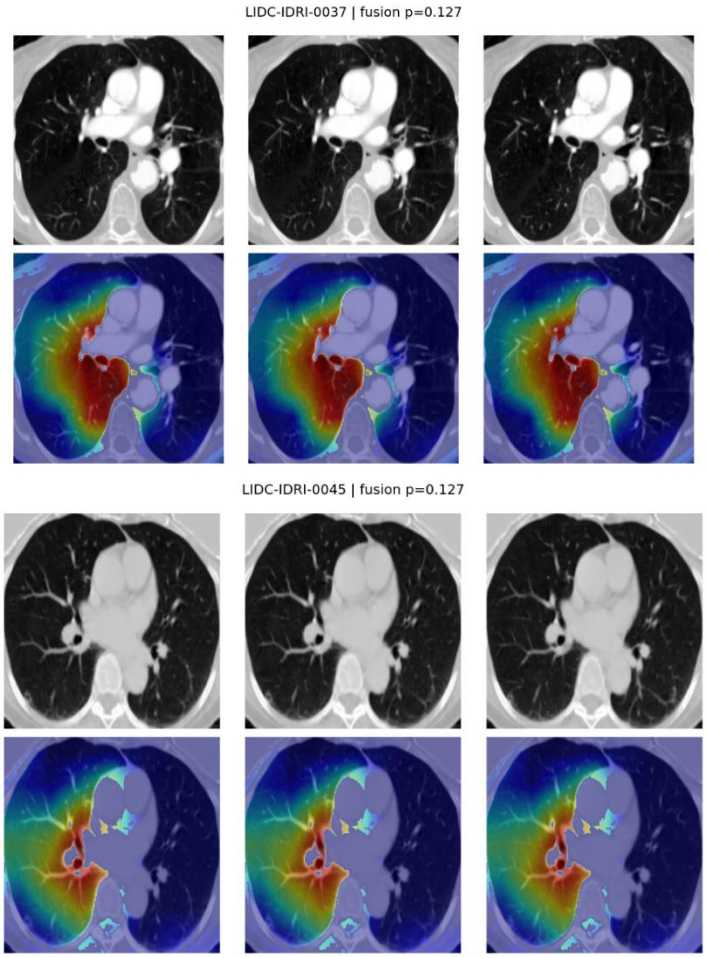


Рисунок 3.2.8-3- Grad-CAM-визуализация случаев с низкой выходной вероятностью положительного класса

Полученные Grad-CAM-карты позволяют сделать два методически значимых вывода. Во-первых, классификационное решение нейросетевого компонента формируется не только за счёт абстрактного числового embedding, но и связано с определёнными областями КТ-изображения. Во-вторых, зоны CNN-активации не полностью тождественны денситометрическим LAA-зонам, поскольку Grad-CAM отражает визуально-текстурную значимость области, а LAA%-950 является физическим показателем доли вокселей ниже заданного порога плотности.

Следовательно, Grad-CAM-анализ подтверждает интерпретируемость нейросетевого компонента, но одновременно подчёркивает необходимость его сопоставления с QCT-показателями. ResNet-152-эмбеддинги в данном контексте следует рассматривать как латентные визуально-текстурные дескрипторы КТ-изображения, тогда как LAA%-950, Perc15 и Mean HU обеспечивают физическую проверяемость результата. Такая комбинация усиливает обоснованность предложенной модели и снижает риск некритической интерпретации CNN-активаций как прямых морфологических маркеров эмфиземы.

Полученные результаты подтверждают, что диагностически значимая информация распределена в эмбеддинговом пространстве ResNet-152 неравномерно. Сокращение признакового пространства до ограниченного подмножества координат позволило сохранить и в отдельных вариантах повысить качество классификации, что подтверждает возможность использования латентного слоя ResNet-152 как источника компактных визуально-текстурных дескрипторов.

3.2.9 Внешняя валидация на данных открытого доступа NLST

В рамках исследования была проведена внешняя валидация предложенной модели на независимом международном датасете NLST (National Lung Screening Trial)[75]. Этот датасет представляет собой уникальную выборку данных КТ-скрининга курильщиков с высоким риском развития рака. Особенностью NLST является наличие у значительной части пациентов сопутствующих патологий, таких как ХОБЛ и эмфизема, что серьезно осложняет визуальную диагностику и делает внешнюю валидацию на этих данных максимально жесткой и приближенной к реальности. Использование данных из разных медицинских центров подтверждает, что система сохраняет свою эффективность вне зависимости от типа оборудования. Чтобы гарантировать научную достоверность, мы применили метод вложенной кросс-валидации (Nested Cross-Validation). Его суть заключается в двойном цикле проверки: внутренний цикл подбирает оптимальные параметры модели, а внешний — оценивает итоговую точность на данных, которые были полностью скрыты в процессе настройки.

Центральным элементом работы является механизм Late Fusion (позднее слияние). В отличие от простого объединения признаков, эта архитектура работает по принципу «консилиума». Выделили три независимых потока данных: клинический (анамнез), морфометрический (точные замеры узла) и интеллектуальный (глубокие признаки нейросети logit_1). Каждый поток обрабатывается отдельно, и только на финальном этапе система объединяет их веса для вынесения итогового вердикта. Это позволяет модели быть гибкой: если один из признаков оказывается неинформативным (например, когда при ХОБЛ границы узла размыты), система опирается на другие источники, такие как текстурные эмбеддинги нейросети.

Результаты тестирования наглядно демонстрируют преимущество такого подхода. На базовом этапе, используя только клинические данные, мы получили точность AUC 0.65. Добавление морфометрии (размера узла) практически не дало прироста (0.67), что подтверждает известную проблему: при наличии сопутствующих легочных заболеваний размер образования часто не коррелирует с его злокачественностью. Однако подключение модуля ИИ позволило достичь уровня AUC 0.9067, что наглядно представлено на рисунке 3.2.9-1.

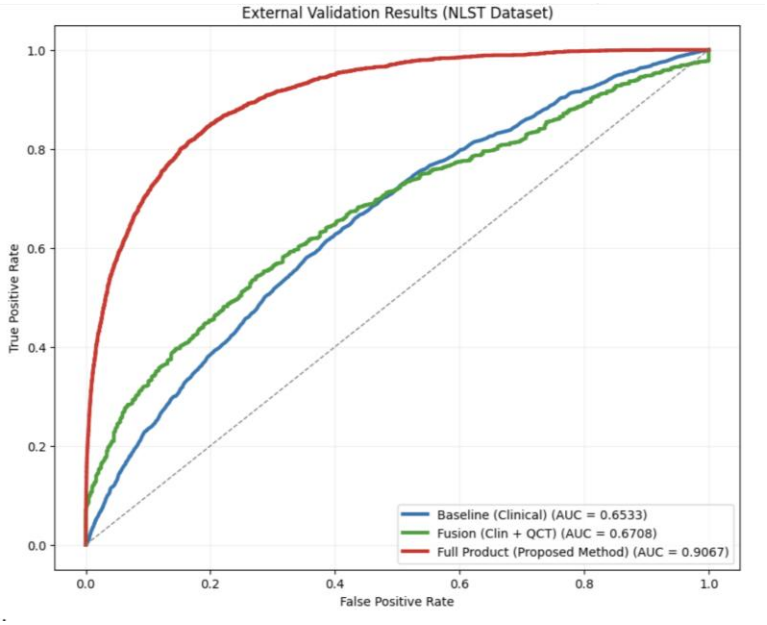


Рисунок 3.2.9-1- Результаты валидации на внешнем датасете

Важно отметить и улучшение показателя Brier Score, который снизился с 0.23 до 0.12. Это означает, что модель стала не просто точнее классифицировать, но и значительно лучше рассчитывать вероятность риска, что критически важно для принятия клинических решений в сложных случаях. В таблице представлены итоговые метрики для каждого этапа.

Таблица 3. 5- Сравнение метрик по этапам

Конфигурация модели	**AUC**	**Brier Score**
Baseline (Клиника)	0.6533	0.23175
Partial Fusion (Клиника + QCT)	0.6708	0.22068
Full Fusion (Proposed)	0.9067	0.12359

3.3 Реализация прототипа системы автоматического анализа КТ-изображений

Для практической проверки предложенного метода был разработан прототип системы автоматического анализа КТ-изображений лёгких. Его назначение состояло не в замене клинического заключения врача, а в реализации полного программного контура, обеспечивающего последовательную обработку DICOM-данных, расчёт количественных КТ-показателей, извлечение нейросетевых признаков и формирование структурированного результата анализа. Таким образом, четвёртая задача исследования была реализована на уровне экспериментального программного pipeline, пригодного для дальнейшей адаптации к задачам популяционного скрининга.

Прототип реализован как монорепозиторий, включающий два основных компонента: серверную часть youscan_backend и клиентскую часть youscan_frontend. Backend-компонент построен на Flask и выполняет функции API-шлюза, управления задачами, запуска ML Pipeline и передачи результатов во

frontend. Клиентская часть реализована на Next.js 14 с использованием React, TypeScript и App Router; она обеспечивает загрузку DICOM-файлов, запуск анализа, отслеживание статуса задачи и отображение итогового дашборда. В техническом описании системы указано, что применяемый стек включает Flask, TensorFlow/Keras, SimpleITK, NumPy, SciPy, ResNet-152 с весами model152.h5, а также возможность развёртывания через Waitress/Gunicorn и Docker-ready конфигурацию.

Пользовательский сценарий построен как последовательность из трёх этапов. На первом этапе выполняется загрузка DICOM-файлов через drag-and-drop интерфейс или стандартный диалог выбора файлов. После передачи данных на сервер создаётся уникальный идентификатор задачи job_id. На втором этапе пользователь инициирует анализ, после чего backend запускает обработку в отдельном рабочем потоке; frontend при этом выполняет периодический polling статуса задачи. На третьем этапе после завершения обработки результат запрашивается через API и отображается в виде интерактивного дашборда. Такая организация разделяет пользовательское взаимодействие и ресурсоёмкую обработку медицинских изображений, не блокируя интерфейс во время выполнения ML Pipeline.

Серверная часть предоставляет набор REST API-эндпоинтов, обеспечивающих полный жизненный цикл задачи. Эндпоинт POST /api/upload принимает DICOM-файлы и создаёт задачу; POST /api/analyze/<job_id> запускает вычислительный pipeline; GET /api/status/<job_id> возвращает текущее состояние анализа; GET /api/result/<job_id> предоставляет результат после завершения обработки; DELETE /api/job/<job_id> удаляет задачу и связанные временные файлы; GET /api/health используется для проверки доступности сервиса. Жизненный цикл задачи включает состояния idle, running, done и error, что позволяет формализованно отслеживать обработку от момента загрузки данных до получения результата. Для предотвращения конфликтов при параллельных запросах применяется потокобезопасное управление состоянием задач.

Центральным элементом системы является ML Pipeline, реализующий полный цикл обработки КТ-исследования. Последовательность включает загрузку DICOM-серии, преобразование изображений в шкалу Хаунсфилда, ресемплинг, сегментацию лёгочной ткани, расчёт QCT-показателей, нормализацию HU-значений, выделение анатомических регионов, извлечение признаков ResNet-152, ансамблирование региональных вероятностей и формирование итоговой структуры результата. В качестве QCT-показателей используются LAA%-950, Mean HU, Perc15 и TLV. Эти признаки выполняют двойную функцию: они участвуют в авторазметке и одновременно обеспечивают физическую интерпретацию результата.

Нейросетевой компонент системы основан на ResNet-152, используемой как экстрактор признаков. После предварительной обработки КТ-объём разделяется на анатомические регионы, для которых формируются входные представления и извлекаются эмбеддинги. Далее региональные вероятности агрегируются, формируя итоговую вероятность риска. Такой подход позволяет

учитывать неоднородность поражения лёгочной ткани, поскольку итоговое решение строится не по одному срезу, а по совокупности региональных признаков.

Результат анализа передаётся во frontend в виде структуры PipelineResult. Она включает вероятность заболевания, ансамблевую вероятность по регионам, текстовый уровень риска, QCT-метку, степень эмфизематозных изменений, региональные оценки, количественные КТ-показатели и флаги контроля качества. На пользовательском уровне эти данные отображаются в виде дашборда, объединяющего нейросетевую оценку и денситометрические параметры. Принципиально важно, что результат сопровождается клиническим предупреждением: система предназначена для поддержки анализа и предварительного выявления признаков риска, но не заменяет спирометрию, заключение рентгенолога и клиническое решение врача.

С технической точки зрения в прототипе реализованы механизмы, необходимые для безопасной и воспроизводимой обработки медицинских данных: ограничение CORS для доверенного frontend-источника, временное хранение загруженных файлов, удаление временных данных после завершения задачи, запрет повторного запуска уже выполняющейся обработки и конфигурирование ключевых параметров через переменные окружения. К числу таких параметров относятся путь к весам модели, директория загрузки DICOM-файлов, порт backend-сервера и число используемых слоёв ResNet-152.

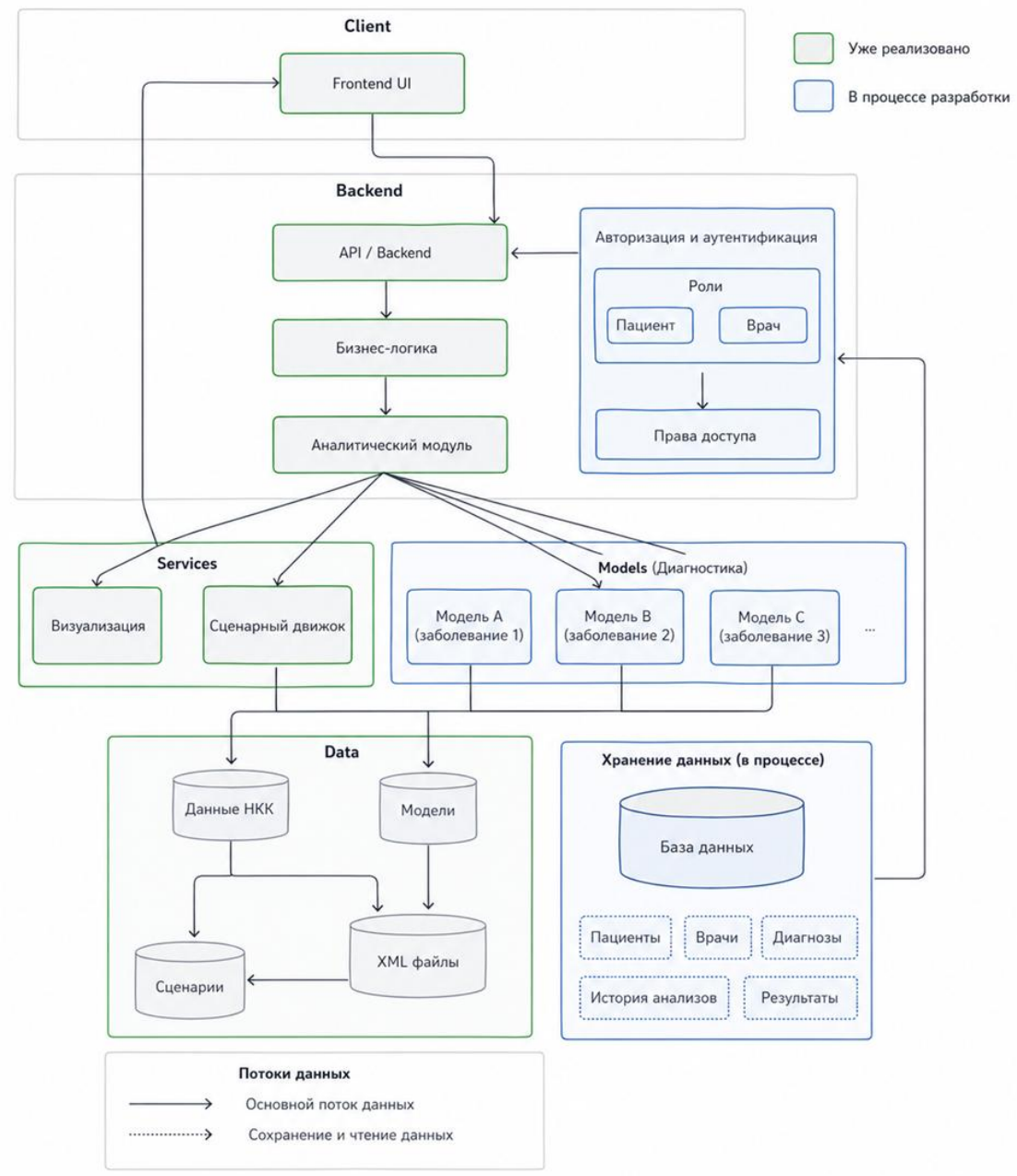


Рисунок 3.2.9-1 - Архитектура прототипа системы с модулем автоматического анализа КТ-изображений

На рисунке 3.2.9-1 - представлена архитектура прототипа системы YouScan.kz с модулем автоматического анализа КТ-изображений. Схема отражает разделение системы на пользовательский интерфейс, backend-компонент, ML Pipeline, модельный блок и модуль формирования результата, что обеспечивает воспроизводимый путь обработки данных от загрузки DICOM-серии до получения диагностического дашборда. Зелёным цветом выделены уже реализованные компоненты: клиентский Frontend UI, API/Backend, бизнес-логика, модель для риска ХОБЛ и аналитический модуль, сервисы визуализации и сценарного движка, а также локальное хранилище данных (данные НКК, модели, XML-файлы, сценарии). Голубым обозначены компоненты в стадии

доработки: подсистема авторизации и аутентификации с ролевой моделью с подтверждением согласия на сбор данных (роли «Пациент» и «Врач», права доступа), мультимодельный блок диагностики (Модель A, B, C для различных заболеваний) и персистентное хранилище данных на базе СУБД (таблицы: пациенты, врачи, диагнозы, история анализов, результаты). Стрелки показывают два вида потоков: основной поток данных (сплошные) и операции сохранения/чтения (пунктирные).

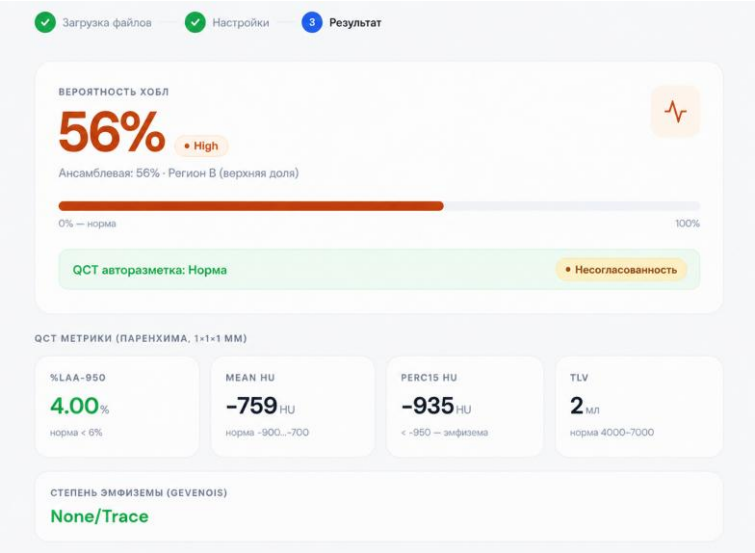


Рисунок 3.2.9-2- Результаты по пациенту X (датасет с данными из Казахстана) [70]

На рисунке представлены результаты анализа пациента в системе YouScan. Система определила уровень риска High (вероятность ХОБЛ - 56%), ансамблевая вероятность по регионам составила 56%, наибольший вклад внёс Регион B (верхняя доля). При этом QCT-авторазметка квалифицирует случай как Норма, что сопровождается предупреждением «Несогласованность»: расхождение между нейросетевым предсказанием и денситометрическими критериями характерно для пограничных случаев, требующих дополнительной клинической верификации.

QCT-метрики паренхимы (воксель 1×1×1 мм): %LAA-950 = 4,00% (норма <6%), Mean HU = −759 HU (в пределах нормы −900...−700), Perc15 HU = −935 HU (ниже порога −950, что соответствует признакам эмфиземы), TLV = 2 мл. Степень эмфиземы по шкале Gevenois: None/Trace (отсутствие или минимальные следовые изменения).

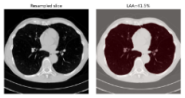
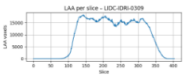
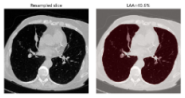
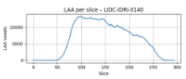
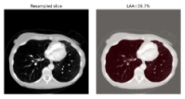
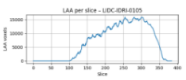
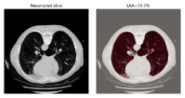
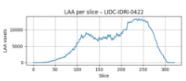
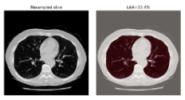
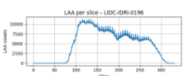
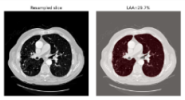
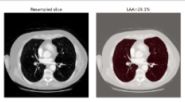
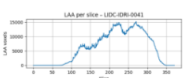
Patient ID	LAA%	Mean Lung HU	Label	Overlay Slice	LAA Curve
LIDC-IDRI-0309	41.49%	-905.0 HU	1		
LIDC-IDRI-0140	40.60%	-888.0 HU	1		
LIDC-IDRI-0105	39.66%	-889.0 HU	1		
LIDC-IDRI-0422	33.97%	-867.3 HU	1		
LIDC-IDRI-0196	33.40%	-865.0 HU	1		
LIDC-IDRI-0467	29.67%	-800.1 HU	0		
LIDC-IDRI-0041	29.06%	-878.7 HU	1		

Рисунок 3.2.9-3- Использование количественной метрики LAA% для автоматической разметки классов (Норма/Эмфизема) и визуальной проверки по КТ-снимкам

По завершении анализа пользователю отображается интерактивный дашборд, объединяющий в единой визуальной форме все ключевые диагностические показатели. Дашборд включает следующие компоненты: уровень риска ХОБЛ, принимающий одно из четырёх значений - Low, Moderate, High или Very High; числовое значение вероятности заболевания, визуализированное посредством прогресс-бара; количественные КТ-показатели - %LAA-950, Mean HU, Perc15 HU, TLV; степень эмфиземы по шкале Gevenois; региональные вероятности paper_A, paper_B и paper_C; предупреждения контроля качества сканирования (QC warnings); а также клинический дисклеймер, информирующий о статусе системы как инструмента поддержки принятия решений.

Таким образом, разработанный прототип демонстрирует прикладную реализуемость предложенного метода. В системе объединены компоненты загрузки DICOM-данных, автоматического препроцессинга, QCT-анализа, извлечения признаков ResNet-152, ансамблирования региональных оценок и формирования структурированного результата. Это позволяет рассматривать прототип как экспериментальную основу для последующего развития системы автоматического выявления КТ-признаков эмфиземы и ХОБЛ-ассоциированных изменений в сценариях массового скрининга.

ЗАКЛЮЧЕНИЕ

В ходе диссертационного исследования была решена задача разработки и экспериментальной проверки системы автоматического выявления КТ-признаков эмфизематозных изменений, ассоциированных с хронической обструктивной болезнью лёгких. Исследование было ориентировано не на замену клинической диагностики ХОБЛ, а на создание воспроизводимого вычислительного подхода, способного автоматически обрабатывать КТ-изображения, рассчитывать количественные денситометрические показатели, извлекать нейросетевые признаки и формировать вероятностную оценку наличия эмфизематозно-подобных изменений лёгочной паренхимы. С учётом того, что исходный датасет LIDC-IDRI не содержит спирометрии, GOLD-стадий и экспертной разметки ХОБЛ, поставленная задача была решена в рамках автоматического выявления морфологических КТ-признаков эмфиземы как одного из значимых компонентов ХОБЛ.

Цель исследования, заключавшаяся в разработке системы автоматического выявления признаков ХОБЛ по данным КТ-изображений с использованием модифицированного метода машинной диагностики, ориентированного на массовый популяционный скрининг, в целом достигнута. В рамках работы был реализован полный экспериментальный pipeline: от выбора DICOM-серий, перевода изображений в шкалу Хаунсфилда, ресемплинга и сегментации лёгких до расчёта QCT-показателей, формирования автоматической разметки, извлечения эмбеддингов ResNet-152, отбора информативных признаков, классификации и визуальной интерпретации результатов с использованием Grad-CAM. Дополнительно был разработан программный прототип системы автоматического анализа КТ-изображений, обеспечивающий загрузку DICOM-данных, запуск ML Pipeline, расчёт метрик и представление результата в структурированном виде.

Основная научная гипотеза исследования получила экспериментальное подтверждение. Предполагалось, что объединение латентных визуально-текстурных признаков, извлечённых сверточной нейронной сетью ResNet-152, с физически интерпретируемыми количественными показателями QCT позволит повысить качество классификации по сравнению с использованием этих групп признаков по отдельности. Полученные результаты подтвердили данное предположение. Изолированное использование эмбеддингов ResNet-152 обеспечивало наличие классификационного сигнала, однако его устойчивость была ограниченной. На тестовой выборке модель Embeddings Probe достигла ROC-AUC = 0,766, PR-AUC = 0,459, F1-score = 0,38 и balanced accuracy = 0,690. В то же время предложенный подход CNN + QCT, объединяющий нейросетевые признаки с LAA%-950, Perc15, Mean HU и TLV, продемонстрировал существенно более высокие показатели: ROC-AUC = 0,996, PR-AUC = 0,962, F1-score = 0,86 и balanced accuracy = 0,919. Это свидетельствует о том, что совместное использование визуально-текстурного и денситометрического уровней описания

КТ-изображения повышает согласованность классификационного решения и усиливает интерпретируемость результата.

В рамках первой задачи был проведён анализ латентного слоя предобученной ResNet-152. Было показано, что 2048-мерное пространство эмбеддингов содержит признаки, связанные с автоматически сформированной LAA%-разметкой, однако диагностически значимая информация распределена в нём неравномерно. Применение методов отбора признаков, включая t-test/FDR, L1-регуляризованную логистическую регрессию и RFE, позволило выделить более компактные подмножества координат эмбеддингового пространства. Это подтверждает, что латентный слой ResNet-152 может использоваться не только как «чёрный ящик», но и как источник признаков, поддающихся последующему статистическому анализу и снижению размерности. Данный результат имеет значение для построения более лёгких и интерпретируемых моделей, пригодных для скрининговых задач.

Вторая задача, связанная с реализацией стратегии автоматической разметки данных на основе QCT, была решена путём расчёта LAA%-950 и дополнительных денситометрических показателей внутри сегментированной маски лёгких. В результате исходный датасет, не содержащий специализированных меток ХОБЛ, был преобразован в аналитический массив с Silver Standard-разметкой. Такая разметка не является клинической истиной и не заменяет экспертное заключение, однако она воспроизводима, формализована и основана на физически измеряемых параметрах КТ. Было показано, что LAA%-950 согласуется со средней плотностью лёгочной ткани, а также связан с общим объёмом лёгких, что подчёркивает необходимость учитывать влияние вдоха и объёмного состояния лёгких при интерпретации денситометрических результатов.

В рамках третьей задачи был разработан модифицированный метод выявления признаков эмфиземы по КТ-изображениям, основанный на объединении эмбеддингов ResNet-152 и QCT-показателей. В отличие от подходов, использующих только пороговую денситометрию или только CNN-классификацию, предложенный подход сохраняет два взаимодополняющих уровня информации. CNN-компонент описывает визуально-текстурный паттерн лёгочной паренхимы, включая неоднородность структуры, распределение низкоплотных участков и особенности лёгочного рисунка. QCT-компонент обеспечивает связь результата с физически измеряемыми характеристиками ткани, выраженными в единицах Хаунсфилда и объёмных показателях. Именно это сочетание повышает не только численное качество классификации, но и объяснимость результата.

Четвёртая задача была реализована через разработку прототипа системы автоматического анализа КТ-изображений. Прототип включает frontend-компонент для загрузки DICOM-файлов и отображения результатов, backend-компонент для управления задачами и ML Pipeline для обработки изображений. В системе реализованы этапы загрузки DICOM-серии, препроцессинга, расчёта QCT-метрик, извлечения признаков ResNet-152, ансамблирования региональных

оценок и формирования итогового результата. Наличие такого прототипа показывает, что предложенный метод может быть встроен в прикладной программный контур и использован как основа для дальнейшего развития систем поддержки принятия решений в задачах КТ-скрининга.

Научная новизна исследования состоит в предложении подхода CNN + QCT для автоматического выявления эмфизематозно-подобных изменений по КТ-изображениям при отсутствии исходной клинической COPD-разметки. В работе объединены три методических элемента: автоматическая Silver Standard-разметка на основе QCT, извлечение латентных признаков ResNet-152 и объединение нейросетевого признакового пространства с физически интерпретируемыми денситометрическими показателями. Такой подход позволяет использовать открытые КТ-датасеты, изначально не предназначенные для ХОБЛ, и адаптировать их к задачам машинного анализа структурных изменений лёгочной ткани.

Практическая значимость исследования заключается в том, что разработанный pipeline может применяться как основа для предварительного автоматизированного анализа КТ-исследований в скрининговых сценариях. Он позволяет без ручной экспертной разметки сформировать исследовательские классы, рассчитать ключевые QCT-показатели, извлечь нейросетевые признаки, оценить вероятность наличия эмфизематозно-подобных изменений и представить результат в структурированной форме. При дальнейшей валидации на специализированных клинических данных такой подход может быть использован в системах поддержки принятия решений, PACS-инфраструктуре, локальных медицинских информационных системах и Edge AI-средах.

Вместе с тем результаты исследования требуют осторожной интерпретации. Главное ограничение связано с использованием LIDC-IDRI, который не содержит клинической разметки ХОБЛ, спирометрии и GOLD-стадий. Следовательно, модель выявляет не ХОБЛ как клинический диагноз, а КТ-признаки низкоплотных изменений лёгочной ткани, ассоциированных с эмфизематозным компонентом заболевания. Второе ограничение связано с тем, что метки формировались на основе QCT-критериев. Поэтому высокие результаты предложенного варианта CNN + QCT отражают внутреннюю согласованность вычислительного pipeline и не могут рассматриваться как окончательное доказательство клинической диагностической точности. Третье ограничение обусловлено чувствительностью LAA%-950 к параметрам сканирования, уровню вдоха, качеству сегментации и реконструкционным характеристикам КТ.

Дополнительное ограничение связано с использованием 2D-патчей и ResNet-152 как экстрактора признаков. Такой подход позволил использовать предобученную архитектуру и снизить вычислительную сложность, однако частично ограничил учёт трёхмерного контекста КТ-исследования. Grad-CAM-анализ показал, что модель действительно использует области внутри лёгочных полей, однако зоны активации CNN не являются прямым аналогом LAA-карт и не должны трактоваться как самостоятельная морфологическая разметка

эмфиземы. Поэтому нейросетевые признаки требуют обязательного сопоставления с QCT-показателями и экспертной клинической оценкой.

Перспективы дальнейшего развития исследования связаны прежде всего с валидацией предложенного подхода на других датасетах. Необходимо проверить модель на специализированных COPD-выборках, содержащих спирометрию, GOLD-стадии, экспертную оценку эмфиземы и клинические данные пациентов. Такая проверка позволит оценить не только внутреннюю согласованность auto-labeling, но и реальную диагностическую значимость модели. Особенно важным направлением является внешняя валидация на данных, полученных в других центрах, на других КТ-сканерах и при иных протоколах реконструкции, поскольку именно устойчивость к доменному сдвигу определяет пригодность модели для популяционного скрининга.

Второе направление дальнейшей работы связано с расширением области применения предложенного подхода. Принцип объединения CNN-признаков и QCT-показателей может быть адаптирован не только к эмфиземе и ХОБЛ, но и к другим заболеваниям лёгких, в которых важны как визуально-текстурные паттерны, так и количественные характеристики ткани. К таким задачам могут относиться интерстициальные заболевания лёгких, поствоспалительные изменения, фиброзные изменения, бронхоэктазы, air trapping, последствия инфекционных поражений и оценка структурных изменений при длительном наблюдении пациента. Для каждой из этих задач потребуется собственная система признаков и критериев разметки, однако общий принцип - объединение нейросетевого представления и физически интерпретируемых количественных показателей - сохраняет методологическую значимость.

Третье направление связано с развитием архитектуры модели. В дальнейшем целесообразно сравнить ResNet-152 с 3D-CNN, ConvLSTM, attention-based MIL и transformer-подходами, способными учитывать объёмный контекст КТ-исследования. Это особенно важно для выявления ранних и регионально ограниченных изменений, которые могут быть недостаточно представлены в отдельных 2D-патчах. Дополнительно перспективным является внедрение механизмов uncertainty estimation, позволяющих сопровождать результат не только вероятностью класса, но и оценкой уверенности модели.

Четвёртое направление связано с подготовкой модели к практическому скрининговому применению. Для этого требуется оптимизация вычислительной сложности, включая pruning, quantization, knowledge distillation и переход к более лёгким архитектурам. Такое развитие позволит адаптировать систему к Embedded AI и Edge AI-сценариям, при которых анализ КТ-исследований выполняется локально - на сервере медицинской организации, диагностической рабочей станции или в составе PACS-инфраструктуры. Для массового скрининга это особенно важно, поскольку обработка больших массивов данных требует не только высокой точности, но и высокой скорости, устойчивости и экономичности вычислений.

Пятое направление связано с развитием программного прототипа. Для перехода от экспериментальной системы к прикладному решению необходимо

реализовать персистентное хранилище результатов, ролевую модель доступа, журналирование действий пользователя, пакетную обработку DICOM-серий, экспорт структурированного отчёта, встроенный DICOM-вьюер и инструменты врачебной аннотации. Кроме того, требуется расширить модуль интерпретации: сопоставлять Grad-CAM-карты с LAA-картами, отображать региональные QCT-показатели и формировать объяснимое заключение, пригодное для последующей экспертной оценки.

Таким образом, проведённое исследование подтвердило возможность построения автоматизированного подхода к выявлению КТ-признаков эмфизематозных изменений на основе совместного использования ResNet-152 и QCT. Цель исследования достигнута, поставленные задачи выполнены, а гипотеза о преимуществе объединения нейросетевых и денситометрических признаков получила экспериментальное подтверждение. Полученные результаты создают основу для дальнейшего развития систем автоматического анализа КТ-изображений, ориентированных на раннее выявление структурных изменений лёгочной ткани и применение в задачах массового популяционного скрининга.

СПИСОК ЛИТЕРАТУРЫ

[1] Bazarbekova, G., Inkarbekov, M., Qumar, A. B., Автор Г.Г., Автор Д.Д. Mortality trends from respiratory disease in Kazakhstan: A 2011–2021 analysis. Journal of Public Health Research. 2025;XX(X):страницы. doi:10.1177/22799036251341526

[2] Sharman, A. T., Автор Б.Б., Автор B.B. Cross-Sectional Study of COPD Prevalence Among Smokers, Ex-smokers and Never-Smokers in Almaty, Kazakhstan. ClinicalTrials.gov: NCT02926534. 2016–2017.

[3] Jung, T., & Vij, N. (2021). Early diagnosis and real-time monitoring of regional lung function changes to prevent chronic obstructive pulmonary disease progression to severe emphysema. *Journal of Clinical Medicine*, *10*(24), 5811. https://doi.org/10.3390/jcm10245811

[4] Kimura, T., Kawakami, T., Kikuchi, A., Ooev, R., Akiyama, M., & Horikoshi, H. (2018). A study on diagnostic assist systems of chronic obstructive pulmonary disease from medical images by deep learning. *J. Comput. Commun*, *6*(1), 21-31. https://doi.org/10.4236/jcc.2018.61003

[5] Bhaskar, R., Singh, S., & Singh, P. (2018). Characteristics of COPD phenotypes classified according to the findings of HRCT and spirometric indices and its correlation to clinical characteristics. *African Health Sciences*, *18*(1), 90-101. https://doi.org/10.4314/ahs.v18i1.13

[6] Kang, H. S., Bak, S. H., Oh, H. Y., Lim, M. N., Cha, Y. K., Yoon, H. J., & Kim, W. J. (2021). Computed tomography-based visual assessment of chronic obstructive pulmonary disease: comparison with pulmonary function test and quantitative computed tomography. *Journal of Thoracic Disease*, *13*(3), 1495. https://doi.org/10.21037/jtd-20-3041

[7] Hashimoto, N., Wakahara, K., & Sakamoto, K. (2021). The importance of appropriate diagnosis in the practical management of chronic obstructive pulmonary disease. *Diagnostics*, *11*(4), 618. https://doi.org/10.3390/diagnostics11040618

[8] Gawlitza, J., Trinkmann, F., Scheffel, H., Fischer, A., Nance, J. W., Henzler, C., ... & Henzler, T. (2018). Time to exhale: additional value of expiratory chest CT in chronic obstructive pulmonary disease. *Canadian respiratory journal*, *2018*(1), 9493504.https://doi.org/10.1155/2018/9493504

[9] Kaya, L., Özel, D., & Özel, B. D. (2017). Evaluating qualitative and quantitative computerized tomography indicators of chronic obstructive pulmonary disease and their correlation with pulmonary function tests. *Polish Journal of Radiology*, *82*, 511-515. https://doi.org/10.12659/PJR.901968

[10] Wesseling, G. J., & Van Schayck, C. P. (2006). Why act early in COPD?. *Monaldi archives for chest disease*, *65*(3). https://doi.org/10.4081/monaldi.2006.561

[11] Choi, J. Y., & Rhee, C. K. (2020). Diagnosis and treatment of early chronic obstructive lung disease (COPD). *Journal of Clinical Medicine*, *9*(11), 3426. https://doi.org/10.3390/jcm9113426

[12] Wanot, B., Stelmach, A., & Biskupek-Wanot, A. MEDICAL SCIENCES. http://dx.doi.org/10.23856/2715

[13] Jokhdar, H. A., Garout, M. A., Alkot, M. M., El-Helbawy, R. H., & Albitar, J. O. (2016). Early Detection of Chronic Obstructive Pulmonary Disease (COPD) in Family Practice. *Journal of Family Medicine and Health Care*, *3*(1), 1-5. https://doi.org/10.11648/j.jfmhc.20170301.11

[14] Yip, K. P., Stockley, R. A., & Sapey, E. (2021). Catching "Early" COPD–the diagnostic conundrum. *International journal of chronic obstructive pulmonary disease*, 957-968. https://doi.org/10.2147/COPD.S296842

[15] Song, J. H., & Kim, Y. (2024). Beyond the spirometry: new diagnostic modalities in chronic obstructive pulmonary disease. *Tuberculosis and Respiratory Diseases*, *88*(1), 1. https://doi.org/10.4046/trd.2024.0040

[16] Kellerer, C., Jörres, R. A., Schneider, A., Alter, P., Kauczor, H. U., Jobst, B., ... & Kahnert, K. (2021). Prediction of lung emphysema in COPD by spirometry and clinical symptoms: results from COSYCONET. *Respiratory research*, *22*(1), 242. https://doi.org/10.1186/s12931-021-01837-2

[17] Fazleen, A., & Wilkinson, T. (2020). Early COPD: current evidence for diagnosis and management. *Therapeutic advances in respiratory disease*, *14*, 1753466620942128. https://doi.org/10.1177/1753466620942128

[18] Lowe, K. E., Regan, E. A., Anzueto, A., Austin, E., Austin, J. H., Beaty, T. H., ... & Walsh, J. W. (2019). COPDGene® 2019: redefining the diagnosis of chronic obstructive pulmonary disease. *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation*, *6*(5), 384. https://doi.org/10.15326/jcopdf.6.5.2019.0149

[19] Kjeldgaard, P., Lykkegaard, J., Spillemose, H., & Ulrik, C. S. (2017). Multicenter study of the COPD-6 screening device: feasible for early detection of chronic obstructive pulmonary disease in primary care?. *International journal of chronic obstructive pulmonary disease*, 2323-2331. https://doi.org/10.2147/COPD.S136244

[20] FDA-NIH Biomarker Working Group. (2016). BEST (Biomarkers, EndpointS, and other Tools) Resource. U.S. Food and Drug Administration; National Institutes of Health. https://www.ncbi.nlm.nih.gov/books/NBK338449/

[21] Sorensen, L., Nielsen, M., Lo, P., Ashraf, H., Pedersen, J. H., & De Bruijne, M. (2011). Texture-based analysis of COPD: a data-driven approach. *IEEE transactions on medical imaging*, *31*(1), 70-78. https://doi.org/10.1109/TMI.2011.2164931

[22] Liu, G., Hu, J., Yang, J., & Song, J. (2024). Predicting early-onset COPD risk in adults aged 20–50 using electronic health records and machine learning. *PeerJ*, *12*, e16950. https://doi.org/10.7717/peerj.16950

[23] Karthik, R., Murugan, R., & Goel, T. (2026). Segmentation and visualization of Emphysema lesions through Computed Tomography images. *Architecture Image Studies*, *7*(1), 1339-1352. https://doi.org/10.62754/ais.v7i1.1020

[24] Oh, A. S., Humphries, S. M., Chung, A., Weigt, S. S., Brown, M., Kim, G. H. J., ... & Goldin, J. G. (2026). Quantitative CT and Artificial Intelligence in Chronic

Lung Disease. *Journal of Thoracic Imaging*, *41*(2), e0867. https://doi.org/10.1097/rti.0000000000000867

[25] Li, F., Choi, J., Zou, C., Newell Jr, J. D., Comellas, A. P., Lee, C. H., ... & Lin, C. L. (2021). Latent traits of lung tissue patterns in former smokers derived by dual channel deep learning in computed tomography images. *Scientific reports*, *11*(1), 4916. https://doi.org/10.1038/s41598-021-84547-5

[26] Lee, S., Lee, I. S., & Kim, S. (2023, July). Predicting development of chronic obstructive pulmonary disease and its risk factor analysis. In *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 1-4). IEEE. https://doi.org/10.1109/EMBC40787.2023.10340286

[27] Zhao, S., Wang, M., Xu, J., Feng, J., Qian, W., Chen, R., ... & Wu, Y. (2026). A self-training framework for semi-supervised pulmonary vessel segmentation and its application in COPD. *Journal of X-ray Science and Technology*, *34*(1), 39-55. https://doi.org/10.1177/08953996251384489

[28] Alve, S. R., Mahmud, M. Z., Islam, S., & Khan, M. M. (2024). Chronic obstructive pulmonary disease prediction using deep convolutional network. *arXiv preprint arXiv:2411.02449*. https://doi.org/10.48550/arXiv.2411.02449

[29] Dhar, J. (2021). Multistage ensemble learning model with weighted voting and genetic algorithm optimization strategy for detecting chronic obstructive pulmonary disease. *IEEE Access*, *9*, 48640-48657. https://doi.org/10.1109/ACCESS.2021.3067949

[30] Смелкина, Н. А., Колсанов, А. В., Чаплыгин, С. С., Зельтер, П. М., & Храмов, А. Г. (2017). Распознавание эмфиземы лёгких по данным компьютерной томографии. *Компьютерная оптика*, *41*(5), 726-731. https://doi.org/10.18287/2412-6179-2017-41-5-726-731

[31] Ibrahim, M. A., & Mukundan, R. (2015). Cascaded techniques for improving emphysema classification in computed tomography images. *Artif. Intell. Res.*, *4*(2), 112-118. https://doi.org/10.5430/air.v4n2p112

[32] Tarando, S. R. (2018). *Quantitative follow-up of pulmonary diseases using deep learning models* (Doctoral dissertation, Institut National des Télécommunications). https://theses.hal.science/tel-01892382

[33] Gawlitza, J., Sturm, T., Spohrer, K., Henzler, T., Akin, I., Schönberg, S., ... & Trinkmann, F. (2019). Predicting pulmonary function testing from quantified computed tomography using machine learning algorithms in patients with COPD. *Diagnostics*, *9*(1), 33. https://doi.org/10.3390/diagnostics9010033

[34] Sarsembayeva, T., Mansurova, M., Oshibayeva, A., & Serebryakov, S. (2026). Interpretable Diagnosis of Pulmonary Emphysema on Low-Dose CT Using ResNet Embeddings. *Journal of Imaging*, *12*(1), 51. https://doi.org/10.3390/jimaging12010051

[35] Xie, W., Jacobs, C., Charbonnier, J. P., Slebos, D. J., & van Ginneken, B. (2023). Emphysema subtyping on thoracic computed tomography scans using deep neural networks. *Scientific Reports*, *13*(1), 14147. https://doi.org/10.1038/s41598-023-40116-6

[36] Wysoczanski, A. (2024). *Machine-Learned Anatomic Subtyping, Longitudinal Disease Evaluation and Quantitative Image Analysis on Chest Computed Tomography: Applications to Emphysema, COPD, and Breast Density*. Columbia University.

[37] Abd elsalam, S. M., Hafez, M., Mohmed, M. F., & Said, A. H. (2020). Correlation between quantitative multi-detector computed tomography lung analysis and pulmonary function tests in chronic obstructive pulmonary disease patients. *Egyptian Journal of Radiology and Nuclear Medicine*, *51*(1), 160. https://doi.org/10.1186/s43055-020-00281-4

[38] Schwartz, A. G., Lusk, C. M., Wenzlaff, A. S., Watza, D., Pandolfi, S., Mantha, L., ... & Gadgeel, S. M. (2016). Risk of lung cancer associated with COPD phenotype based on quantitative image analysis. *Cancer Epidemiology, Biomarkers & Prevention*, *25*(9), 1341-1347. https://doi.org/10.1158/1055-9965.EPI-16-0176

[39] Almeida, S. D., Norajitra, T., Lüth, C. T., Wald, T., Weru, V., Nolden, M., ... & Maier-Hein, K. (2024). Prediction of disease severity in COPD: a deep learning approach for anomaly-based quantitative assessment of chest CT. *European radiology*, *34*(7), 4379-4392. https://doi.org/10.1007/s00330-023-10540-3

[40] Li, Z., Huang, K., Liu, L., & Zhang, Z. (2022). Early detection of COPD based on graph convolutional network and small and weakly labeled data. *Medical & Biological Engineering & Computing*, *60*(8), 2321-2333. https://doi.org/10.1007/s11517-022-02589-x

[41] Ahmed, J., Vesal, S., Durlak, F., Kaergel, R., Ravikumar, N., Rémy-Jardin, M., & Maier, A. (2020, February). COPD classification in CT images using a 3D convolutional neural network. In *Bildverarbeitung für die Medizin 2020: Algorithmen–Systeme–Anwendungen. Proceedings des Workshops vom 15. bis 17. März 2020 in Berlin* (pp. 39-45). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-29267-6_8

[42] Gomes, P., Bastos, H. N. E., Carvalho, A., Lobo, A., Guimarães, A., Rodrigues, R. S., ... & Carvalho, A. R. S. (2021). Pulmonary emphysema regional distribution and extent assessed by chest computed tomography is associated with pulmonary function impairment in patients with COPD. *Frontiers in medicine*, *8*, 705184. https://doi.org/10.3389/fmed.2021.705184

[43] Dai, Q., Zhu, X., Hua, Q., Zhang, J., Dong, Z., Zheng, J., & Shi, J. (2026). Single-inspiratory quantitative CT nomogram for enhanced PRISm and COPD differentiation: a cross-sectional study with interpretable diagnostic boundaries. *PeerJ*, *14*, e20575. https://doi.org/10.7717/peerj.20575

[44] Fu, L., Cui, Y., Wang, X., Luo, H., Wu, Y., Wei, Q., ... & Long, L. (2025). Machine learning models using dual-phase CT radiomics for early detection of PRISm. *Scientific Reports*, *15*(1), 39409. https://doi.org/10.1038/s41598-025-21995-3

[45] Zhang, X., Zhao, M., Yao, F., Ma, W., Zhang, J., Li, Y., ... & Liu, S. (2026). A self-supervised framework for emphysema anomaly detection and staging in computed tomography scans. *Patterns*, *7*(2). https://doi.org/10.1016/j.patter.2025.101426

[46] Humphries, S. M., Notary, A. M., Centeno, J. P., Strand, M. J., Crapo, J. D., Silverman, E. K., ... & Genetic Epidemiology of COPD (COPDGene) Investigators. (2020). Deep learning enables automatic classification of emphysema pattern at CT. *Radiology*, *294*(2), 434-444. https://doi.org/10.1148/radiol.2019191022

[47] Sørensen, L., Loog, M., Lo, P., Ashraf, H., Dirksen, A., Duin, R. P., & De Bruijne, M. (2010, September). Image dissimilarity-based quantification of lung disease from CT. In *International Conference on medical image computing and Computer-Assisted Intervention*(pp. 37-44). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-15705-9_5

[48] Bodduluri, S., Newell Jr, J. D., Hoffman, E. A., & Reinhardt, J. M. (2013). Registration-based lung mechanical analysis of chronic obstructive pulmonary disease (COPD) using a supervised machine learning framework. *Academic radiology*, *20*(5), 527-536. https://doi.org/10.1016/j.acra.2013.01.019

[49] Armato III, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., ... & Clarke, L. P. (2011). The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical physics*, *38*(2), 915-931. https://doi.org/10.1118/1.3528204

[50] Cheplygina, V., Pena, I. P., Pedersen, J. H., Lynch, D. A., Sørensen, L., & De Bruijne, M. (2017). Transfer learning for multicenter classification of chronic obstructive pulmonary disease. *IEEE journal of biomedical and health informatics*, *22*(5), 1486-1496. https://doi.org/10.1109/JBHI.2017.2769800

[51] Almeida, S. D., Norajitra, T., Lüth, C. T., Wald, T., Weru, V., Nolden, M., ... & Maier-Hein, K. (2024). How do deep-learning models generalize across populations? Cross-ethnicity generalization of COPD detection. *Insights Into Imaging*, *15*, 198. https://doi.org/10.1186/s13244-024-01781-x

[52] Yang, Y., Wang, S., Zeng, N., Duan, W., Chen, Z., Liu, Y., ... & Kang, Y. (2022). Lung radiomics features selection for COPD stage classification based on auto-metric graph neural network. *Diagnostics*, *12*(10), 2274. https://doi.org/10.3390/diagnostics12102274

[53] Zhou, T., Zhou, X., Ni, J., Guan, Y., Jiang, X. A., Lin, X., ... & Fan, L. (2024). A CT-Based Lung Radiomics Nomogram for Classifying the Severity of Chronic Obstructive Pulmonary Disease. *International journal of chronic obstructive pulmonary disease*, 2705-2717. https://doi.org/10.2147/COPD.S483007

[54] Cheplygina, V., Sørensen, L., Tax, D. M., Pedersen, J. H., Loog, M., & De Bruijne, M. (2014, August). Classification of COPD with multiple instance learning. In *2014 22nd International Conference on pattern recognition* (pp. 1508-1513). IEEE. https://doi.org/10.1109/ICPR.2014.268

[55] Braman, N., Beymer, D., & Dehghan, E. (2018). Disease detection in weakly annotated volumetric medical images using a convolutional LSTM network. *arXiv preprint arXiv:1812.01087*. https://doi.org/10.48550/arXiv.1812.01087

[56] Nyboe Ørting, S., Petersen, J., Cheplygina, V., Thomsen, L. H., Wille, M. M., & De Bruijne, M. (2018, September). Feature learning based on visual similarity triplets in medical image analysis: A case study of emphysema in chest CT scans.

In *International Workshop on Large-scale Annotation of Biomedical data and Expert Label Synthesis* (pp. 140-149). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-01364-6_16

[57] Yang, J., Angelini, E. D., Smith, B. M., Austin, J. H., Hoffman, E. A., Bluemke, D. A., ... & Laine, A. F. (2016). Explaining radiological emphysema subtypes with unsupervised texture prototypes: MESA COPD study. In *Bayesian and grAphical Models for Biomedical Imaging* (pp. 69-80). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-61188-4_7

[58] Zhang, L., Jiang, B., Wisselink, H. J., Vliegenthart, R., & Xie, X. (2022). COPD identification and grading based on deep learning of lung parenchyma and bronchial wall in chest CT images. *The British Journal of Radiology*, *95*(1133), 20210637. https://doi.org/10.1259/bjr.20210637

[59] Sørensen, L. (2010). Pattern Recognition-Based Analysis of COPD in CT. *Faculty of Science*.

[60] Moslemi, A. (2022). *Quantitative Computed Tomography Imaging and Machine Learning for Evaluating Chronic Obstructive Pulmonary Disease* (Doctoral dissertation, Toronto Metropolitan University).

[61] Nagaraj, Y., Wisselink, H. J., Rook, M., Cai, J., Nagaraj, S. B., Sidorenkov, G., ... & van Ooijen, P. (2022). AI-driven model for automatic emphysema detection in low-dose computed tomography using disease-specific augmentation. *Journal of Digital Imaging*, *35*(3), 538-550. https://doi.org/10.1007/s10278-022-00599-7

[62] Toews, M., Wachinger, C., Estepar, R. S. J., & Wells III, W. M. (2015, June). A feature-based approach to big data analysis of medical images. In *International Conference on Information Processing in Medical Imaging* (pp. 339-350). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-19992-4_26

[63] Wang, J. M., Labaki, W. W., Murray, S., Martinez, F. J., Curtis, J. L., Hoffman, E. A., ... & Hatt, C. (2023). Machine learning for screening of at-risk, mild and moderate COPD patients at risk of FEV1 decline: results from COPDGene and SPIROMICS. *Frontiers in physiology*, *14*, 1144192. https://doi.org/10.3389/fphys.2023.1144192

[64] Ørting, S. N., Petersen, J., Thomsen, L. H., Wille, M. M., & de Bruijne, M. (2019). Learning to quantify emphysema extent: What labels do we need?. *IEEE Journal of Biomedical and Health Informatics*, *24*(4), 1149-1159. https://doi.org/10.1109/JBHI.2019.2932145

[65] Argus, M., Schaefer-Prokop, C., Lynch, D. A., & van Ginneken, B. (2019). Function follows form: regression from complete thoracic computed tomography scans. *arXiv preprint arXiv:1909.12047*. https://doi.org/10.48550/arXiv.1909.12047

[66] Guan, Y., Zhang, D., Zhou, X., Xia, Y., Lu, Y., Zheng, X., ... & Fan, L. (2024). Comparison of deep-learning and radiomics-based machine-learning methods for the identification of chronic obstructive pulmonary disease on low-dose computed tomography images. *Quantitative Imaging in Medicine and Surgery*, *14*(3), 2485-2498. https://doi.org/10.21037/qims-23-1307

[67] Bortsova, G., Dubost, F., Ørting, S., Katramados, I., Hogeweg, L., Thomsen, L., ... & de Bruijne, M. (2018, September). Deep learning from label proportions for emphysema quantification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 768-776). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-00934-2_85

[68] Pino Peña, I., Cheplygina, V., Paschaloudi, S., Vuust, M., Carl, J., Weinreich, U. M., ... & de Bruijne, M. (2018). Automatic emphysema detection using weakly labeled HRCT lung images. *PloS one*, *13*(10), e0205397. https://doi.org/10.1371/journal.pone.0220873

[69] Mondal, S., Sadhu, A. K., & Dutta, P. K. (2021). Adaptive local ternary pattern on parameter optimized-faster region convolutional neural network for pulmonary emphysema diagnosis. *IEEE Access*, *9*, 114135-114152. https://doi.org/10.1109/ACCESS.2021.3105114

[70] Tang, L. Y., Coxson, H. O., Lam, S., Leipsic, J., Tam, R. C., & Sin, D. D. (2020). Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT. *The Lancet Digital Health*, *2*(5), e259-e267. https://doi.org/10.1016/S2589-7500(20)30064-9

[71] Fuhrman, J., Yip, R., Zhu, Y., Jirapatnakul, A. C., Li, F., Henschke, C. I., ... & Giger, M. L. (2023). Evaluation of emphysema on thoracic low-dose CTs through attention-based multiple instance deep learning. *Scientific Reports*, *13*(1), 1187. https://doi.org/10.1038/s41598-023-27549-9

[72] Armato, S. G., III, McLennan, G., Bidaut, L., … Clarke, L. P. (2015). Data from LIDC-IDRI [Data set]. The Cancer Imaging Archive. https://doi.org/10.7937/K9/TCIA.2015.LO9QL9SX

[73] Sarsembayeva T., Mansurova M., Abdildayeva A., Serebryakov S. (2025) Enhancing U-Net Segmentation Accuracy Through Comprehensive Data Preprocessing. *Journal of Imaging.* 11; (2): 50. https://doi.org/10.3390/jimaging11020050

[74] Sarsembayeva T., Oshibayeva A., Shayakhmetova A., Ospan A. (2026) ResNet embedding-based pipeline for transparent diagnosis of pulmonary emphysema on low-dose CT // *Journal of Problems in Computer Science and Information Technologies.* –. – Vol. 4, No. 1. DOI: https://doi.org/10.26577/jpcsit4120262

[75] https://www.cancerimagingarchive.net/collection/nlst/

ПРИЛОЖЕНИЕ А- Ссылка программного кода

https://github.com/SarsembayevaTalshyn/COPD_AI_in_Med.git



ПРИЛОЖЕНИЕ Б – Список публикаций

Scopus/ WoS:

1. Sarsembayeva T., Mansurova M., Oshibayeva A., Serebryakov S. Interpretable Diagnosis of Pulmonary Emphysema on Low-Dose CT Using ResNet Embeddings. Journal of Imaging. 2026; 12(1):51. (Q1, 87th percentile) https://doi.org/10.3390/jimaging12010051

2. Sarsembayeva T., Mansurova M., Abdildayeva A., Serebryakov S. Enhancing U-Net Segmentation Accuracy Through Comprehensive Data Preprocessing. Journal of Imaging. 11 2025; (2): 50. https://doi.org/10.3390/jimaging11020050 (Q1, 87th percentile)

КОКСНВО РК:

1. Sarsembayeva T., Oshibayeva A., Shayakhmetova A., Ospan A. ResNet embedding-based pipeline for transparent diagnosis of pulmonary emphysema on low-dose CT // Journal of Problems in Computer Science and Information Technologies. – 2026. – Vol. 4, No. 1. DOI: https://doi.org/10.26577/jpcsit4120262

Международные конференции:

1. Sarsembayeva, T., Mansurova, M., Shomanov, A., Sarsembayev, M., Sagyzbayeva, S., Rakhimzhanov, G. (2022). Pre-processing of CT Images of the Lungs 14th Asian Conference on Intelligent Information and Database Systems , ACIIDS 2022. // Lecture Notes in Computer Science(), vol 13758. Springer, Cham. https://doi.org/10.1007/978-3-031-21967-2_39 (Scopus Proceedings)

2. Sarsembayeva T., Shomanov A., Sarsembayev M., Mansurova M., Zhunussova A., Rakhimzhanov G. UNet model for segmentation of COPD lung lesions on computed tomography images // CEUR Workshop Proceedings. – 2022. – Vol. 3382.URL:https://ceur-ws.org/Vol-3382/Short5.pdf (Scopus Proceedings)

3. Sarsembayeva T., Mansurova M., Essengaliyeva N., Koshanova A. Өкпенің созылмалы обструктивті ауруы белгілерінің машиналық диагностикасы үшін өкпенің КТ кескінін өңдеу жүйесін жасау // Материалы IX Международной научно-практической конференции «Computer Science and Applied Mathematics» (CSAM 2024). – Алматы, 2024. – С. 419–428. – URL: https://conf.iict.kz/wp-content/uploads/2025/01/collection_CSAM_IX_2024.pdf