

Разработка интеллектуальной системы для анализа рентгенограмм грудной клетки на основе сверточных нейронных сетей: сегментация анатомических структур и автоматический расчет кардиоторакального индекса

П. М. Нурахунов^{1✉}, А. А. Мелдо^{1,2}, И. Ю. Прохоров²

¹ Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина), Санкт-Петербург, Россия

² Клиника высоких технологий «Белоостров» (ООО «ММЦ ВТ»), Санкт-Петербург, Россия

✉ pankrat.nurakhunov@mail.ru

Аннотация. Представлены результаты разработки интеллектуальной системы для автоматизированного анализа рентгенограмм органов грудной клетки, включая этапы семантической сегментации анатомических структур и расчета кардиоторакального индекса. Проведен сравнительный анализ 10 архитектур сверточных нейронных сетей на основе U-Net и SegNet с варьированием слоев повышения разрешения, типов нормализации и функций активации, а также с применением трансферного обучения с использованием предобученной модели ResNet18. Обучение и оценка моделей выполнены на открытом наборе данных miniJSRT (247 изображений). Наилучшие результаты сегментации достигнуты с помощью модели ResUNet (средний IoU = 0.9382). Для автоматического расчета кардиоторакального индекса по полученным маскам разработан алгоритм, определяющий срединную линию и поперечные размеры срединной тени и грудной клетки. Валидация системы на независимом наборе из 31 рентгенограммы высокого разрешения показала, что наибольшую точность относительно экспертных измерений обеспечивает модель UNetModInstanceNormLeakyReLU (средняя абсолютная ошибка MAE = 0.0169, относительная ошибка 4.0 %, коэффициент корреляции Пирсона = 0.898). Полученные данные свидетельствуют о перспективности использования нейросетевой сегментации для автоматизации расчета кардиоторакального индекса и служат основой для дальнейшей адаптации метода к задачам клинической диагностики.

Ключевые слова: сверточные нейронные сети, семантическая сегментация, рентгенография грудной клетки, кардиоторакальный индекс, медицинская диагностика, глубокое обучение, U-Net, SegNet, трансферное обучение, ResNet18

Для цитирования: Нурахунов П. М., Мелдо А. А., Прохоров И. Ю. Разработка интеллектуальной системы для анализа рентгенограмм грудной клетки на основе сверточных нейронных сетей: сегментация анатомических структур и автоматический расчет кардиоторакального индекса // Изв. СПбГЭТУ «ЛЭТИ». 2026. Т. 19, № 7. С. 55–71. doi: 10.32603/2071-8985-2026-19-7-55-71.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Original article

Development of an Intelligent System for Chest X-ray Analysis Based on Convolutional Neural Networks: Segmentation of Anatomical Structures and Automatic Calculation of the Cardiothoracic Index

P. M. Nurakhunov^{1✉}, A. A. Meldo^{1,2}, I. Yu. Prokhorov²

Saint Petersburg Electrotechnical University, Saint Petersburg, Russia

✉ eaburkov@gmail.com

Abstract. The article presents the results of developing an intelligent system for automated analysis of chest radiographs, including the stages of semantic segmentation of anatomical structures and calculation of the cardiothoracic index. A comparative analysis of 10 convolutional neural network architectures based on U-Net

and SegNet was performed with varying resolution enhancement layers, normalization types, and activation functions, as well as with the application of transfer learning using the pretrained ResNet18 model. Training and evaluation of the models were performed on the open miniJSRT dataset (247 images). The best segmentation results were achieved using the ResUNet model (average IoU = 0.9382). To automatically calculate the cardiothoracic index from the obtained masks, an algorithm was developed that determines the midline and transverse dimensions of the median shadow and the chest. Validation of the system on an independent dataset of 31 high-resolution radiographs showed that the UNetModInstanceNormLeakyReLU model provides the highest accuracy relative to expert measurements (mean absolute error MAE = 0.0169, relative error 4.0 %, Pearson correlation coefficient = 0.898). The obtained data demonstrate the potential of using neural network segmentation to automate cardiothoracic index calculations and serve as the basis for further adaptation of the method to clinical diagnostic tasks.

Keywords: convolutional neural networks, semantic segmentation, chest X-ray, cardiothoracic index, medical diagnostics, deep learning, U-Net, SegNet, transfer learning, ResNet18

For citation: Nurakhunov P. M., Meldo A. A., Prokhorov I. Yu. Development of an Intelligent System for Chest X-ray Analysis Based on Convolutional Neural Networks: Segmentation of Anatomical Structures and Automatic Calculation of the Cardiothoracic Index // LETI Transactions on Electrical Engineering & Computer Science. 2026. Vol. 19, no. 7. P. 55–71. doi: 10.32603/2071-8985-2026-19-7-55-71.

Conflict of interests. The authors declare no conflicts of interests.

Введение. Рентгенография органов грудной клетки (ОГК) – один из наиболее распространенных и доступных инструментов первичной диагностики патологий легких и сердечно-сосудистой системы [1]. Это исследование позволяет визуализировать анатомические структуры срединной тени, оценить размеры и контуры сердца, а также состояние легочного рисунка, что имеет ключевое значение для диагностики широкого спектра заболеваний [2]. Один из ключевых количественных показателей, используемых в клинической практике для оценки размеров сердца, – кардиоторакальный индекс (КТИ), отношение наибольшего поперечного размера срединной тени к наибольшему поперечному размеру грудной полости. КТИ служит важным диагностическим критерием для выявления кардиомегалии – патологического увеличения сердца, сопутствующего ряду серьезных заболеваний [3].

Однако традиционный анализ рентгенограмм, включающий визуальную оценку контуров и ручной расчет КТИ, представляет собой рутинную задачу для врача, сопряженную с субъективностью измерений (как внутриоператорской, так и межоператорской вариабельностью) и значительными временными затратами [4]. Это обуславливает актуальность разработки автоматизированных систем, призванных повысить скорость и точность диагностики, а также обеспечить воспроизводимость результатов [5].

Современные методы глубокого обучения, в частности сверточные нейронные сети (СНС), демонстрируют высокую эффективность в задачах

анализа медицинских изображений, включая семантическую сегментацию [6], [7]. Качественная сегментация областей срединной тени и легких создает основу для последующей автоматизации расчета диагностических параметров, таких как КТИ. Рассмотрение существующих работ показывает их ограниченность применением единичных архитектур на датасетах фиксированного разрешения, без анализа влияния гиперпараметров сетей и без количественной оценки погрешностей, возникающих при переходе к реальным клиническим изображениям высокого разрешения [3], [8].

Настоящая статья направлена на восполнение указанных пробелов. Научная новизна исследования заключается в проведении многофакторного сравнительного анализа 10 архитектур СНС (на основе U-Net и SegNet) с варьированием слоев повышения разрешения, типов нормализации и функций активации; оценена погрешность, вносимая изменением разрешения при переносе моделей, обученных на изображениях низкого разрешения (256×256), на реальные клинические данные высокого разрешения; разработан алгоритм расчета КТИ, дополненный визуализацией трех анатомических ориентиров для контроля ротации пациента.

Цель исследования. Разработка интеллектуальной системы для автоматизированного анализа рентгенограмм ОГК, включающей два этапа:

1) семантическую сегментацию областей легких и срединной тени с использованием различных архитектур СНС;

2) автоматический расчет кардиоторакального индекса на основе полученных сегментационных масок.

Методы и материалы. Методику работы автоматизированной системы можно описать следующим образом:

1. Предобработка входного рентгеновского изображения.
2. Подача изображения на вход модели СНС.
3. Сегментация моделью СНС 4 классов областей, получение предсказанной маски.
4. Применение к предсказанной маске алгоритма расчета КТИ.
5. Визуализация сегментации и расчета КТИ.

Наборы данных. Для реализации моделей СНС был выбран набор данных для сегментации из открытой базы данных miniJSRT_database, содержащий 247 рентгеновских изображений с масками разрешением 256×256 для следующих классов: фон, тело (область тела пациента на снимке), легкие, срединная тень. Используемый набор данных основан на наборе JSRT Database – The Japanese Society of Radiological Technology (JSRT) Database, и предназначен для исследования моделей машинного обучения для сегментации органов, тестирования и сравнения алгоритмов. JSRT Database – база цифровых изображений рентгенограмм органов грудной клетки с узловыми образованиями в легких и без них, разработанная в рамках исследовательской деятельности Академического комитета Японского общества радиологических технологий [9]. Рентгеновские снимки грудной клетки представлены в формате BMP с одним цветовым каналом. Изображения масок, имеющие значение пикселей области легких – 255, области сердца – 85, области тела –

170, области фона – 0, записаны в формате PNG. Примеры рентгеновских снимков и соответствующих им масок представлены на рис. 1, для визуализации масок выбрана цветовая палитра tab10 (одна из стандартных палитр библиотеки Matplotlib), содержащая контрастные и легко различимые цвета.

Набор данных разделен на 199 примеров в обучающем наборе и 48 – в тестовом. В качестве предобработки данных обучающего набора для рентгеновских изображений применяется изменение яркости и контрастности на 20 % и нормализация на основе предварительно рассчитанных на обучающем наборе данных статистик среднего значения и среднеквадратичного отклонения (СКО), для рентгеновских изображений и соответствующих им масок применяется случайное горизонтальное отражение. Для тестового набора применяются аналогичные операции, за исключением случайного изменения яркости и контрастности рентген-изображений. Для повышения эффективности обработки изображений применяется пакетная обработка (mini-batching) с размером пакета (batch size), равным 8.

Для повышения надежности оценки моделей и алгоритма расчета КТИ был сформирован дополнительный независимый набор данных, включающий 31 пару обезличенных рентгенограмм высокого разрешения (порядка 3000×3000 пикселей) в прямой и левой боковой проекциях в формате DICOM. Сегментационные маски областей легких и срединной тени для 31 рентгенограммы в прямой проекции формировались при

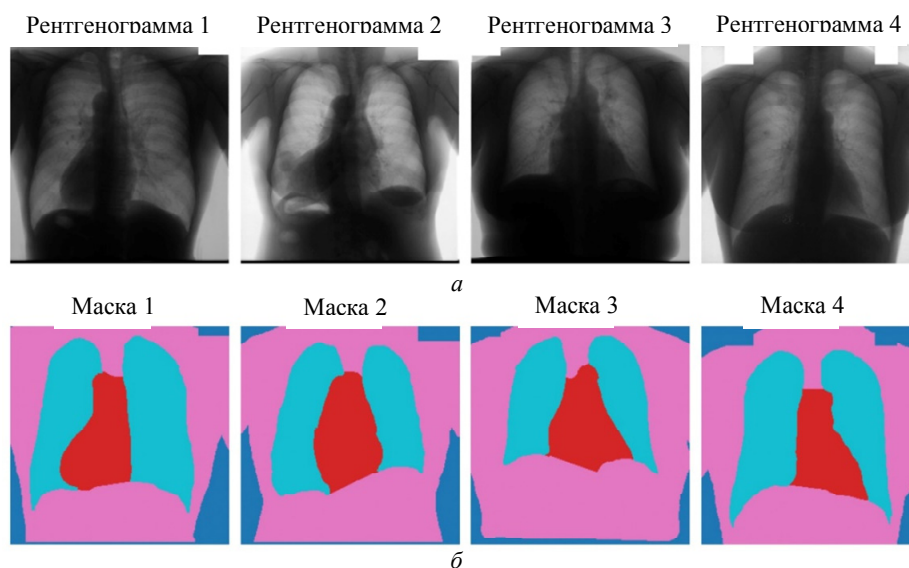


Рис. 1. Примеры рентгенограмм (а) и соответствующих им масок (б) из набора данных miniJSRT
Fig. 1. Examples of radiographs (a) and their corresponding masks (b) from the miniJSRT dataset

участии практикующего рентгенолога с опытом работы более 5 лет. Итоговые маски содержат три класса: фон, срединная тень, легкие.

Программно-аппаратное окружение. Все вычисления выполнялись на персональном компьютере с процессором AMD Ryzen 5 5600 и видеокартой NVIDIA RTX 4060 (8 ГБ) под управлением ОС Windows 10.

Наиболее мощным аппаратом для построения и настройки сверточных нейронных сетей обладает язык программирования Python с использованием фреймворков Tensorflow и PyTorch [6]. Это объясняется наличием предобученных моделей, набора инструментов для аугментации данных и развертывания моделей, готовых датасетов и слоев для СНС, поддержкой автоматического расчета градиентов для обновления весов, работы с использованием GPU/TPU для ускорения вычислений. Разница заключается в том, что Pytorch имеет более гибкий код, подходящий под исследовательские проекты, проще в отладке, в то время как TensorFlow более адаптирован для интегрирования модели в реальные системы. Программная реализация в данной работе выполнена на Python 3.13 в среде Visual Studio Code с использованием фреймворка PyTorch 2.10.0 для создания и обучения моделей, библиотек OpenCV и PIL для обработки изображений, NumPy и Pandas для численных расчетов и анализа данных, Matplotlib и Seaborn для визуализации результатов.

Для сегментации дополнительного набора данных использовалась программа 3D Sliser. Измерения, необходимые для расчета эталонного КТИ, производились с использованием IMV DICOM Viewer.

Методы сегментации. Под сегментацией изображений понимается процесс выделения на изображении значимых областей, соответствующих искомым объектам. В настоящей работе решается задача выделения срединной тени и легких на рентгенограммах с использованием подхода семантической сегментации, при котором каждому пикселю изображения присваивается метка соответствующего класса.

Сверточная нейронная сеть, решающая задачу сегментации, строится из последовательности сверточных и пулинговых (подвыборочных) слоев. Базовым элементом такой сети является сверточный слой, выполняющий операцию свертки. Эта операция заключается в скольжении окна (фильтра/ядра) по входному изображению с заданным шагом, при этом для каждой позиции

окна производится поэлементное умножение весов фильтра на значения соответствующего фрагмента изображения с последующим суммированием результатов [6]. Выходом сверточного слоя служит карта признаков, отражающая степень активации фильтра в различных областях входных данных.

Следующий ключевой компонент – пулинговый слой (слой подвыборки), который уменьшает пространственную размерность карт признаков, агрегируя информацию в локальных областях (например, взятие максимума или среднего). Это позволяет снизить вычислительную сложность и обеспечить инвариантность к небольшим пространственным искажениям.

Между сверточным и пулинговым слоями, как правило, размещаются слои нормализации (например, Batch Normalization) и активации. Нормализация ускоряет сходимость модели, стабилизирует распределение активаций и служит регуляризатором, снижая риск переобучения. Нелинейная функция активации (такая, как ReLU) вносит в модель нелинейность, что позволяет ей обучаться сложным иерархическим закономерностям.

В данной статье разработаны и протестированы модели на основе архитектур U-Net и SegNet с вариациями слоев активации, нормализации и повышения размерности. Выбор базовых архитектур связан с тем, что U-Net стала стандартным инструментом обработки биомедицинских изображений [10], а SegNet, в свою очередь, активно применяется и развивается в задачах медицинской сегментации [11], [12].

Ключевая особенность архитектуры U-Net заключается в ее симметричной структуре, состоящей из энкодера (путь сжатия) и декодера (путь восстановления). На этапе «кодирования» происходит извлечение иерархических признаков из входного изображения посредством последовательного снижения пространственного разрешения карт признаков за счет операций свертки и пулинга. Декодер, в свою очередь, восстанавливает исходное разрешение изображения, используя повышающие операции, такие как транспонированные свертки. Для сохранения пространственных деталей в архитектуре используются пропускные соединения (skip connections), передающие карты признаков от слоев энкодера к соответствующим слоям декодера. Это позволяет компенсировать потерю информации при субдискретизации и улучшает качество сегментации, особенно для мелких объектов и границ. Общий

вид архитектуры U-Net, представленный на рис. 2 в виде последовательности карт признаков и выполняемых над ними операций, иллюстрирует описанные принципы построения сети. Число каналов каждой карты признаков соответствует архитектурам разработанных моделей.

SegNet – архитектура сверточной нейронной сети, разработанная для эффективной сегментации изображений. Главная особенность SegNet – это механизм декодирования, при котором восстановление пространственного разрешения осуществляется с использованием индексов пулинга (pooling indices), запомненных на соответствующих слоях энкодера. В отличие от U-Net, где пропускные соединения передают полные

карты признаков от энкодера к декодеру, SegNet передает только индексы позиций максимальных элементов, полученных при операции пулинга. Это позволяет снизить объем передаваемой информации и, как следствие, уменьшить вычислительную сложность модели, сохраняя при этом высокую точность локализации границ объектов. Общий вид архитектуры SegNet представлен на рис. 3 в виде последовательности слоев и операций. Количество слоев соответствует архитектурам разработанных моделей.

Также в работе применяется трансферное обучение (transfer learning), заключающееся в использовании предварительно обученной на наборе данных ImageNet модели ResNet18 с последу-

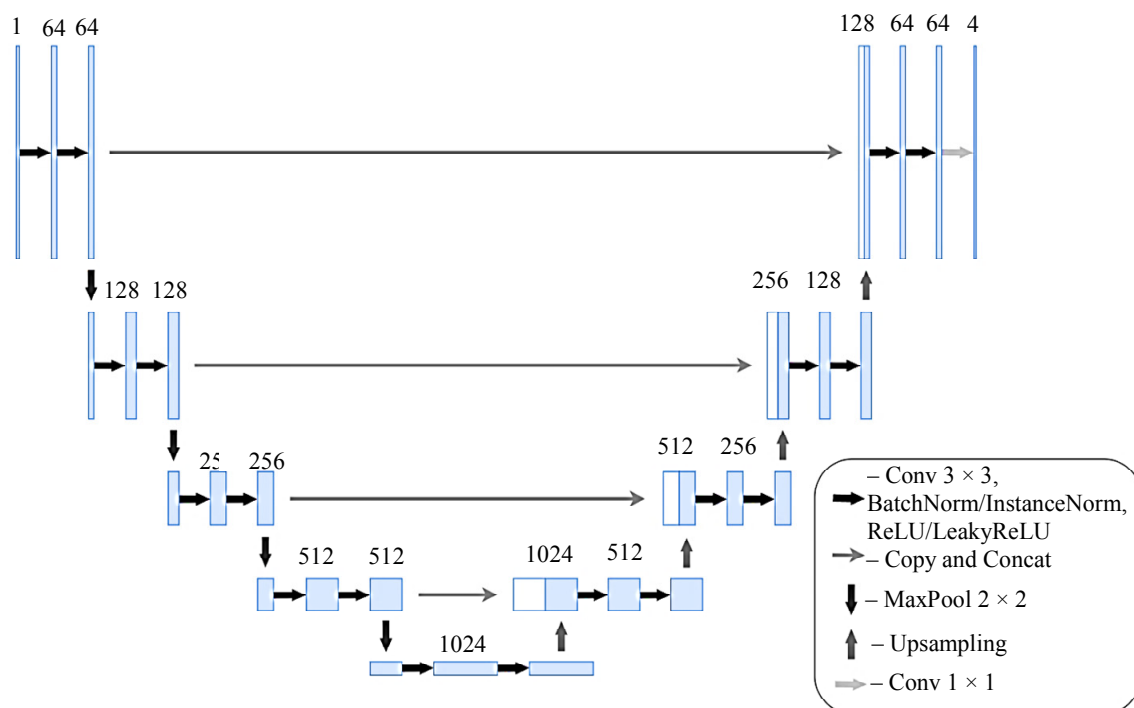


Рис. 2. Общий вид архитектуры U-Net
 Fig. 2. General view of the U-Net architecture

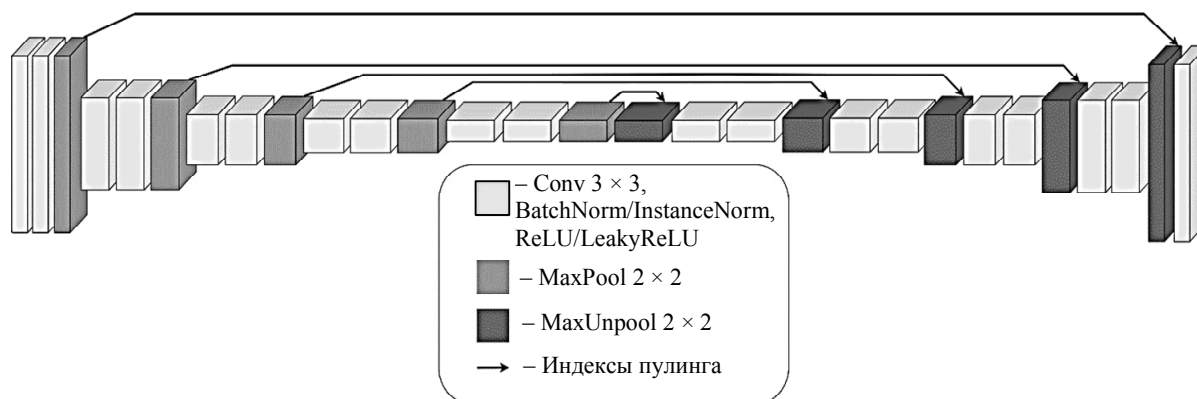


Рис. 3. Общий вид архитектуры SegNet
 Fig. 3. General view of the SegNet architecture

ющим дообучением (fine-tuning) на применяемом в работе наборе данных. Использование предобученной модели позволяет задействовать уже сформированные на масштабном наборе данных признаки низкого и среднего уровня – контуры, текстурные и структурные особенности и т. п., что важно при ограниченном объеме медицинской обучающей выборки [7]. Дальнейшее дообучение модели на целевом наборе данных обеспечивает адаптацию весов сети к специфике рентгенографических изображений органов грудной клетки.

В основе архитектуры ResNet18 лежат четыре остаточных блока, каждый из которых содержит по два базовых блока BasicBlock. При этом в зависимости от необходимости уменьшения пространственной размерности входной карты признаков базовый блок может иметь два варианта строения. Детальная структура базового блока, а также полная схема архитектуры ResNet18 представлены на рис. 4.

Архитектуры моделей CHC. В ходе работы было разработано и протестировано 10 архитектур моделей CHC.

Модели архитектуры U-Net.

1. Модель UNet. Энкодер модели состоит из 4 последовательных блоков. Каждый блок включает два сверточных слоя (Conv) с ядром 3×3 , за каждым из которых следуют пакетная нормализация (BatchNorm) и функция активации ReLU. Схематично последовательность слоев в блоке можно представить в следующем виде:

Conv $\rightarrow$ BatchNorm $\rightarrow$ ReLU $\rightarrow$ Conv $\rightarrow$ BatchNorm $\rightarrow$ ReLU.

Между блоками используется слой подвыборки MaxPool для уменьшения размерности карт признаков в два раза. Число каналов после каждого блока энкодера увеличивается по схеме:

$1 \rightarrow 64 \rightarrow 128 \rightarrow 256 \rightarrow 512$.

После добавляется центральный блок (bottleneck) с аналогичным строением, итоговое число каналов становится равным 1024.

В декодере симметрично на каждом из 4 уровней применяется следующая последовательность операций:

Повышение пространственного разрешения признаков карты (Upsampling, рис. 2) в 2 раза с использованием слоя Upsample с билинейной интерполяцией $\rightarrow$ конкатенация с соответствующей картой признаков из энкодера, переданной через пропускное соединение (Copy and Concat, рис. 2) $\rightarrow$ Два сверточных слоя с BatchNorm и ReLU (подобно блоку энкодера).

Число каналов уменьшается по схеме:

$(1024 + 512) \rightarrow 512 \rightarrow (512 + 256) \rightarrow 256 \rightarrow (256 + 128) \rightarrow 128 \rightarrow (128 + 64) \rightarrow 64$.

На выходе сети используется сверточный слой 1×1 с 4 каналами по количеству классов.

Для сохранения пространственной совместности все сверточные слои с ядром 3×3 используют padding=1.

2. Модель UNetMod. Данная архитектура имеет ту же структуру, что и предыдущая, но в

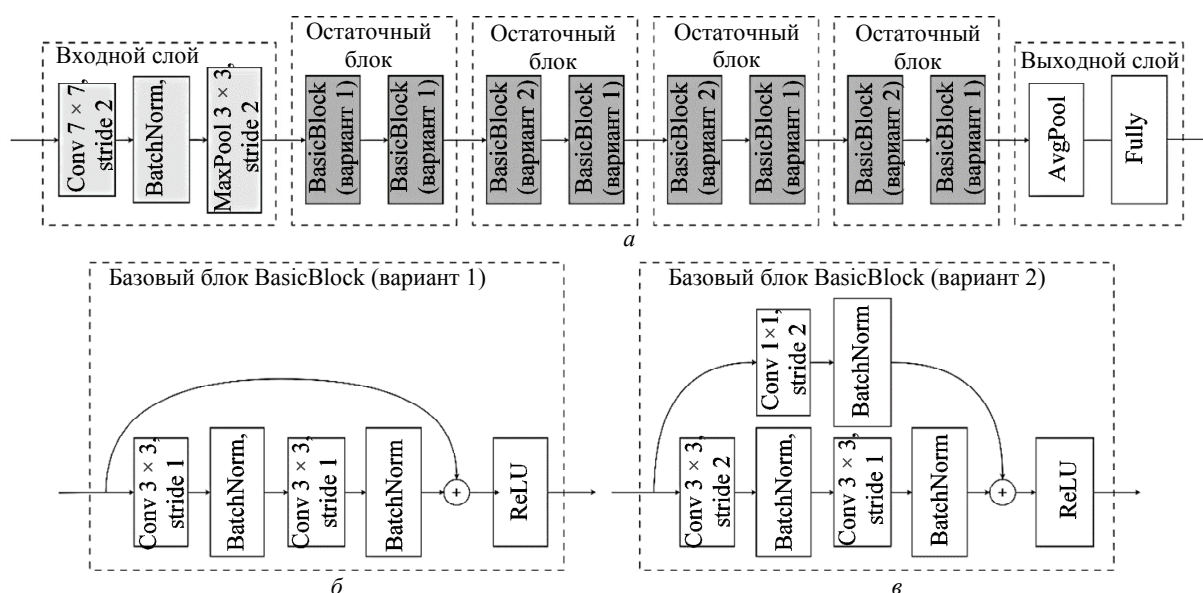


Рис. 4. Архитектура ResNet18: а – общая структура; б – строение базового блока для случая без изменения размерности

Fig. 4. ResNet18 architecture: а – general structure; б – structure of the base unit for the case without changing the dimension; в – structure of the base unit for the case with dimension reduction

качестве слоя повышения пространственной размерности (Upsampling) в декодере используется слой транспонированной свертки (ConvTranspose) с размером ядра 2×2 и шагом 2. Данный слой уменьшает количество каналов вдвое (например, с 1024 до 512), одновременно увеличивая пространственное разрешение карт признаков в 2 раза, после чего выполняется конкатенация с соответствующей картой из энкодера. Преимущество применения данного слоя относительно слоя Upsample с билинейной интерполяцией состоит в наличии обучаемых параметров, адаптирующихся к данным, однако это требует больше вычислительных затрат.

3. Модели UNetLeakyReLU и UNetModLeakyReLU – модификации архитектур UNet и UNetMod соответственно, с применением функции активации Leaky ReLU в блоках энкодера и декодера вместо стандартной ReLU. В отличие от ReLU, обнуляющей отрицательные значения, Leaky ReLU пропускает их с малым коэффициентом (обычно 0.01), что позволяет избежать проблемы «умирающих нейронов»: даже при отрицательных входах градиент не равен нулю, поэтому веса продолжают обновляться.

4. Модели UNetInstanceNormLeakyReLU и UNetModInstanceNormLeakyReLU – модификации архитектур UNet и UNetMod соответственно, в которых слои пакетной нормализации BatchNorm заменены на нормализацию по экземплярам InstanceNorm. В отличие от BatchNorm, вычисляющего статистики распределения (среднее и дисперсию) по всей выборке в пределах батча, InstanceNorm выполняет нормализацию независимо для каждого канала каждого отдельного изображения. Такой подход позволяет устранить зависимость качества обучения от размера мини-пакета и снижает влияние контрастных различий между входными данными, что актуально для рентгенограмм.

5. Модель ResUNet. В архитектуре данной модели применяется трансферное обучение с использованием предобученной на наборе данных ImageNet модели ResNet18 в качестве энкодера. Для работы с одноканальными рентгенограммами входной сверточный слой ResNet18 заменен на слой с одним входным каналом. Его веса инициализированы посредством поэлементного суммирования весов оригинального слоя (рассчитанного на RGB-изображения) по трем каналам, что позволяет сохранить суммарный эффект от всех каналов. Далее используются четыре остаточных

блока предобученной сети, формирующих карты признаков с числом каналов, возрастающим по схеме: $1 \rightarrow 64 \rightarrow 128 \rightarrow 256 \rightarrow 512$.

После энкодера добавлен центральный блок из двух сверточных слоев с BatchNorm и ReLU, увеличивающий число каналов до 1024.

Строение декодера аналогично архитектурному решению в модели UNetMod. Число каналов в декодере изменяется следующим образом:

$(1024 + 512) \rightarrow 512 \rightarrow (512 + 256) \rightarrow 256 \rightarrow (256 + 128) \rightarrow 128 \rightarrow (128 + 64) \rightarrow 64$.

Поскольку после всех операций пространственное разрешение составляет половину от исходного, на выходе сети применяется билинейная интерполяция для восстановления полного размера изображения. Завершающий сверточный слой 1×1 формирует выходную карту с 4 каналами.

Модели архитектуры SegNet.

1. Модель SegNet. В данной архитектуре используется 5 блоков энкодера с последовательностью слоев, аналогичной модели UNet:

$\text{Conv} \rightarrow \text{BatchNorm} \rightarrow \text{ReLU} \rightarrow \text{Conv} \rightarrow \text{BatchNorm} \rightarrow \text{ReLU}$.

Между блоками используется слой подвыборки MaxPool с размером ядра 2 и шагом 2, который запоминает индексы максимальных элементов для последующего восстановления пространственной структуры на этапе декодирования. Число каналов между блоками увеличивается по схеме:

$1 \rightarrow 64 \rightarrow 128 \rightarrow 256 \rightarrow 512 \rightarrow 1024$.

После пятого блока также применяется слой MaxPool.

Декодер построен симметрично: на каждом уровне сначала выполняется операция обратного пулинга (слой MaxUnpool), восстанавливающая пространственное разрешение с использованием запомненных на этапе кодирования индексов, а затем последовательность операций слоев блока. Число каналов уменьшается по схеме: $1024 \rightarrow 512 \rightarrow 256 \rightarrow 128 \rightarrow 64$.

На выходе сети используется сверточный слой с ядром 3×3 и паддингом 1, формирующий карту признаков с 4 каналами.

2. Модель SegNetLeakyReLU. Модификация архитектуры SegNet с применением функции активации Leaky ReLU в блоках энкодера и декодера.

3. Модель SegNetInstanceNormLeakyReLU. Модификация архитектуры SegNet с комбинированием слоев InstanceNorm и функции активации Leaky ReLU в блоках энкодера и декодера.

Обучение моделей СНС. Обучение моделей проводилось с использованием оптимизатора AdamW и функции потерь перекрестной энтропии (Cross-Entropy Loss). Скорость обучения составляла 0.001, количество эпох – 30.

Метрики оценки качества сегментации. Оценка проводилась по метрикам Intersection over Union (IoU), Dice Coefficient (DC), Precision и Recall.

Алгоритм расчета КТИ. Общий алгоритм расчета кардиоторакального индекса можно описать в виде последовательности следующих этапов [1]:

1. Определение поперечного размера срединной тени.

1.1. Определение срединной линии. Относительно срединной линии определяются размеры правой и левой частей сердечной тени. Важное условие получения диагностически значимых рентгенограмм – правильная укладка пациента. Ротация туловища относительно вертикальной оси приводит к проекционным искажениям, которые могут повлиять на точность диагностики.

В связи с этим интерес представляет выявление ротаций на рентгенограмме и принятие мер по компенсации вносимых искажений. Особое значение это имеет при анализе серии снимков одного пациента, сделанных в разные временные

периоды, так как для корректного сравнения и выявления динамики лечения необходимо предварительное совмещение изображений [5], [13].

Для определения положения срединной линии применяется метод, основанный на анализе вертикальной проекции яркости инвертированного рентгеновского изображения [4], [14]. Для каждого столбца изображения вычисляется сумма значений яркости пикселей, в результате чего формируется профиль, отражающий распределение суммарной яркости L вдоль горизонтальной оси X . Участок профиля, соответствующий позвоночнику, выделяется как локальный экстремум (рис. 5).

Однако, как видно на рис. 5, $\bar{b}$, срединная линия (штриховая) не совпадает с точкой максимума графика. Это объясняется тем, что при расчете используется взвешенное среднее значений яркости в окрестности пика. Данное решение обусловлено особенностями рентгенологической картины: позвоночный столб имеет меньшую яркость по сравнению с прилегающими вертикальными областями вследствие затемнений в зоне межпозвонковых дисков.

Дополнительно для контроля ротации выполняется построение срединных линий по симметрии легких, а также на основе взвешенного среднего вертикальных проекций области срединной тени.

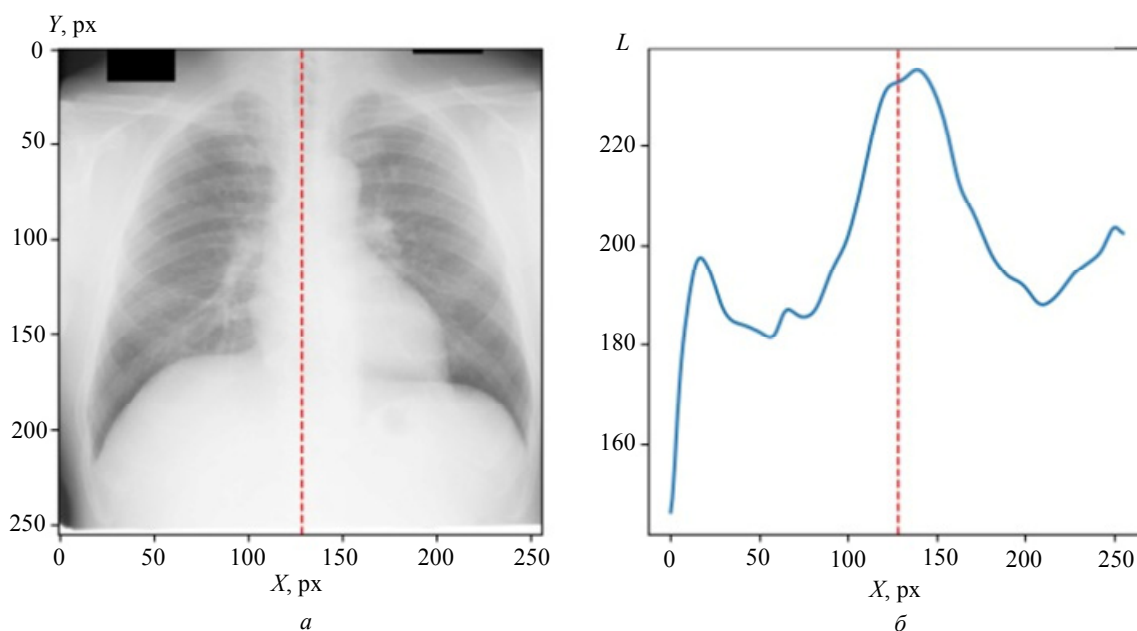


Рис. 5. Выделение срединной линии методом вертикальной проекции яркости: a – инвертированная рентгенограмма; $\bar{b}$ – вертикальная проекция распределения яркости. По осям X и Y отложены номера пикселей изображения по горизонтали и вертикали, по оси L – значения суммы яркостей пикселей каждого столбца изображения в диапазоне от 0 до 255

Fig. 5. The selection of the median line by the vertical brightness projection method: a – inverted radiograph; $\bar{b}$ – vertical projection of the brightness distribution. The X and Y axes display the pixel numbers of the image horizontally and vertically, and the L axis shows the sum of the pixel brightness values of each column of the image in the range from 0 to 255

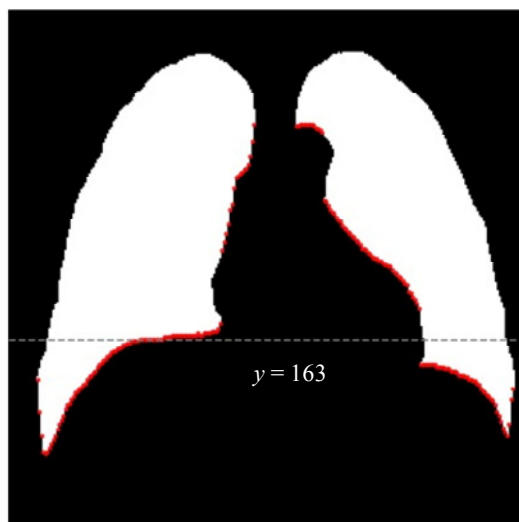
Взаимное расположение этих линий позволяет судить о наличии поворота пациента относительно вертикальной оси и может быть использовано для последующей коррекции изображения. Следует отметить, что смещение центров симметрии легких и срединной тени не всегда однозначно свидетельствует о ротации – оно может быть обусловлено патологическими изменениями, что необходимо учитывать при интерпретации результатов.

1.2. Определение расстояний от срединной линии до контуров сердца. Для найденной срединной линии вычисляются максимальные расстояния до правого и левого контуров сердечной тени.

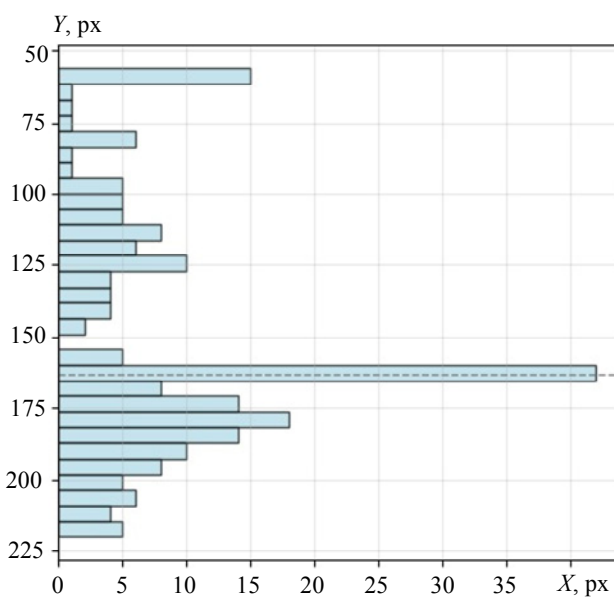
1.3. Вычисление поперечного размера сердца. Поперечный размер сердца определяется как сумма найденных расстояний.

2. Определение поперечного размера грудной клетки.

2.1. Выбор уровня измерения. Согласно стандартной методике, поперечный размер грудной клетки измеряется между внутренними краями ребер на уровне правого купола диафрагмы, что обычно соответствует наиболее широкой части грудной клетки.



а



б

Рис. 6. Определение уровня измерения поперечного размера грудной клетки: а – маска легких; б – распределение нижних границ по столбцам, штриховая линия – медианное значение, выбранное в качестве уровня измерения. На фрагменте а $y = 163$ – найденный уровень измерения ширины грудной клетки. На фрагменте б по оси Y отложены номера пикселей по вертикали изображения, по оси X – количество нижних точек в строке изображения

Fig. 6. Determination of the level of measurement of the transverse size of the chest: а – lung mask; б – distribution of lower boundaries in columns, gray line – the median value selected as the measurement level. In fragment а $y = 163$ is the found level of chest width measurement. In fragment б the Y axis shows the number of pixels along the vertical of the image, and the X axis shows the number of lower points in the image row

В данной работе для автоматизации этого этапа используется анализ маски легких. Для каждого столбца маски находится самая нижняя точка, принадлежащая легким, а уровень измерения ширины грудной клетки выбирается как медианное значение координаты Y этих границ (рис. 6).

3. Расчет КТИ.

Итоговое значение КТИ вычисляется по формуле:

$$\text{КТИ} = \frac{d_{\text{пр}} + d_{\text{лев}}}{D_{\text{ГК}}},$$

где $d_{\text{пр}}$ – наибольший поперечный размер правой части сердечной тени (правее срединной линии); $d_{\text{лев}}$ – наибольший поперечный размер левой части сердечной тени (левее срединной линии); $D_{\text{ГК}}$ – наибольший поперечный размер грудной клетки.

КТИ не зависит от единиц измерения размеров, которые в данном случае представляют собой пиксели изображения. При этом, при известном значении шага изображения, может быть осуществлен перевод в метрическую систему для дальнейшей оценки размеров.

Метрики оценки качества расчета КТИ.

Для количественной оценки точности разрабо-

танного алгоритма расчета КТИ использовались средняя абсолютная ошибка (MAE) и средне-квадратичная ошибка (RMSE), позволяющие оценить отклонение автоматического расчета от эталонных значений. Дополнительно для оценки согласованности автоматического и эталонного методов измерений применялся коэффициент корреляции Пирсона. Для интерпретации результатов с клинической точки зрения рассчитывалась относительная ошибка как отношение MAE к среднему значению эталонного КТИ, выраженное в процентах.

Обсуждение результатов. Сравнительный анализ моделей сегментации. В табл. 1 представлены значения метрик качества и функции потерь на тестовой выборке, а также временные затраты на обучение для каждой из рассмотренных моделей. В табл. 2 представлены значения метрики IoU по классам на тестовой выборке для каждой модели. Серым цветом в таблицах выделены три лучших значения по каждому столбцу, максимальное значение выделено жирным шрифтом.

Наилучшие результаты по большинству метрик (табл. 1) достигнуты моделью ResUNet, ис-

пользующей предобученный энкодер ResNet18. Высокое качество сегментации и минимальное время обучения обусловлены использованием предобученных слоев с оптимизированными весами, что позволяет модели быстрее адаптироваться к новым данным.

Применение транспонированных сверток вместо билинейного апсемплинга в модели UNetMod не дало ожидаемого повышения точности, однако позволило сократить время обучения. Отсутствие выигрыша в качестве может быть связано с большим числом обучаемых параметров при ограниченном наборе данных – дополнительные параметры не имеют достаточного количества примеров для эффективного обучения.

Влияние замены функции активации ReLU на LeakyReLU оказалось неоднозначным и зависящим от типа декодера. Для базовой архитектуры UNet с билинейным апсемплингом замена активации не привела к значимому изменению качества (средний IoU 0.9310 у UNet против 0.9317 у UNetLeakyReLU). Однако в модификации UNetMod, использующей транспонированные свертки, применение LeakyReLU вызвало существенное снижение метрик (с 0.9332 до 0.9210). Это позво-

Табл. 1. Сравнительная таблица по результатам обучения моделей CNN
Tab. 1. Comparative table based on the results of training CNN models

Модель	Метрика				Значение функции потерь	Время обучения, с
	IoU	DC	Precision	Recall		
UNet	0.9310	0.9640	0.9657	0.9628	0.0744	465.7
UNetMod	0.9332	0.9652	0.9712	0.9598	0.0700	421.1
UNetLeakyReLU	0.9317	0.9644	0.9650	0.9641	0.0724	473.2
UNetModLeakyReLU	0.9210	0.9584	0.9623	0.9552	0.0782	430.5
UNetInstanceNormLeakyReLU	0.9380	0.9678	0.9715	0.9644	0.0661	506.8
UNetModInstanceNormLeakyReLU	0.9351	0.9662	0.9694	0.9634	0.0745	469.1
ResUNet	0.9382	0.9679	0.9682	0.9678	0.0732	141.3
SegNet	0.9053	0.9497	0.9552	0.9449	0.0969	264.6
SegNetLeakyReLU	0.9124	0.9537	0.9562	0.9516	0.0906	268.8
SegNetInstanceNormLeakyReLU	0.9153	0.9553	0.9596	0.9515	0.0925	271.4

Табл. 2. Сравнительная таблица метрики IoU моделей на тестовом наборе по классам
Tab. 2. Comparison table of IoU metric for models on the test set by class

Модель	Значение метрики IoU по классам			
	Фон	Срединная тень	Тело	Легкие
UNet	0.9118	0.9030	0.9596	0.9490
UNetMod	0.9238	0.9003	0.9619	0.9487
UNetLeakyReLU	0.9135	0.9028	0.9595	0.9519
UNetModLeakyReLU	0.8813	0.8987	0.9540	0.9535
UNetInstanceNormLeakyReLU	0.9283	0.9073	0.9629	0.9564
UNetModInstanceNormLeakyReLU	0.9273	0.9047	0.9595	0.9499
ResUNet	0.9235	0.9120	0.9623	0.9532
SegNet	0.8627	0.8755	0.9457	0.9400
SegNetLeakyReLU	0.8696	0.8863	0.9503	0.9469
SegNetInstanceNormLeakyReLU	0.8838	0.8859	0.9515	0.9428

ляет предположить, что комбинация обучаемых повышающих слоев с LeakyReLU требует более тщательной настройки гиперпараметров или большего объема данных для стабильного обучения. При этом комбинация InstanceNorm и LeakyReLU, напротив, продемонстрировала высокую эффективность для обоих семейств архитектур: UNetInstanceNormLeakyReLU достигла сопоставимое с ResUNet среднее IoU (0.9380), а SegNetInstanceNormLeakyReLU показала наибольшее качество среди моделей SegNet.

Модели семейства SegNet показали наиболее низкие результаты среди всех архитектур. Несмотря на более высокую скорость обучения по сравнению с базовыми UNet (264...271 с против 421...506), модели SegNet значительно уступают по качеству. Таким образом, для данной задачи применение архитектуры SegNet представляется нецелесообразным.

Таким образом, можно выделить три лучших модели: UNetInstanceNormLeakyReLU, UNetMod InstanceNormLeakyReLU и ResUNet. Данные модели использованы для валидации интеллектуальной системы.

Валидация интеллектуальной системы.

Для оценки функционирования разработанной интеллектуальной системы автоматического рас-

чета КТИ была проведена валидация на сформированном независимом наборе данных высокого разрешения. Поскольку обучение моделей выполнялось на изображениях размером 256×256 пикселей, исходные изображения и маски потребовали предварительной обработки, связанной с изменением разрешения для подачи на вход моделей. При этом для практического использования разрабатываемой системы необходима визуализация рассчитанных контуров и линий (рис. 7), что требует последующего увеличения предсказанных масок до исходного разрешения. Такая последовательность процедур уменьшения и последующего восстановления изображений неизбежно вносит определенные искажения, оценке которых посвящен отдельный этап валидации.

Оценка погрешности масштабирования.

Для контроля погрешностей, возникающих при изменении разрешения, истинные маски высокого разрешения последовательно уменьшались до размера 256×256 пикселей, после чего вновь увеличивались до исходного размера. Полученные таким образом маски сравнивались с оригинальными эталонными масками с использованием метрик IoU, Dice Coefficient, Precision и Recall, рассчитанных как в среднем по классам, так и для каждого класса отдельно.

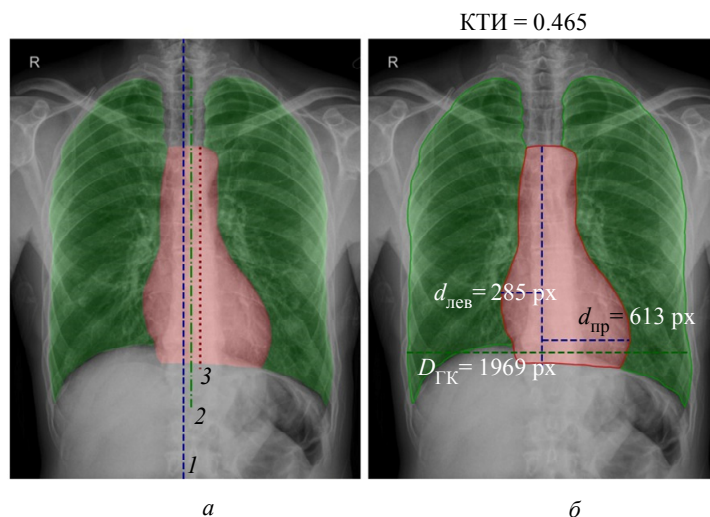


Рис. 7. Пример работы алгоритма расчета КТИ на истинных масках независимого набора данных: а – визуализация рассчитанных срединных линий (1 – на основе вертикального профиля яркости рентген-изображения, 2 – на основе симметрии легких, 3 – на основе взвешенного среднего вертикальных проекций сердечной тени); б – визуализация основных измерений для расчета КТИ в пикселях

Fig. 7. An example of an algorithm for calculating CTI on true masks of an independent data set: а – visualization of the calculated median lines (1 – based on the vertical brightness profile of the X-ray image, 2 – based on the symmetry of the lungs, 3 – based on the weighted average of vertical projections of the cardiac shadow); б – visualization of the main measurements for calculating CTI in pixels

Представленные результаты (табл. 3) демонстрируют, что полный цикл изменения разрешения вносит минимальную погрешность в структуру масок. Для фоновго класса, занимающего большую площадь изображения, значения метрик составляют $IoU = 0.9942$ и $Dice = 0.9971$, что соответствует погрешности менее 1 %. Для анатомических структур, имеющих более сложную конфигурацию и меньший размер (легкие и срединная тень), значения метрик ожидаемо несколько ниже: $IoU = 0.986 - 0.987$, $Dice = 0.993$. Полученные результаты подтверждают, что искажения, вносимые операциями масштабирования, являются пренебрежимо малыми и не оказывают существенного влияния на точность масок.

Табл. 3. Метрики качества масок после полного цикла масштабирования

Tab. 3. Metrics of mask quality after a full scaling cycle

Класс	IoU	DC	Precision	Recall
Фон	0.9942	0.9971	0.9971	0.9971
Средостение	0.9863	0.9931	0.9931	0.9931
Легкие	0.9874	0.9936	0.9937	0.9936
Среднее по классам	0.9893	0.9946	0.9946	0.9946

Верификация алгоритма расчета КТИ. Для оценки соответствия автоматически рассчитанных значений КТИ ручным измерениям эксперта, алгоритм был последовательно применен к трем вариантам истинных масок: исходным маскам высокого разрешения (ИМВР), уменьшенным до 256×256 пикселей (УИМ), и прошедшим цикл уменьшения до 256×256 и последующего восстановления (ВИМ).

Результаты, представленные в табл. 4, демонстрируют высокую корреляцию (около 0.91) и низкую относительную ошибку (5.1 % и ниже) для всех вариантов автоматического расчета, что подтверждает корректность математической реализации алгоритма и его устойчивость к измене-

нию разрешения. Наилучшие показатели достигнуты при расчете на уменьшенных масках (УИМ): $MAE = 0.0195$, $RMSE = 0.0284$, относительная ошибка 4.6 %. Интересно, что расчет на исходных масках высокого разрешения (ИМВР), несмотря на отсутствие каких-либо трансформаций, демонстрирует несколько более высокую ошибку (5.1 %). Это может быть связано с тем, что потеря детализации при уменьшении изображения приводит к сглаживанию мелких ступенчатых неровностей на границах масок, что приводит к лучшему соответствию ручным измерениям эксперта.

На скаттерограмме (рис. 8) представлено распределение значений КТИ, рассчитанных алгоритмом на различных версиях масок, относительно ручных измерений эксперта. Можно наблюдать систематическое смещение облака точек выше линии идеального соответствия ($y = x$), что указывает на тенденцию алгоритма к завышению значений КТИ по сравнению с ручными измерениями эксперта. Отклонения работы алгоритма от ручной разметки могут быть связаны с погрешностью определения эталонных сегментационных масок; субъективностью выделения границ сердца, нечеткость которых более выражена при увеличенных размерах; чувствительностью алгоритма к выбору уровня измерения ширины грудной клетки.

В качестве эталонных при оценке работы интеллектуальной системы далее используются значения КТИ, полученные на исходных масках высокого разрешения.

Оценка работы системы на предсказанных масках. Для трех наиболее зарекомендовавших себя моделей были получены предсказания масок с последующим восстановлением до исходного разрешения и расчетом КТИ на их основе. Полученные значения КТИ сравнивались с двумя эта-

Табл. 4. Сопоставление автоматического расчета КТИ на различных версиях масок с ручными измерениями эксперта

Tab. 4. Comparison of automatic calculation of CTI on different versions of masks with manual measurements of an expert

Способ измерения	Среднее значение КТИ	СКО	Медиана	Диапазон значений	MAE	RMSE	Корреляция Пирсона	Относительная ошибка, %
Вручную экспертом	0.426	0.047	0.430	[0.347; 0.516]	—	—	—	—
Алгоритмом на ИМВР	0.448	0.051	0.450	[0.355; 0.540]	0.0216	0.0299	0.912	5.1
Алгоритмом на ВИМ	0.448	0.051	0.450	[0.355; 0.545]	0.0219	0.0301	0.913	5.1
Алгоритмом на УИМ	0.446	0.051	0.448	[0.352; 0.543]	0.0195	0.0284	0.913	4.6

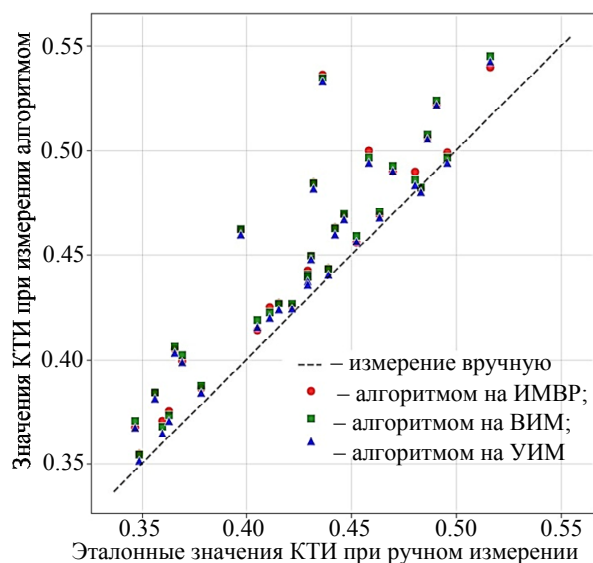


Рис. 8. Распределение значений КТИ, полученных алгоритмом на различных версиях масок, в сравнении с экспертными измерениями

Fig. 8. Distribution of CTI values obtained by the algorithm on different versions of masks, in comparison with expert measurements

Табл. 5. Сравнение статистических распределений КТИ, полученных вручную и алгоритмически на истинных масках и на предсказанных масках моделей
 Tab. 5. Comparison of statistical distributions of CTI obtained manually and algorithmically on true masks and on preset masks of models

Способ измерения	Среднее значение КТИ	СКО	Медиана	Диапазон значений
Вручную экспертом	0.426	0.047	0.430	[0.347; 0.516]
Алгоритмом на ИМБР	0.448	0.051	0.450	[0.355; 0.540]
Алгоритмом на масках модели UNetInstanceNormLeakyReLU	0.435	0.062	0.430	[0.347; 0.642]
Алгоритмом на масках модели UNetModInstanceNormLeakyReLU	0.435	0.057	0.429	[0.358; 0.593]
Алгоритмом на масках модели ResUNet	0.405	0.040	0.408	[0.333; 0.486]

Табл. 6. Сравнение результатов расчета КТИ на предсказанных масках трех моделей ручными измерениями эксперта и с расчетом алгоритма на ИМБР

Tab. 6. Comparison of the results of the CTI calculation on the predicted masks of the three models by manual measurements of the expert and with the calculation of the algorithm on the true masks of high resolution

Способ измерения	Источник эталонных значений КТИ	MAE	RMSE	Корреляция Пирсона	Относительная ошибка, %
Алгоритмом на масках модели UNetInstanceNormLeakyReLU	Измерение вручную	0.0220	0.0368	0.819	5.2
	Алгоритмически на ИМБР	0.0235	0.0313	0.891	5.2
Алгоритмом на масках модели UNetModInstanceNormLeakyReLU	Измерение вручную	0.0169	0.0271	0.898	4.0
	Алгоритмически на ИМБР	0.0231	0.0308	0.871	5.2
Алгоритмом на масках модели ResUNet	Измерение вручную	0.0228	0.0295	0.906	5.4
	Алгоритмически на ИМБР	0.0430	0.0492	0.886	9.6

лонами: расчетом алгоритма на ИМБР и ручными измерениями эксперта. Статистические характеристики результатов представлены в табл. 5, а метрики качества – в табл. 6. На рис. 9 приведены скаттерграммы распределения значений КТИ, рассчитанных по предсказанным маскам, относительно обоих эталонов.

Анализ статистических характеристик (табл. 5) показывает, что модели UNetInstanceNormLeakyReLU и UNetModInstanceNormLeakyReLU дают близкие средние значения КТИ (0.435), занимающие промежуточное положение между ручными измерениями (0.426) и расчетом на ИМБР (0.448). Модель ResUNet демонстрирует существенно

более низкое среднее значение (0.405), что указывает на выраженную тенденцию к занижению результатов. Обращает на себя внимание расширенный диапазон значений для модели UNet InstanceNormLeakyReLU, верхняя граница которого достигает 0.642, что свидетельствует о наличии отдельных выбросов с сильным завышением КТИ.

При сопоставлении с расчетом на ИМБР (табл. 6) модели UNetInstanceNormLeakyReLU и UNetModInstanceNormLeakyReLU показывают близкие значения MAE (0.0235 и 0.0231 соответственно) при высокой корреляции (0.891 и 0.871). Для модели ResUNet фиксируется резкий рост ошибки (MAE = 0.0430, относительная ошибка – 9.6 %), что подтверждает наличие сильного систематического смещения относительно автоматического эталона.

При сопоставлении с ручными измерениями эксперта (табл. 6) наилучшие результаты демонстрирует модель UNetModInstanceNormLeakyReLU: минимальные значения MAE (0.0169) и RMSE (0.0271), максимальная корреляция Пирсона (0.898) и наименьшая относительная ошибка (4.0 %). Модели UNetInstanceNormLeakyReLU и ResUNet показывают сопоставимые между собой, но более высокие значения MAE (0.0220...0.0228) и отно-

сительной ошибки (5.2...5.4 %). При этом модель ResUNet неожиданно демонстрирует наивысшую корреляцию с экспертом (0.906) при сопоставимой с другими моделями абсолютной ошибке (MAE = 0.0228).

На скаттерограммах (рис. 9) визуализированные распределения подтверждают описанные выше закономерности. При сопоставлении с расчетом на ИМБР (рис. 9, б) для всех моделей наблюдается преимущественное расположение точек ниже линии идеального соответствия, что подтверждает тенденцию к занижению значений. Наиболее выражено это смещение для модели ResUNet. В области высоких значений КТИ фиксируются отдельные выбросы в сторону завышения результатов. При сопоставлении с ручными измерениями эксперта (рис. 9, а) наблюдается значительно более плотная концентрация точек вдоль линии идеального соответствия. Наилучшая группировка достигается для модели UNetModInstanceNormLeakyReLU, точки которой равномерно распределены вокруг диагонали с минимальным разбросом, что согласуется с ее наилучшими метриками.

Таким образом, по совокупности показателей наилучший результат демонстрирует модель

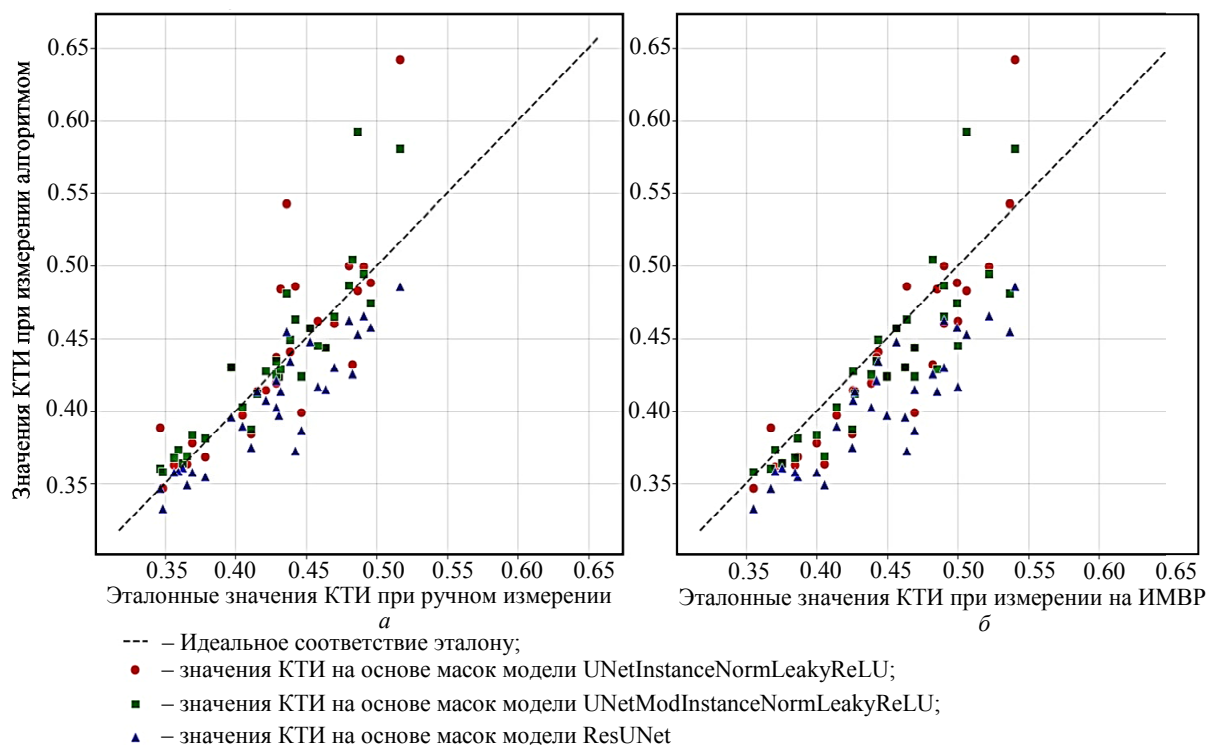


Рис. 9. Скаттерограммы значений КТИ, рассчитанных по предсказанным маскам моделей:
а – относительно ручных измерений эксперта; б – относительно расчета алгоритма на ИМБР
Fig. 9. Scatterograms of CTI values calculated from predicted model masks: а – regarding the expert's manual measurements; б – regarding the calculation of the algorithm of the true masks of high resolution

UNetModInstanceNormLeakyReLU, которая обеспечивает минимальную ошибку относительно экспертных измерений ($MAE = 0.0169$, относительная ошибка – 4.0 %), высокую корреляцию (0.898) и стабильность результатов. Это делает ее наиболее предпочтительной для практического применения в составе разработанной интеллектуальной системы.

Выводы и заключение. В работе представлена разработка интеллектуальной системы для анализа рентгенограмм грудной клетки на основе сверточных нейронных сетей, решающей задачи сегментации анатомических структур и автоматического расчета кардиоторакального индекса.

На первом этапе были разработаны и протестированы 10 моделей сверточных нейронных сетей на основе архитектур U-Net и SegNet с варьированием слоев повышения разрешения, типов нормализации и функций активации, а также с применением трансферного обучения с использованием предобученной модели ResNet18. Сравнительный анализ показал, что наилучшие результаты сегментации на тестовой выборке набора miniJSRT продемонстрировала модель ResUNet (средний IoU = 0.9382).

Для оценки работоспособности системы в условиях, приближенных к реальным, была проведена валидация на независимом наборе данных высокого разрешения. Установлено, что модель UNetModInstanceNormLeakyReLU обеспечивает наиболее точный расчет КТИ при сравнении с ручными измерениями эксперта: относительная ошибка составила 4.0 %, коэффициент корреляции Пирсона – 0.898.

При интерпретации полученных результатов необходимо учитывать ряд ограничений исследования: относительно небольшой объем и фиксированное разрешение обучающей выборки (256×256), что может ограничивать обобщающую способность моделей; неполнота исследования современных архитектур, альтернативных функций потерь, а также возможностей трансферного обучения; недостаточное исследование альтернативных методов для использования моделей с фиксированным входом для обработки изображений высокого разрешения; алгоритм расчета КТИ предполагает типичную анатомию грудной клетки и наличие выраженных патологических изменений может существенно влиять на точность его работы.

Указанные ограничения определяют направления дальнейшего развития работы: расширение обучающего набора данных; более детальное рассмотрение применения трансферного обучения; оптимизацию гиперпараметров и тестирование различных функций потерь; изучение альтернативных архитектур и методов обработки изображений высокого разрешения; исследование методов предобработки предсказанных масок для повышения качества расчета КТИ на их основе; исследование методов для реализации компенсации ротации пациента в алгоритме расчета КТИ.

Полученные результаты подтверждают возможность автоматизации расчета КТИ на основе нейросетевой сегментации и создают основу для последующей адаптации предложенного подхода к практическому клиническому применению без существенной потери диагностической значимости.

Список литературы

1. Лучевые методы исследования в диагностике, оценке объема резекции, планировании торакопластики и контроле эффективности хирургического лечения распространенного деструктивного туберкулеза органов дыхания с множественной лекарственной устойчивостью (МЛУ) / Д. Б. Гиллер, Г. В. Ратобильский, М. М. Никитин, В. В. Короев, О. П. Фролова, А. И. Шехтер, О. Н. Отс, Ю. Г. Григорьев, В. Н. Лавров // REJR. 2018. Т. 8, № 3. С. 249–258. doi: 10.21569/2222-7415-2018-8-3-249-258.

2. Топография отделов средостения в рентгеновском изображении в контексте дифференциации объемных медиастинальных образований / А. В. Леншин, А. В. Ильин, Е. А. Игнатьева, С. А. Крайнов, Ю. М. Перельман // Бюл. физиологии и патологии дыхания. 2022. Вып. 84. С. 100–115. doi: 10.36604/1998-5029-2022-84-100-115.

3. Deep learning-based cardiothoracic ratio measurement on chest radiograph: accuracy improvement

without self-annotation / K. Yoshida, A. Takamatsu, T. Matsubara, T. Kitagawa, F. Toshima, R. Tanaka, T. Gabata // Quantitative Imaging in Med. and Surgery. 2023. Vol. 13, no. 10. P. 6547–6553. doi: 10.21037/qims-23-187.

4. Кораблев В. М. Измерение кардиоторакального индекса на флюорографах серии ПроСкан. М.: ЗАО РЕНТГЕНПРОМ, 2016. С. 33–38.

5. Ограничения при применении сервисов искусственного интеллекта для анализа рентгенограмм органов грудной клетки / Ю. А. Васильев, А. В. Владимирский, К. М. Арзамасов, И. М. Шулькин, Е. В. Астапенко, Л. Д. Пестренин // Оригинальные исследования. 2024. Т. 5, № 3. С. 407–420. doi: 10.17816/DD626310.

6. Интеллектуальные методы анализа данных в биомедицинских исследованиях: сверточные нейронные сети / А. Н. Наркевич, К. А. Виноградов, К. М. Параскевопуло, Т. Х. Мамедов // Экология человека. 2021. № 5. С. 53–64.

7. Демин Н. С., Ильясова Н. Ю. Семантическая сегментация рентгенограмм легких на основе нейронной сети класса U-net // Сб. тр. по мат. VIII Междунар. конф. и молодежной шк. «Информационные технологии и нанотехнологии (ИТНТ-2022)». Самара: Изд-во СамНИУ им. акад. С. П. Королева, 2022. Т. 4. С. 53–54.

8. Deep learning in cardiothoracic ratio calculation and cardiomegaly detection / J. Kufel, I. Paszkiewicz, S. Kocot, A. Lis, P. Dudek, L. Czogalik, M. Janik, K. Bargiel-Laczek, W. Bartnikowska, M. Kozlik, M. Cebula, K. Gruszczynska, Z. Nawrat // J. of Clin. Med. 2024. Vol. 13, no. 14. P. 4180. doi: 10.3390/jcm13144180.

9. Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules / J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Kodera, K. Doi // AJR. 2000. Vol. 174, no. 1. P. 71–74. doi: 10.2214/ajr.174.1.1740071.

10. Сравнительный анализ модификаций нейросетевых архитектур U-Net в задаче сегментации медицинских изображений / А. М. Достовалова, А. К. Горшенин, Ю. В. Старичкова, К. М. Арзамасов // Digital

Diagnostics. 2024. Т. 5, No 4. С. 833–853. doi: 10.17816/DD629866.

11. Bala G., Mondal H., Sethi A. Brain MRI segmentation using U-Net and SegNet: A comparative study across modalities with robust loss functions // Proc. of the 18th Intern. Joint Conf. on Biomed. Engin. Syst. and Technol. (BIOSTEC). Porto, Portugal: SciTePress, 2025. Vol. 1: BIOIMAGING. P. 246–254. doi: 10.5220/0013302900003911.

12. Shafi S. M., Chinnappan S. K. Segmenting and classifying lung diseases with M-Segnet and Hybrid SqueezeNet-CNN architecture on CT images // PLoS One. 2024. Vol. 19, no. 5. Art. e0302507. doi: 10.1371/journal.pone.0302507.

13. Deep learning for automated triaging of stable chest radiographs in a follow-up setting / J. Yun, Y. Ahn, K. Cho, S. Y. Oh, S. M. Lee, N. Kim, J. B. Seo // Radiology. 2023. Vol. 309, № 1. Art. e230606. doi: 10.1148/radiol.230606.

14. Automatic heart localization and radiographic index computation in chest X-rays / S. Candemir, S. Jaeger, W. Lin, Z. Xue, S. Antani, G. Thoma // Proc. of SPIE – The Intern. Society for Opt. Engin. San Diego, CA, USA: SPIE, 2016. Vol. 9785. P. 1–8. doi: 10.1117/12.2217209.

Информация об авторах

Нурахунов Панкрат Максимович – магистрант кафедры вычислительной техники СПбГЭТУ «ЛЭТИ».

E-mail: pankrat.nurakhunov@mail.ru

Мелдо Анна Александровна – д-р мед. наук, профессор кафедры вычислительной техники СПбГЭТУ «ЛЭТИ».

E-mail: anna.meldo@yandex.ru

Прохоров Иван Юрьевич – врач-рентгенолог отделения радиотерапии клиники высоких технологий «Белоостров» (ООО «ММЦ ВТ»), тер. «Клиника «Белоостров», д. 1, корп. 1, Южковское с/п, Всеволожский муниципальный р-н, Ленинградская обл., 188632, Россия.

E-mail: x.ray.prokhorov@gmail.com

References

1. Luchevye metody issledovaniya v diagnostike, ocenke obshchego rezekcii, planirovaniy torakoplastiki i kontrole jeffektivnosti hirurgicheskogo lecheniya rasprostranennogo destruktivnogo tuberkuleza organov dyhaniya s mnozhestvennoj lekarstvennoj ustojchivost'ju (MLU) / D. B. Giller, G. V. Ratobyl'skij, M. M. Nikitin, V. V. Koroev, O. P. Frolova, A. I. Shehter, O. N. Ots, Ju. G. Grigor'ev, V. N. Lavrov // REJR. 2018. Т. 8, № 3. С. 249–258. doi: 10.21569/2222-7415-2018-8-3-249-258. (In Russ.)

2. Topografija otделov sredosteniya v rentgenovskom izobrazhenii v kontekste differenciacii obshchih meditsinal'nyh obrazovaniy / A. V. Lenshin, A. V. Il'in, E. A. Ignat'eva, S. A. Krajnov, Ju. M. Perel'man // Bjul. fiziologii i patologii dyhaniya. 2022. Vyp. 84. S. 100–115. doi: 10.36604/1998-5029-2022-84-100-115. (In Russ.)

3. Deep learning-based cardiothoracic ratio measurement on chest radiograph: accuracy improvement

without self-annotation / K. Yoshida, A. Takamatsu, T. Matsubara, T. Kitagawa, F. Toshima, R. Tanaka, T. Gabata // Quantitative Imaging in Med. and Surgery. 2023. Vol. 13, no. 10. P. 6547–6553. doi: 10.21037/qims-23-187. (In Russ.)

4. Korablev V. M. Izmerenie kardiorakal'nogo indeksa na fljuorografah serii ProScan. M.: ZAO RENTGENPROM, 2016. S. 33–38. (In Russ.)

5. Ogranicheniya pri primenenii servisov iskusstvennogo intellekta dlja analiza rentgenogramm organov grudnoj kletki / Ju. A. Vasil'ev, A. V. Vladimirovskij, K. M. Arzamasov, I. M. Shul'kin, E. V. Astapenko, L. D. Pestrenin // Original'nye issledovaniya. 2024. Т. 5, № 3. С. 407–420. doi: 10.17816/DD626310. (In Russ.)

6. Intellektual'nye metody analiza dannyh v biomeditsinskih issledovaniyah: svertochnye nejronnye seti / A. N. Narkevich, K. A. Vinogradov, K. M. Paraskevopulo,

T. H. Mamedov // Jekologija cheloveka. 2021. № 5. S. 53–64. (In Russ.)

7. Demin N. S., Il'jasova N. Ju. Semanticheskaja segmentacija rentgenogramm legkih na osnove nejronnoj seti klassa U-net // Sb. tr. po mat. VIII Mezhdunar. konf. i molodezhnoj shkoly «Informacionnye tehnologii i nanotehnologii (ITNT–2022)». Samara: Izd-vo Sam-NIU im. akad. S. P. Koroleva, 2022. T. 4. S. 53–54. (In Russ.)

8. Deep learning in cardiothoracic ratio calculation and cardiomegaly detection / J. Kufel, I. Paszkiewicz, S. Kocot, A. Lis, P. Dudek, L. Czogalik, M. Janik, K. Bargiel-Laczek, W. Bartnikowska, M. Kozlik, M. Cebula, K. Gruszczynska, Z. Nawrat // J. of Clin. Med. 2024. Vol. 13, no. 14. P. 4180. doi: 10.3390/jcm13144180.

9. Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules / J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Kodera, K. Doi // AJR. 2000. Vol. 174, no. 1. P. 71–74. doi: 10.2214/ajr.174.1.1740071.

10. Sravnitel'nyj analiz modifikacij nejrosetevyh arhitektur U-Net v zadache segmentacii medicinskih izobrazhenij / A. M. Dostovalova, A. K. Gorshenin, Ju. V. Stari-

chkova, K. M. Arzamasov // Digital Diagnostics. 2024. T. 5, no. 4. S. 833–853. doi: 10.17816/DD629866. (In Russ.)

11. Bala G., Mondal H., Sethi A. Brain MRI segmentation using U-Net and SegNet: A comparative study across modalities with robust loss functions // Proc. of the 18th Intern. Joint Conf. on Biomed. Engin. Syst. and Technol. (BIOSTEC). Porto, Portugal: SciTePress, 2025. Vol. 1: BIOIMAGING. P. 246–254. doi: 10.5220/0013302900003911.

12. Shafi S. M., Chinnappan S. K. Segmenting and classifying lung diseases with M-Segnet and Hybrid Squeezenet-CNN architecture on CT images // PLoS One. 2024. Vol. 19, no. 5. Art. e0302507. doi: 10.1371/journal.pone.0302507.

13. Deep learning for automated triaging of stable chest radiographs in a follow-up setting / J. Yun, Y. Ahn, K. Cho, S. Y. Oh, S. M. Lee, N. Kim, J. B. Seo // Radiology. 2023. Vol. 309, № 1. Art. e230606. doi: 10.1148/radiol.230606.

14. Automatic heart localization and radiographic index computation in chest X-rays / S. Candemir, S. Jaeger, W. Lin, Z. Xue, S. Antani, G. Thoma // Proc. of SPIE – The Intern. Society for Opt. Engin. San Diego, CA, USA: SPIE, 2016. Vol. 9785. P. 1–8. doi: 10.1117/12.2217209.

Information about the authors

Pankrat M. Nurakhunov – master's degree student, Department of Computer Engineering, Saint Petersburg Electrotechnical University.

E-mail: pankrat.nurakhunov@mail.ru

Anna A. Meldo – Dr Sci. (Med.), Professor of the Department of Computer Engineering, Saint Petersburg Electrotechnical University.

E-mail: anna.meldo@yandex.ru

Ivan Yu. Prokhorov – Radiologist at the Department of Radiotherapy at the «Beloostrov» High Technology Clinic (MMC HT LLC). Territory «Clinic «Beloostrov», block 1, building 1, Yukkovskoye rural settlement, Vsevolozhsky municipal district, Leningrad region, 188632, Russia.

E-mail: x.ray.prokhorov@gmail.com

Статья поступила в редакцию 11.03.2026; принята к публикации после рецензирования 01.06.2026; опубликована онлайн 28.09.2026.

Submitted 11.03.2026; accepted 01.06.2026; published online 28.09.2026.