

ПРИМЕНЕНИЕ ТРЁХКЛАССОВОЙ АРХИТЕКТУРЫ CNN ДЛЯ МИНИМИЗАЦИИ ЛОЖНЫХ ДИАГНОЗОВ В ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКЕ РАКА И ТУБЕРКУЛЕЗА ЛЕГКИХ: РЕЗУЛЬТАТЫ ПРЕДВАРИТЕЛЬНЫХ ИССЛЕДОВАНИЙ

И.В. Молодкин^{1,2}, С.А. Рыжкин^{2,3,4,5}, А.Г. Якунин⁶

¹ ООО “ОнкоСтол”, Москва

² Российская медицинская академия непрерывного постдипломного образования
Минздрава России, Москва

³ Казанский ГМУ Минздрава России, Казань

⁴ Казанский (Приволжский) федеральный университет, Казань

⁵ Академия наук РТ, Казань

⁶ Алтайский государственный технический университет им. И.И. Ползунова,
Барнаул

MINIMIZING FALSE DIAGNOSES IN THE DIFFERENTIAL DIAGNOSIS OF LUNG CANCER AND TUBERCULOSIS USING A THREE-CLASS CNN: PRELIMINARY RESULTS

I.V. Molodkin^{1,2}, S.A. Ryzhkin^{2,3,4,5}, A.G. Yakunin⁶

¹ LLC OncoStop, Moscow, Russia

² Russian Medical Academy of Continuous Postgraduate Education, Moscow, Russia

³ Kazan State Medical University, Kazan, Russia

⁴ Kazan (Volga region) Federal University, Kazan, Russia

⁵ Academy of Sciences of the Republic of Tatarstan, Kazan, Russia

⁶ I.I. Polzunov Altai State Technical University, Barnaul, Russia

Реферат

Цель: Получить предварительную оценку возможностей применения трёхклассовой архитектуры сверточной нейронной сети (CNN) для снижения вероятности получения ложноположительных и ложноотрицательных заключений при дифференциальной диагностике туберкулеза и рака легких за счет выделения в отдельный класс нетипичных для данных заболеваний изображений.

Материал и методы: Исследования проводились путем вычислительного эксперимента с использованием разработанного для этих целей программного комплекса на основе сверточной нейронной сети. На выходе нейросети формировалась диаграмма, показывающей вероятности соответствия снимка раку, туберкулёзу, либо иному, не содержащему признаки данных заболеваний, изображению. Для обучения нейросети было отобрано 108 изображений, количество которых было доведено до 512 за счет использования увеличения. Работа модели оценивалась по тестовой части выборки и по изображениям, взятым из внешних источников.

Результаты: Разработанная нейросетевая модель позволяет не только идентифицировать типовые КТ-изображения или рентгеновские снимки рака и туберкулеза легких. Вероятность отнесения исследуемого снимка к верифицированному классу (рак, туберкулез или иное не содержащее признаков данных заболеваний изображение), лежит в пределах 80–95 %, тогда как на вероятность его отнесения к альтернативным классам не превышает 5–20 %. И даже для изображе-

ний, взятых из сторонних источников, существенно отличающихся от изображений датасетов, результат распознавания был положительным. Представление результатов работы нейросети в виде вероятности отнесения изображения к тому или иному классу заболевания существенно повышает степень доверия врача к получаемым результатам и минимизирует вероятность получения как ложноположительных, так и ложноотрицательных заключений.

Заключение: Проведенные исследования показали, что разработанная нейросетевая модель верно классифицирует изображения и за счет введения дополнительного класса повышает надёжность и степень доверия к получаемому результату.

Ключевые слова: КТ-изображения, рентгеновские снимки дифференциальная диагностика, рак, туберкулёз, программный комплекс, трёхклассовая идентификация, сверточные нейронные сети

Abstract

Purpose: To conduct a preliminary assessment of the potential of a three-class convolutional neural network (CNN) architecture to reduce the likelihood of false-positive and false-negative diagnoses in the differential diagnosis of tuberculosis and lung cancer by classifying images atypical for these diseases into a separate class.

Materials and methods: The study was conducted through a computational experiment using a specially developed software package based on a convolutional neural network. The neural network output generated a diagram showing the probabilities of an image matching cancer, tuberculosis, or another image not containing signs of these diseases. 108 images were selected for training the neural network, which was increased to 512 through the use of augmentation. Model performance was evaluated using a test sample and images taken from external sources.

Results: The developed neural network model enables the identification of not only typical CT images and X-rays of lung cancer and tuberculosis. The probability of classifying a scanned image into a verified class (cancer, tuberculosis, or another image containing no signs of these diseases) ranges from 80–95 %, while the probability of classifying it into alternative classes does not exceed 5–20 %. Even for images taken from third-party sources that differed significantly from the dataset images, the recognition result was positive. Presenting the neural network's results as the probability of classifying an image into a particular disease class significantly increases the physician's confidence in the obtained results and minimizes the likelihood of both false positives and false negatives.

Conclusion: The studies showed that the developed neural network model correctly classifies images and, by introducing an additional class, increases the reliability and confidence in the obtained results.

Key words: CT images, X-ray images, differential diagnosis, cancer, tuberculosis, software package, three-class identification, convolutional neural networks

E-mail: tian.08@mail.ru

<https://doi.org/10.52775/1810-200X-2025-108-4-112-129>

Введение

По данным Всемирной организации здравоохранения (ВОЗ), легочные заболевания занимают третье место среди всех причин смертности во всем мире после сердечно-сосудистых и онкологических заболеваний. Согласно статистике, хроническая обструктивная болезнь легких является четвертой ведущей причиной смерти в мире и вызывает около 3,5 млн смертей в год, а хронические респираторные заболевания в целом являются третьей по распространенности причиной смертности [1]. И хотя смертность от обструкции составляет 60 %, считается, что наиболее опасным являет-

ся рак легких. Именно он является самой частой причиной смерти от онкозаболеваний, и смертность от него постоянно растет. Рак легких остается наиболее часто диагностируемым раком и ведущей причиной смерти от рака в мире. Исследования эпидемиологии рака показали, что вклад рака легкого в общую онкологическую заболеваемость составляет 12,4 %, а смертность – 18,7 % [2]. Рак легких остается доминирующей формой рака с выраженными географическими, гендерными и социально-экономическими различиями. Так, смертность от рака легких в развитых странах в 5 раз превышает среднестатистическое мировое значение [2]. Анамнез данного заболевания, методы

его диагностики и виды, клинико-анатомическая и гистологическая классификация, а также основы диагностики достаточно подробно представлены, например, в [3].

Наряду с раком, туберкулез также является одним из серьезных легочных заболеваний. Так, из отчета ВОЗ за 2024 г. следует, что туберкулез вернул себе позицию ведущей инфекционной причины смерти в мире, представляющей существенную угрозу общественному здоровью. В 2023 г. было зарегистрировано 8,2 млн новых случаев туберкулеза и 1,28 млн смертей от него – почти вдвое больше, чем от СПИДа и больше, чем от COVID-19, достигнув максимального числа с начала глобального мониторинга ВОЗ в 1995 г. [4, 5].

Анализ методов диагностики легочных заболеваний в целом и диагностики рака легких и туберкулеза в частности показал, что симптоматика рака легких и туберкулеза во многом схожа, что вызывает значительную проблему, и патогенетическая связь туберкулеза и рака легких все еще остается предметом дискуссии. [6–8]. По этой причине для обоих заболеваний используются во многом сходные методы морфологической диагностики, из которых наиболее информативными и наиболее распространенными являются такие рентгенологические методы исследования, как рентгенография и компьютерная томография (КТ) и менее распространенными из-за меньшей доступности – многосрезовая компьютерная томография (МСКТ) и позитронно-эмиссионная томография (ПЭТ), часто совмещаемая с КТ (ПЭТ/КТ). Результатом перечисленных рентгенологических и радионуклидных обследований являются изображения проекций или поперечных срезов легких, которые передаются для анализа врачу-рентгенологу или врачу-радиологу для составления описания и постановки диагноза.

Однако, несмотря на высокую информативность перечисленных методов структурно-анатомической диагностики, сложность интерпретации получаемых с их применением изображений остаётся значительной, их интерпретация во многом субъективна и может приводить к диагностическим ошибкам [9, 10]. Поэтому для минимизации субъективности принятия решения в помощь врачу-диагносту в последние годы стали предоставляться различные программные средства для автоматизированного распознавания изображений, позволяющие, например, выделять и анализировать

конкретные области интереса путем исследования статистических свойств границ новообразований, их фрактальной размерности, либо вариации плотности изображения в области интереса [11–13].

В последние годы для автоматизации формирования диагностического заключения при исследовании органов дыхания, включая легкие, стали широко применяться различные виды искусственных нейронных сетей [14, 15] и даже искусственный интеллект [16].

Для обработки визуальной информации при этом применяются следующие методы и подходы:

- ✓ *Методы глубокого обучения* (Deep Learning), позволяющие нейросети решать поставленные задачи не по конкретным алгоритмам, а учиться на опыте, по множеству предъявляемых для накопления этого опыта изображений-эталонов и оценки получаемого результата с целью корректировки параметров сети [17].
- ✓ *Метод переноса обучения* (Transfer Learning, TL), суть которого состоит в том, что “знания” и “опыт”, полученные при решении одной конкретной задачи, берутся за основу для решения другой задачи, близкой по содержанию к изначально решаемой. Это значительно экономит время и аппаратные ресурсы на обучение сети при ее применении для решения новой задачи. Кроме того, при решении задач анализа медицинских изображений метод переноса обучения позволяет также преодолеть проблему дефицита необходимых для машинного обучения данных [18].
- ✓ *Метод диагностики на основе сиамских нейронных сетей* [19] – такая сеть представляет собой две идентичные сверточные нейронные сети (convolutional neural network, CNN), на вход одной из которых подается исследуемое изображение, на вход второй – изображение с известным диагнозом, а на выходе формируется заключение о степени сходства данных изображений. Следует отметить, что такие сети больше подходят для систем допускового контроля, где требуется идентифицировать личность входящего, чем для целей медицинской диагностики, когда для одного и того же диагноза как эталонные, так и анализируемые изображения могут сильно варьироваться.
- ✓ *Использование в нейросетях (преимущественно рекуррентных) механизма внима-*

ния. Он позволяет нейросети в процессе поиска решения идти от общего к частному, имитируя человеческий когнитивный процесс. Различают канальное и пространственное внимание (Channel Attention и Spatial Attention). Первый механизм определяет наиболее важные признаки изображения, а второй находит местоположение ключевой информации [20, 21].

- ✓ *Применение для диагностики объяснимого искусственного интеллекта (Explainable Artificial Intelligence, XAI) [22, 23], представляющего собой набор процессов и методов, позволяющих наряду с конечным результатом анализа входных данных предоставить врачу объяснение того, как этот результат был получен. Это дает врачу возможность проконтролировать процесс получения результата и тем самым получить возможность исключения критических ошибок при постановке диагноза искусственным интеллектом. Таким образом, понимание того, как был получен результат, может не только повысить доверие, но и избежать опасных для жизни ошибок диагностики.*
- ✓ *Метод аугментации хранимых в датасетах данных [24, 25]. Суть метода состоит в том, что для обучения нейросети используются не только содержащиеся в датасетах изображения, но и изображения, полученные из исходных путём их ограниченной трансформации, например, за счет использования аффинных преобразований. Это позволяет кратно увеличить число используемых при обучении и тестировании сети изображений, что существенно повышает качество обучения.*
- ✓ *Методы компенсации дисбаланса классов [26, 27]. Суть дисбаланса классов состоит в том, что в используемых для обучения датасетах изображения здоровых и пораженных легких представлены неравномерно. Неравномерно могут быть представлены изображения с разными видами заболеваний. В [27] также отмечается, что на медицинских изображениях поражения могут занимать всего несколько пикселей изображения, что приводит к значительному дисбалансу классов между поражением и фоном. Одним из путей ликвидации дисбаланса является применение аугментации. Другой, предлагаемый в [27], подход состоит в применении двухэтапной структуры глубокого обучения, когда на первом этапе на изображении выде-*

ляются области интересов и минимизируются фоновые участки, а на втором этапе производится финальное обучение.

Анализ перечисленных выше работ показал, что все рассмотренные в них подходы широко используются для решения задачи диагностики рака легких и туберкулеза.

При этом следует отметить следующее.

1. Большинство работ посвящено выявлению только одного конкретного заболевания – рака (например, [28, 29]) или туберкулеза (например, [30, 31]), и лишь небольшое число работ решает задачу их дифференциальной диагностики (например, [32, 33]), хотя именно такая диагностика крайне важна в силу схожести обусловленных туберкулезом и раком легких изображений пораженных участков легкого.
2. В подавляющем большинстве случаев для анализа изображений применяются CNN глубокого обучения.
3. Программные средства автоматизированной диагностики, особенно основанные на применении искусственного интеллекта, достаточно требовательны к используемым вычислительным ресурсам, что ограничивает возможности их применения в неспециализированных лечебных учреждениях для проведения скрининг-обследований.
4. В ряде публикаций, например, [34, 35], отмечается, что даже прошедшие сертификацию программные комплексы на основе искусственного интеллекта не всегда устанавливают верный диагноз. Отчасти это связано с тем, что в них не предусмотрена возможность выявления ситуации, когда предоставленные для анализа результаты не содержат признаков заболеваний, для которых разрабатывались данные комплексы.

Отсюда вытекает цель данной работы – исследовать возможность повышения качества диагностирования туберкулеза и рака легкого за счет введения в модель нейросети в дополнение к двум базовым классам (рак и туберкулез) третьего класса для изображений, не содержащих признаки данных заболеваний, либо когда эти признаки выражены недостаточно явно и для установления диагноза следует привлечь квалифицированного специалиста или даже собрать консилиум.

Положительный результат таких исследований позволит повысить достоверность работы систем автоматизированной диагностики за счет уменьшения вероятности отнесения

снимка к одному из перечисленных заболеваний из-за ложноположительных или ложноотрицательных диагнозов, а значит и расширить сферу их применения. В перспективе данный подход может быть развит для выявления более широкого спектра легочных заболеваний, основанных на результатах КТ или других рентгенологических исследований.

Для достижения данной цели был выбран тип и архитектура нейросети, разработан программный комплекс для ее реализации, сформирована обучающая выборка, с её помощью проведено обучение сети, и на завершающей стадии проведены экспериментальные исследования для оценки эффективности ее применения. Остановимся на перечисленных этапах реализации проекта более подробно.

Выбор и описание структуры нейросетевой модели

Как следует из рассмотрения работ, посвященных автоматизации процесса формирования диагностических заключений, для этих целей используются в основном многослойные перцептроны (Multilayered Perceptron, MLP), рекуррентные нейронные сети (Recurrent Neural Network, RNN) и сверточные нейронные сети (CNN) (рис. 1).

Каждая из этих архитектур обладает определенными преимуществами и ограничениями применительно к типу решаемой ими задачи. MLP, как правило, используется для анализа табличных или одномерных данных, но слабо подходит для обработки изображений из-за отсутствия пространственной инвариантности. RNN применяется преимущественно в задачах, связанных с временными рядами и последовательностями, например, при анализе ЭКГ. CNN является наиболее эффективным инструментом при работе с изображениями (рис. 2). Она способна извлекать иерархические признаки из изображений, начиная с базовых (границы, контуры) и заканчивая высокоуровневыми абстракциями (формы, текстуры). Поэтому при разработке необходимого для проведения исследований программного комплекса был выбран именно такой тип нейросети.

Основу любой архитектуры сверточной нейронной сети составляют блоки свертки (Conv2D) и подвыборки или пулинга (MaxPooling2D), за которыми следует слой вы-

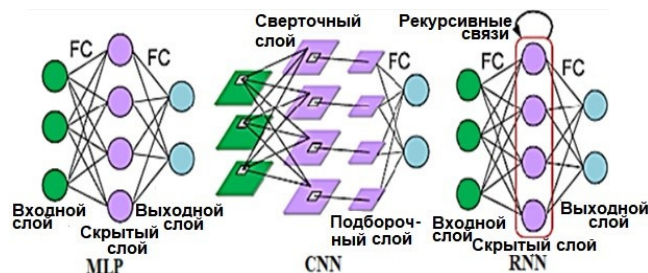


Рис. 1. Архитектуры нейронных сетей MLP и RNN применяемых в автоматизированных системах медицинской диагностики, где FC (fully connected) – полносвязные слои

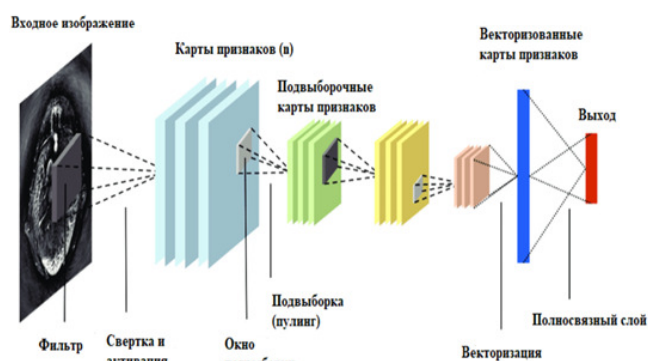


Рис. 2. Структура классической сверточной нейросети

прямления (векторизации) и полносвязные (Dense) слои.

Конкретная архитектура реализованной в программном комплексе структуры сверточной нейросети содержит следующие слои и их специфические настройки.

Первый сверточный слой Conv2D с 32 фильтрами размером 3×3 извлекает базовые признаки (границы, текстуры, простые формы) из изображения. Он проходит сверткой по всему изображению, создавая карты признаков, где каждый фильтр формирует своё представление об определённой особенности на изображении. Активация ReLU применяется для устранения отрицательных значений, вводя в модель необходимую нелинейность.

За первым слоем Conv2D следует слой MaxPooling2D, который уменьшает размерность карт признаков, оставляя только наиболее значимые элементы из каждого окна 2×2, что снижает вычислительные затраты и делает извлечённые признаки инвариантными к сдвигу.

Далее используется второй Conv2D слой с 64 фильтрами, за которым также следует пулинг, что углубляет уровень извлечённых при-

знаков – теперь сеть может распознавать более сложные структуры и формы. Затем данные проходят через слой Flatten, переводящий их в одномерный вектор, который подаётся в полносвязный слой Dense с 64 нейронами. Перед этим используется Dropout слой с параметром `drop_rate=0.3`, отключающий случайные нейроны на каждой итерации, что предотвращает переобучение и заставляет сеть учиться более устойчивым представлениям.

Заключительный полносвязный Dense-слой соединяет все признаки, собранные с предыдущих уровней, и отправляет их на softmax-слой, который возвращает вероятностное распределение по классам. Таким образом, сеть принимает решение на основе совокупной информации, извлечённой со всего изображения.

Функция потерь `categorical_crossentropy` сравнивает предсказанные вероятности с истинными метками классов и возвращает значение ошибки – чем ближе предсказание к истинной метке, тем меньше значение функции. Во время обратного распространения ошибки (backpropagation) вычисляется градиент этой ошибки по отношению к каждому весу сети, и веса обновляются с помощью оптимизатора Adam, который адаптивно регулирует шаг обучения на основе первого и второго моментов градиента.

Гиперпараметры модели были взяты типовые: размер ядра свёртки – 3×3 , пулинг – 2×2 , значение коэффициента Dropout, равное 0,3, количество эпох обучения – от 10. Для контроля процесса обучения использовались обратные вызовы `EarlyStopping` и `ModelCheckpoint`, автоматически сохраняющие наилучшую по валидационной точности модель. Предусмотрено использование регуляризации через Dropout, что помогает бороться с переобучением.

Используемая на выходе нейросети softmax-активация необходима для реализации многоклассовой идентификации. Softmax обеспечивает, чтобы каждый соответствующий конкретному классу выход представлял вероятность принадлежности результата диагностики к конкретному классу, и сумма вероятностей всех выходов равнялась единице. В данной модели было реализовано три выхода для идентификации трех классов-диагнозов: “Туберкулез”, “Рак легкого” и “Что-то иное”. Наличие последнего класса, предполагающего, что на вход сети подано любое иное изображение, не содержащее признаков перечисленных за-

болеваний, принципиально отличает предлагаемое решение от других аналогичных решений, позволяя пользователю не предположить, а наглядно увидеть, в какой степени исследуемый снимок не относится ни к туберкулезу, ни к раку легкого.

Описание программного комплекса

Построение интеллектуальной системы автоматизированной диагностики лёгочных заболеваний требует поэтапной реализации программного решения, охватывающего весь жизненный цикл обработки медицинских изображений – от их загрузки и предобработки до обучения с последующей классификацией и представлением диагностических выводов при применении модели во врачебной практике.

В качестве основной среды разработки использовался язык программирования Python, благодаря его широким возможностям в сфере анализа данных и глубокого обучения. При разработке программного комплекса использовались следующие инструменты (фреймворки и библиотеки) с открытым исходным кодом.

- ✓ TensorFlow / Keras – написанные на Python фреймворки глубокого обучения, необходимые для построения, обучения и сохранения нейросетевой модели. TensorFlow предлагает комплексные инструменты для работы с нейросетями, а Keras обеспечивает масштабируемость и надёжное развёртывание приложения.
- ✓ NumPy – библиотека для Python, использованная для работы с большими многомерными числовыми массивами и матрицами в процессе обработки преобразуемых в них изображений.
- ✓ OpenCV и PIL – библиотеки, использованные для обработки и преобразования изображений.
- ✓ PyDICOM – Python-библиотека, необходимая для чтения файлов медицинского формата DICOM и извлечения из них графической и текстовой информации.
- ✓ Matplotlib и Seaborn – библиотеки для Python, использованные для построения графической информации, в частности для построения графиков обучения и представления результатов анализа DICOM-изображений.

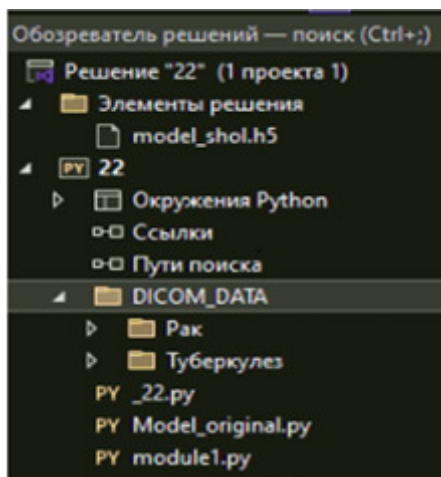


Рис. 3. Файловая структура проекта

- ✓ Scikit-learn – построенная на основе NumPy и SciPy библиотека Python, предназначенная для машинного обучения и анализа данных. В процессе исследований использовалась для расчёта точности, полноты, F1-меры, а также построения матриц ошибок.

Программный комплекс построен на модульной архитектуре, позволяющей разделить весь процесс на независимые блоки: предварительная обработка данных, построение и обучение модели, сохранение результатов, визуализация. Файловая структура проекта отражена на рис. 3. Модульный подход обеспечивает гибкость и масштабируемость решения, а также упрощает отладку и возможную доработку программного продукта в будущем. В программный комплекс входит также модуль применения модели для решения практических задач медицинской диагностики по КТ и рентгеновским снимкам. Его интерфейс позволяет пользователю загрузить снимок (в формате DICOM, PNG или JPEG), после чего изображение автоматически нормализуется, масштабируется, подается на вход модели и на её выходе рассчитываются вероятности принадлежности изображения к каждому из вышеперечисленных диагностических классов. Результаты расчета отображаются в понятной визуальной форме, отражающей степени уверенности по каждому диагнозу.

Таким образом, разработанный программный комплекс представляет собой завершённый инструмент, способный не только решать задачи дифференциальной диагностики, но и стать основой для последующего его расширения для решения задач сегментации

изображений для выделения областей интереса, поиска очагов патологии и построения объяснимых моделей. Простота предложенной архитектуры CNN позволила использовать для проведения предварительных исследований обычный персональный компьютер на базе CPU Intel® Core (TM) i7-10750H @2.6GHz с объемом оперативной памяти 16 ГБ.

Подготовка исходных данных для построения и тестирования нейросетевой модели

Основными этапами отбора и подготовки данных для обучения нейросети и проведения последующих исследований являются выбор подходящего датасета, извлечение из него необходимых данных, преобразование этих данных в формат изображений, нормализация полученных изображений и их преобразование в единообразный формат, репродуцирование полученного множества изображений до количества, минимально достаточного для обучения сети, и разбиение этого множества на подмножества, необходимые для проведения обучения и тестирования нейросети. Рассмотрим эти этапы более подробно.

Получение набора исходных изображений

Для получения исходных изображений, необходимых для обучения нейронной сети, был выбран один из известных ресурсов с открытым доступом, Cancer Imaging Archive (TCIA), представляющий собой базу данных с открытым доступом, содержащую медицинские изображения с исследованиями рака лёгкого. Данные в базе организованы в обширные коллекции медицинских изображений, которые сгруппированы по типу рака и его анатомической локализации [36, 37]. Большая часть коллекций состоит из изображений КТ, МРТ и ядерной медицины (например, ПЭТ), которые хранятся в формате DICOM, но для повышения исследовательской ценности также предоставляются или приводятся ссылки на многие другие типы вспомогательных данных. Все данные обезличены в соответствии с Законом о переносимости и подотчётности медицинского страхования и политикой обмена данными Национальных институтов здравоохранения.

Для обучения нейросети диагностике туберкулеза был использован датасет с известного ресурса Kaggle, содержащего в откры-

том доступе 700 рентгенограмм различных форм туберкулеза и пациентов с нормальными показателями, не имеющих признаков заболеваний (набор Tuberculosis Chest X-ray Dataset), а также 165 КТ-срезов с бинарными масками (CT Tuberculosis Dataset) [38].

Кроме того, для обучения сети по классу туберкулеза и прочих заболеваний был задействован датасет LIDC-IDRI (The Lung Image Database Consortium – Image Database Resource Initiative), содержащий 1018 случаев КТ-исследований [39, 40].

Поскольку файлы изображения в датасетах хранятся в специальном медицинском формате DICOM (Digital Imaging and Communications in Medicine, расширение файлов “dcm”), предназначенном для хранения медицинских изображений и сопровождаемых их метаданных, для извлечения из них изображений использовались соответствующие функции библиотеки PyDICOM, представляющей собой мощный инструмент для работы с медицинскими изображениями в формате DICOM. Библиотека предоставляет удобные возможности не только для извлечения данных из DICOM-файлов, но и для проведения разнообразных манипуляций с изображениями, анализа метаданных и их конвертации в различные форматы.

Всего из упомянутых датасетов было выбрано 108 изображений с диагнозами “туберкулез” (54 снимка), “рак легкого” (54 снимка), а также дополнительно 54 снимка легких из последнего датасета, формирующих третий класс. Выбор одинакового количества снимков для каждого класса позволил исключить ранее упомянутую проблему дисбаланса выборок. Здесь стоит отметить, что столь малый объем выборок объясняется тем, что изображения извлекались из архивов вручную, что было достаточно трудоемко. Кроме того, в работе ставилась цель не проведения исследований по оценке статистических свойств получаемых оценок, а проведения исследований только для качественной оценки достоинств и недостатков применения нового подхода при проведении дифференциальной диагностики.

Предварительная обработка исходных изображений

Предварительная обработка DICOM-изображений играет ключевую роль как на стадии подготовки данных для обучения нейронных сетей, так и при их использовании для ре-

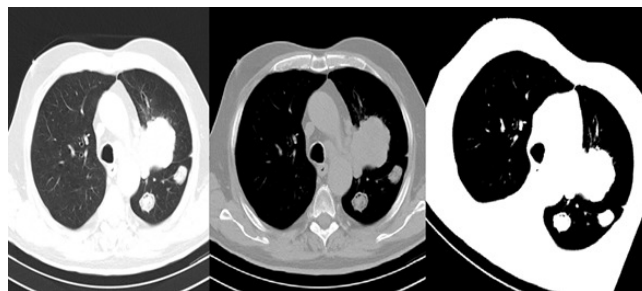


Рис. 4. Пример изображения до и после нормализации и после бинаризации

шения практических задач, и содержит несколько этапов.

На первом этапе изображения в формате DICOM с применением библиотеки PyDICOM конвертировались в одноканальные (grayscale) изображения в формате PNG. На втором этапе они масштабировались до стандартизированного разрешения 512×512 пикселей для совместимости с большинством современных нейросетевых архитектур. На третьем этапе производилась коррекция контрастности (VOI LUT) и выполнялась нормализация их яркости, что особенно важно для медицинских изображений, варьирующихся по уровню экспозиции.

Особое значение имеет преобразование на следующем этапе обработки полученных черно-белых изображений в восьмибитный формат, обеспечивающий диапазон яркости пиксела в виде 256 градаций серого. Это позволило снизить требования к вычислительным ресурсам, а также сократить время обучения нейросети. Для повышения скорости обучения и снижения требований к вычислительным ресурсам рассматривалась также возможность применения бинаризации изображений, однако исследования показали, что бинаризация снижает точность идентификации. Пример исходного, нормализованного и бинаризованного изображения показаны на рис. 4.

Аугментация исходного набора данных

Точность предсказания нейросетевой модели во многом определяется объемом и разнообразием обучающей выборки – чем она больше, тем выше точность и тем больше обобщающая способность и устойчивость нейросетевой модели к вариативности данных, а также меньше риск переобучения из-за ограниченности объема обучающей выборки. Поэтому следующим этапом со взятыми из датасетов изображениями стала их аугментация, позволяющая

увеличить объем обучающей выборки. Она позволяет создавать дополнительные изображения путем внесения различных изменений в имеющиеся в наличии оригиналы и широко используется для решения проблемы нехватки данных в обучающих наборах. В данных исследованиях использовалось два подхода с аугментацией исходных изображений.

Первый основан на применении к изображению простейших операций, к которым относятся:

- ✓ горизонтальное и вертикальное отражение, позволяющее создать зеркальные отражения изображений, увеличивая разнообразие данных;
- ✓ случайные повороты и сдвиги, позволяющие модели обучаться на различных ракурсах объектов и сцен, улучшая ее устойчивость к вариациям;
- ✓ изменение яркости и контраста, помогающее модели адаптироваться к изменениям освещения;
- ✓ масштабирование и обрезка, создающие дополнительные варианты изображений без внесения в них геометрических искажений формы;
- ✓ добавление шума, улучшающее устойчивость модели к артефактам;
- ✓ эластичные искажения, помогающие модели учитывать разнообразие форм и структур объектов.

Реализация данных операций была выполнена на базе библиотеки `torchvision.transforms`, а их результат сохранялся в директории `augmented_dataset`, разбитой на подпапки по классам ("Рак" и "Туберкулез") (рис. 3). Примеры применения к изображению части из перечисленных операций, относящихся к аффинным преобразованиям, представлены на рис. 5.

Второй подход основан на создании множества вариаций одного и того же DICOM-изображения путем использования технологии Stable Diffusion и вариационных автоэнкодеров (VAE). Для реализации этого подхода использовался ImageDataGenerator из Keras, позволяющий эффективно организовать подачу данных в модель за счет включения в них параметра `validation_split` для встроенного разделения в обучающей выборке различных типов изображений.

В результате применения аугментации общее количество изображений по двум основным классам возросло до 512. Каждый исходный снимок был обработан и дополнен не-

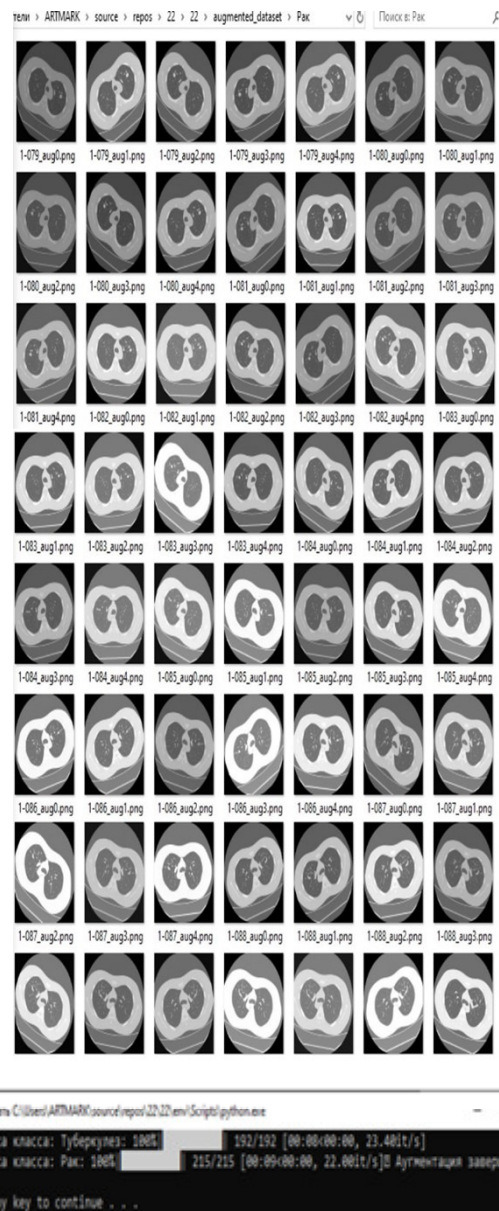


Рис. 5. Примеры применения аффинных преобразований к исходному изображению (вверху) и сообщение о завершении процесса аугментации (внизу)

сколькими вариациями, отличающимися по наклону, яркости и контрастности, что позволило искусственно расширить и разнообразить датасет. Использование аугментации вместо расширения числа исходных изображений объясняется тем, что аугментация выполнялась автоматически программными средствами, тогда как отбор снимков делался вручную. Кроме того, её применение оправдано и для проведения последующих, более масштабных исследований, так как позволяет в разы увели-

чить объем обучающей выборки и тем самым повысить качество обучения нейросетевой модели.

Организация процесса обучения нейросети

После построения и формирования исходных данных следует собственно этап обучения. Для этого весь набор данных был разбит на тренировочную, валидационную и тестовую части, которые в работе были выбраны в соотношении 80:10:10. Для такого разбиения на этапе подготовки к обучению, как уже отмечалось, использовался функционал ImageData-Generator с параметром `validation_split`, позволяющий автоматически делить данные на обучающую и валидационную выборки.

Все предварительно обработанные (масштабированные, нормализованные и преобразованные в восьмибитный формат greyscale) изображения подавались на вход модели в виде тензоров с одним каналом, что позволяло сохранить их визуальные особенности. Метка класса для каждого изображения формировалась на основе названия папки и кодировалась в формате one-hot: [1, 0] – для изображений с диагнозом “рак”, [0, 1] – для “туберкулёза” и [0, 0] – для прочих изображений. Далее снимок проходил через архитектуру свёрточной нейросети, реализованной с использованием модели Sequential.

Обучение происходило поэтапно, батчами, по 32 изображения в батче. Это позволяло

обрабатывать данные по частям, не загружая всю выборку в память сразу. Одна эпоха представляет собой полный проход по обучающему датасету, а итерация – это количество шагов внутри одной эпохи, равное числу батчей. На каждой итерации сеть делала прогноз, сравнивала его с истинной меткой (target), рассчитывала ошибку (loss) и обновляла внутренние параметры с использованием метода обратного распространения ошибки и оптимизатора Adam с начальной скоростью обучения 0,001. За счёт такой обратной связи сеть в процессе обучения постепенно минимизировала функцию потерь и улучшала свои предсказания, делая акцент на признаках, оказавшихся значимыми для постановки диагноза.

Для отслеживания динамики и качества обучения применялись такие метрики, как accuracy, precision, recall, F1-score. Для исключения переобучения использовался механизм early stopping и строились графики изменения валидационной точности и функции потерь, позволяющие визуально оценить, как быстро и стабильно обучается модель (рис. 6). Благодаря этим визуализациям было установлено, что уже на 5–6 эпохе точность стабилизируется, а потери на валидации перестают снижаться, что свидетельствовало о достижении оптимума.

Результатом процесса обучения явилась модель CNN, способная различать патологические образования на изображениях лёгких с высокой степенью достоверности.

Для обеспечения дальнейшего использования модели без необходимости повторного

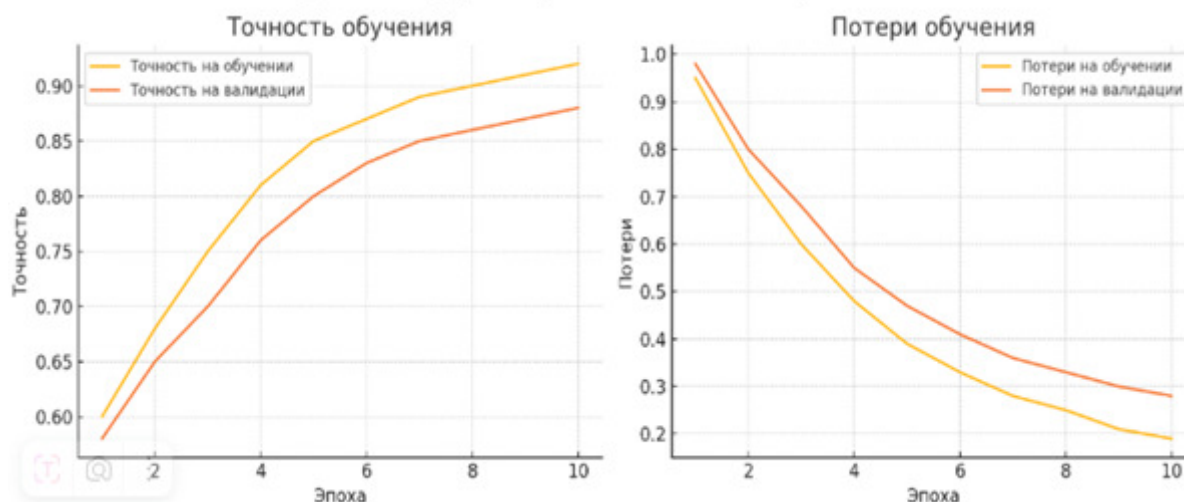


Рис. 6. Зависимость точности и потери в зависимости от эпохи обучения

обучения или для её применения по технологии Transfer Learning для диагностики других близких по внешним проявлениям заболеваний она была сохранена в формате .h5 в соответствующем каталоге файловой структуры проекта (рис. 3).

Пример работы нейросетевой модели и полученные результаты

Для количественной и качественной оценки работы модели на тестовой выборке и сравнительного анализа с экспертными заключениями использовался программный модуль с пользовательским интерфейсом (на базе Tkinter). Он позволяет загружать рентгеновские и КТ-снимки, выполнять их классификацию и получать результат в виде диаграммы, показывающей вероятности принадлежности снимка к одному из классов диагностических классов (рак, туберкулёз, иное изображение), что делает систему удобной и интерпретируемой в условиях клинической практики.

Далее приводится демонстрация работы разработанной нейросетевой модели на примере нескольких снимков легких, не входивших в обучающую выборку, с целью оценки поведения модели на новых данных, а также оценки интерпретируемости и достоверности полученных результатов классификации. Все снимки до подачи на вход сети подвергались ранее описанной предварительной обработке.

Так, на рис. 7 показан КТ срез легкого пациента с диагнозом “Рак легкого”. Соответствующий данному изображению результат работы нейросети в виде распределения вероятностей (или степени уверенности в диагнозе) заболеваний по их классам представлено на рис. 8.

Как видно из представленной на этом рисунке диаграммы, модель продемонстрировала высокую степень уверенности при классификации данного изображения, с явным преобладанием вероятности класса “рак лёгких”. При этом вероятность отнесения к другим классам была минимальной, что свидетельствует об однозначности предсказания. На самом снимке, который использовался для диагностики, хорошо видна округлая область – характерное образование, соответствующее опухоли в лёгком. Именно наличие этой заметной круглой зоны позволило нейросети правильно распознать и классифицировать изображение. Таким обра-

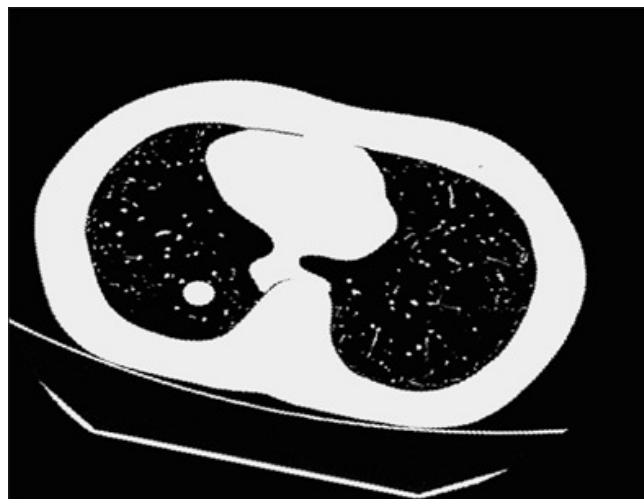


Рис. 7. Нормализованное изображение среза легкого с признаками онкологического заболевания

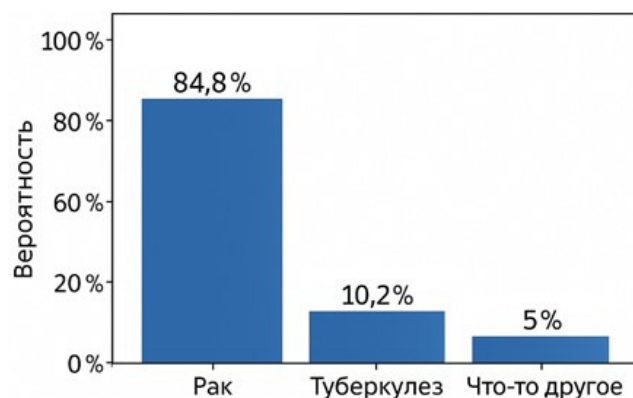


Рис. 8. Распределение вероятностей по диагностическим классам для рака легких

зом, модель эффективно выделяет ключевые признаки, присутствующие на изображении, и на их основе строит своё заключение, что подтверждает высокую точность и надёжность её работы в диагностике данного заболевания.

Высокая вероятность принадлежности снимка к классу “Рак” подтверждает способность модели точно выделять характерные визуальные признаки злокачественного образования. Тем не менее, оставшиеся вероятности (на уровне 10,2 % и 5,0 %) говорят о том, что модель полностью не исключает альтернативные диагнозы, что особенно важно в условиях реальной медицинской практики.

Аналогичные результаты были получены и для десяти других подобных снимков. При этом усреднённые вероятностные показатели



Рис. 9. Пример изображения туберкулеза после его предобработки

составили 82,3 % для рака легкого, 12,4 % для туберкулеза и 5,3 % – для иных снимков.

Другой пример демонстрирует ситуацию, когда модель уверенно отнесла изображение к классу “туберкулез”. Визуальный анализ снимка (рис. 9) показывает наличие плотной округлой тени с четкими контурами и однородной внутренней структурой – именно такие характеристики свойственны туберкуломам, возникающим при локализованных формах туберкулеза легких. При обработке этого изображения система определила наиболее вероятным именно данный класс (рис. 10).

Такая картина вероятностного распределения свидетельствует о высокой чувствительности модели к особенностям, характерным для туберкулезного поражения. При этом, несмотря на доминирующий класс, сохраняется небольшая вероятность принадлежности снимка альтернативным вариантам, что подчеркивает осторожность алгоритма при принятии решений и снижает риск ложноположительных предсказаний. В этом отмечается некоторое отдаленное сходство предложенного решения с объяснимыми системами диагностики.

Ближние результаты были получены и для других подобных снимков. При их общем количестве, равном девяти, усредненные вероятностные показатели по туберкулезу составили 12,3 % для рака легкого, 84,5 % для туберкулеза и 3,2 % – для иных снимков.

В рамках устойчивости и обобщающей способности нейросетевой модели был проведен эксперимент с использованием не КТ, а

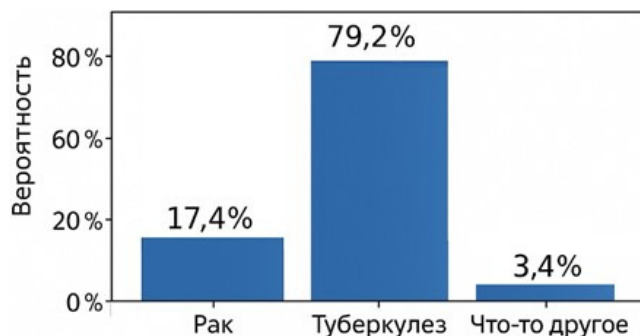


Рис. 10. Распределение вероятностей по диагностическим классам для туберкулеза

рентгеновского снимка грудной клетки, полученного из открытых интернет-источников, на котором визуализируются признаки туберкулезного поражения легких и отчетливо прослеживается замкнутая форма очага, типичная для инкапсулированных воспалительных процессов (рис. 11). Это позволило модели опираться на перечисленные диагностически значимые визуальные признаки и принять решение, близкое к клинической интерпретации (рис. 12).

Несмотря на то, что изображение не входило в обучающую выборку и имело более низкое разрешение, чем изображения, использованные при обучении, модель продемонстрировала способность к корректной интерпретации. Это объясняется тем, что сверточные нейронные сети опираются преимущественно на пространственные и морфологические признаки изображений – такие как контуры, плот-

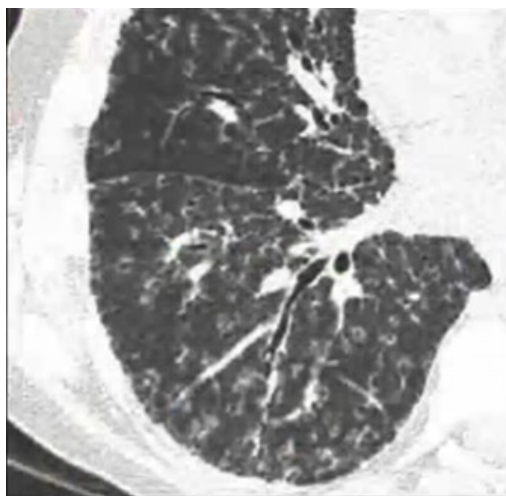


Рис. 11. Взятая из открытых источников рентгенограмма туберкулеза легких

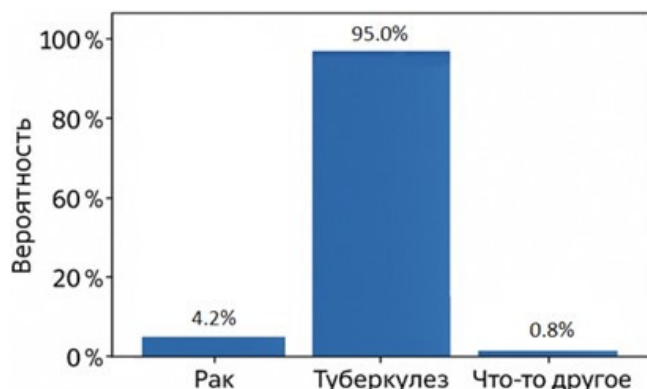


Рис. 12. Распределение по диагностическим категориям для изображения, приведенного на рис. 11

ность, текстура и границы объектов – а не на абсолютные значения пикселей, что делает их устойчивыми к вариациям масштаба и качества изображений.

Обученная модель была натренирована на классе изображений легкого для двух патологических классов – рака лёгкого и туберкулёза и одного класса, не содержащих характерных для данных заболеваний патологических изменений, либо вообще не содержащих патологий. Рак лёгкого и туберкулёза могут визуализироваться схожим образом и оба заболевания, как правило, сопровождаются появлением одиночных или множественных очагов округлой формы (так называемых шаровидных образований) в лёгочной ткани. Однако патоморфологические различия отражаются в специфике рентгенологических проявлений. Так, при туберкулёзе лёгких, особенно в фазе диссеминации, характерно множественное очаговое поражение с преимущественно чёткими границами и гомогенной внутренней структурой, тогда как при злокачественных новообразованиях чаще визуализируются одиночные узлы с неровными или нечеткими контурами, возможными признаками инфильтрации и перифокального воспаления.

Анализ приведенного на рис. 11 изображения проводился на сегментированной области левого лёгочного поля. Нейросетевая модель классифицировала данное изображение с высокой вероятностью как соответствующее туберкулёзу лёгких. Основным критерием, использованным моделью, стало наличие множественных узелков по всей паренхиме лёгкого – типичный паттерн для гематогенно-диссеминированной формы заболевания. Согласно клинико-рентгенологическим данным, такой

характер распространения очагов (мелкие фокусные тени диаметром 1–3 мм с равномерной диссеминацией) наблюдается у более чем 85 % пациентов с генерализованным туберкулёзом.

Была также исследована возможность применения разработанной модели к сегменту рентгеновского изображения, содержащему шарообразное новообразование в лёгких, с клинически подтверждённым диагнозом – немелкоклеточный рак лёгкого (рис. 13). Данное изображение было взято из открытых медицинских источников и не входило в обучающую выборку модели, что позволяет рассматривать его как внешний (вневыборочный) тест.

Для проведения данных исследований архитектура модели была модифицирована: количество выходных классов было ограничено только раком и туберкулёзом, чтобы сфокусировать классификацию на наиболее релевантных для данного эксперимента патологиях и устранить влияние других классов.

Полученное в модели вероятностное распределение по этим двум классам отражено на рис. 14. Модель неверно классифицировала изображение как «туберкулёз» с вероятностью 58 %, в то время как вероятность отнесения к классу «рак» составила 42 %. Такой результат можно объяснить рядом факторов. Во-первых, визуальные признаки, представленные на изображении (в частности, малоконтрастные сероватые контуры вокруг образования), ха-

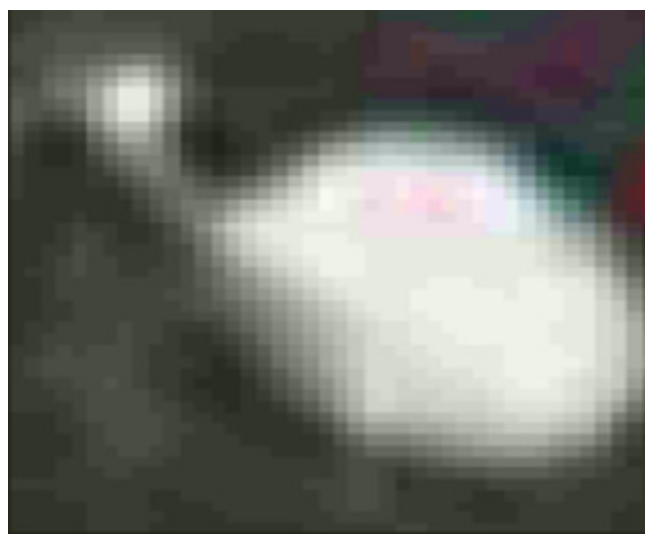


Рис. 13. Сегмент рентгенограммы легких с изображением шарообразной опухоли

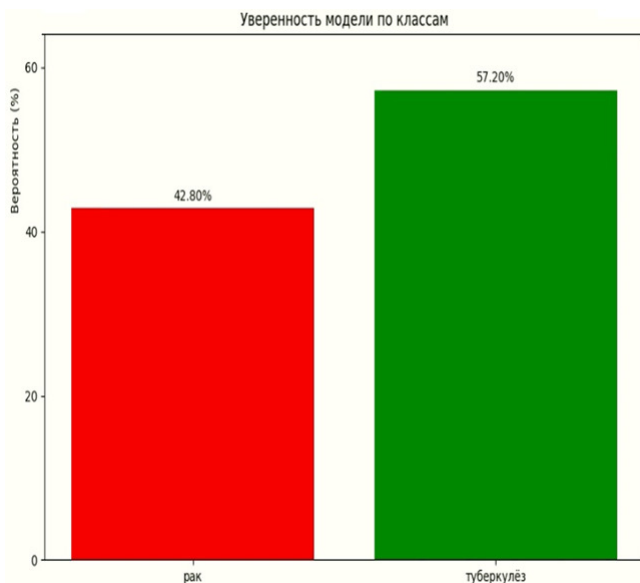


Рис. 14. Распределение по диагностическим категориям рак и туберкулез

рактарны как для инфильтративных форм туберкулёза, так и для ранних стадий злокачественных новообразований. В литературе отмечается, что на ранних стадиях рак лёгких часто имеет форму одиночного узла с умеренно выраженными краями, что затрудняет визуальную дифференциацию. Во-вторых, ошибка во многом была обусловлена малым объёмом и разнообразием обучающей выборки. Ограниченный набор изображений снижает способность модели корректно интерпретировать ранее не встречавшиеся снимки, особенно если они сильно отличаются от изображений, использованных при обучении.

Дополнительно стоит отметить отсутствие модуля локализации очага патологии в текущей архитектуре. Модель анализирует всё изображение целиком, что может вносить в классификацию нерелевантные признаки, находящиеся вне зоны патологического образования. Это снижает точность принятия решения, особенно при наличии артефактов или сопутствующих изменений в лёгочной ткани.

Тем не менее, тот факт, что модель присвоила изображению 42 % вероятности принадлежности к классу “рак”, свидетельствует о том, что она зафиксировала характерные признаки злокачественного процесса, но не смогла окончательно дифференцировать их от признаков туберкулёза. Такой уровень неопределённости (близкий к 50/50) может служить индикатором необходимости привлечения экс-

пертного заключения, а также свидетельствует о потенциале модели к обучению на более широком спектре данных.

Таким образом, проведённый эксперимент демонстрирует как текущие ограничения модели, так и направления для её дальнейшего совершенствования – в частности, расширение обучающей выборки, внедрение механизмов сегментации патологических очагов и использование методов интерпретируемого машинного обучения.

Ограничения модели и перспективы её развития

Несмотря на то, что в большинстве тестов модель однозначно принимала правильное решение, тем не менее, как видно из рис. 14, возможны случаи принятия моделью ложных диагнозов (ложноположительного для туберкулёза и ложноотрицательного для рака легкого). Поэтому было бы логично при ситуациях, когда перепад вероятностей (или долей уверенности) между классами оказывается ниже некоторого заданного порога, модель должна выдавать предупреждение о высокой вероятности ошибочности поставленного диагноза. Отсюда вытекает необходимость проведения дальнейших исследований, направленных на установление такого порога, что, в свою очередь, требует существенного увеличения объема тестирования, причем лучше за счет привлечения датасетов, не задействованных в обучении модели. Стоит также увеличить и объем исходной обучающей выборки за счет автоматизации процесса отбора снимков из имеющихся датасетов. И только после таких исследований и работ, позволяющих составить матрицу несоответствия по всем трем классам, построить соответствующие этим классам ROC-кривые и найти другие используемые в многоклассовой классификации и машинном обучении метрики, можно будет провести объективное сравнение с имеющимися аналогами. И лишь затем можно будет поднимать вопрос о возможности практического применения разработанной модели.

В дальнейшем для развития модели целесообразно будет рассмотреть вопрос о добавлении в нее новых диагностических классов, использовании для диагностики других видов медицинских исследований (например, ПЭТ/КТ или МРТ, МСКТ). Поскольку рентгеновские

снимки существенно отличаются от снимков при других видов обследований, возможно, есть смысл выделить диагностику по рентгеновским снимкам в отдельную нейросетевую модель. Кроме того, в качестве альтернативных изображений лучше подбирать изображения, на которых диагностические признаки подобны признакам туберкулеза и рака легких, что потребует привлечения к работе медиков – специалистов в области диагностики легочных заболеваний.

Поскольку главной задачей данной работы являлась демонстрация преимуществ использования для дифференциальной диагностики трехклассовой модели, была использована простейшая структура сети и восьмибитная градация серого на изображениях, чтобы можно было выполнять исследования при минимальных требованиях к вычислительным ресурсам. Поэтому в перспективе стоит рассмотреть необходимость применения более сложной структуры нейросети и увеличения числа градаций серого для повышения точности прогнозирования. Стоит также добавить в модель модуль сегментации изображений с целью выделения в них значимых для диагностики областей интересов и модуля объяснимости получаемых результатов. Всё это сделает предложенный подход перспективным направлением для внедрения в клиническую практику и дальнейшего научного развития.

Заключение

В результате проведенных исследований была разработана и реализована нейросетевая модель, предназначенная для диагностики таких лёгочных заболеваний, как рак и туберкулез, что было обусловлено их высокой распространённостью и сложностью дифференциальной диагностики по визуальным признакам. Применение методов искусственного интеллекта, в частности свёрточных нейронных сетей с идентификацией изображений по трем классам позволило повысить объективность диагностики и уменьшить влияние человеческого фактора на принятие решения.

Были проанализированы современные подходы к визуальной диагностике, разработан и обучен прототип нейросетевой модели, выполнена аугментация обучающих данных, обеспечивающая устойчивость модели к вариативности медицинских изображений. Для про-

ведения исследований был создан программный комплекс, способный не только классифицировать изображения по диагностическим классам, но и визуализировать их вероятностное распределение. Это является важным достоинством разработанного комплекса, позволяя наглядно видеть, насколько уверенно был поставлен данный диагноз и какое из возможных альтернативных решений преобладает. Кроме того, введение класса для прочих патологий уменьшает “искушение” нейросети отнести изображение к одному из двух основных классов, то есть снижает вероятность получения ложных диагностических заключений. Это делает систему более надежной, прозрачной и понятной для конечного пользователя – врача-рентгенолога.

Экспериментальное тестирование показало высокую точность работы модели. На примерах изображений, не входивших в обучающую выборку, она продемонстрировала уверенное определение патологий с распределением вероятностей, позволяющим делать верные выводы. Особенно важно, что модель сохраняет чувствительность к альтернативным диагнозам, что повышает её надёжность в условиях клинической неопределённости.

Однако, на данном этапе, несмотря на в целом положительный результат исследований, рекомендовать разработанный комплекс в клиническую практику преждевременно в силу крайне ограниченного объема как обучающей, так и тестовой выборки и отсутствия оценки статистических свойств модели. Поэтому только после проведения дополнительных исследований и вышераассмотренных доработок можно рассматривать применение предложенного решения в качестве вспомогательного инструмента при проведении первичной диагностики.

Благодарности

Авторы благодарят бывшего магистранта Алтайского государственного технического университета Шавандина Артёма Александровича за неоценимую помощь в написании программного кода для разработанного комплекса и непосредственного участия в проведении экспериментальных исследований.

Список литературы

- World Health Organization. Chronic obstructive pulmonary disease (COPD) [Internet]. Geneva: World Health Organization; 2023 Jul 25 [cited 2024 Nov 20]. Available from: [https://www.who.int/ru/news-room/factsheets/detail/chronic-obstructive-pulmonary-disease-\(copd\)](https://www.who.int/ru/news-room/factsheets/detail/chronic-obstructive-pulmonary-disease-(copd)).
- Zhou J, Xu Y, Liu J, Feng L, Yu J, Chen D. Global burden of lung cancer in 2022 and projections to 2050: Incidence and mortality estimates from GLOBOCAN. *Cancer Epidemiol*. 2024; 93: 102693. <https://doi.org/10.1016/j.canep.2024.102693>.
- Рак легкого. Под ред. Лактионов КК, Бредер ВВ, М.: ГРАНАТ; 2020. 151 с. Доступно по: https://rsmu.ru/fileadmin/templates/DOC/Faculties/LF/oncology/DO/rak_legkogo_Kniga.pdf.
Lung Cancer. Editors Laktionov KK, Breder VV, Moscow: GRANAT; 2020. 151 p. Available from: https://rsmu.ru/fileadmin/templates/DOC/Faculties/LF/oncology/DO/rak_legkogo_Kniga.pdf (In Russ).
- Goletti D, Meintjes G, Andrade BB, Zumla A, Lee SS. Insights from the 2024 WHO Global Tuberculosis Report – More Comprehensive Action, Innovation, and Investments required for achieving WHO End TB goals. *Int J Infect Dis*. 2025 Jan; 150: 107325. <https://doi.org/10.1016/j.ijid.2024.107325>.
- Chen Z, Wang T, Du J, Liu Y, Fu R, Wang L, et al. Decoding the WHO Global Tuberculosis Report 2024: A Critical Analysis of Global and Chinese Key Data. *Zoonoses (Burlingt)*. 2025; 5 (1). <https://doi.org/10.15212/ZOONOSSES-2024-0061>
- Каращук НП, Киселева МВ. Рак и туберкулез легкого. Научный медицинский вестник Югры. 2014; (5-6): 71-3.
Karashchuk NP, Kiseleva MV. Lung cancer and tuberculosis. *Nauchnyi Meditsinskii Vestnik Yugry*. 2014; (5-6): 71-3. (In Russ).
- Shu CC, Chang SC, Lai YC, Chang CY, Wei YF, Chen CY. Factors for the Early Revision of Misdiagnosed Tuberculosis to Lung Cancer: A Multicenter Study in A Tuberculosis-Prevalent Area. *J Clin Med*. 2019 May 17; 8 (5): 700. <https://doi.org/10.3390/jcm8050700>.
- Новицкая ТА, Ариэль БМ, Двораковская ИВ, Аветисян АО, Яблонский ПК. Морфологические особенности сочетания туберкулеза и рака легких. *Архив патологии*. 2021; 83 (2): 19-24. <https://doi.org/10.17116/patol20218302119>.
- Novitskaya TA, Ariel BM, Dvorakovskaya IV, Avetisyan AO, Yablonsky PK. Morphological features of the combination of tuberculosis and lung cancer. *Pathology Archive*. 2021; 83 (2): 19-24. <https://doi.org/10.17116/patol20218302119> (In Russ).
- Frederick ED. Computer aided diagnosis of acute pulmonary embolism [PhD Thesis]. Durham: Duke University; 2001.
- Коновалов ВК, Горбунова ВА, Таричко ЮВ, Соколова ЕБ. Особенности селективной денситометрии шаровидных образований лёгких при мультиспиральной компьютерной томографии. *Российский онкологический журнал*. 2014; 19 (5): 14-23.
Konovalov VK, Gorbunova VA, Tarichko IuV, Sokolova EB. Features of selective densitometry of spherical lung formations at multispiral computed tomography. *Russian Oncology Journal*. 2014; 19 (5): 14-23. (In Russ).
- Ницын АЮ, Поворозняк АИ, Ницын ДА. Оценка вероятности диагноза по фрактальной размерности медицинского изображения. *Вестник НТУ "ХПИ"*. 2008; (24): 107-12.
Nitsyn Alu, Povorozniuk AI, Nitsyn DA. Estimation of diagnosis probability by fractal dimension of a medical image. *Vestnik NTU "KhPI"*. 2008; (24): 107-12. (In Russ).
- Molodkin IV, Yakunin AG, Shayiduk AM. Method of calculation of diagnostic signs based on the results of densitometry of spherical lung formations. *J Phys Conf Ser*. 2019; 1327: 012045. <https://doi.org/10.1088/1742-6596/1327/1/012045>.
- Molodkin IV, Yakunin AG, Shayduk AM, Konovalov VK, Ostanin SA, Leonov SL, et al. The constitution system of diagnostic characters using outcome analysis of boundary circular-shaped masses in lungs. *Radiol Electron J*. 2019; 9 (3): 79-85. <https://doi.org/10.21569/2222-7415-2019-9-3-79-85>.
- Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging*. 2018 Aug; 9 (4): 611-29. <https://doi.org/10.1007/s13244-018-0639-9>.
- Вишняков ВА, Хэ Т. Сверточная нейронная сеть для ИТ-диагностики легких. *Системный анализ и прикладная информатика*. 2024; (1): 59-64. <https://doi.org/10.21122/2309-4923-2024-1-59-64>.

- Vishnyakov VA, Khe T. Convolutional neural network for CT diagnostics of lungs. *Systems Analysis and Applied Informatics*. 2024; (1): 59-64. <https://doi.org/10.21122/2309-4923-2024-1-59-64> (In Russ).
16. Selivanov A, Rogov OY, Chesakov D, Gusev A, Lukashenko A, Dylov DV. Medical image captioning via generative pretrained transformers. *Sci Rep*. 2023 Mar 13; 13 (1): 4171. <https://doi.org/10.1038/s41598-023-31223-5>.
 17. Гудфеллоу Я, Бенджио И, Курвилль А. Глубокое обучение. 2-е изд. М.: ДМК Пресс; 2018. 652 с.
Goodfellow I, Bengio Y, Courville A. *Glubokoe obuchenie [Deep Learning]*. 2nd ed. Moscow: DMK Press; 2018. 652 p. (In Russ).
 18. Kim HE, Cosa-Linan A, Santhanam N, Janesari M, Maros ME, Ganslandt T. Transfer learning for medical image classification: a literature review. *BMC Med Imaging*. 2022 May 4; 22 (1): 69. <https://doi.org/10.1186/s12880-022-00793-7>.
 19. Уткин АА, Мелдо ОС, Ипатов МА, Рябинин МА. Медицинские интеллектуальные системы на примере диагностики рака легкого. *Известия ЮФУ. Технические науки*. 2018; (8): 241-9. <https://doi.org/10.23683/2311-3103-2018-8-241-249>.
Utkin AA, Meldo OS, Ipatov MA, Riabinin MA. Medical intellectual systems on the example of lung cancer diagnosis. *News of the Southern Federal University. Tekhnicheskie Nauki*. 2018; (8): 241-9. <https://doi.org/10.23683/2311-3103-2018-8-241-249> (In Russ).
 20. UrRehman Z, Qiang Y, Wang L, Asghar W, Chen K. Effective lung nodule detection using deep CNN with dual attention mechanisms. *Sci Rep*. 2024 Feb 17; 14 (1): 3934. <https://doi.org/10.1038/s41598-024-51833-x>.
 21. Uddin J. Attention-Based DenseNet for Lung Cancer Classification Using CT Scan and Histopathological Images. *Designs*. 2024 Apr 2; 8 (2): 27. <https://doi.org/10.3390/designs8020027>.
 22. Hammad M, ElAffendi M, El-Latif AAA, Hussain A. Explainable AI for lung cancer detection via a custom CNN on CT images. *Sci Rep*. 2025 Jul 2; 15 (1): 12707. <https://doi.org/10.1038/s41598-025-97645-2>.
 23. van der Velden BHM, Kuijf HJ, Gilhuijs KGA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med Image Anal*. 2022 Jul; 79: 102470. <https://doi.org/10.1016/j.media.2022.102470>.
 24. Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. *J Big Data*. 2019; 6: 60. <https://doi.org/10.1186/s40537-019-0197-0>.
 25. Goceri E. Medical image data augmentation: techniques, comparisons and interpretations. *Artif Intell Rev*. 2023; 56: 12561-605. <https://doi.org/10.1007/s10462-023-10453-z>.
 26. Salmi M, Atif D, Oliva D, Ettaouil M, Hmimou A. Handling imbalanced medical datasets: review of a decade of research. *Artif Intell Rev*. 2024; 57 (7): 273. <https://doi.org/10.1007/s10462-024-10884-2>.
 27. Bria A, Marrocco C, Tortorella F. Addressing class imbalance in deep learning for small lesion detection on medical images. *Comput Biol Med*. 2020 Jun; 120: 103735. <https://doi.org/10.1016/j.compbiomed.2020.103735>.
 28. Jian W, Haq AU, Afzal N, Alarfaj M, Almusallam N. Developing an innovative lung cancer detection model for accurate diagnosis in AI healthcare systems. *Sci Rep*. 2025 Nov 14; 15 (1): 22945. <https://doi.org/10.1038/s41598-025-03960-2>.
 29. Crasta LJ, Neema R, Pais AR. A novel Deep Learning architecture for lung cancer detection and diagnosis from Computed Tomography image analysis. *Health Anal*. 2024 Jun; 5: 100316. <https://doi.org/10.1016/j.health.2024.100316>.
 30. Li X, Zhou Y, Du P, Wu G, Wang L, Li J, et al. A deep learning system that generates quantitative CT reports for diagnosing pulmonary Tuberculosis. *Appl Intell*. 2021; 51: 4082-93. <https://doi.org/10.1007/s10489-020-02051-1>.
 31. Yan C, Wang L, Lin J, Xu J, Zhang T, Qi J, et al. A fully automatic artificial intelligence-based CT image analysis system for accurate detection, diagnosis, and quantitative severity evaluation of pulmonary tuberculosis. *Eur Radiol*. 2022 Apr; 32 (4): 2188-99. <https://doi.org/10.1007/s00330-021-08365-z>.
 32. Tan H, Bates JHT, Kinsey CM. Discriminating TB lung nodules from early lung cancers using deep learning. *BMC Med Inform Decis Mak*. 2022 Jun 23; 22 (1): 161. <https://doi.org/10.1186/s12911-022-01904-8>.
 33. Zhang G, Shang L, Li S, Zhang X, Li W, Wang X, et al. Non-enhanced CT deep learning model for differentiating lung adenocarcinoma from

- tuberculoma: a multicenter diagnostic study. Eur Radiol. 2025 Jan 29. <https://doi.org/10.1007/s00330-025-11721-y>.
34. Muralidharan V, Adewale BA, Huang CJ, Tran T, Mello-Thoms C, Yi PH. A scoping review of reporting gaps in FDA-approved AI medical devices. npj Digit Med. 2024 Oct 2; 7 (1): 273. <https://doi.org/10.1038/s41746-024-01270-x>.
35. Khunte M, Chae A, Wang R, Jain R, Sun Y, Sollee JR, et al. Trends in clinical validation and usage of US Food and Drug Administration-cleared artificial intelligence algorithms for medical imaging. Clin Radiol. 2023 Feb; 78 (2): 123-9. <https://doi.org/10.1016/j.crad.2022.09.122>.
36. National Cancer Institute. The Cancer Imaging Archive (TCIA) [Internet]. Bethesda, MD: The Cancer Imaging Archive; [cited 2025 Mar 17]. Available from: <https://www.cancerimagingarchive.net>.
37. Prior F, Smith K, Sharma A, Kirby J, Tarbox L, Clark K, et al. The public cancer radiology imaging collections of The Cancer Imaging Archive. Sci Data. 2017 Sep 12; 4: 170124. <https://doi.org/10.1038/sdata.2017.124>.
38. Mader K. CT Tuberculosis (CT-TB) Dataset. Kaggle. 2023 [cited 2025 Mar 17]. Available from: <https://www.kaggle.com/datasets/kmader/ct-tuberculosis>.
39. National Cancer Institute. The Cancer Imaging Archive (TCIA): LIDC-IDRI Collection [Internet]. Bethesda, MD: The Cancer Imaging Archive; [cited 2025 Mar 17]. Available from: <https://www.cancerimagingarchive.net/collection/lidc-idri/>.
40. Wang W, Luo J, Yang X, Lin H. Data analysis of the lung imaging database consortium and image database resource initiative. Acad Radiol. 2015; 22 (4): 488-95. <https://doi.org/10.1016/j.acra.2014.12.004>.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Финансирование. Исследование не имело спонсорской поддержки.

Участие авторов. Статья подготовлена с равным участием авторов.

Поступила: 13.11.2025. **Принята к публикации:** 15.12.2025.

Conflict of interest. The authors declare no conflict of interest.

Financing. The study had no sponsorship.

Contribution. Article was prepared with equal participation of the authors.

Article received: 13.11.2025. **Accepted for publication:** 15.12.2025.