

<https://doi.org/10.29296/25877305-2020-05-07>

Система компьютерного зрения для анализа обзорных рентгенограмм органов грудной клетки: возможности искусственного интеллекта в обнаружении патологических изменений и инородных тел

Е.А. Жуков¹, Д.С. Блинов¹, доктор медицинских наук, В.С. Леонтьев², П.В. Гаврилов³, кандидат медицинских наук, У.А. Смольникова³,

Е.В. Блинова⁴, доктор медицинских наук, профессор,

И.Г. Камышанская⁵, кандидат медицинских наук

¹ООО «КэреМенторЭйАй», Москва

²Городская клиническая больница им. Ф.И. Иноземцева Департамента здравоохранения Москвы

³Санкт-Петербургский НИИ фтизиопульмонологии Минздрава России

⁴Первый МГМУ им. И.М. Сеченова Минздрава России (Сеченовский Университет)

⁵Санкт-Петербургский государственный университет

E-mail: d.blinov@cmai.team

Цель – разработать эффективную систему компьютерного зрения для обнаружения признаков патологии и инородных тел бытового и медицинского происхождения на обзорных рентгенограммах (РГ) органов грудной клетки (ОГК).

Материал и методы. Для построения модели использовался ансамбль сверточных искусственных нейронных сетей архитектур Inception V3, ResNet-50 и Global Average Pooling. Выходы со всех моделей объединялись в один вектор и использовались как вход для модели бустинга, в качестве которой применялась XGBoost модель. Для обучения и тестирования системы использовали 276 840 обезличенных обзорных РГ ОГК в прямой проекции.

Результаты. Был разработан ряд моделей компьютерного зрения для анализа рентгенологических исследований легких. Для достижения удовлетворительного баланса между показателями точности предсказания эмпирическим путем был подобран порог принятия решения, равный 0,4. Такой баланс позволяет добиться снижения числа ложноотрицательных прогнозов модели и увеличения количества случаев, где подозревается наличие патологических изменений.

Выводы. Разработанная модель компьютерного зрения может рассматриваться как эффективный ассистент рентгенолога при анализе РГ-снимков ОГК, позволяющая формировать список приоритетных изображений для немедленного и отсроченного анализа и описания.

Ключевые слова: нейронная сеть, искусственный интеллект, патология легких, инородное тело, обнаружение.

Для цитирования: Жуков Е.А., Блинов Д.С., Леонтьев В.С. и др. Система компьютерного зрения для анализа обзорных рентгенограмм органов грудной клетки: возможности искусственного интеллекта в обнаружении патологических изменений и инородных тел. Врач. 2020; 31 (5): 34–41. <https://doi.org/10.29296/25877305-2020-05-07>

На долю рентгенологического (РГ) метода исследования приходится до 45% всех диагностических исследований в мире, что ставит его на первое место по распространенности [1]. Широкая доступность РГ-диагностики обусловлена ее относительно невысокой стоимостью, хорошей обеспеченностью соответствующим оборудованием и большим диагностическим потенциалом в отношении таких социально значимых заболеваний, как рак и туберкулез легких, пневмония и др. [2]. Вместе с тем двумерные рентгеновские изображения являются и неисчерпаемым источником информации неоднозначности, обусловленной наложением теней различных по своей анатомической и гистологической структуре, составу и плотности органов и тканей. Как результат – рентгеновский снимок может содержать десятки визуальных паттернов, присущих сотням патологических процессов и состояний [2]. Все это приводит к вполне объяснимым сложностям при прочтении и интерпретации РГ-изображения, возникновению расхождений экспертных оценок и зачастую необоснованному назначению дополнительных лабораторно-инструментальных исследований.

В настоящее время врачебные ошибки наряду с сердечно-сосудистыми и онкологическими заболеваниями входят в тройку лидирующих причин смерти населения на планете [3]. По данным различных источников, только в США ежегодно насчитывается от 44 до 400 тыс. смертей вследствие врачебных ошибок [4, 5], что по подсчетам экспертов ВОЗ, обходится бюджетам всех уровней от 17 до 29 млрд долларов в год. При этом частота ложноотрицательных результатов при анализе РГ органов грудной клетки (ОГК) в развитых странах Запада составляет около 4% [6–8], тогда как частота выявления отдельных рентгенологических феноменов не опускается ниже 30%, начиная с 1949 г., когда Garland опубликовал результаты своего первого наблюдательного исследования [9].

Последние годы были ознаменованы прорывными решениями в области машинной обработки медицинских данных, в том числе диагностических изображений. Различные модели применялись для диагностики заболеваний дыхательной системы инфекционной и злокачественной природы [10–12]. Вместе с тем прогностическая точность предлагаемых методов уступает точности опытного диагноста [13, 14]. Мы полагаем, что технологические возможности машинного обучения могут и должны применяться не вместо, а вместе с врачом, использоваться в качестве эффективных ассистентов, избавляющих специалистов от рутинных процедур, не требующих высокой врачебной квалификации. Такие возможности открывает перед современной медициной использование сверточных нейронных сетей [15, 16], на базе которых становится возможным создание эффективных систем компьютерного зрения.

Целью настоящего исследования явилась разработка эффективной системы компьютерного зрения для обна-

ружения признаков патологии и инородных тел бытового и медицинского происхождения на обзорных РГ ОГК.

МАТЕРИАЛ И МЕТОДЫ

Источники получения рентгенологических изображений, обучающая разметка. Протокол исследования прошел независимую этическую экспертизу в Этическом комитете СПбНИИФ Минздрава России (Протокол от 23.03.2019 19/10-534) и Локальном этическом комитете СПбГУ (Протокол от 17.02.2019 №15561). Добровольное информированное согласие на использование рентгенологических изображений в научных целях было получено у каждого индивида, чей снимок использовался для обучения или тестирования системы компьютерного зрения.

В работе использовали 276 840 обезличенных обзорных РГ ОГК в прямой проекции. Часть из них, включавшая 112 120 цифровых изображений, была размещена в открытом доступе в виде базы данных National Institutes of Health Clinical Center (США) [17]. Большая часть изображений была любезно предоставлена архивами СПбГУ, СПбНИИФ Минздрава России и ГКБ им. Ф.И. Иноземцева Департамента здравоохранения Москвы. В табл. 1 представлено распределение цифровых ресурсов в зависимости от источника.

Все полученные изображения случайным образом были распределены в соотношении 7:3 для формирования обучающих и тестовых совокупностей. Каждая рентгенограмма, использованная для обучения нейронной сети, была размечена одним из пяти высококвалифицированных рентгенологов, участвовавших в исследовании, в специально разработанной для этого программной среде.

Для формирования тестового data-set каждое изображение размечалось независимо двумя специалистами, что позволяло при анализе точности работы системы компьютерного зрения осуществлять корректное сравнение.

АЛГОРИТМ СИСТЕМЫ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Описание алгоритма. Блок-схема предлагаемого алгоритма представлена на рис. 1. Как правило, разработ-

ка моделей машинного и глубокого обучения начинается с предобработки исходных данных. Предобработка данных может отличаться для различных этапов разработки модели, таких как обучение модели и ее оценка. Например, изменение исходного размера изображения (приведение изображений к одному размеру) применяется как на этапе обучения, так и на этапе оценки модели, но на этапах обучения также могут потребоваться дополнительные преобразования изображений (так называемые аугментации) с целью увеличить разнообразие обучающего набора данных, тем самым сделать модель более устойчивой при работе с новыми данными.

Стоит отметить, что процесс предобработки данных также включает в себя разбиение исходного множества данных на обучающий и тестовый подвыборки.

После того, как сформированы и обработаны обучающая и тестовая выборки, наступает этап непосредственно обучения модели. Для начала необходимо обучить набор нейросетевых моделей. Затем обученные нейросетевые модели используются в обучении бустинг-модели. И, наконец, ответы от нейросетевых моделей и модели бустинга используются для формирования итогового ответа.

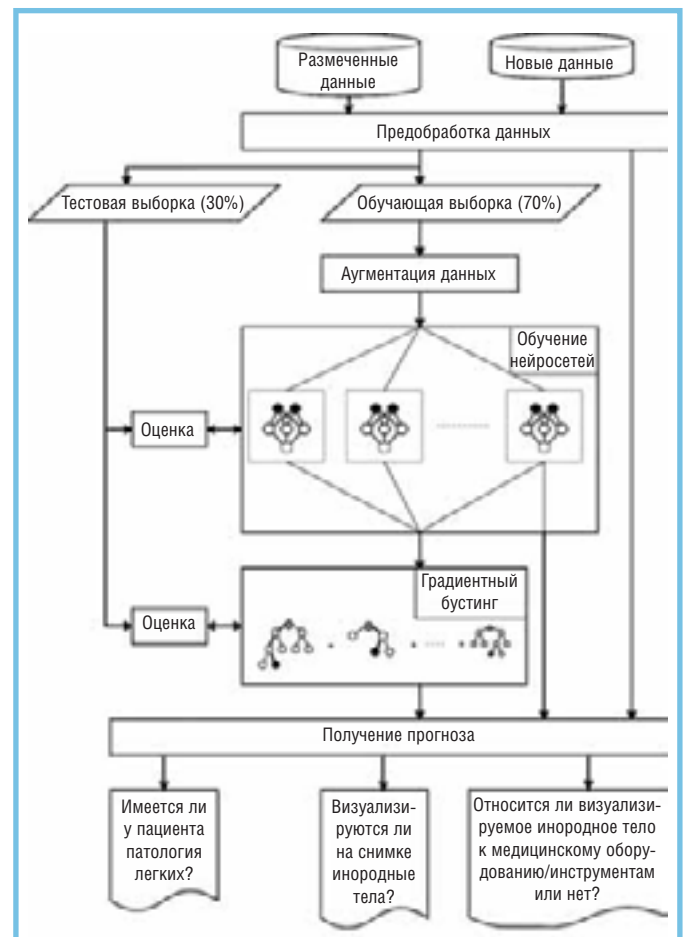


Рис. 1. Описание процесса обучения и применения модели
Fig. 1. Description of the learning process and application of the model

Таблица 1

Источники архивных данных, использованных для обучения и тестирования модели компьютерного зрения

Table 1

Sources of archival data used for training and testing a computer vision model

Источник изображений	Количество снимков	Количество размеченных изображений
СПбГУ	77 439	77 439
СПбНИИФ	53 514	5689
ГКБ им. Ф.И. Иноземцева	33 767	10 096

Предобработка данных. Все нейросетевые модели обучались независимо друг от друга. Таким образом, процесс обучения каждой модели также может разли-

чаться. Так, различные группы моделей обучались на изображениях размерами 224×224; 299×299; 512×512 (табл. 2).

Таблица 2

Table 2

Ансамбль нейросетевых моделей			
Ensemble of neural network models			
№	Назначение нейросети	Используемая пред-обработка данных	Архитектура модели
1	Бинарная классификация: • патологические изменения в проекции легких не визуализируются; • есть вероятность наличия у пациента каких-либо патологических изменений легких	Размер: 224×224 Нормализация: inpt = img / 255.0	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Полносвязный слой (1024 нейрона и ReLU в качестве функции активации); Полносвязный слой (1 нейрон и сигмоида в качестве функции активации)
2	Бинарная классификация: Положение пациента: • стоя • лежа	Размер: 224×224 Нормализация: inpt = img / 255.0	Основная архитектура: ResNet-50 до полносвязного слоя и без pooling-слоя на выходе; «Выпрямляющий» слой – для того, чтобы при классификации использовать весь выход с базовой архитектуры без pooling-слоя; Полносвязный слой (512 нейронов и ReLU в качестве функции активации); Полносвязный слой (1 нейрон и сигмоида в качестве функции активации)
3	Бинарная классификация: Качество рентгеновского снимка: • удовлетворительное • неудовлетворительное	Размер: 512×512 Нормализация: inpt = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (2 нейрона и softmax в качестве функции активации)
4	Бинарная классификация: Визуализируются ли инородные тела: • нет • да	Размер: 512×512 Нормализация: inpt = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (2 нейрона и softmax в качестве функции активации)
5	Бинарная классификация: Имеются ли патологические изменения плевральной полости: • нет • да	Размер: 512×512 Нормализация: inpt = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (2 нейрона и softmax в качестве функции активации)
6	Бинарная классификация: Имеются ли патологические изменения в проекции легочных полей: • нет • да	Размер: 512×512 Нормализация: inpt = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (2 нейрона и softmax в качестве функции активации)
7	Многоклассовая классификация: Наличие патологических изменений аорты и их тип: • патология не визуализируется • развернута дуга аорты • аорта расширена • аорта склерозирована • аорта уплотнена	Размер: 512×512 Нормализация: inpt = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (5 нейронов и softmax в качестве функции активации)
8	Многоклассовая классификация: Изменения диафрагмы: • не выявлены • границы четко не визуализируются • изменены	Размер: 512×512 Нормализация: inpt = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (3 нейрона и softmax в качестве функции активации)
9	Многоклассовая классификация: Патологии сердца: • не выявлены • границы четко не визуализируются • границы поперечно расширены • смещение границы влево	Размер: 512×512 Нормализация: inpt = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (4 нейрона и softmax в качестве функции активации)
10	Многоклассовая классификация: Патологические изменения корней легких: • не выявлены • нечеткие • неструктурные • расширены	Размер: 512×512 Нормализация: inpt = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (4 нейрона и softmax в качестве функции активации)

Продолжение табл. 2
Continuing of Table 2

№	Назначение нейросети	Используемая пред-обработка данных	Архитектура модели
11	Бинарная классификация: Визуализируется ли плевральные спайки: • нет • да	Размер: 512×512 Нормализация: input = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (2 нейрона и softmax в качестве функции активации)
12	Бинарная классификация: Визуализируются ли изменения рисунка легких: • нет • да	Размер: 512×512 Нормализация: input = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (2 нейрона и softmax в качестве функции активации)
13	Бинарная классификация: Обнаружена ли у пациента инфильтрация: • нет • да	Размер: 512×512 Нормализация: input = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (2 нейрона и softmax в качестве функции активации)
14	Многоклассовая классификация: Выявлены ли следующие патологические процессы: • изменений не выявлено • пневмосклероз • пневмоторакс • эмфизема • пневмофиброз • гидроторакс	Размер: 512×512 Нормализация: input = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (6 нейронов и softmax в качестве функции активации)
15	Бинарная классификация: Визуализируются ли очаговые изменения в проекции легочных полей: • нет • да	Размер: 512×512 Нормализация: input = img / 255.0	Основная архитектура: InceptionV3 до полносвязного слоя и без pooling-слоя на выходе; Dropout с вероятностью 0,4; Полносвязный слой (1024 нейрона и ReLU в качестве функции активации); Dropout с вероятностью 0,4; Полносвязный слой (512 нейрона и ReLU в качестве функции активации); Dropout с вероятностью 0,3; Полносвязный слой (1 нейрона и сигмоида в качестве функции активации)
16	Бинарная классификация: Относится ли инородное тело к медицинскому оборудованию/инструменту: • нет • да	Размер: 512×512 Нормализация: input = img / 127.5 – 1	Основная архитектура: InceptionV3 до полносвязного слоя и с GlobalAveragePooling на выходе; Dropout с вероятностью 0,5; Полносвязный слой (2 нейрона и softmax в качестве функции активации)

Также применяется процедура нормирования изображения. Так, в зависимости от модели, значения пикселей изображений нормируются на диапазон [-1, 1] или на [0, 1]. Таким образом, используется два способа нормализации (1):

$$\text{input} = \text{source_image} / 127.5 - 1 \text{ (для } [-1, 1]); \text{ (1)}$$

$$\text{input} = \text{source_image} / 255.0 \text{ (для } [0, 1]),$$

где input – это нормированное изображение, а source_image – исходное.

Входные изображения могут иметь различное исходное разрешение, уровни яркости и контраста, степень детализации и шума. Некоторые изображения могут быть предварительно сконвертированы в grayscale-формат с расширением *.png или *.jpg, некоторые же представлены в исходном DICOM-формате и нуждаются в дополнительном шаге, заключающемся в их конвертации. Примеры исходных изображений представлены на рис. 2.

В то время как изображения требуют изменения исходного размера при обучении различных моделей, ряд этапов предобработки будет одинаковым для всех мо-

делей. Это касается именно аугментации данных – смещение по вертикали или горизонтали, поворот, искажения по типу аффинного преобразования, случайные изменения яркости и контраста, наложение шума. Аугментация необходима для повышения многообразия исходных данных в процессе обучения.

Как указывалось ранее, набор данных также подвергается разбиению на обучающую и тестовую подвыборки в отношении 70:30%.

Архитектура модели. Предлагаемый подход заключается в применении ансамбля нейросетевых моделей совместно с моделью градиентного бустинга, обученного на выходах с этого ансамбля. Полный перечень используемых нейросетевых архитектур и их особенностей представлен в табл. 2.

Выходы со всех моделей объединяются в один вектор и используются как вход для модели бустинга, в качестве которой применяется модель XGBoost.

Процесс обучения модели состоит из 2 этапов:

- независимое обучение каждой нейросетевой модели из ансамбля;
- обучение модели бустинга.

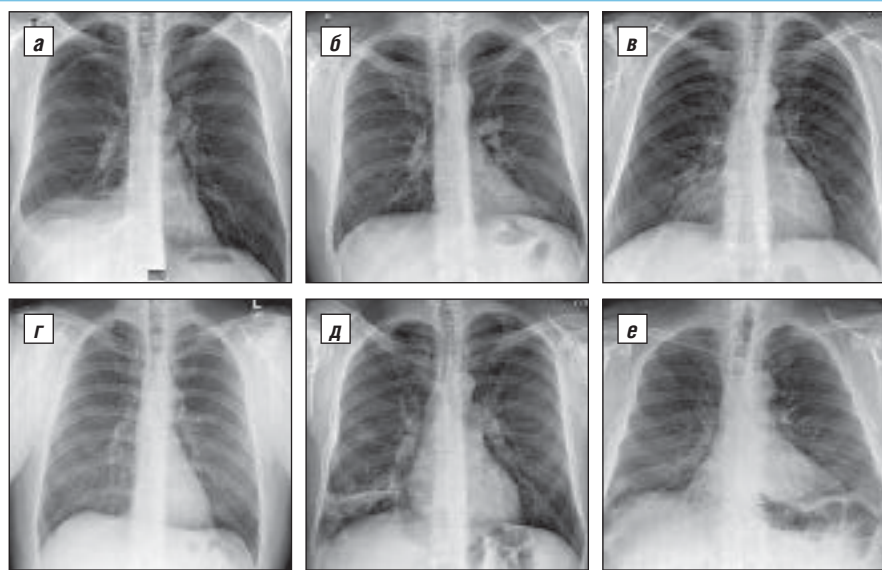


Рис. 2. Примеры исходных данных: *а* – гидроторакс и пневмоторакс; *б* – деформация легкого; *в* – пневмония; *г* – без патологий; *д* – ателектаз легких; *е* – инородные тела
Fig. 2. Examples of input data: *a* – hydrothorax and pneumothorax; *б* – lung deformation; *в* – pneumonia; *г* – without pathologies; *д* – atelectasis of the lungs; *е* – foreign bodies

Выходом модели является степень уверенности, что у пациента имеется патология. Параметры модели подбирались с применением кросс-валидации. В результате были получены следующие параметры: $\text{max_depth}=7$ (максимальная глубина дерева) и $\text{eta}=0,1$ (шаг обучения). Остальные параметры использованы по умолчанию. Больше информации о параметрах XGBoost модели можно узнать в документации [18].

Постобработка. Выходом модели градиентного бустинга является значение из диапазона $[0, 1]$, где 0 – легочных патологий не выявлено, 1 – вероятно наличие патологии легких. Для того чтобы принять итоговое решение по прогнозу модели, необходимо установить специальное пороговое значение. Если выход модели превышает такое пороговое значение, то делается вывод о том, что имеется легочная патология. В рамках данной

Каждая модель требует для обучения свою собственную размеченную подвыборку данных. Все нейросети обучались с применением оптимизации Адама с параметрами $\beta_1=0,9$ и $\beta_2=0,999$. Начальный шаг обучения был равен 0,001 и уменьшался в 10 раз, если функция потерь переставала уменьшаться на протяжении 10 эпох. Обучение прекращалось, если функция потерь переставала уменьшаться на протяжении 15 эпох.

В качестве функции потерь использовалась бинарная кросс-энтропия (2), если активационной функцией в классифицирующем слое является сигмоида, и категориальная кросс-энтропия (3) для тех сетей, где в качестве активационной функции классифицирующего слоя используется softmax:

$$\text{loss}_{\text{binary_crossentropy}} = - [y \cdot \log(p(y)) + (1 - y) \cdot \log(1 - p(y))], \quad (2)$$

где y – истинное значение класса (положительный или отрицательный); $p(y)$ – ответ нейросети (значение из диапазона $[0; 1]$ – вероятность положительного значения класса).

$$\text{loss}_{\text{categorical_crossentropy}} = - \sum_{c=1}^C y_c \cdot \log(p(y_c)), \quad (3)$$

где $y_c = 1$, если изображение имеет класс c , 0 – в противном случае; $p(y_c)$ – ответ нейросети – вероятность того, что изображение относится к классу c .

XGBoost модель используется в качестве модели градиентного бустинга. В качестве входа для обучения модели используется 38-мерный вектор, сформированный из выходов ансамбля нейросетевых моделей.

работы порог подбирается таким образом, чтобы подобрать удовлетворительный баланс между чувствительностью и специфичностью. В результате был выбран порог, равный 0,4.

Оценка качества. Для оценки качества модели использовались следующие метрики: ROC AUC, Recall, Precision.

В задачах бинарной классификации precision (точность) (4) – это доля объектов, для которых класс предсказан верно, относительно всего количества объектов, для которых предсказан данный класс. Recall (5) (также известна, как чувствительность) – это доля объектов, для которых класс предсказан верно, относительно всего количества объектов с данным классом:

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4),$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5),$$

где TP – количество объектов, для которых положительный класс предсказан корректно; FP – количество объектов, для которых положительный класс предсказан некорректно; FN – количество объектов, для которых отрицательный класс предсказан некорректно.

Если модель оценивается относительно класса «патология», то класс «патология» рассматривается как «положительный». В противном случае «положительным» классом является класс «отсутствие патологии», если модель оценивается относительно класса «отсутствие патологии».

В процессе настройки модели ключевое внимание уделялось значению recall относительно «патологического» класса и значения precision относительно «непатологического» класса в целях уменьшить количество ложноотрицательных прогнозов.

AUC ROC – площадь под ROC (Receiver Operating Characteristics) кривой – еще одна метрика, часто используемая при оценке моделей бинарной классификации. В общем виде значение метрики отражает способность модели различать «положительный» и «отрицательный» классы. Чем выше значение AUC ROC, тем лучше модель определяет «положительный» класс как «положительный», а «отрицательный» – как «отрицательный». Расчет метрики AUC ROC состоит из 2 этапов:

- расчет True Positive Rate (TPR, доля истинно положительных прогнозов) и False Positive Rate (FPR, доля ложноположительных прогнозов) при различных значениях порога (от 0 до 1) – расчет координат ROC;
- расчет площади фигуры, образованной ROC-кривой и осями X и Y (FPR и TPR).

Также анализ ROC-кривой помогает подобрать порог для достижения желаемого баланса между FPR и TPR.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В результате обучения моделей были достигнуты следующие показатели метрик качества – табл. 3.

В табл. 3 представлены показатели метрик качества, рассчитанные с учетом порога 0,5 для моделей для инородных тел, и 0,4 – для модели градиентного бустинга. Для модели градиентного бустинга считается, что

пациент имеет патологию, если выход модели $>0,4$. На рис. 3 представлены графики ROC-кривых для предлагаемых моделей.

В рамках исследования был разработан ряд моделей компьютерного зрения для анализа рентгенологических исследований легких. Для достижения удовлетворительного баланса между recall для «положительного» класса и precision – для «отрицательного» эмпирическим путем был подобран порог принятия решения, равный 0,4. Такой баланс позволяет добиться снижения числа ложноотрицательных прогнозов модели и увеличение количества случаев, где подозревается наличие патологических изменений. Несмотря на то, что такая балансировка приводит к увеличению ложноположи-

Значения метрик для предлагаемых моделей

Таблица 3

Table 3

Metric Values for Suggested Models

Метрика	Модель		
	Бинарная модель (наличие/отсутствие патологических изменений) – градиентный бустинг. Положительный: изменения имеются. Отрицательный: изменения отсутствуют	Бинарная модель – имеются ли инородные тела. Положительный: инородные тела визуализируются. Отрицательный: инородные тела не визуализируются	Тип инородного тела – бинарная классификация. Положительный: визуализируется инородное тело медицинского происхождения (инструменты, кардиостимуляторы и т.д.). Отрицательный: визуализируется инородное тело немедицинского происхождения (например, пуговицы, цепочки, и т.д.)
Recall (для «положительного»)	0,7912	0,9489	0,9278
Recall (для «отрицательного»)	0,7411	0,9374	0,9231
Precision (для «положительного»)	0,6777	0,969	0,9570
Precision (для «отрицательного»)	0,8376	0,8989	0,8898
AUC ROC	0,7662	0,9432	0,9377

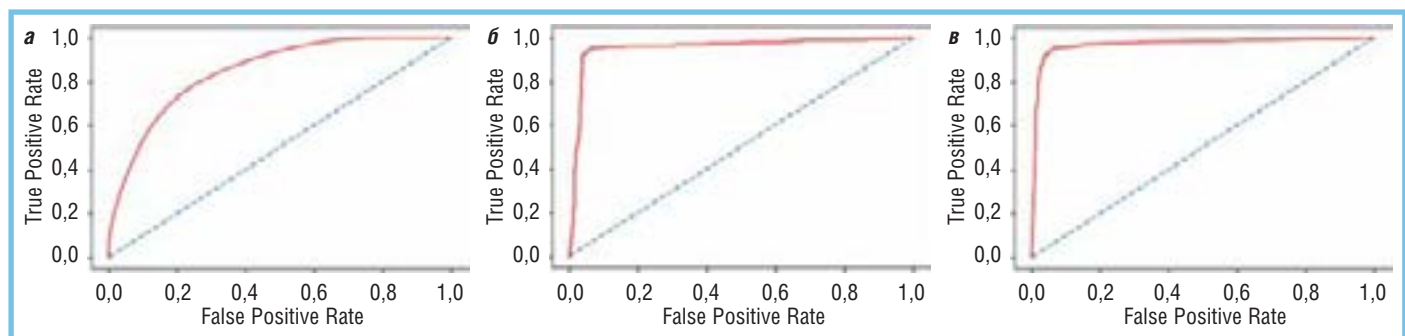


Рис. 3. ROC-кривые для предлагаемых моделей: а – бинарная модель (наличие/отсутствие патологии) – градиентный бустинг; б – бинарная модель – имеются ли инородные тела; в – тип инородного тела – бинарная классификация

Fig. 3. ROC curves for the proposed models: а – binary model (presence / absence of pathology) – gradient boosting; б – binary model – whether there are foreign bodies; в – foreign body type – binary classification

тельных прогнозов (гипердиагностики), данный подход все равно позволяет избежать случаев, когда пациент с «патологией» ошибочно классифицируется моделью как «здоровый».

Стоит отметить две модели из ансамбля нейросетей, которые используются для обнаружения и классификации инородных тел. Данные модели показали неплохие результаты согласно рассчитанным метрикам. Слабым местом моделей оказалась метрика precision для «отрицательного» класса, но для данного класса задач это не является критичным показателем.

Оценивая качество моделей с точки зрения формы характеристической кривой, мы наблюдаем высокое качество моделей для обнаружения и классификации инородных тел. Более сглаженная ROC-кривая для модели по определению факта наличия изменений легочных полей показывает не такие превосходные результаты.

В качестве визуализации логики работы нейросетей на рис. 4 приводится пример карты активации, построенной для модели по обнаружению инородных тел.

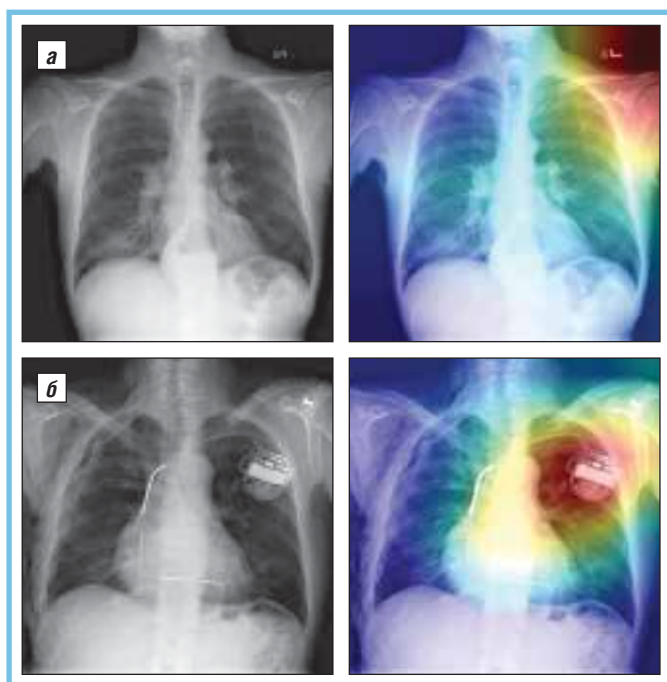


Рис. 4. Пример карт активации, построенных на случайных рентгеновских исследованиях, на которых присутствуют и отсутствуют инородные тела: *a* – можно заметить «повышенное внимание» нейросети на символе «L», но для модели это не повод сделать заключение о том, что присутствует инородное тело; *б* – повышенные значения карты активации можно наблюдать именно в области расположения инородного тела

Fig. 4. An example of activation maps built on random X-ray studies in which foreign bodies are present and absent: *a* – we may notice «increased attention» to the neural network on the «L» symbol, but for the model, this is not a reason to conclude that a foreign body is present; *б* – here we can observe increased values of the activation map precisely in the area of the foreign body location

Как можно заметить на примерах, нейросетевые модели компьютерного зрения «проявляют большую активность» на тех сегментах легких, в которых может наблюдаться отклонение от нормального состояния, т.е. присутствие чего-то иного, чего нет в проекции легочных полей, размеченных как легкие здорового человека.

Для анализа наличия инородных тел и определения их типа для начала применяется нейросетевая модель, которая определяет, имеется ли какое-либо инородное тело на снимке, а затем при положительном ответе на предыдущем этапе другая нейросетевая модель определяет природу выявленного инородного тела – является ли это предметом медицинского обследования, либо нет.

Таким образом, разработанная модель компьютерного зрения может рассматриваться как эффективный ассистент рентгенолога при анализе РГ ОГК, позволяющая формировать список приоритетных изображений для немедленного и отсроченного анализа и описания.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Признательность. Авторы чрезвычайно признательны коллегам, врачам-рентгенологам, оказавшим неоценимую помощь при осуществлении разметки рентгенологических изображений, а также техническим специалистам, поддерживавшим бесперебойное функционирование всех сервисов разрабатываемой системы.

Литература/Reference

1. Yao L., Poblenz E., Dagunts D. et al. Learning to diagnose from scratch by exploiting dependencies among labels. arXiv:1710.10501v2 [cs.CV]. 2018; 1: 1–18.
2. Sabih D.E., Sabih A., Sabih Q. et al. Image perception and interpretation of abnormalities; can we believe our eyes? Can we do something about it? *Insights Imaging*. 2011; 2: 47–55. <https://doi.org/10.1007/s13244-010-0048-1>
3. Makary M.A., Daniel M. Medical error: the third leading cause of death in the US. *BMJ*. 2016; 353: i2139. <https://doi.org/10.1136/bmj.i2139>
4. Kohn L.T., Corrigan J.M., Donaldson M.S. et al. To err is human: building a safer health system. Washington, DC: National Academies Press, 2000. pp. 287. <https://doi.org/10.17226/9728>
5. Busby L.P., Courtier J.L., Glastonbury C.M. Bias in radiology: the How and Why of misses and misinterpretation. *Radiographics*. 2018; 38: 236–47. <https://doi.org/10.1148/rg.2018170107>
6. Waite S., Scott J., Gale J. et al. Interpretative error in radiology. *AJR*. 2017; 208: 739–49. <https://doi.org/10.2214/ajr.16.16963>
7. Ropp A., Waite S., Reede D. et al. Did I miss that: subtle and commonly missed findings on chest radiographs. *Curr Probl Diagn Radiol*. 2015; 44: 277–89. <https://doi.org/10.1067/j.cpradiol.2014.09.003>
8. Del Ciello A., Franchi D., Contegiacomo A. et al. Missed lung cancer: when, where, and why? *Diagn Interv Radiol*. 2017; 23 (2): 118–26. <https://doi.org/10.5152/dir.2016.16187>
9. Garland L.H. On the scientific evaluation of diagnostic procedures. *Radiology*. 1949; 52: 309–28. <https://doi.org/10.1148/52.3.30910>
10. Er O., Yumusak N., Temurtas F. Chest diseases diagnosis using artificial neural networks. *Expert Sys Appl*. 2010; 37 (12): 7648–55. <https://doi.org/10.1016/j.eswa.2010.04.078>
11. Er O., Sertkaya C., Temurtas F. et al. A comparative study on chronic obstructive pulmonary and pneumonia diseases diagnosis using neural networks and artificial immune system. *J. Med. Sys*. 2009; 33 (6): 485–92. <https://doi.org/10.1007/s10916-008-9209-x>

12. Khobragade S., Tiwari A., Pati C.Y. et al. Automatic detection of major lung diseases using chest radiographs and classification by feed-forward artificial neural network. Proceedings of 1st IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES-2016) 2016 IEEE, p. 1–5. <https://doi.org/10.1109/icpeices.2016.7853683>

13. Litjens G., Kooi T., Bejnordi E.B. et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*. 2017; 42: 60–88. <https://doi.org/10.1016/j.media.2017.07.005>

14. Albarqouni S., Baur C., Achilles F. et al. Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE Transactions on Medical Imaging*. 2016; 35 (5): 1313–21. <https://doi.org/10.1109/tmi.2016.2528120>

15. Avendi M.R., Kheradvar A., Jafarkhani H.A. combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Medical Image Analysis*. 2016; 30: 108–19. <https://doi.org/10.1016/j.media.2016.01.005>

16. Shin H.-C., Roberts K., Lu L. et al. Learning to read chest X-rays: recurrent neural cascade model for automated image annotation. *Cornel University library*. 2016. <https://arxiv.org/abs/1603.08486>.

17. Wang X.S., Peng Y.F., Lu L. et al. ChestX-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). Honolulu, HI, USA: IEEE. 2017. p. 3462–71. <http://dx.doi.org/10.1109/CVPR.2017.369>

18. XGBoost Documentation, 2019, accessed 29 September 2019, <https://xgboost.readthedocs.io/en/latest/index.html>

SYSTEM OF DIGITAL VISION FOR X-RAY LUNG PATHOLOGY AND FOREIGN BODY DETECTION

E. Zhukov¹; D. Blinov¹, MD; V. Leontiev²; P. Gavrilov³, Candidate of Medical Sciences; U. Smolnikova³; Professor E. Blinova⁴, MD; I. Kamishanskaya⁵, Candidate of Medical Sciences

¹Care Mentor AI, Moscow

²Inozemtzev Moscow State Clinical Hospital, Moscow

³Saint-Petersburg Research Institute of Phthisiopulmonology

⁴I.M. Sechenov First Moscow State Medical University (Sechenov University)

⁵Saint-Petersburg State University

Goal and objectives: to develop an effective computer vision system for detecting pathology and foreign bodies of medical and unmedical origin on plain chest radiographs.

Material and methods: in order to build the model, aggregation of convolutional artificial neural networks of the InceptionV3, ResNet-50 and GlobalAveragePooling architectures was used. The outputs from all models were combined into a single vector and used as input for the boosting model, which was used as the XGBoost model. For training and testing the system, 276840 anonymized chest x-ray in a frontal view were used.

Results. A number of computer vision models have been developed for the analysis of X-ray examinations of the lungs. To achieve a satisfactory balance between the prediction accuracy indicators, a decision threshold of 0.4 was chosen empirically. Such a balance makes it possible to reduce the number of false-negative model predictions and increase the number of cases where pathological changes are suspected.

Conclusions. The developed model of computer vision can be considered as an effective assistant to the radiologists in the analysis of chest x-ray images, allowing them to create a list of priority images for immediate and delayed analysis and description.

Key words: neural network, artificial intelligence, lung pathology, foreign body, detection.

For citation: Zhukov E., Blinov D., Leontiev V. et al. System of Digital Vision for X-Ray Lung Pathology and Foreign Body Detection. *Vrach*. 2020; 31 (5): 34–41. <https://doi.org/10.29296/25877305-2020-05-07>

Об авторах/About the Authors: Zhukov E.A. ORCID: 0000-0003-0965-5268; Blinov D.S. ORCID: 0000-0002-8385-4356; Leontiev V.S. ORCID: 0000-0001-7527-3861; Gavrilov P.V. ORCID: 0000-0003-3251-4084; Smolnikova U.A. ORCID: 0000-0001-9568-3577; Blinova E.V. ORCID: 0000-0003-0050-0251; Kamishanskaya I.G. ORCID: 0000-0002-8351-9216