

© 2022 г. А.С. МАРКОВ (to.asmarkov@gmail.com),
Е.Ю. КОТЛЯРОВ (tyztot@gmail.com),
Н.П. АНОСОВА (anosova-np@rudn.ru),
В.А. ПОПОВ, канд. физ.-мат. наук (popov-va@rudn.ru)
(Российский университет дружбы народов, Москва),
Я.М. КАРАНДАШЕВ, канд. физ.-мат. наук (karandashev@niisi.ras.ru)
(Российский университет дружбы народов, Москва;
Научно-исследовательский институт
системных исследований РАН, Москва),
Д.Е. АПУШКИНСКАЯ, д-р физ.-мат. наук (apushkinskaya@gmail.com)
(Российский университет дружбы народов, Москва)

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ВЫЯВЛЕНИЯ АНОМАЛИЙ НА РЕНТГЕНОВСКИХ СНИМКАХ, ПОЛУЧЕННЫХ НА СКАНЕРАХ ПЕРСОНАЛЬНОГО ДОСМОТРА

В данной работе рассматривается решение задачи выявления аномалий на рентгеновских снимках, полученных сканерами персонального досмотра (СПД). В работе описана последовательность и описание методов предобработки изображений, с помощью которых оригинальные изображения, полученные на СПД, преобразуются к изображениям с визуально различимыми аномалиями. Приведены примеры обработанных снимков. Показаны первые (предварительные) результаты использования нейронной сети для выделения аномалий.

Ключевые слова: сканеры персонального досмотра, рентгеновские снимки, выделение аномалий, выравнивание гистограммы изображения, нейронные сети, U-2-Net.

DOI: 10.31857/S0005231022100038, **EDN:** AJXBSP

1. Введение

На объектах, требующих повышенного контроля безопасности, часто используются сканеры персонального досмотра. Они позволяют быстро сделать снимок человека в рентгеновском диапазоне излучения, на котором оператор сканера персонального досмотра (СПД) может увидеть все объекты на теле человека и визуально подтвердить или опровергнуть наличие запрещенных среди них.

Процесс обладает рядом существенных недостатков, в том числе связанных с человеческим фактором: для качественного анализа снимка требуются существенное время и повышенное внимание, что приводит к быстрой утомляемости оператора СПД и может негативно сказаться на качестве анализа снимков. В данный процесс можно внести существенную долю автоматиза-

ции, сделав его более дешевым для организации и более комфортным для человека.

Для решения поставленной задачи использовались глубокие нейронные сети. В этой работе представлены способы предобработки данных и применение сети U-2-Net, а также анализ результатов.

2. Постановка задачи

Требуется разработать решение, ставящее в соответствие каждому рентгеновскому снимку булеву маску, единичные значения которой соответствовали бы пикселям, на которых присутствовали инородные объекты, такие как телефоны, оружие, металлические предметы и прочее. Эти объекты в дальнейшем будем называть аномалиями.

Набор данных, предоставленный компанией, специализирующейся на разработке сканеров персонального досмотра, состоит из оригинальных снимков людей, полученных с аппарата персонального досмотра человека «Express inspection» в рентгеновском спектре волн. Всего предоставлено четыре набора данных с различных сканеров персонального досмотра. Каждый снимок представляет из себя одноканальное 16-битное изображение в формате tiff размера 1600×500 пикселей. На снимках присутствуют различные аномалии, такие как: элементы одежды, аксессуары, оружие, протезы и прочее. Всего набор данных содержит 1654 снимка.

Выделим две проблемы с данными:

1. На оригинальных снимках аномалии визуально не различимы, что затрудняет их визуальный анализ.
2. Различные аппараты выдают снимки с различным распределением значений интенсивности пикселей. По этой причине исходные данные не подходят для автоматической обработки нейронной сетью.

Таким образом, во-первых, необходимо было разработать алгоритм для предобработки снимков, который решает обе поставленные задачи. После этого требовалось осуществить разметку аномалий на снимках. Наконец, для автоматизации этого процесса нужно было обучить нейронную сеть выделению аномалий.

3. Методы

3.1. Существующие подходы

Сегментация объектов на рентгеновских снимках является весьма распространенной задачей. Сначала эта задача решалась с помощью классических методов обработки изображений [1, 2]. Но со временем эта задача стала решаться преимущественно с помощью сверточных нейронных сетей.

В [3] произведена модификация архитектуры SegNet [4], упрощающая оригинальную сеть и позволяющая проводить обучение на небольшом наборе данных. Дальнейшее развитие сети SegNet — архитектура XNet [5], специализирующаяся на сегментации рентгеновских изображений, а именно на разделении мягкий тканей и костей.

Задача сегментации получила особое распространение в сфере медицины. Огромное количество работ было посвящено сегментации клеточных струк-

тур [6–8]. Среди множества подобных работ выделяется сеть U-Net [6], подходящая для более широкого класса задач. Эта сеть входит в число базовых архитектур для сегментации. В ней используются только сверточные слои, что позволяет подавать на вход изображения произвольного размера и получать маску с классами на выходе.

Следующем этапом развития U-Net является нейронная сеть U-2-Net [9]. Она показывает лучшие результаты на различных наборах данных по сравнению с U-Net. При этом она так же осталась применима для весьма широкого спектра задач, в том числе для анализа рентгеновских снимков. В данной работе использовалась именно эта архитектура.

3.2. Предобработка изображений

Для исправления описанных выше недостатков данных требовалось создать алгоритм предобработки снимков, состоящий из последовательности преобразований, решающий обе поставленные задачи.

В рамках этого направления произведена поэтапная обработка снимков, в результате которой снимки приведены к единому формату. Использованы алгоритмы увеличения контрастности снимков и удаления аномальных значений интенсивности пикселей. В частности, для повышения визуальной видимости применены методы трансформации гистограммы распределений пикселей к более светлым тонам, что позволяет оставить на снимке только самые темные объекты.

Для начала посмотрим, как выглядит гистограмма распределений пикселей в оригинальном изображении. Значение 0 — это черный цвет пикселя, 255 — белый (в дальнейшем для улучшения визуализации все гистограммы нормализуются от 0 до 255).

Из рис. 1 видно, что большинство пикселей оригинального изображения имеют яркость в окрестности значений 25 (черный цвет — человек) и 240 (белый цвет — фон). Цель состоит в уменьшении количества темных пикселей

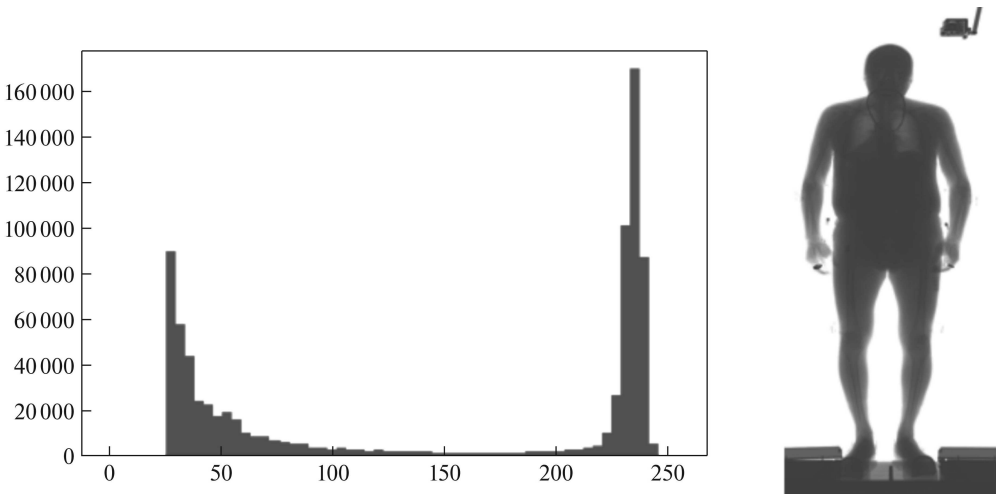


Рис. 1. Гистограмма оригинального изображения.

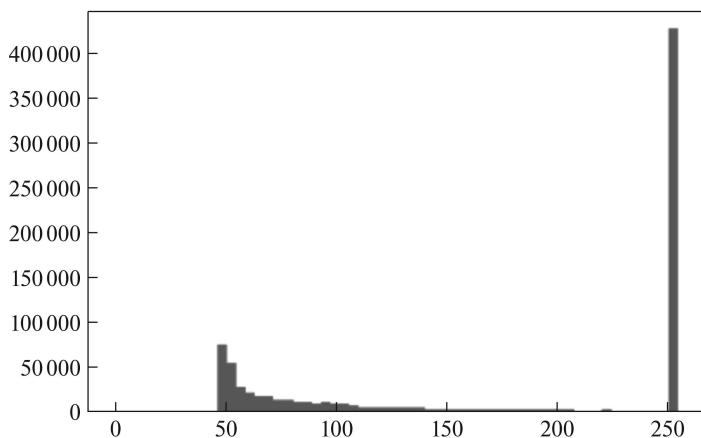


Рис. 2. Гистограмма после ThreshTrunc.

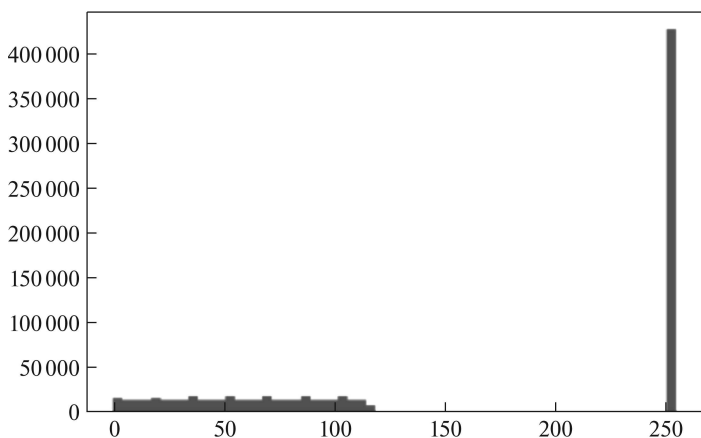


Рис. 3. Гистограмма после EqualizeHist.

так, чтобы менее темные участки (мягкие ткани) перешли в более светлый цвет, а самые темные участки (кости и металлические объекты) оставались темными. К оригинальным снимкам применяется следующая последовательность процедур обработки:

- 1) ThreshTrunc (обрезание по порогу)
- 2) EqualizeHist (выравнивание гистограммы)
- 3) ThreshTrunc
- 4) EqualizeAdaphthist (адаптивное выравнивание гистограммы)
- 5) ThreshTrunc.

Ниже каждая из этих операций описана подробнее.

3.2.1. ThreshTrunc. ThreshTrunc [10] — операция изменения гистограммы распределения таким образом, что значения интенсивности всех пиксе-

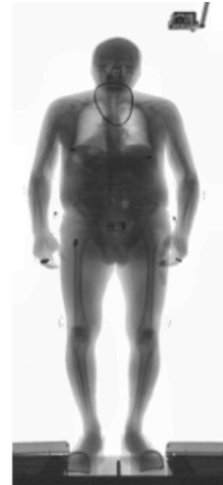
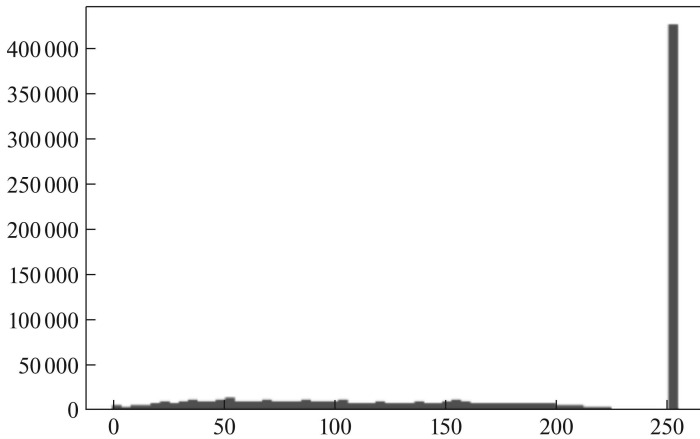


Рис. 4. Гистограмма после EqualizeAdaphthist.

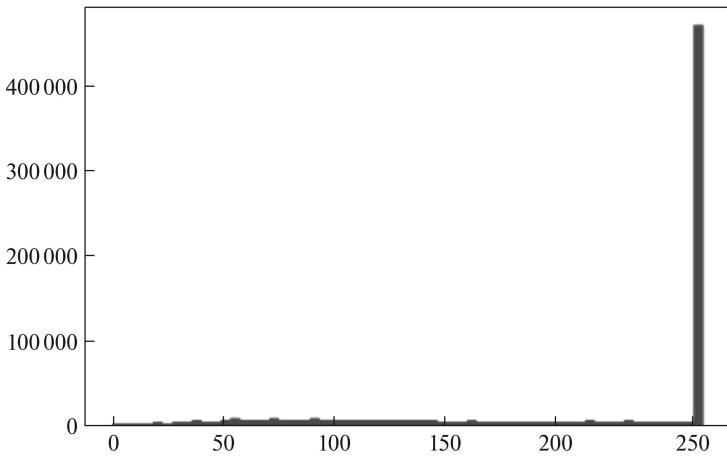


Рис. 5. Итоговая гистограмма.

лей, большие среднего значения интенсивности на снимке, заменяются на это среднее значение. Это влечет за собой смещение гистограммы распределения интенсивности пикселей в более светлую часть.

$$\text{ThreshTrunc}(x, y) = \begin{cases} \text{src}(x, y), & \text{если } \text{src}(x, y) < \text{threshol}, \\ \text{threshol}, & \text{если } \text{src}(x, y) \geq \text{threshol}, \end{cases}$$

где threshol — порог, равный среднему значению интенсивности всех пикселей снимка, $\text{src}(x, y)$ — значение интенсивности пикселя с координатами (x, y) .

3.2.2. EqualizeHist. EqualizeHist — процедура выравнивания гистограммы [11], позволяющая увеличить общий контраст изображения. Данное пре-

образование особенно эффективно, когда изображение представлено узким диапазоном значений интенсивности. Благодаря этому преобразованию можно лучше распределить интенсивности на гистограмме, равномерно используя весь диапазон кодирования интенсивности цвета.

$H(i)$ — гистограмма для значения интенсивности каждого пикселя. Находим кумулятивное распределение

$$H'(i) = \sum_{0 \leq j < i} H(j).$$

Затем заменяем значения интенсивности пикселя в изображении на полученное значение из распределения:

$$\text{EqualizeHist}(x, y) = H'(\text{src}(x, y)).$$

3.2.3. EqualizeAdapthist. EqualizeAdapthist — адаптивное выравнивание гистограммы. В отличие от EqualizeHist, данный метод вычисляет несколько гистограмм, каждая из которых соответствует отдельному участку изображения, и использует их для перераспределения значений интенсивности (по алгоритму EqualizeHist). Поэтому он подходит для улучшения локального контраста и улучшения четкости краев в каждой области изображения.

Для применения функций ThreshTrunc, EqualizeHist и EqualizeAdapthist использовался язык программирования Python3 и библиотека OpenCV [12].

Проведя все преобразования, мы получаем распределение пикселей, представленное на рис. 5. Как следует из этой гистограммы, большинство пикселей являются светлыми (т.е. фоном и участками мягких тканей), а остальные участки снимка распределены по градациям интенсивности более равномерно.

После применения всех преобразований металлические объекты на снимке становятся визуально различимы. Это позволяет выделять инородные тела объекты, не прибегая к специализированному программному обеспечению для индивидуальной предобработки каждого изображения. На рис. 6 показан процесс улучшения снимка после каждого этапа обработки.

До обработки средние значения интенсивности пикселей на разных сканерах сильно различаются (см. таблицу). После обработки средние значения становятся практически равными, т.е. изображения приводятся к единому формату, и их можно подавать на вход нейронной сети для выделения аномалий.

Таким образом, применение данного алгоритма предобработки изображений позволяет нивелировать различия в конструкции и настройках сканеров персонального досмотра.

Средние значения интенсивности пикселей

	Набор А	Набор В	Набор С	Набор D
До обработки	176,71	180,41	173,61	158,81
После обработки	201,36	201,16	203,36	201,62

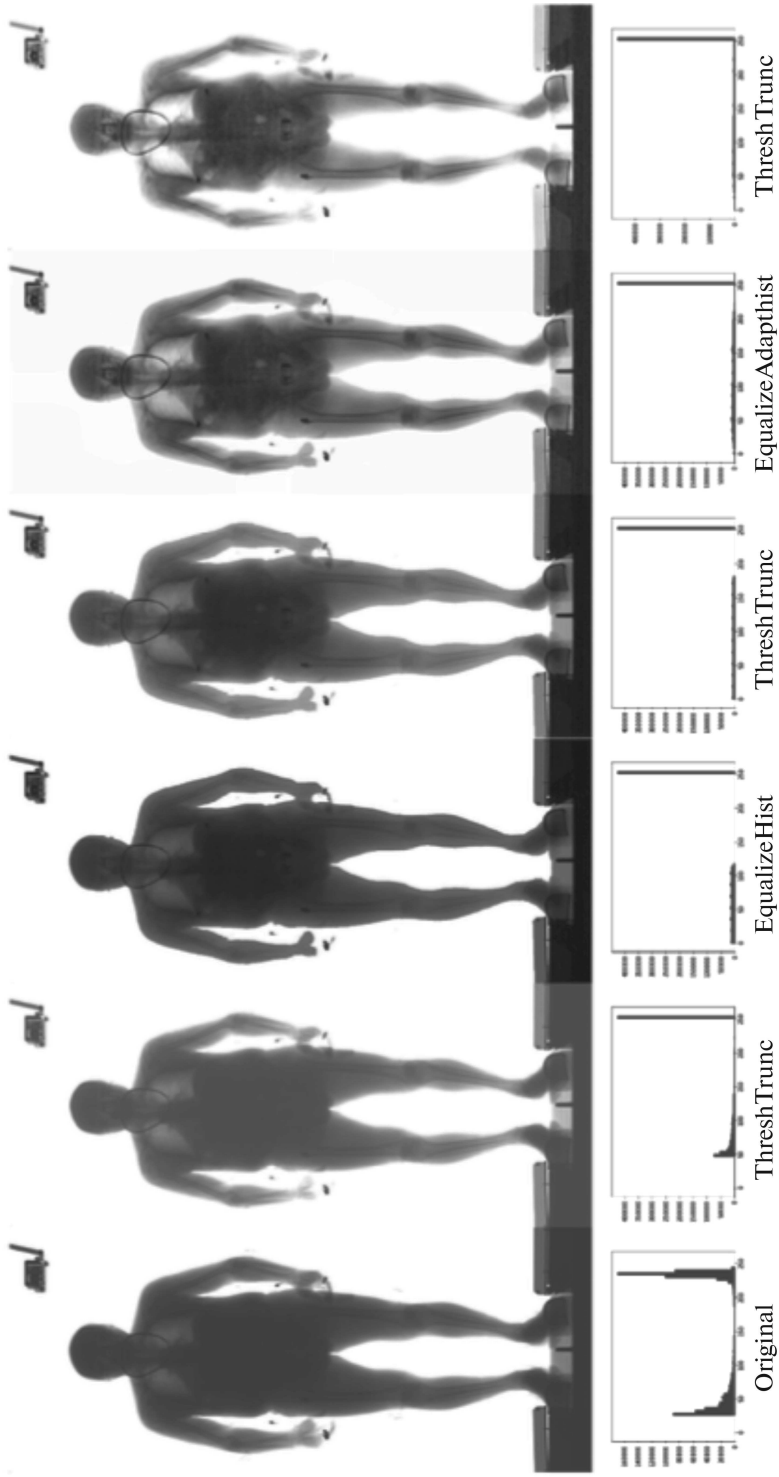


Рис. 6. Результат обработки после каждого преобразования.

3.3. Разметка

После обработки набора данных аномалии стали визуально различимыми. Благодаря этому с помощью программы labelme [13] на снимках вручную были выделены инородные объекты. Ручная разметка производилась в том числе с помощью сервиса Yandex Toloka [14]. Всего было выделено около 18 000 аномалий на 1654 изображениях. В результате были получены обработанные изображения и маски инородных объектов для них (пример приведен на рис. 7).

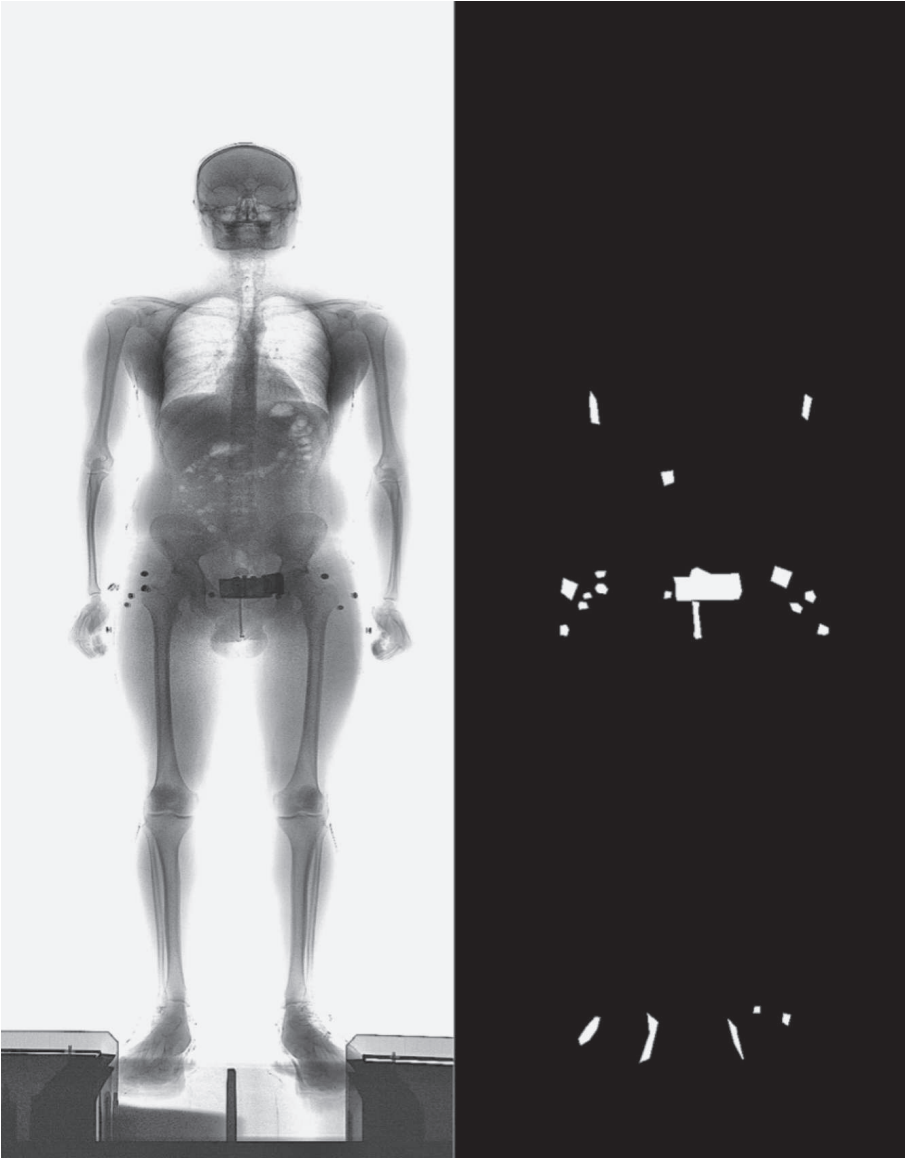


Рис. 7. Пример размеченной вручную маски.

3.4. Нейронная сеть

Для решения задачи выявления аномалий использовалась сеть U-2-Net [9], которая была представлена в 2020 г. Для реализации нейронной сети использовались язык программирования Python и фреймворк PyTorch. Архитектура использовалась без дополнительных модификаций. В качестве оптимизатора был выбран алгоритм Adam [15] с параметром скорости обучения (learning rate), равным 0,001. Обучение проводилось на графическом процессоре Nvidia gtx 1080ti с 11Gb памяти. Размер изображений на входе сети составляет 512×512 пикселей типа float32. В связи с ограничением памяти и большим количеством слоев нейронной сети размер батча был выбран равным десяти.

На вход модели подавались одноканальные снимки с указанной выше предобработкой. В процессе аугментации для увеличения вариативности данных использовалась операция случайной обрезки изображения (cropping), что позволило сгенерировать большое количество изображений, содержащих различные части тела. Модель отображает снимок в матрицу, характеризующую вероятности нахождения аномалий в соответствующих участках снимка. Далее матрица вероятностей приводится к бинарному виду. Таким образом, на выходе имеем булеву маску. В качестве функции потерь была использована функция

$$L(A, B) = 1 - \text{IoU}(A, B),$$

где A — оригинальная маска, B — маска, полученная из нейронной сети.

В качестве метрики была выбрана функция IoU (intersection over union) [16], которая характеризует, насколько предсказанная маска покрывает настоящую маску:

$$\text{IoU}(A, B) = \frac{A \cap B}{A \cup B}.$$

Обучение проводилось на 1454 снимках. Тестовый набор данных содержал 200 изображений.

4. Заключение

В результате проведенной на первом этапе работы разработан универсальный алгоритм, который позволяет обрабатывать снимки и приводить их к виду, подходящему для обучения нейронной сети. Была обучена нейронная

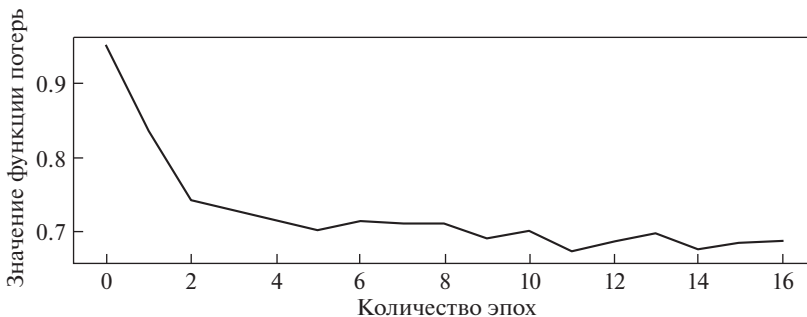


Рис. 8. Значение функции потерь на тестовой выборке в процессе обучения.

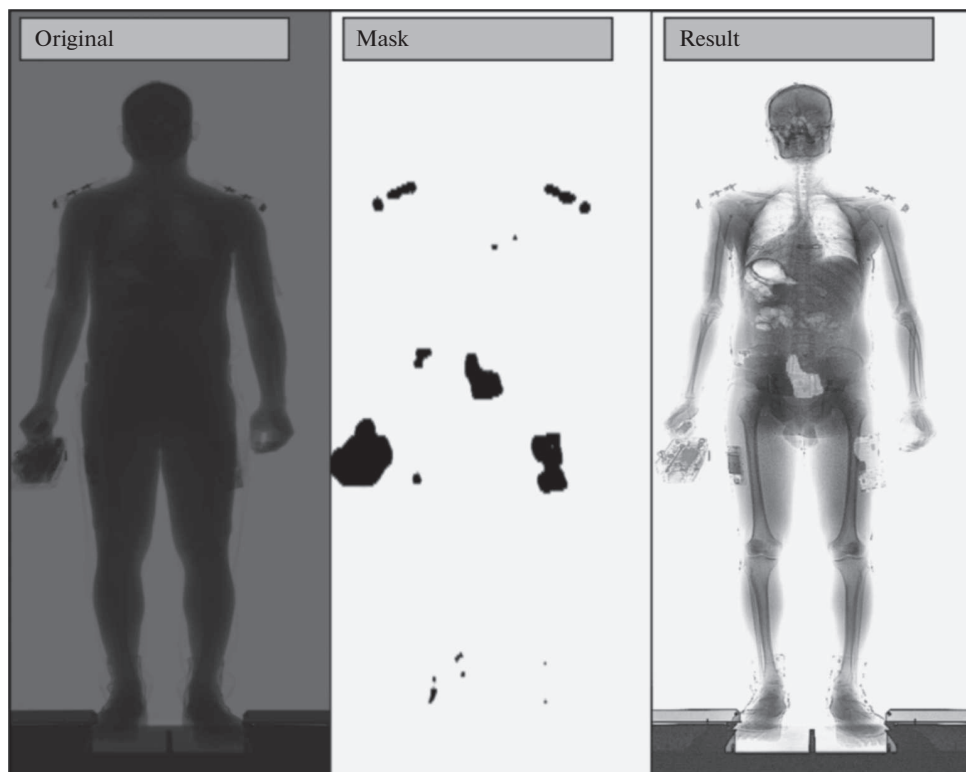


Рис. 9. Слева оригинальный снимок, в центре маска, полученная с помощью нейронной сети, справа маска, наложенная на обработанный снимок.

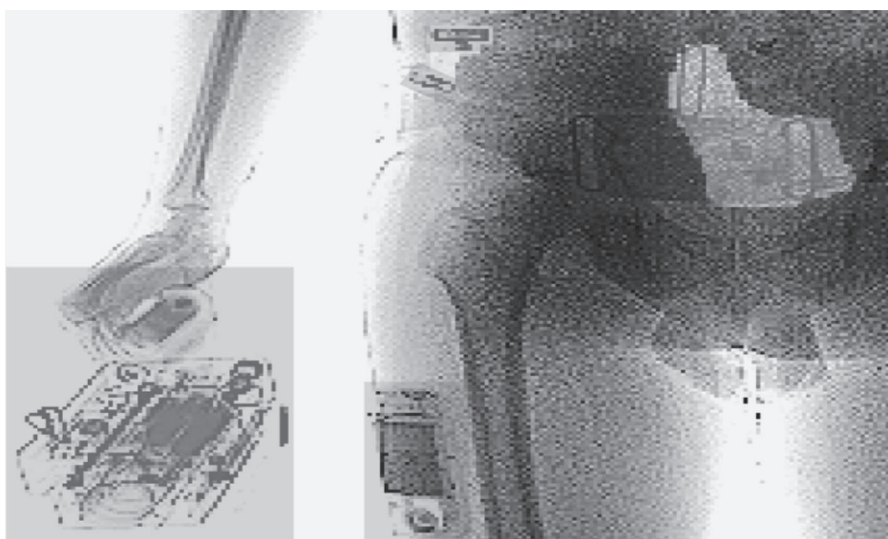


Рис. 10. Детальное рассмотрение полученной маски.

сеть (см. рис. 8), выполняющая поставленную задачу. В результате проверки модели на тестовом наборе данных с помощью метрики IoU сеть показала точность в 0,838.

Результаты работы этой модели приведены ниже (см. рис. 9 и 10). Можно видеть, что нейронная сеть научилась выделять большие инородные объекты, но границы этих объектов пока выделяются плохо. Получение маски на выходе нейронной сети одного снимка на CPU в среднем занимает 4 с.

Обычно сети типа U-2-net используются для сегментации цветных фотографий и выделения объектов на них [9]. К сожалению, не удалось найти открытых источников, где бы подобные сети применялись для рентгеновских снимков. По этой причине нет возможности сравнить качество полученной модели.

5. Выводы

Предложена универсальная схема предобработки изображений, которая позволяет приводить данные с различных аппаратов персонального досмотра (или с различными настройками излучения) к нормализованному формату, в котором средние значения интенсивности пикселей совпадают, и при этом все объекты становятся различимы человеком.

С использованием предложенного подхода обработки из четырех различных наборов данных сформирована обучающая выборка. На основе полученной выборки обучена нейронная сеть U-2-net. Качество сегментации обученной модели позволяет распознавать аномалии большого и среднего размера. На объектах сравнительно меньшего размера качество сегментации существенно снижается. Несмотря на это, модель может быть использована на промышленных объектах, в качестве средства автоматизации работы СПД для поиска объектов среднего размера, таких как оружие, телефоны, слитки металлов и прочее, значительно увеличивая скорость работы оператора.

СПИСОК ЛИТЕРАТУРЫ

1. *Sharma N., Aggarwal L.M.* Automated medical image segmentation techniques // *J. Med. Phys.* 2010. V. 35. No. 1. P. 3–14.
2. *Mansoor A., Bagci U., Foster B., et. al.* Segmentation and image analysis of abnormal lungs at CT: current approaches, challenges, and future trends // *Radiographics.* 2015. V. 35. No. 4. P. 1056–1076.
3. *Badrinarayanan V., Handa A., Cipolla R.* Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labeling. arXiv preprint. arXiv:1505.07293, 2015.
4. *Badrinarayanan V., Kendall A., Cipolla R.* Segnet: a deep convolutional encoder-decoder architecture for image segmentation // *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2017. V. 39. No. 12. P. 2481–2495.
5. *Aylett-Bullock J., Cuesta-Lázaro C., Quera-Bofarull A.* XNet: a convolutional neural network (CNN) implementation for medical X-Ray image segmentation suitable for small datasets // *Proc. SPIE. Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging.* 2019. V. 10953.

6. *Ronneberger O., Fischer P., Brox T.* U-Net: convolutional networks for biomedical image segmentation / Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science. 2015. V. 9351. P. 234–241.
7. *Ciresan D., Giusti A., Gambardella L.M., Schmidhuber J.* Deep neural networks segment neuronal membranes in electron microscopy images / Advances in Neural Information Processing Systems 25, Pereira F., Burges C.J.C., Bottou L., and Weinberger K.Q., eds., 2843-2851, Curran Associates, Inc., 2012.
8. *Arganda-Carreras I., Turaga S.C., Berger D.R., et. al.* Crowdsourcing the creation of image segmentation algorithms for connectomics // Front. Neuroanat. 2015. V. 9. No. 142.
9. *Xuebin Q., Zhang Z., Huang C., et. al.* U2-Net: Going deeper with nested U-structure for salient object detection // Pattern Recognition. 2020. V. 106. P. 107404.
10. https://docs.opencv.org/3.4/db/d8e/tutorial_threshold.html
11. https://docs.opencv.org/3.4/d6/dc7/group_imgproc_hist.html
12. OpenCV. <https://opencv.org/>
13. LabelMe. <http://labelme.csail.mit.edu/Release3.0/>
14. Яндекс Толока. <https://toloka.ai/>
15. *Kingma D.P., Ba J.L.* Adam: a method for stochastic optimization. arXiv preprint. arXiv:1412.6980, 2017.
16. https://id.wikipedia.org/wiki/Indeks_Jaccard

Статъа представена к публикации членом редколлегии А.А. Лазаревым.

Поступила в редакцию 01.02.2022

После доработки 31.05.2022

Принята к публикации 29.06.2022