

УДК 004.89 ГРНТИ 28.23.15

Толкачев Алексей Юрьевич, a.tolkachev@innopolis.ru**Кулеев Рамиль Фуатович, к.т.н., r.kuleev@innopolis.ru**АНО ВО «Университет Иннополис», Республика Татарстан, г.Иннополис,
ул.Университетская, д.1, тел. +7 (843) 203-92-53

Об опыте применения технологий искусственного интеллекта для автоматического распознавания рентгеновских изображений органов грудной полости

Аннотация

В работе описан наш опыт решения задач по автоматизированной диагностике заболеваний органов грудной полости современными методами глубокого обучения. Приведено краткое описание подходов и результатов подавления костного каркаса, сегментации легочных полей, средостения и ключиц, обнаружения и классификации патологических изменений. Показано, что результаты сопоставимы или превышают известные по научным публикациям результаты.

Ключевые слова:

заболевания органов грудной полости, рентген легких, флюорография, глубокое обучение, нейронные сети, автоматизация диагностики, классификация, сегментация, обнаружение патологии.

Введение

Заболевания грудной полости являются одной из главных причин смертей в мире (20%), они способны быстро развиваться и приводить к летальному исходу, если не обнаружены на ранних стадиях. К таким заболеваниям относятся рак легких (занимает лидирующие позиции по смертности среди онкологических заболеваний у мужчин), туберкулез, хроническая обструктивная болезнь легких (ХОБЛ) и другие. Флюорография органов грудной клетки является наиболее распространенным инструментом скрининга во многих странах. В России проводится 72 миллиона исследований в год. Обычно, флюорография направлена на определение туберкулёза, поэтому зачастую информация о возможном наличии других заболеваний остаётся неиспользованной и пропадает навсегда. Более того, замена устаревших аппаратов плёночной флюорографии на новые цифровые

и, как следствие, улучшение качества снимков, делает возможным не только скрининг, но и полное диагностическое исследование изображений.

За последние десятилетия накопился огромный объем рентгенографических изображений органов грудной полости. Интеллектуальный анализ данных с целью автоматизации диагностики заболеваний органов грудной полости стал одним из основных предметов обсуждения специалистов в области здравоохранения и информационных технологий. В последние годы внимание сообщества специалистов в области анализа данных (data science) привлекает глубокое обучение, так как оно позволяет выполнять автоматическую обработку изображений без использования вручную выбранных признаков. Также стоит отметить, что на всемирно известном соревновании по распознаванию изображений ILSVRC, проводимом с 2010 по

2017 годы, методы, основанные на глубоком обучении, занимали первую строчку с 2012 года [14], когда победу одержала глубокая сверточная нейронная сеть AlexNet, разработанная А.Крижевским и др. [15].

Наша команда нацелена на применение техник глубокого обучения для анализа рентгеновских изображений органов грудной полости. Основной целью является создание сервиса автоматизированной диагностики по рентгеновским изображениям органов грудной полости с решением задач классификации исследований на норму и патологию и обнаружением конкретных рентгенологических синдромов и нозологий. Для достижения этой цели решается несколько задач, каждая из которых имеет самостоятельную ценность для врача-рентгенолога. Данные задачи включают предобработку изображения с подавлением костного каркаса, сегментацию легочных полей и средостения, классификацию с определением рентгенологических синдромов. Кроме того, были сформулированы и решены отдельные задачи обнаружения рентгенологических симптомов пневмонии и сегментации области пневмоторакса.

Целью данной работы стало проведение исследования на предмет выявления наиболее эффективных методов решения перечисленных выше задач анализа рентгеновских изображений органов грудной полости. Для реализованных алгоритмов была проведена оценка их эффективности путем сравнения, где возможно, полученных результатов с описанными в научных публикациях.

Методы и результаты

В данном разделе приведено краткое описание использованных подходов и результатов для решения задач подавления костного каркаса, сегментации легочных полей и средостения, обнаружения пневмонии, сегментации области пневмоторакса и классификации.

1. Подавление костного каркаса.

По данным врачей-рентгенологов, большинство случаев пропуска патологии по рентгеновским изображениям органов грудной полости касается небольших очагов, скрытых тенями ребер и ключиц. Подавление костного каркаса на рентгеновских изображениях потенциально может помочь упростить обнаружение такой патологии врачом-рентгенологом. Также есть предположение, что подавление костного каркаса в качестве одного из этапов автоматической обработки изображений может повысить точность классификации изображений.

Одним из наиболее популярных методов решения

поставленной задачи является использование автокодировщика [1], который и был нами использован. В этом случае костная ткань определяется на изображении как шум, который модель пытается устранить посредством кодирования наиболее важных признаков и последующего восстановления изображения из заданного набора признаков.

Поскольку открытых наборов данных для решения задач по подавлению костного каркаса на рентгеновских изображениях нам неизвестно, мы были вынуждены собрать собственный набор данных. Всего было получено 35 пар изображений, полученных на двух энергетических рентгеновских аппаратах, которые были разделены для обучения, валидации и тестирования на 24, 4 и 7 пар соответственно. Далее изображения были аугментированы 99 раз грубым набором преобразований (аффинные преобразования, цветовые, оптические искажения, добавление шума и размытия) для обучающего набора данных. В то же время, для валидации и тестирования были применены достаточно простые аугментации, чтобы новые преобразованные снимки визуально не сильно отличались от тех, что могут быть использованы в реальных условиях.

Субъективный анализ результатов с привлечением врачей-рентгенологов позволяет говорить о том, что использование пар изображений (исходное и с подавлением костного каркаса) позволяет лучше различать некоторые патологии.

2. Сегментация легких

Сегментация органов является, как правило, первым шагом в задачах по анализу медицинских изображений, поскольку определение расположения целевого органа может существенно повысить точность, например, обнаружения патологии конкретного органа. Решалась задача сегментации легочных полей, средостения и ключиц. Использовалась автокодировочная нейронная сеть, в основе которой лежит архитектура U-Net [2] с предобученным кодировщиком семейства ResNeXt [3] на наборе данных ImageNet.

Обучение и оценка алгоритмов проводились на наборе данных JSRT [4]. Он состоит из 247 рентгеновских изображений органов грудной полости, на части которых присутствуют узловые образования в легких. Все изображения имеют разрешение 2048 x 2048 пикселей, с размером пикселя равным 0.175мм. Также каждому снимку соответствуют три вручную размеченные маски: легких, сердца и ключицы [5].

Полученные результаты показали, что применение глубокого обучения позволяет повысить точность сегментации по сравнению с широко

применяемыми методами ASM, AAM и другими.

3. Обнаружение пневмонии

Пневмония является одной из наиболее частых причин смертности среди пожилых людей и детей. По данным ВОЗ, в 2017 году у 15% детей до 5 лет в мире причиной смерти стало это заболевание. Диагностика пневмонии является сложной задачей, так как помимо анализа рентгенограммы требуется подтверждение диагноза на основании данных из истории болезни.

Пневмония на рентгеновском изображении обычно проявляется как область или области повышенной непрозрачности [6]. Однако ее диагностика сложна из-за ряда других состояний в легких, таких как наличие жидкости (отек легких), кровотечение, потеря объема (ателектаз или коллапс), рак легких, пострадиационные или хирургические изменения.

Для того, чтобы помочь радиологам диагностировать пневмонию, Североамериканская радиологическая ассоциация (RSNA) в сотрудничестве с Национальными институтами здравоохранения США, Обществом торакальной радиологии и MD.ai организовало на платформе Kaggle соревнование по обнаружению пневмонии по рентгеновским изображениям.

Предоставленный набор данных содержал 26,684 уникальных рентгеновских изображения, содержащие 3 класса меток (без патологии - 29%, с патологией, но без затемнений - 40%, с затемнениями - 31%). Все изображения с затемнениями были размечены прямоугольными областями, указывающими на область, свидетельствующую о наличии пневмонии. Изображения были разделены на обучающую (23,115), валидационную (2569) и тестовую (1000) выборки. С целью снижения влияния переобучения полученных моделей на тестовой выборке из первой стадии, в конце соревнования организаторы добавили закрытую тестовую часть, содержащую дополнительные 3000 изображений.

Наша команда приняла участие в соревновании и предложила в качестве решения ансамбль из двух нейросетевых моделей: RetinaNet [7] и Mask R-CNN [8]. В основе этих сетей используются классические сверточные нейронные сети с “остаточными соединениями” (residual connections) - ResNet-50 и ResNet-101 [9].

Объединение моделей в ансамбль происходило следующим образом: сначала обе модели предсказывали области, содержащие пневмонию, затем применялся механизм подавления немаксимумов (Non-Maximum Suppression). Если какое-либо предсказание RetinaNet пересекалось с предсказанием Mask R-CNN

и при этом значение IoU превышало 0.5, итоговые предсказания усреднялись, при этом веса предсказаний RetinaNet и Mask R-CNN брались в пропорции 3:1. Оптимальная пропорция весов из набора 4:1, 3:1, 2:1 была подобрана эмпирическим путем на валидационном наборе данных. Если предсказанная одной моделью прямоугольная область с патологией не имела соответствующего пересекающегося с ней предсказания, полученного от другой модели, эта область была использована в результатах без каких-либо изменений.

Полученное решение позволило занять в итоговой таблице лидеров 30 место из 1499 участников, что эквивалентно попаданию в топ-3% соревнующихся.

4. Сегментация пневмоторакса

Изначально, соревнование по сегментации пневмоторакса, проведенное на платформе Kaggle, стало возможностью реализовать и сравнить с другими участниками решение нашей команды проекта. В качестве обучающих данных было предоставлено более 10 тысяч рентгеновских снимков вместе с соответствующими им бинарными масками с пропорцией патологических снимков к условно-нормальным приблизительно как один к пяти.

В качестве сегментационной модели была выбрана нейросетевая архитектура, представляющая одну из разновидностей известной архитектуры U-Net [2], работающей по принципу автокодировщика. Важной особенностью стало использование в кодирующей части предобученного на задаче классификации по набору данных ImageNet кодировщика, именуемого SE-ResNeXt-50 [10].

Результаты работы данной модели позволили расположиться на 16 месте в итоговой таблице лидеров из 1475 участников (топ-2%).

5. Классификация патологии

Задача классификации патологии на рентгеновских изображениях органов грудной полости заключалась в определении отклонений, дифференцируемых на 14 различных типов, представленных в известном наборе данных Chest-14 [11]. Если на изображении ни один из вариантов патологии не был обнаружен, пациент относился к здоровым.

Всего набор данных состоит из 112,120 рентгеновских изображений, принадлежащих 30,805 пациентам. Все данные были разделены на обучающую (70%) и тестовую (30%) выборку. В свою очередь 10% от обучающей выборки было отведено под локальную валидацию получаемого решения.

Решение представляет из себя сверточную нейросетевую модель, архитектура которой носит

название Inception-ResNet-v2, продемонстрировавшую высокие результаты на соревновании ILSVRC [12].

В процессе обучения использовался специфический набор аугментаций из общего набора аффинных преобразований, цветовых искажений, зашумления и размытия. Лучшая комбинация данных аугментаций позволила превзойти доступные результаты на том же разбиении от Вонга и др. [11] и Яо и др. [13].

Обсуждение и выводы

В результате выполненной работы были разработаны методы и алгоритмы для решения отдельных задач анализа рентгеновских изображений органов грудной полости. В целом, удалось достичь показателей точности на уровне или даже превосходящих результаты опубликованных решений. Оценка результата подавления костного каркаса является субъективной, так как сложно подобрать адекватные метрики ее качества, и формализованных результатов точности других исследователей нам не известно. Сегментация легочных полей позволит исключить ряд явных ошибок при обнаружении патологии, а сегментация средостения открывает возможности для автоматического определения различной патологии сердечно-сосудистой системы по конфигурации сердца и крупных сосудов. Участие в конкурсах по обнаружению пневмонии и сегментации пневмоторакса на площадке Kaggle позволило сравнить разработанные командой проекта алгоритмы с решениями других команд по всему миру. Для сегментации пневмоторакса отрыв победителя конкурса от нашего решения (16-е место из 1475 участников) составил порядка 0,01, что говорит об очень высокой плотности результатов и практически одинаковых ожидаемых клинических результатах в случае применения моделей на практике. Точность разработанного алгоритма классификации на уровне решений ведущих исследователей позволяет рассчитывать на возможность ее применения в рамках сервиса автоматизированной диагностики.

В то же время, для обеспечения возможности применения полученных результатов в клинической практике и внедрения сервиса на основе

разработанных алгоритмов и моделей необходимо решить ряд серьезных задач. Прежде всего, речь идет о качестве данных, на которых проводилось обучение моделей. Открытые датасеты зачастую представляют большой интерес для научного сообщества и позволяют сравнивать решения различных исследователей, но не подходят для создания моделей, способных эффективно работать на практике. Датасет Chest-14 не является исключением, так как точность его разметки в некоторых случаях (особенно для случаев присутствия ряда патологических изменений) сомнительна. Кроме того, существенную его часть составляют случаи тяжелой патологии у пациентов в реанимации или, как минимум, лежащих в клинике. Поэтому обученные на данном датасете модели не могут быть использованы для автоматизации диагностики в рамках скрининговых флюорографических исследований, где подавляющее большинство пациентов не имеют отклонений от нормы, и наиболее важно обнаружить патологии на ранней стадии. Поэтому важнейшей задачей является дообучение разработанных моделей на данных из медицинских организаций, где и планируется применять сервис автоматизированной диагностики в дальнейшем.

Но самое главное, требуется провести полноценные клинические испытания разработанных решений. При этом измеряемые метрики будут отличаться от метрик точности, которые оценивались разработчиками в ходе «лабораторных» исследований. Сервис автоматизированной диагностики планируется использовать в качестве помощника врача, и важно оценивать изменения показателей выявляемости, числа ошибочных госпитализаций, времени до вызова пациента для госпитализации и другие на уровне медицинской организации, а в идеале, и в масштабах региона в целом. Это позволит определить социально-экономический эффект и обосновать затраты на внедрение и масштабирование сервисов автоматизированной диагностики в дальнейшем.

Работа выполнена при финансовой поддержке РФФИ и Правительства Республики Татарстан в рамках научного проекта №18-47-160015.

СПИСОК ЛИТЕРАТУРЫ:

1. Maxim Gusarev, Ramil Kuleev, Adil Khan, Adin Ramirez Rivera, and Asad Masood Khattak. Deep learning models for bone suppression in chest radiographs // Proceedings of the 2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). – 2017. – P.1–7.
2. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation // Proceedings of the International Conference on Medical image computing and computer-assisted intervention. – Springer, 2015. – P.234–240.
3. Xie S. et al. Aggregated residual transformations for deep neural networks // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – P. 1492-1500.

4. Shiraishi, Junji, et al. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules//American Journal of Roentgenology, 2000. – 174.1. – P.71-74.
5. Van Ginneken, Bram, Mikkel B. Stegmann, and Marco Loog. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database//Medical image analysis, 2006. – 10.1. – P.19-40.
6. Franquet T. Imaging of community-acquired pneumonia//Thorac Imaging, 2018 (epub ahead of print). – PMID 30036297.
7. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. Focal loss for dense object detection//Proceedings of the IEEE international conference on computer vision. – 2017. – P.2980-2988.
8. He, K., Gkioxari, G., Dollár, P., & Girshick, R. Mask r-cnn//Proceedings of the IEEE international conference on computer vision. – 2017. – P.2961-2969.
9. He, K., Zhang, X., Ren, S., & Sun, J. Deep residual learning for image recognition//Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P.770-778.
10. Hu, J., Shen, L., & Sun, G. Squeeze-and-excitation networks//Proceedings of the IEEE conference on computer vision and pattern recognition. – 2018. – P.7132-7141.
11. Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – P.3462–3471.
12. Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning// Thirty-First AAAI Conference on Artificial Intelligence. – 2017.
13. Li Yao, Jordan Prosky, Eric Poblenz, Ben Covington, and Kevin Lyman. Weakly supervised medical diagnosis and localization from multiple resolutions//arXiv preprint, 2018. –1803.07703.
14. Tweedale J. W. An Application of Transfer Learning for Maritime Vision Processing Using Machine Learning //International Conference on Intelligent Decision Technologies. – Springer, Cham, 2018. – P. 87-97.
15. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks // Advances in neural information processing systems. – 2012. – P. 1097-1105.

UDC 004.89 GRNTI 28.23.15

Aleksey Yu. Tolkachev, a.tolkachev@innopolis.ru

Ramil F. Kuleev, PhD, r.kuleev@innopolis.ru

Innopolis University, Republic of Tatarstan, Innopolis, Universitetskaya street, 1, +7 (843) 203-92-53

Experience of artificial technologies usage for chest x-ray images automatic recognition

Abstract. The paper describes our experience in solving problems of automated diagnosis of thorax diseases using modern methods of deep learning. A brief description of approaches and results of the bone suppression, segmentation of the pulmonary fields, mediastinum and clavicles, detection of pathologies and classification is given. It is shown that the results are comparable or exceed the results known from scientific publications.

Keywords: Thorax diseases, chest x-ray images, fluorography, deep learning, neural networks, diagnostics automatization, classification, segmentation, pathologies detection