

УДК 004.93'12

А.Д. Кульневич, Н.Д. Сергеева, Р.А. Чугунов

**СИСТЕМА РАННЕГО ДЕТЕКТИРОВАНИЯ ПНЕВМОНИИ НА ОСНОВЕ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ**

*В статье рассматривается проектирование архитектуры нейронной сети для распознавания визуальных признаков заболеваний органов дыхания на медицинских изображениях. В частности, алгоритм автоматически обнаруживает снижение прозрачности тканей легких на рентгенограммах грудной клетки.*

*Ключевые слова: распознавание образов, нейронные сети, обработка медицинских изображений, детекция пневмонии.*

**SYSTEM OF EARLY PNEUMONIA DETECTION BASED ON DEEP LEARNING**

*The article discusses the neural network architecture designed to recognize visual signs of respiratory diseases in medical images. In particular, the algorithm automatically recognizes reduced lung tissue transparency in chest radiographs.*

*Key words: pattern recognition, neural network, medical image processing, pneumonia detection.*

**Введение**

Для точной постановки диагноза или проверки наличия заболеваний на ранней стадии назначаются лучевые и функциональные тесты – такие как рентгеновский снимок, электрокардиограмма, компьютерная томография, магниторезонансная терапия и т.п. Проводить тест на медицинском оборудовании может технический специалист, однако анализировать результаты исследований способен только квалифицированный узкоспециализированный врач. В малонаселенных пунктах наблюдаются дефицит специалистов, в результате нет возможности выполнить исследование, а в больницах крупных городов – большая нагрузка на врачей, в результате может снизиться качество анализа из-за непреднамеренных ошибок.

С 1999 г. в здравоохранении развивается практика электронной медицины – eHealth, обеспечиваемая электронными процессами и коммуникациями. eHealth охватывает целый ряд услуг или систем, которые находятся на грани медицины, здравоохранения и информационных технологий; она также включает понятие мобильной медицины – mHealth, под которым подразумевают использование мобильных технологий в области медицинских услуг и медицинских данных [1-4].

Проекты eHealth направлены на облегчение доступа граждан к медицинским услугам и повышение информированности населения по вопросам здоровья, расширение возможностей диагностики и мониторинга заболеваний и т.д. Данное исследование релевантно двум направлениям разработки – телемедицина и использование технологии распределенных вычислений в медицинских исследованиях. Телемедицина представляет собой дистанционную диагностику и назначение лечения, а технологии распределенных вычислений позволяют обрабатывать большое количество гетерогенных

данных [5]. Особый интерес вызывает слияние этих направлений, сопровождаемое исследованием возможностей использования алгоритмов машинного обучения в системе удаленной постановки диагноза на основе анализа медицинских тестов [6-12].

В России болезни органов дыхания занимают второе место в структуре общей заболеваемости, после заболеваний кровеносной системы [13, 14]. Согласно статистике Всемирной организации здравоохранения, одной из самых распространенных причин смертности от заболеваний органов дыхания является пневмония, значительная доля всех случаев болезни заканчивается летальным исходом – порядка 75% всех смертей от болезней органов дыхания. Пневмония – острое поражение легких инфекционно-воспалительного характера, в которое вовлекаются все структурные элементы легочной ткани, преимущественно альвеолы и интерстициальная ткань легких. Одним из основных методов диагностики пневмонии является рентгенологическое исследование легких.

При диагностике заболеваний по медицинским соображениям вероятность врачебной ошибки довольно высока: для пневмонии около трети поставленных диагнозов оказываются неверными. Поэтому так актуальна разработка сервисов для поддержки принятия врачебных решений [15]. Алгоритмы машинного обучения рассматриваются как часть системы предварительной оценки тестовых медицинских данных пациентов. В таких системах используется расширяемый набор алгоритмов обучения с учителем, каждый из которых используется для предварительной оценки вероятности того или иного диагноза. Система автоматически анализирует доступные ей данные и выдает рекомендации для специалиста.

### **Особенности применения машинного обучения для обработки медицинских изображений**

На протяжении десятилетий применялись различные методы машинного обучения, – например, линейный дискриминантный анализ, метод опорных векторов, деревья решений и случайные леса, нейронные сети, алгоритмы глубокого обучения. По сравнению с другими методами машинного обучения в медицине глубокое обучение имеет значительный успех. Глубокое обучение – это подкатегория машинного обучения, в которой многоуровневые сети используются для оценки сложных шаблонов в данных исходного изображения. В первую очередь к алгоритмам глубокого обучения, применяемым для распознавания образов, относят сверточные нейронные сети.

Глубокое обучение позволяет напрямую анализировать визуальные данные, без дополнительного извлечения признаков изображений. Тем не менее, качество алгоритмов глубокого обучения зависит от количества размеченных данных. В большинстве случаев для достижения необходимого уровня распознавания требуется большое число изображений (десятки и сотни тысяч). Однако объем наборов медицинских изображений редко превышает 1000. Поэтому основной проблемой при применении глубокого обучения к медицинским изображениям является ограниченное количество образцов для обучения, доступных для создания глубоких моделей, не подверженных переобучению [16]. Для решения этой проблемы исследовательские группы разработали различные стратегии – такие как: а) извлечение признаков из изображений, позволяющих использовать для обучения меньше данных и более простые модели машинного обучения [17, 21, 22]; б) расширение набора данных путем искусственного генерирования выборок посредством аффинного преобразования (т.е. увеличение количества данных), а затем обучение сети с нуля с помощью расширенного набора данных [18-20]; в) инициализация параметров модели с параметрами предварительно обученных моделей из немедицинских или естественных изображений, с дальнейшим дообучением (fine-tuning) параметров сети с помощью изображений, соответствующих данной задаче [23, 24].

Несмотря на то, что ежегодно появляется множество публикаций, посвященных машинному обучению, до сих пор существует лишь несколько методов, способных обрабатывать широкий спектр визуальных признаков схожих состояний болезней. Например, использование сверточных нейронных

сетей для различения интерстициальных заболеваний легких включает слабые изменения в структурах текстурного типа, что сильно отличается от классификации обычных фотографий кошек и собак. Диапазон представлений различных состояний болезни, потребность в больших размеченных наборах клинических данных и сложная структура многих методов машинного обучения означают необходимость в продолжении исследований и разработок, прежде чем будет возможно внедрение и использование их в клинических условиях.

Использование глубокого обучения для медицинских задач также вызывает беспокойство из-за его принципа «черного ящика»; однако существуют методы оценки обученных параметров в нейронных сетях для раскрытия методологии принятия решений.

### Предлагаемый алгоритм обнаружения снижения прозрачности тканей легких на рентгенограммах грудной клетки



Рис. 1. Пример изображения, анализируемого моделью.

В данной работе предложен алгоритм детектирования снижения прозрачности легких (рис. 1). Сниженная прозрачность легких на рентгеновских снимках может сигнализировать о развитии пневмонии. Для обучения все снимки сжимались до разрешения  $256 \times 256$  пикселей.

Предлагается использовать нейронную сеть Masc R-CNN [25] со следующими нововведениями.

1. Для улучшения качества распознавания используется предобучение на изображениях набора данных Common Objects in Context (COCO) (рис. 2), затем нейронная сеть обучается на имеющихся данных для поставленной задачи. Это позволяет компенсировать проблему сравнительно небольшого количества изображений в выборке. При обучении веса никакие слои не замораживались, сеть дообучалась целиком.

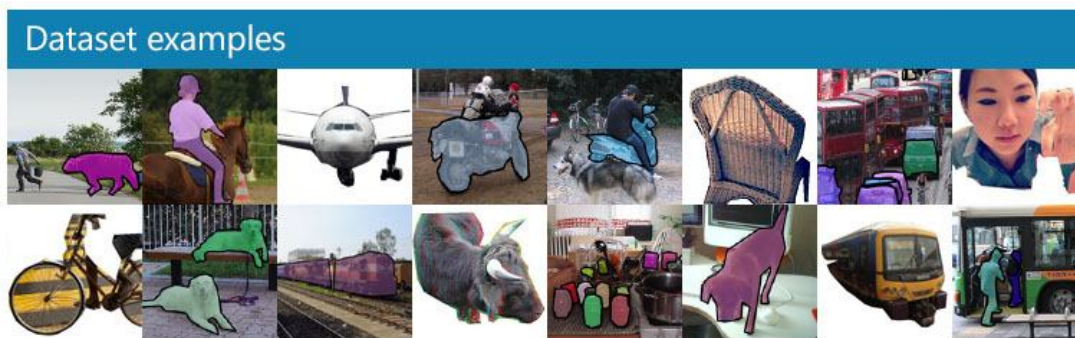


Рис. 2. Пример изображений датасета COCO.

2. Для улучшения сходимости алгоритма коэффициент обучения изменяется циклично (рис. 3).

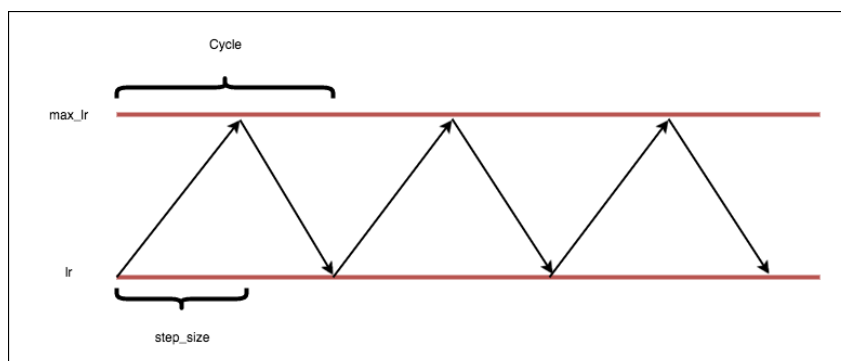


Рис. 3. Циклический коэффициент обучения.

Использование подхода Transfer Learning позволяет уменьшить количество размеченных рентгеновских снимков, необходимых для решения.

#### Описание используемых данных

Данные были собраны в Сибирском федеральном округе. Размер собранной выборки – 4387 изображений. Данные обезличены и размечены экспертами. Данные хранятся в формате DICOM (Digital Imaging and Communications in Medicine). 70% набора данных (3071 изображение) использовалось в качестве обучающих, 20% (877 изображений) – для настройки алгоритма (валидационная выборка) и 10% (439 изображений) – для тестирования.

#### Оценка результатов работы модели

Работа модели, описываемой в этой статье, оценивалась по метрике, вычисляемой как средняя точность на разных уровнях меры intersection over union (IoU). IoU набора обнаруженных контуров областей и истинных контуров областей вычисляется как:

$$IoU(A, B) = \frac{A \cap B}{A \cup B} \quad (1)$$

IoU является мерой величины перекрытия между двумя контурами (или, в более общем случае – двумя объектами). Она вычисляет размер перекрытия между двумя объектами, деленный на общую площадь двух объединенных объектов.

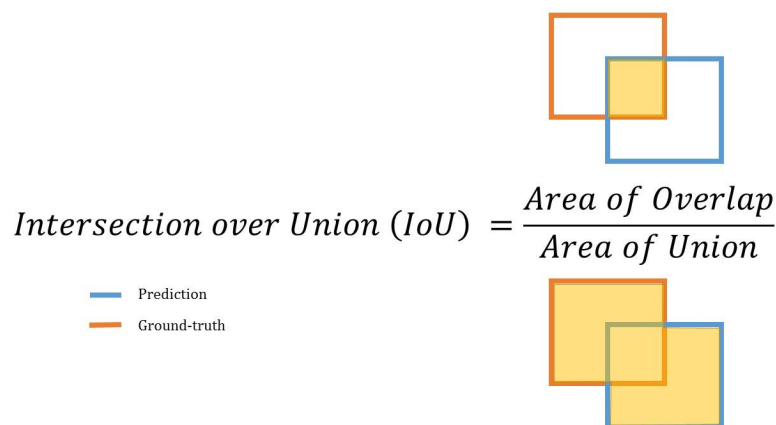


Рис. 4. Визуальное представление меры intersection over union (IoU).

На рис. 4 два контура перекрываются, но область перекрытия несущественна по сравнению с областью, занятой обоими объектами вместе. IoU будет низкой и, вероятно, не будет считаться «попаданием» при более высоких пороговых значениях IoU.

Метрика измеряет диапазон пороговых значений IoU в каждой точке, вычисляя среднее значение точности. Пороговые значения находятся в диапазоне от 0,4 до 0,75, с размером шага 0,05: (0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75). Другими словами, при пороге 0,5 предсказание объекта считается успешным, если пересечение его области с истинным объектом больше 0,5.

При каждом пороговом значении  $t$  значение точности вычисляется на основе количества истинных положительных (TP), ложных отрицательных (FN) и ложных положительных (FP) результатов, возникающих при сравнении прогнозируемого объекта со всеми истинными объектами:

$$\frac{TP(t)}{TP(t) + FP(t) + FN(t)} \quad (2)$$

Истинный положительный результат засчитывается, когда обнаруженный объект соответствует истинному объекту с уровнем IoU выше порога. Ложный положительный результат указывает

на то, что обнаруженный объект не имел соответствующего ему истинного объекта. Ложный отрицательный результат означает, что истинный объект не был обнаружен.

Если для какого-либо изображения вообще нет истинных объектов, то любое количество предсказаний (ложных срабатываний) приведет к получению значения метрики, равной нулю, которое будет включено в усредненную точность.

Средняя точность одного изображения вычисляется как среднее из указанных значений точности на каждом пороге IoU:

$$\frac{1}{|\text{thresholds}|} \sum_t \frac{TP(t)}{TP(t) + FP(t) + FN(t)}. \quad (3)$$

Наконец, оценка, возвращаемая метрикой, является средним значением, полученным по индивидуальным средним значениям каждого изображения в тестовом наборе данных.

### Результаты

В таблице представлены значения метрики качества при оценке работы алгоритма на тестовых данных. Результаты были усреднены по итогам 10 попыток. В каждой попытке деление выборок выполнялось заново случайным образом.

Алгоритм	Обуч.	Вал.	Тест
Masc R-CNN	0.150	0.148	0.142
Masc R-CNN + COCO	0.170	0.169	0.167
Masc R-CNN + COCO + циклический коэффициент обучения	0.179	0.176	0.175

Качество распознавания сниженной прозрачности легких пока мало в сравнении с результатами, получаемыми экспертом. В дальнейшей работе предполагается улучшить качество распознавания за счет увеличения размера выборки и оптимизации алгоритма. Однако текущая разработка также может быть полезна исследователям и практикующим врачам как дополнительное средство для проверки результатов, поскольку может способствовать предотвращению врачебной ошибки.

1. Обзор современного состояния e-health / А.А. Ахмед, М.А. Блинников, Р.Я. Пирмагомедов и др. // Информационные технологии и телекоммуникации. – 2017. – Т. 5, № 3. – С. 1-13.

2. Рынок мобильной медицины в России и в мире: основные тенденции и прогнозы / J'son & Partners Consulting, Russian Analytics. March 10, 2014. URL: [http://json.tv/en/ict\\_telecom\\_analytics\\_view/rynok-mobilnoy-meditsiny-v-rossii-i-v-mire-osnovnye-tendentsii-i-prognozy](http://json.tv/en/ict_telecom_analytics_view/rynok-mobilnoy-meditsiny-v-rossii-i-v-mire-osnovnye-tendentsii-i-prognozy)

3. Della, Mea V. What is e-Health (2): The death of telemedicine? // Journal of medical Internet research. – 2001. – Vol. 3, № 2.

4. Oh, H. et al. What is eHealth (3): A systematic review of published definitions // Journal of medical Internet research. – 2005. – Vol. 7, № 1.

5. Ahern, D.K., Kreslake, J.M., Phalen, J.M. What is eHealth (6): perspectives on the evolution of eHealth research // Journal of medical Internet research. – 2006. – Vol. 8, № 1.

6. Demner-Fushman, D., Chapman, W.W., McDonald, C.J. What can natural language processing do for clinical decision support? // Journal of biomedical informatics. – 2009. – Vol. 42, № 5. – P. 760-772.

7. Yuan, M.J. et al. Watson and healthcare, How natural language processing and semantic search could revolutionize clinical decision support // IBM developerWorks, IBM Corporation. – 2011.

8. Арутюнян, Г. Об использовании когнитивной системы IBM Watson Analytics для анализа данных о работе сердца / Блог компании IBM. 2016. URL: <https://habr.com/company/ibm/blog/400493/>

9. Kontzer, T. Deep Learning Drops Error Rate for Breast Cancer Diagnoses by 85% / The NVIDIA Blog: Deep Learning. 2016. URL: <https://www.nvidia.ru/object/deep-learning-breast-cancer-diagnosis-blog-ru.html>

10. Kubota, T. Deep learning algorithm does as well as dermatologists in identifying skin cancer / Stanford News. 2017. URL: <https://news.stanford.edu/2017/01/25/artificial-intelligence-used-identify-skin-cancer/>

11. IBM Debuts Watson Imaging Clinical Review, the First Cognitive Imaging Offering / IBM News. 2017. URL: <https://www-03.ibm.com/press/us/en/pressrelease/51643.wss>

12. Xu, J. et al. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images // Neurocomputing. – 2016. – Vol. 191. – P. 214-223.

13. Иванова, Е.В., Биличенко, Т.Н., Чучалин, А.Г. Заболеваемость и смертность населения трудоспособного возраста России по причине болезней органов дыхания в 2010-2012 гг. // Пульмонология. – 2015. – Т. 25, № 3. – С. 291-297.
14. Таубалдинова, Н.А. Распространенность заболеваемости и применение физиотерапии в лечении пневмонии // Вестник Казахского национального медицинского университета. – 2013. – №. 1.
15. Metwally, O. Why isn't machine learning more widely used for medical diagnoses? / Quora. 2013. URL: <https://www.quora.com/Why-isn%E2%80%99t-machine-learning-more-widely-used-for-medical-diagnoses/answer/Omar-Metwally>
16. Shen, D., Wu, G., Suk, H.I. Deep learning in medical image analysis // Annual review of biomedical engineering. – 2017. – Vol. 19. – P. 221-248.
17. Suk, H.I. et al. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis // NeuroImage. – 2014. – Vol. 101. – P. 569-582.
18. Shen, W. et al. Multi-scale convolutional neural networks for lung nodule classification // International Conference on Information Processing in Medical Imaging. – Springer, Cham, 2015. – P. 588-599.
19. Cheng, J.Z. et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans // Scientific reports. – 2016. – Vol. 6. – P. 24454.
20. Roth, H.R. et al. Improving computer-aided detection using convolutional neural networks and random view aggregation // IEEE transactions on medical imaging. – 2016. – Vol. 35, № 5. – P. 1170-1181.
21. Li, R. et al. Deep learning based imaging data completion for improved brain disease diagnosis // International Conference on Medical Image Computing and Computer-Assisted Intervention. – Springer, Cham, 2014. – P. 305-312.
22. Hoo-Chang, S. et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning // IEEE transactions on medical imaging. – 2016. – Vol. 35, № 5. – P. 1285.
23. Gupta, A., Ayhan, M., Maida, A. Natural image bases to represent neuroimaging data // International conference on machine learning. – 2013. – P. 987-994.
24. Brosch, T. et al. Manifold learning of brain MRIs by deep learning // International Conference on Medical Image Computing and Computer-Assisted Intervention. – Springer, Berlin, Heidelberg, 2013. – Vol. 633-640.
25. Kaiming, H., Gkioxari, G., Dollár, P., Girshick, R. Mask R-CNN // 2017 IEEE International Conference on Computer Vision (ICCV). – 2017. – P. 2980-2988.