

# Семантическая сегментация рентгенограмм легких на основе нейронной сети класса U-net

Н.С. Демин

*Институт систем обработки изображений - филиал  
ФНИЦ «Кристаллография и фотоника» РАН  
Самарский национальный исследовательский университет  
им. академика С.П. Королева  
Самара, Россия  
volfgunus@gmail.com*

Н.Ю. Ильясова

*Институт систем обработки изображений - филиал  
ФНИЦ «Кристаллография и фотоника» РАН  
Самарский национальный исследовательский университет  
им. академика С.П. Королева  
Самара, Россия  
ilyasova.nata@gmail.com*

**Аннотация**—В данной работе рассматривается семантическая сегментация рентгенограмм легких с использованием Transfer learning. Используется сеть типа U-net, у которой в качестве энкодера были взяты веса сети EfficientNet, предобученные на наборе ImageNet. В результате дообучения сети в течении 50 эпох была достигнута точность сегментации выше 90% для классов легкие и ребра, и 80% для класса сердце.

**Ключевые слова**—рентгенограммы легких, сегментация, трансферное обучение.

## 1. ВВЕДЕНИЕ

В настоящее время несмотря на то, что в основном применяются такие методы лучевой диагностики, как МРТ, имеющие в свою очередь высокую диагностическую информативность, обследования этими методами имеют определенные недостатки: дорогое оборудование и специализированное программное обеспечение. Однако, методы лучевой диагностики закрепились как дешевая и надежная процедура в начальных этапах исследований патологий. Согласно медицинским исследованиям, с помощью рентгенографии грудной клетки, можно выявлять патологические пертурбации анатомических структур, расположенных в грудной полости, а также исследовать изменения костей грудной клетки [1]. Применение методов машинного обучения и нейронных сетей в сфере анализа биомедицинских изображений облегчает раннюю диагностику и лечение заболеваний [2]. Согласно результатам медицинских исследований, на основе отслеживания изменения размеров легких, сердце и параметров костной части грудной клетки как ребра и ключицы, многие заболевания, в том числе хронические, могут быть диагностированы уже на их ранних этапах [3]. Выявлено, что путем наблюдения масса легких неравномерно распределена в органах, а именно с большей плотностью в нижних лепестках. На основе отслеживания изменения размеров легких и параметров ребер многие заболевания могут быть диагностированы уже на их ранних стадиях [4]. Сегментация рентгеновских изображений с помощью нейронных сетей играет большую роль в анализе заболеваний легких [5]. Прорыв в глубоком обучении с использованием сверточных нейронных сетей можно считать разработку AlexNet [6]. Эта сеть, победившая в конкурсе ImageNet, состояла из сверточных, MaxPool и полностью связанных слоев, а также Dropout, который использовался как регуляризатор, помогающий бороться с переобучением. Вслед за этим в статье [7] авторы представили U-Net, архитектуру нейронной сети, основанную на

сужающемся пути для захвата контекста изображения, за которым следует симметричный расширяющийся путь для обеспечения точной локализации. Сети типа U-Net нашли широкое применение в сегментации биомедицинских изображений.

В рамках данной работы предлагается применение Transfer learning для семантической сегментации снимков рентгенограмм.

## 2. СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ РЕНТГЕНОГРАММ

Задача семантической сегментации заключается в разбиении исходного растрового изображения или группы изображений на представляющие интереса области или сегменты и последующая одновременная классификация этих сегментов по выбранному классом. В качестве исходных данных был использован набор из 214 размеченных снимков рентгенограмм в формате bmp с размерами 512×512 пикселей. Изображения были размечены на четыре класса: сердце, легкие, ребра и ключицы. Данные были разбиты на обучающую и валидационную выборку в соотношении 200:14. Для увеличения количества данных было проведено их аугментирование. Использовались такие методы, как: HorizontalFlip, OneOf, RandomContrast, RandomGamma, Resize, RandomBrightness, ElasticTransform, GridDistortion, Compose, OpticalDistortion, PadIfNeeded, RandomSizedCrop, Normalize. Что позволило увеличить размеры выборки в 10 раз.

Для получения высоких результатов основных метрик семантической сегментации модели применен подход, который называется трансферным обучением, в обиходе более известен Transfer learning. Этот подход был предложен еще в 1992 [8]. Сама методика трансфер лернинга заключается в передаче знаний нейросети, которая была обучена на данных одной задачи (обычно используется нейронные сети, которые обучались на большом количестве данных) и применение этих знаний для решения другой задачи. В задачах компьютерного зрения данный подход, состоящий из двух этапов, нашел свое широкое применение в различных задачах. Свою оценку исследования трансфер лернинга дала команда из Google AI, где в своей работе [9] оценили вклад этой техники в задачах компьютерного зрения, в частности для задач, связанных с медицинскими изображениями. Результаты исследования показали, что использование данного подхода помогает улучшить, хоть и не значительно, результаты предсказаний нейросети для медицинских задач. Поэтому чаще используется двухэтапный подход трансфер лернинга состоящий из предобучение, когда сеть обучается на большом датасете

с последующей множественной классификацией, и Fine-tuning, когда используется веса из этапа предобучения для целевой нейросети, которая впоследствии дообучается на данных поставленной задачи.

В рамках данной работы в качестве предобученной сверточной нейронной сети был использован класс моделей называемые как EfficientNets, которые в свою очередь, как предполагают ее авторы, порождаются после изменения разрешение изображений в сетке, масштабирование моделей (Скейлинг) и балансирование количество каналов сети [10]. Под операцией скейлинга понимают процесс, когда фиксируются производимые внутри архитектуры сверточной нейронной сети операции и меняются лишь глубина (количество повторений одних и тех же модулей), ширина (количество каналов в свёртках) и разрешение. В работе [11] была предложена улучшенная методология, названная как Noisy-student, для получения более высоких результатов метрик на наборе изображений из датасета ImageNet. Веса данной модели использовались для дообучения целевой задачи семантической сегментации рентгенограмм. Для неё, которая по сути является разновидностью задачи классификации, для выходного слоя была выбрана функция активации Softmax. В качестве метрик оценки работы использовались Dice (Коэффициент Соренсена – Дайса) и IoU (Коэффициент Жаккара).

### 3. РЕЗУЛЬТАТЫ СЕГМЕНТАЦИИ

С целью сопоставления предсказаний и реальности в области машинного обучения используется матрица ошибок или confusion matrix – это таблица с различными комбинациями прогнозируемых и фактических значений. Прогнозируемые значения описываются как положительные и отрицательные, а фактические – как истинные и ложные. Обычно матрица ошибок используется для оценки точности моделей в задачах классификации. Но задачу семантической сегментации можно рассматривать как частный случай этой проблемы, поэтому confusion matrix актуальна и для измерения точности предсказаний пикселей. Результаты представлены в таблице 1.

Таблица 1. МАТРИЦА ОШИБКИ

		Фактический класс				
		Фон	Легкие	Ребра	Ключицы	Сердце
Предсказанный класс	Фон	0.7560	0.0232	0.0516	0.3680	0.0339
	Легкие	0.0352	0.9290	0.0936	0.0892	0.0064
	Ребра	0.1763	0.0433	0.8410	0.2044	0.0113
	Ключицы	0.0104	0.0015	0.0045	0.3382	0
	Сердце	0.0219	0.0028	0.0091	0	0.9521

В таблице 1 показаны основные классы сегментации, так же указан класс фон – все что не попадает ни под один из классов. Как видно из таблицы проблемным классом являются ключицы, это связано с тем, что это класс, который представлен на изображениях в меньшем размере, чем остальные и часто его пиксели относятся к фону или к ребрам. У остальных классов полученная точность выше 0,8.

Таблица 2. ЗНАЧЕНИЯ МЕТРИК

Метрики	Датасет	Ребра	Ключицы	Легкие	Сердце
IoU	Training	0.8302	0.7169	0.9244	0.7733
	Validation	0.8472	0.7248	0.9401	0.7950
Dice loss	Training	0.1247	0.2072	0.0626	0.1902
	Validation	0.1183	0.2140	0.0558	0.1807

В таблице 2 представлены значения основных метрик.

### 4. ЗАКЛЮЧЕНИЕ

В данной работе рассматривалось применение transfer learning для задачи семантической сегментации рентгенограмм легких. Для нейронной сети класса U-net, у которой в качестве энкодера использовалась сеть EfficientNets, предобученная на наборе ImageNet, было проведено дообучение для сегментации рентгеновских снимков на 4 класса: сердце, легкие, ребра и ключицы. В результате была получена нейронная сеть, позволяющая выделять классы легкие и ребра с точностью выше 90%.

### БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке РФФИ (грант № 19-29-01135), Министерства науки и высшего образования РФ, в рамках выполнения работ по государственному заданию ФНИЦ «Кристаллография и фотоника» РАН.

### ЛИТЕРАТУРА

- [1] Линденбратен, Л.Д. Медицинская рентгенология / Л.Д. Линденбратен, Ф.М. Лясс. – М.: Рипол Классик, 1986. – 368 с.
- [2] Дюдин, М.В. Нейросетевые модели принятия решений для диагностики заболеваний легких на основе анализа флюорограмм грудной клетки / М.В. Дюдин, Р.А. Томакова, М.В. Томаков // Биомедицинская радиоэлектроника. – 2014. – № 9. – С. 12-16.
- [3] Mathieson, J.R. Chronic diffuse infiltrative lung disease: comparison of diagnostic accuracy of CT and chest radiography / J.R. Mathieson, J.R. Mayo, C.A. Staples, N.L. Müller // Radiology. – 1989. – Vol. 171(1). – P. 111-116.
- [4] Verschakelen, J.A. The role of high-resolution computed tomography in the work-up of interstitial lung disease // Curr Opin Pulm Med. – 2010. – Vol.16(5). – P. 503-510.
- [5] Zhou, B. Lung mass density analysis using deep neural network and lung ultrasound surface wave elastography / B. Zhou, X. Zhang // Ultrasonics. – 2018. – Vol. 89. – P. 173-177.
- [6] Krizhevsky, A. ImageNet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // Commun. ACM. – 2017. – Vol. 60(6). – P. 84-90.
- [7] Ronneberger, O. U-Net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // International Conference on Medical image computing and computer-assisted intervention. – 2015. – P. 234-241.
- [8] Pratt, L.Y. Discriminability-based transfer between neural networks / L.Y. Pratt, S.J. Hanson, J.D. Cowan, C.L. Gile // Advances in Neural Information Processing Systems. – 1992. – Vol. 5. – P. 204-211.
- [9] Raghu, M. Transfusion: Understanding transfer learning for medical imaging / M. Raghu, C. Zhang, J. Kleinberg, S. Bengio // Advances in Neural Information Processing Systems. – 2019. – P. 3342-3352.
- [10] Mingxing, T. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks / T. Mingxing, Q. Le // Proceedings of the 36th International Conference on Machine Learning. – 2019. – P. 6105-6114.
- [11] Xie, Q. Self-training with noisy student improves imagenet classification / Q. Xie, M.T. Luong, E. Hovy, Q.V. Le // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2020 – P. 10687-10698.