

КОМПЬЮТЕРНАЯ ГРАФИКА И ВИЗУАЛИЗАЦИЯ

УДК 004.421.6

УЛУЧШЕНИЕ СЕГМЕНТАЦИИ ПАТОЛОГИЙ ЛЕГКИХ И ПЛЕВРАЛЬНОГО ВЫПОТА НА КТ-СНИМКАХ ПАЦИЕНТОВ С COVID-19

© 2021 г. Д. С. Лашенкова^{a,*}, А. М. Громов^{b,**},
А. С. Конушин^{a,***}, А. М. Мещерякова^{b,****}

^a Московский государственный университет имени М.В. Ломоносова,
119991 Москва, Ленинские горы, д. 1, Россия

^b ООО «Платформа Третье Мнение»,
121205 Москва, территория Сколково инновационного центра, ул. Нобеля, д. 7, Россия

*E-mail: daria.laschenova@graphics.cs.msu.ru

**E-mail: alexander.gromov@3opinion.ai

***E-mail: anton.konushin@graphics.cs.msu.ru

****E-mail: ceo@3opinion.ai

Поступила в редакцию 10.10.2020 г.

После доработки 20.10.2020 г.

Принята к публикации 12.01.2021 г.

В 2020 пандемия коронавируса затронула миллиарды людей по всему свету и заставила пересмотреть отношение к системам здравоохранения и к методам, используемым в современной медицине. Ввиду высокой нагрузки на радиологов и врачей появилась необходимость автоматических систем выявления патологий на медицинских исследованиях. Множество работ, посвященных работе с КТ-снимками пациентов с Covid-19, предполагают внедрение в системы медицинской помощи. Но улучшение по “классическим” метрикам вроде mAP или IoU по всем исследованиям не всегда отображает улучшение модели с точки зрения врачей. В данной работе было предложено считать метрики, усредняя не по всем исследованиям, а по группам в зависимости от размера патологий, а также оценивать количество ложноположительных участков найденных вне легких, поскольку наличие таких участков очень негативно оценивается врачами. Так же был предложен метод, улучшающий сегментацию патологий легких и плеврального выпота, с учетом замечаний, которые были высказаны выше.

DOI: 10.31857/S0132347421030067

1. ВВЕДЕНИЕ

Covid-19 — болезнь, вызываемая вирусом SARS-CoV-2. Часто встречающееся осложнение при болезни — вирусная пневмония.

Ранняя диагностика осложнений может уменьшить время и интенсивность лечения. Обычно для нее используется компьютерная томография. КТ-исследование — это радиологическое трехмерное исследование части тела, сконструированное компьютером из последовательности плоскостных поперечных срезов сделанных вдоль одной оси. Несмотря на то, что КТ-исследование не является основным в диагностике Covid-19, оно позволяет осуществлять диагностику осложнений, приоритизацию пациентов по степени тяжести осложнений, оценивать динамику болезни. Радиологи находят на исследовании различные патологии, оценивают их тип и размер. Затем эта информация используется для того, чтобы оценить состояние пациента и назначить лечение. Выделяют 5 степе-

ней тяжести пациента: КТ-0, КТ-1, КТ-2, КТ-3, КТ-4. Они определяются по доле объема пораженных легких. Это позволяет назначить лечение и понять, нужна ли госпитализация пациенту. В редких случаях при тяжелой степени появляется плевральный выпот — избыточное скопление жидкости в плевральной полости. Для врачей важно обнаружить данную патологию и оценить объем жидкости и ее тип.

Целью этого исследования является улучшение системы, автоматически сегментирующей легкие, патологии и плевральный выпот на КТ-снимке, чтобы можно было оценить процент поражения легких и объем плеврального выпота.

2. ОБЗОР ОБЛАСТИ

Поскольку Covid-19 стремительно распространился и еще продолжает распространяться, специалисты в компьютерном зрении начали поиск

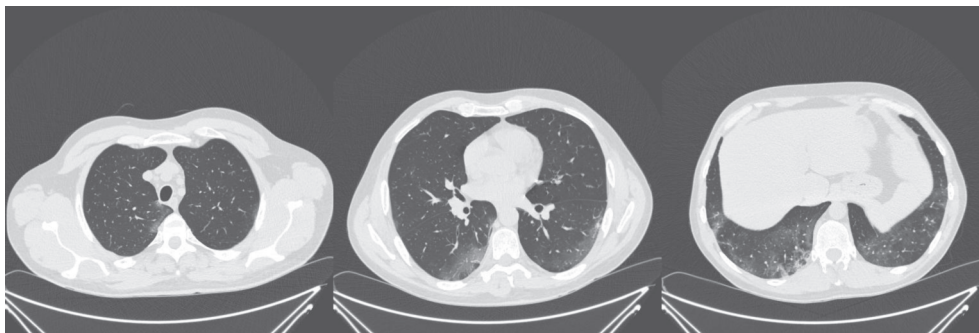


Рис. 1. Пример трех срезов из одного КТ-исследования.

решений, которые могут помочь радиологам в их работе.

В самом начале, до массового производства тестов на вирус SARS-CoV-2, было важно предсказать, насколько вероятно, что эта пневмония вызвана именно коронавирусом. В связи с этим появлялось много работ по задаче классификации: Linda Wang [1] представила сеть COVID-Net, которая определяла, болен ли пациент и болен ли он Covid-19. Lin Li предложил сеть COVNet [2] для решения аналогичной задачи. Но подобные решения потеряли релевантность в связи с тем, что доступность теста на вирус стала выше, чем возможность делать КТ-исследование.

Для того, чтобы систему можно было использовать для задач маршрутизации и отслеживания динамики состояния пациента, необходимо, чтобы она решала задачу сегментации патологий и определенных типов патологий. Fei Shan предложил 3D-модель VB-Net [3] для сегментации патологий, долей легкого и сегментов легкого, используя стратегию human-in-the-loop. Parham Yazdekhashty, Ali Zindar [4] используют U-Net-подобную сеть, но используют два декодера для предсказания класса легких и класса патологий, а затем объединяют результаты для более качественного поиска патологий. Кто-то объединяет задачи классификации и сегментации, как, например, Amine Amyar [5].

3. ВЫБОРКА

Выборка была предоставлена компанией ООО «Платформа Третье Мнение» [6]. Она содержит 938 КТ-исследований легких пациентов с Covid-19. Исследования были получены в формате dicom. Предварительная обработка данных не производилась. Примеры срезов можно увидеть на рис. 1.

10–20 срезов из каждого исследования были отданы на разметку радиологам, что дало 18673 изображения со срезами легких. Радиологи с помощью полигонов выделили на них области, содержащие легкие, патологии легких и плевральный выпот.

Выборка была разделена на 2 части: тренировочную и тестовую, содержащие 85 и 15% соответственно.

Стоит отметить, что радиологи обычно рисуют полигоны с более сглаженными границами, а модель обычно дает более точные границы. Также некоторые патологии имеют размытые границы на изображении, что усложняет построение точной маски. Это приводит к тому, что значения метрик будут не близки к идеальным.

4. МЕТРИКИ

4.1. Общие метрики

В качестве метрик используются IoU и среднее AP по исследованиям.

Полнота определяется как количество верно определенных положительных пикселей, деленное на общее количество положительных пикселей. Точность определяется как количество верно определенных положительных пикселей, деленное на общее количество пикселей, отмеченных, как положительные. AP вычисляется как площадь под графиком точность–полнота, где каждая точка графика показывает значение точности и полноты при различных порогах, по которым отсекаются пиксели, определенные моделью как положительные.

IoU показывает, насколько сильно пересекается маска в разметке и полученная моделью:

$$IoU = \frac{TP}{FN + TP + FP}$$

Стоит также заметить, что при выбранных метриках ошибка в малом количестве пикселей может приводить к различному падению метрик в зависимости от того, сколько истинных пикселей было в размеченной маске. Сложная ситуация возникает с исследованиями, на которых нет пикселей определенного класса. По ним нельзя вычислить mAP. Если пикселей этого класса не нашлось моделью, то не вычисляется и IoU (или его можно принять 1). Если же нашелся хотя бы один, то IoU

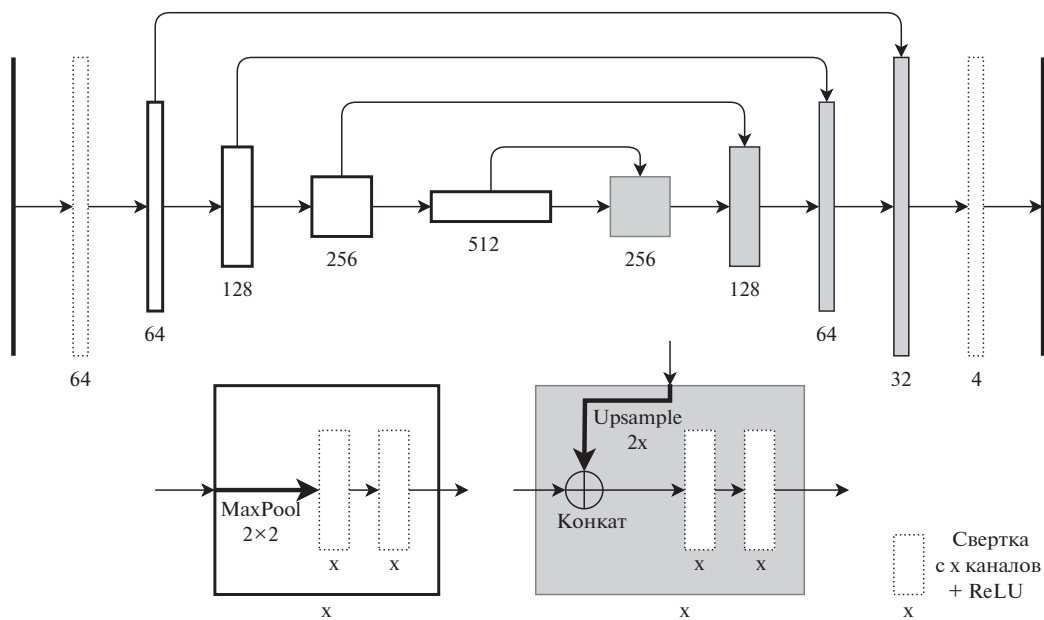


Рис. 2. Базовая архитектура сети.

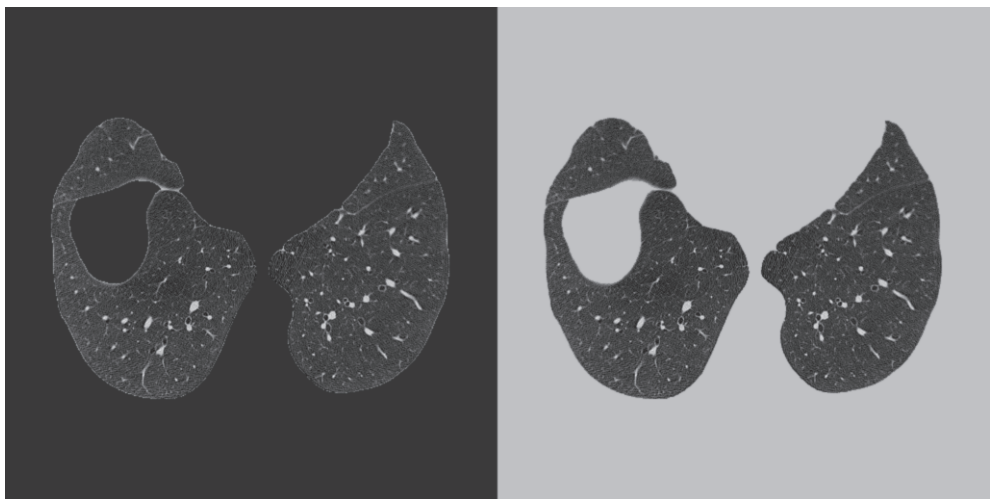


Рис. 3. Пример среза маскированного константой, соответствующей воздуху (слева) и ткани (справа).

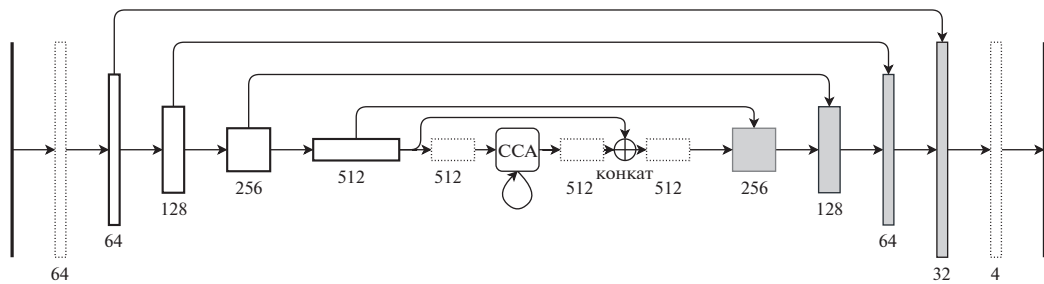


Рис. 4. Архитектура сети с модулем RCCA.

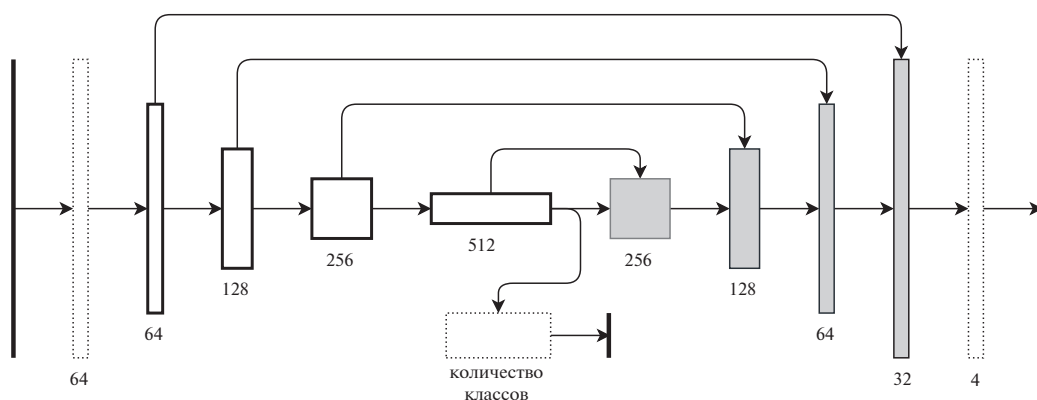


Рис. 5. Архитектура сети со стабилизирующей функцией потерь.

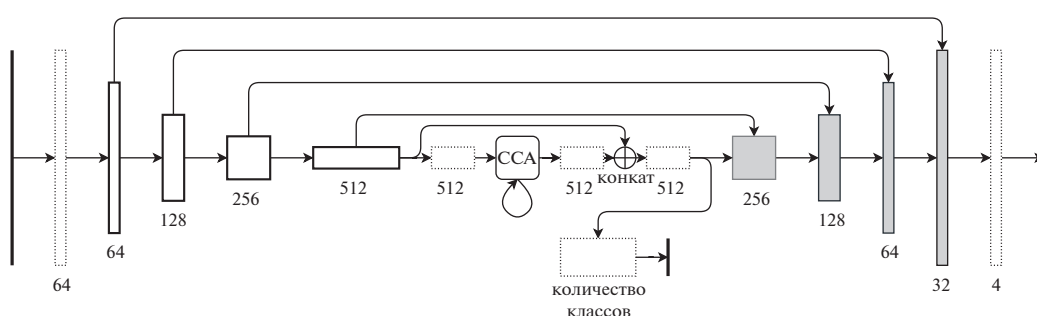


Рис. 6. Архитектура сети с модулем RCCA и стабилизирующей функцией потерь.

становится 0, что плохо отражает качество работы модели. Также небольшие ошибки в ответах будут ухудшать метрики на исследованиях с маленьким количеством истинно положительных пикселей гораздо сильнее, чем с большим.

Поэтому было принято решение считать общие метрики по 3 группам исследований: на исследованиях с отсутствующим классом считать среднее количество ложноположительных пикселей на исследование. На исследованиях, где пиксели класса есть, но их меньше 10000 на все исследование и на остальных считать mAP и IoU раздельно.

4.2. Метрики областей

В результате диалога с радиологами оказалось, что даже при наличии улучшения модели по общим метрикам, врачи иногда не только не замечали этого улучшения, но еще и считали, что модель работает менее точно. Одной из причин этого является то, что модель находила участки легких и патологий вне легких, например, выделяла области кишечника. Для этого для каждого класса выделялись связанные компоненты и для каждой считалось, пересекается ли она с выделенными радиологами участками легких. Если не пересекается,

то считаем, что компонента была определена вне легких. В качестве метрики будем использовать количество компонент вне легких разного размера: маленьких (до 100 пикселей), средних (от 100 до 500) и больших (более 500). Больше внимания стоит уделять именно большим компонентам, поскольку они заметны радиологам больше всего.

5. ПРЕДЛОЖЕННЫЙ МЕТОД

В данной работе решается задача уменьшения компонент, найденных вне легких, поэтому будут предложены способы устранения таких компонент. Базовым решением является сеть, подобная U-Net [7], архитектура представлена на рис. 2 обученная на классах “фон”, “легкое”, “патология”, “плевральный выпот” с функцией потерь отрицательная кросс-энтропия.

5.1. Предварительная сегментация легких

Одним из подходов является предварительная сегментация участков легких, а затем поиск патологий в маскированных легких. Плюсом такого подхода будет то, что при корректной сегментации легких компоненты вне них не найдутся. Также сегментацию легких можно проводить на

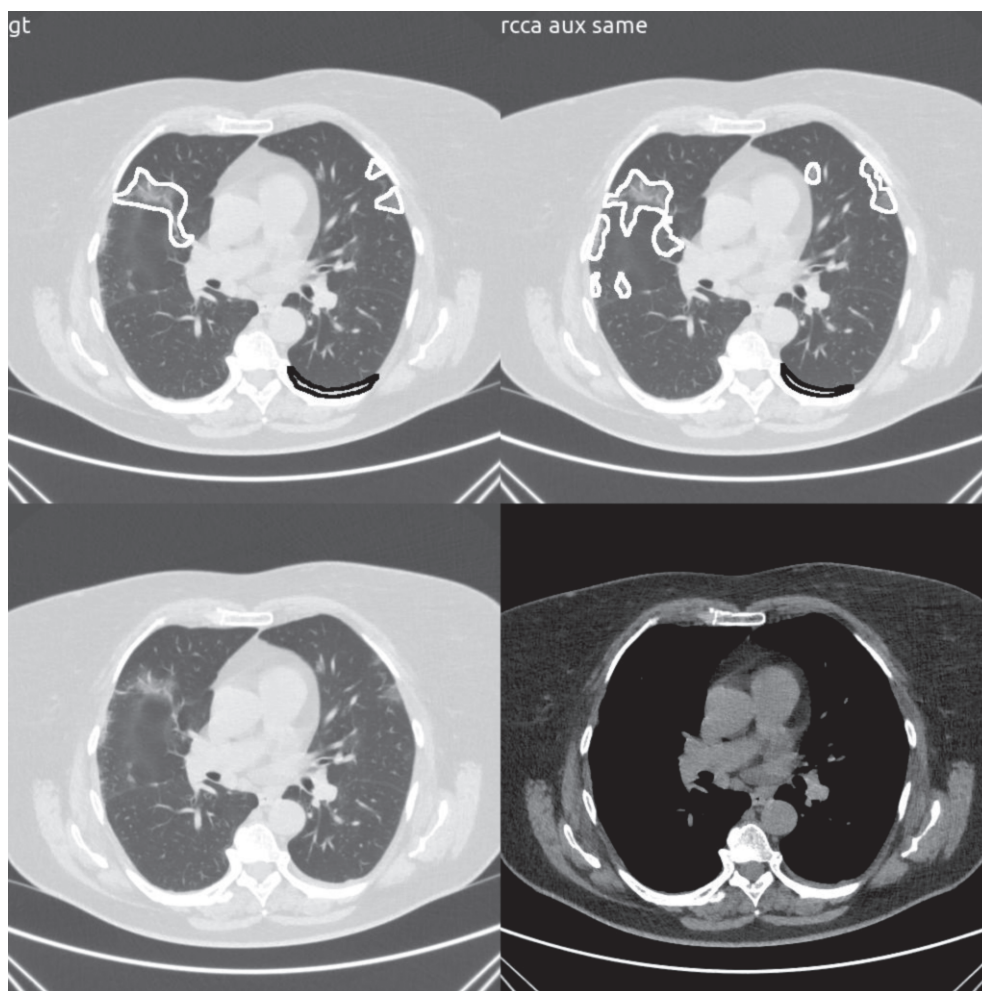


Рис. 7. Пример работы алгоритма на изображении с плевральным выпотом.

изображениях меньшего масштаба, что повысит точность сегментации, поскольку легкие занимают достаточно большую часть среза, в отличие от патологий. Минусами такого подхода является то, что в таком случае плевральный выпот будет маскирован, поскольку он не является частью легкого, а также то, что возрастут вычислительные и временные расходы на сегментацию легких.

Отдельным вопросом стоит то, какой константой следует заполнять маскированную область. С одной стороны ее можно заполнить значением, которое соответствует воздуху, а с другой стороны ткани, которая окружает легкие. Было опробовано оба варианта, пример можно видеть на рис. 3.

5.2. Введение модуля “RCCA”

Этот модуль был описан в статье CCNet: Criss-Cross Attention for Semantic Segmentation [8]. Основная идея внедрения этого модуля заключается в том, чтобы в какой-то момент информация перераспределялась со всего изображения, чтобы

сеть смогла более корректно отличать патологии и легкие от похожих на них структур по информации из контекста со всего изображения. В этой работе модуль помещен между энкодером и декодером сети U-Net, как показано на рис. 4.

5.3. Введение стабилизирующей функции потерь

Также было предложено вставить дополнительную голову сети после энкодера или после модуля RCCA (при наличии) и вычислять грубую сегментацию в маленьком размере для того, чтобы сеть научилась не выделять большие компоненты в неправильных местах. Предложено предсказывать два варианта сегментаций и сравнить их между собой. Можно предсказывать ту же самую карту, что и после декодера, но в меньшем разрешении и использовать ту же функцию потерь, а можно предсказывать вероятность того, что пиксель принадлежит к интересующим нас классам и использовать бинарную кросс-энтропию.

Таблица 1. Сравнение базового (baseline) метода и методов с предварительной сегментацией для класса “патология”. segm air – сегментация с заполнением константой, соответствующей воздуху, segm body – сегментация с заполнением константой, соответствующей плотности тела

	FP (0)	IoU (pat, s)	IoU (pat, l)	mAP (pat, s)	mAP (pat, l)
baseline	575.6	0.413	0.603	0.656	0.856
segm air	641.5	0.402	0.596	0.626	0.840
segm body	783.6	0.374	0.584	0.604	0.833

Таблица 2. Сравнение модулей по количеству обнаруженных связанных компонент вне легких

	Легкие			Патологии		
	small comp	med comp	big comp	small comp	med comp	big comp
baseline	4718	207	78	1051	393	52
rcca	5822	181	68	862	339	55
aux fore	6085	120	21	1144	237	20
aux same	6379	140	42	952	291	37
rcca aux fore	6356	152	30	835	354	47
rcca aux same	5788	128	31	736	330	41

6. РЕЗУЛЬТАТЫ

В таблицах ниже class_name, 0 означает исследования, на которых класс отсутствует, class_name, s – исследования, на которых в разметке содержится менее 10000 пикселей класса, class_name, l – исследования, на которых в разметке содержится более 10000 пикселей класса.

Так же приняты обозначения baseline – базовый метод, rcca в названии означает методы, использующие rcca, aux в названии означает наличие стабилизирующей функции потерь: aux fore для бинарной кросс-энтропии, различающей фон и интересные классы, aux same для кросс-энтропии после софтмакс, для предсказания аналогичных требуемым классам, но меньшего размера.

6.1. Исследование влияния предварительной сегментации легких

Исследования проводились для моделей, которые предсказывали класс “патология” на предварительно маскированных легких.

Как видно из табл. 1, сегментация ухудшает ситуацию и на тех срезах, где патологий нет, поскольку начинает выдавать больше ложноположительных результатов, и на срезах, где патологии присутствовали.

Отдельно стоит отметить, что на результат может сильно влиять выбранная константа для заполнения маскированных регионов, поэтому стоит заострять на этом внимание, если используется подобная техника.

6.2. Исследование влияния модуля “RCCA” и стабилизирующей функции потерь

Исследования проводились для моделей, которые предсказывали классы “легкое”, “патология”, “плевральный выпот”.

Для начала проанализируем то, как повлияло включение модуля на нахождение участков классов вне легких. Результаты представлены в табл. 2.

А также проанализируем влияние на общие метрики качества сегментации для классов “патология” (табл. 3) и “плевральный выпот” (табл. 4).

Сравнивая значения метрик, получаем, что применение стабилизирующих функций потерь уменьшает количество участков, найденных вне легких. При этом она чуть понижает значения метрик, которые были получены на изображениях с малым количеством патологий, зато увеличивает значения метрик, которые были получены на изображениях с большим количеством патологий. Также видно, что модуль RCCA сам по себе ухудшает значения метрик, но если поставить на его выходы стабилизирующую функцию потерь, то по общим метрикам он выигрывает на изображениях с большим количеством патологий и не сильно проигрывает на изображениях с малым количеством патологий, а также уменьшает количество ложных участков вне легких.

Примеры работы модели можно увидеть на рис. 7. В левой верхней строке слева разметка радиологов, справа результат работы модели, в нижней строке оригинальный срез с разными настройками просмотра.

Таблица 3. Сравнение модулей по общим метрикам для класса патологий

	FP (pat, 0)	IoU (pat, s)	IoU (pat, l)	mAP (pat, s)	mAP (pat, l)
baseline	575.6	0.413	0.603	0.656	0.856
rcca	525.9	0.412	0.615	0.627	0.856
aux fore	635.0	0.408	0.610	0.636	0.850
aux same	547.1	0.413	0.604	0.643	0.848
rcca aux fore	618.8	0.400	0.620	0.638	0.856
rcca aux same	613.9	0.409	0.627	0.656	0.861

Таблица 4. Сравнение модулей по общим метрикам для класса плеврального выпота

	FP (pl eff, 0)	IoU (pl eff, s)	IoU (pl eff, l)	mAP (pl eff, s)	mAP (pl eff, l)
baseline	961.6	0.383	0.632	0.590	0.803
rcca	407.9	0.320	0.636	0.556	0.807
aux fore	837.2	0.333	0.624	0.547	0.798
aux same	757.3	0.374	0.652	0.524	0.814
rcca aux fore	385.9	0.326	0.630	0.572	0.805
rcca aux same	698.7	0.369	0.660	0.596	0.812

7. ЗАКЛЮЧЕНИЕ

При оценке автоматической системы, которой будут пользоваться люди, необходимо пользоваться не только “классическими метриками” сегментации, но и метриками, которые будут оценивать применимость моделей для использования их людьми. Обычно при сравнении новых моделей, полученных в результате экспериментов, не получается модели, которая лучше остальных моделей во всем, поэтому стоит искать компромиссы.

СПИСОК ЛИТЕРАТУРЫ

1. Wang L., Lin Z. Q., Wong A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images // Scientific Reports. 2020. V. 10. № 1. P. 1–12.
2. Li, L., Qin, L., Xu, Z. et al. Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT // Radiology. 2020.
3. Shan F., Gao Y. et al. Lung infection quantification of covid-19 in ct images with deep learning, arXiv preprint arXiv:2003.04655, 2020.
4. Yazdekhasty P., Zindar A., Nabizadeh-ShahreBabak Z., Roshandel R., Khadivi P., Karimi N., Samavi S. Bifurcated Autoencoder for Segmentation of COVID-19 Infected Regions in CT Images, arXiv preprint arXiv:2011.00631, 2020.
5. Amyar A., Modzelewski R., Li H., Ruan S. Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation // Computers in Biology and Medicine. 2020. V. 126. P. 104037.
6. Third Opinion Platform, Limited Liability Company, <https://thirdopinion.ai/>.
7. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical image computing and computerassisted intervention. Springer, 2015. P. 234–241.
8. Huang Z., Wang X., Huang L., Huang Ch., Wei Y., Liu W. Ccnet: Criss-cross attention for semantic segmentation // Proceedings of the IEEE International Conference on Computer Vision. 2019. P. 603–612.