

Зависимость качества работы нейронных сетей от характеристик обучающих данных при работе со снимками УЗИ щитовидной железы

Э.А. Ехлаков, К.В. Цыгулева, М.Е. Дунаев, С.М. Захарова

Аннотация. Целью настоящей работы является проверка гипотезы о том, что качество работы нейросетевых моделей обнаружения и сегментации узловых образований на снимках УЗИ щитовидной железы практически не зависит от числа анализируемых снимков одного пациента, полученных за одно исследование. Для проверки использовались две глубоких архитектуры: YOLOv5 при решении задачи обнаружения и DeepLabV3 при решении задачи сегментации. При проведении экспериментов были использованы кинопетли (последовательности кадров) УЗИ щитовидной железы, содержащие более 7000 снимков поперечной и продольной проекций 166 пациентов. Оценка работы глубоких архитектур проводилась как на этапе их обучения и валидации, так и на этапе их тестирования. По результатам проведенных экспериментов было выявлено, что увеличение количества однотипных снимков в кинопетлях при неизменном количестве исследуемых пациентов не влияет на работу глубоких архитектур.

Ключевые слова – нейронные сети, машинное обучение, глубокие архитектуры, ультразвуковое изображение щитовидной железы, сегментация железы, обнаружение железы

I. ВВЕДЕНИЕ

В последние годы глубокие нейросетевые архитектуры обеспечили уникальные возможности и заметные прорывы при решении задач в различных областях [1], [2]. В клинической медицине было показано, что системы компьютерной диагностики, основанные на глубоком обучении, обеспечивают конкурентную, а иногда даже превосходящую диагностическую точность и эффективность по сравнению с опытными клиницистами [3], [4], [5], [6]. Хотя применение методов глубокого обучения к

Статья получена 19 ноября 2022.

Ехлаков Эдуард Владимирович, Национальный Исследовательский Ядерный Университет МИФИ, студент, fruha1980@gmail.com

Цыгулева Ксения Владимировна, Национальный Исследовательский Ядерный Университет МИФИ, студент, k151201@yandex.ru

Дунаев Максим Евгеньевич, Национальный Исследовательский Ядерный Университет МИФИ, аспирант, Max.dunaev@mail.ru
Захарова Светлана Михайловна, Национальный медицинский исследовательский центр эндокринологии, ст.н.с., Smzakharova@mail.ru

медицинским изображениям обычно используется для повышения диагностической эффективности, постоянно растет спрос на более совершенные глубокие нейронные сети для решения новых более сложных сценариев.

В настоящее время многие существующие исследования подчеркивают конкурентоспособность глубоких сверточных нейронных сетей (CNN) при их применении для диагностики различных заболеваний. Эти модели дают достойные результаты в задачах классификации, обнаружения и сегментации для выявления конкретного заболевания; однако во многих случаях они все еще уступают диагнозу человеческого уровня. Это происходит потому, что опытные клиницисты обычно используют дополнительные знания предметной области, чтобы делать диагностический вывод, а не полагаться только на медицинские изображения. Поэтому, глубокие нейронные сети, анализирующие снимки, могут делать лишь ассистирующие выводы в области слепой диагностики, но могут успешно обнаруживать интересные области на снимке или сегментировать узлы и другие новообразования.

Одной из важных проблем применения глубоких архитектур в медицине является подбор значительных объемов обучающих датасетов, т.к. в большинстве клиник наборы таких данных не велики. Поэтому есть искушение включить в обучающие выборки клиник много снимков из кинопетель своих пациентов. Чтобы понять насколько это эффективно мы в настоящем исследовании провели серию экспериментов.

Автоматизация процесса проведения слепой диагностики узловых образований щитовидной железы – комплексная задача. Данный вывод

основан на методе обнаружения врачами узловых образований: для постановки диагноза им необходимо сначала обнаружить узел, затем выделить и проанализировать его очертания, и наконец, классифицировать [7].

Поскольку ключевым этапом в процессе диагностики является выделение очертаний узлового образования – задача сегментации снимков УЗИ должна быть решена в полной мере.

Реализация алгоритма сегментации должна соответствовать следующим требованиям:

1. Независимость результата от входных данных. Модель должна быть обучена таким образом, чтобы формат, расположение, кадрирование и разрешение входных данных не влиял как на дополнительное обучение модели, так и на предсказания при тестировании и использовании.

2. Независимость результата от возможностей диагностических аппаратов. Поскольку в разных медицинских учреждениях используется разное оборудование, модель сегментации должна быть мобильной. Алгоритм необходимо обучить таким образом, чтобы он был готов к любым изменениям в данных. Для соответствия данному требованию предлагается расширить набор обучающих данных для алгоритма аугментированными данными.

3. Комплексное решение проблемы. Задача сегментации таких сложных изображений, как снимки УЗИ, не является тривиальной. В силу того, что на результат работы может влиять не только вид снимка при исследовании (продольный или поперечный), но и вид узлового образования, ширина шеи и т.д., поэтому необходимо обеспечить такой подход к реализации сегментации, чтобы минимизировать влияние описанных обстоятельств процесса диагностики.

4. Минимизирование требований к ручной разметке данных.

Поскольку на соответствие описанным выше требованиям влияет множество факторов, в этой статье описывается исследование влияния характеристик обучающих данных на результаты работы алгоритмов обнаружения и сегментации.

II. ПОСТАНОВКА ЗАДАЧИ ОБНАРУЖЕНИЯ И СЕГМЕНТАЦИИ УЗЛОВЫХ ОБРАЗОВАНИЙ ЩИТОВИДНОЙ ЖЕЛЕЗЫ

Формальные постановки задач обнаружения и сегментации приведены ниже.

Обнаружение.

Пусть дано X – множество признаковых описаний объектов. Отображение $f: X \rightarrow D_f$ является признаком f объекта a . Здесь D_f – множество допустимых значений признака. В рассматриваемом методе обнаружения объектом a является элемент разбиения изображения на ячейки определенного масштаба.

Каждый образ объекта $a \in X$ может быть охарактеризован значениями признаков $f_i, i = 1, \dots, r$, наборы которых одинаковы для всех объектов. Следовательно, вектор признаков объекта $a \in X$ может быть определен по $x = (f_1(a), \dots, f_r(a))$. Здесь вектор признаков можно отождествлять с самими объектами.

Для реализации задачи обнаружения необходимо построить функцию $F: X \rightarrow Y$, которая будет осуществлять сопоставление класса $y_i \in Y$ произвольному объекту из множества X с определенной вероятностью из пространства заранее определенного пространства распределения вероятностей. Здесь $Y = \{y_1, y_2, \dots, y_k\}$ является конечным множеством классов, разбиение на которое существует на всем множестве X [8].

Сегментация.

Задача семантической сегментации (пиксельной классификации изображений) может быть рассмотрена как задача нахождения функции оценки $h: X \rightarrow Y$ для каждого пикселя из пространства входного изображения X на пространство меток Y . Пространство меток может включать в себя семантические карты или теги классификации.

Рассматривая размеченные на классы изображения $(x, y) \in X \times Y$, можно предположить, что они принадлежат фиксированному неизвестному вероятностному распределению D , определенному на $X \times Y$.

Таким образом, задача нахождения h сводится к нахождению лучшего показателя из заранее заданного функционального пространства H (класс гипотез), ограничивающего h и выбранного на основе знаний о решаемой проблеме сегментации [9].

III. ИССЛЕДУЕМЫЕ АРХИТЕКТУРЫ

В настоящем исследовании рассматриваются две нейросетевые архитектуры: YOLOv5 для решения задачи обнаружения и DeepLabV3 для решения задачи сегментации.

YOLOv5.

Эта архитектура строится на подходе одноступенчатого детектора (One-Stage detector), т.е. предсказывает координаты ограничительных рамок (bounding box) и классифицирует найденные объекты в рамках конечного числа классов (рис. 1).

В качестве *BackBone* уровня используется CSPDarknet53 (сравнивалась с CSPResNext50, CSPDarknet53 и EfficientNet-B3), состоящая из сверточного базового слоя и CSP (Cross Stage Partial) блока, разделяющего входные данные на две половины и объединяющего их обратно. За счет такого устройства решается проблема исчезающего градиента.

На уровне *Neck* используется PANet, которая агрегирует признаки. Основная роль PANet заключается в улучшении процесса сегментации экземпляров за счет сохранения пространственной информации, которая, в свою очередь, помогает в правильной локализации пикселей для предсказания маски. Здесь модифицированный PANet объединяет соседние слои вместо добавления при использовании адаптивного объединения функций.

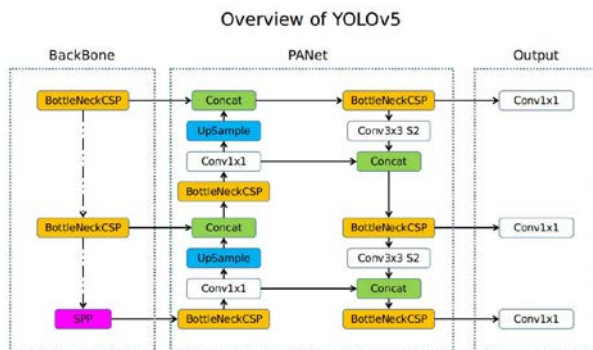


Рис. 1. Архитектура YOLOv5 [8].

Между слоем *BackBone* и *Neck* располагается дополнительный блок SPP (Spatial Pyramid Pooling), который необходим для связи плотных и сверточных слоев CSPDarknet53.

На уровне *Head* основная функция — определение границ боксов и выполнение классификации [8].

DeepLabV3.

Эта архитектура относится к семейству сетей DeepLab, представленных как инструмент эффективной семантической сегментации [10]. Отличительная черта DeepLab – слои жесткой свертки (Atrous Convolution).

Такой вид свертки (см. рис. 2) способен результативно разбирать поступающую многомасштабную контекстную информацию за счет использования фильтров, прореженных нулями, получая информацию как бы в разных масштабах изображения. Это позволяет контролировать поле зрения свертки и увеличивать его без использования дополнительных вычислительных ресурсов и без изменения количества параметров.

Архитектура DeepLabV3 состоит из нескольких convolutional-блоков, ASPP-блока и блока объединения. Сверточные блоки, не использующие atrous-алгоритм («block1», «block2», «block3») отвечают за выборку ценной информации из исходного изображения, теряя при этом информацию в силу уменьшения размерности изображения.

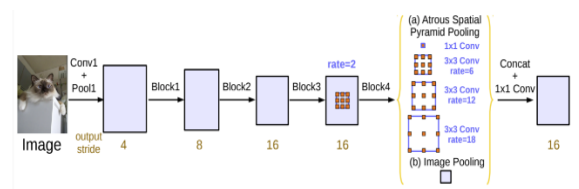


Рис. 2. Архитектура DeepLabV3 [10].

Основной особенностью DeepLabV3 является использование расширенной свертки (ASPP – Atrous Spatial Pyramid Pooling) с дополнительным Batch Normalization. Это позволяет использовать atrous-слои с разными показателями скоростей фильтров параллельно, а после обработки обобщать результаты в единую карту. ASPP в DeepLabV3 состоит из 256 фильтров и BN-слоев (одна свертка 1×1 и три свертки 3×3 со скоростями = (6, 12, 18) при выходном шаге равным 16).

IV. ДАННЫЕ ДЛЯ ОБУЧЕНИЯ

Для проведения экспериментов с описанными выше архитектурами, были использованы 261 кинопетля УЗИ щитовидной железы 166 пациентов. Эти кинопетли были записаны на аппаратах ультразвукового исследования LogiqE и GE Voluson E8. Количественные данные по распределению кинопетель по собственным их характеристикам представлены в таблицах 1 и 2.

Таблица 1 – Распределение снимков по аппаратам УЗИ и по виду проекции

Аппарат УЗИ Logiq	Аппарат УЗИ GE Voluson	Поперечный срез ЩЗ	Продольный срез ЩЗ
-------------------------	------------------------------	-----------------------	-----------------------

	E8		
113 видео-петель	53 видео-петли	148 видео-петель	113 видео-петель

21 видео-петля	53 видео-петли	35 видео-петель	54 видео-петли
----------------	----------------	-----------------	----------------

Перед началом обучения все кинопетли, предоставленные для обучения, были разделены на отдельные снимки. Из полученных снимков путем анализа кинопетли были выделены только те снимки, на которых присутствовал хотя бы один узел. Примеры снимков и их масок, подаваемых в качестве обучающей выборки, представлены на рисунке 3.

Таблица 2 – Распределение снимков классу TI-RADS

2 TI-RADS	3 TI-RADS	4 TI-RADS	5 TI-RADS
-----------	-----------	-----------	-----------

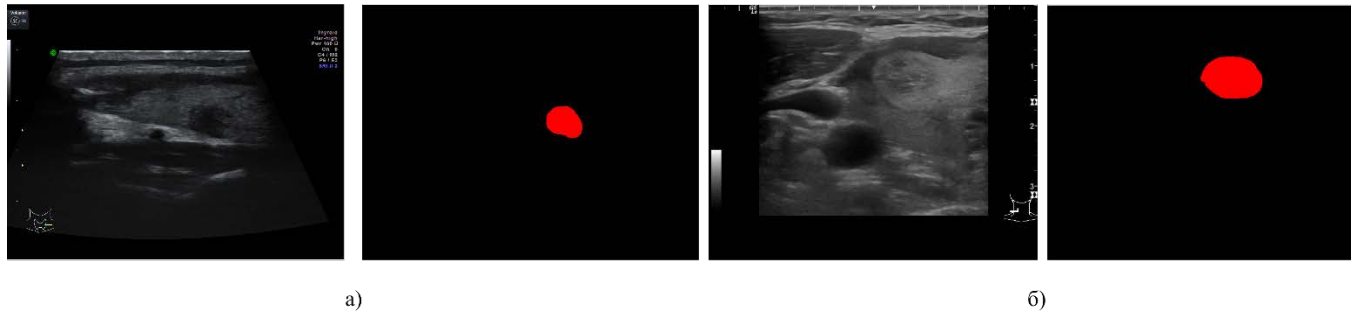


Рис. 3. Пример снимков из видео-петель: а) аппарат GE Voluson E8, продольная проекция б) аппарат Logiq, поперечная проекция.

V. ОБУЧЕНИЕ СЕТЕЙ

При проведении эксперимента на предоставленных снимках были обучены модели сетей YOLOv5 и DeepLabV3.

Обучение YOLOv5.

Для тренировки сетей использовалось 15 сетов данных. Датасеты подразделялись на 3 основные категории: поперечные изображения, продольные изображения и все изображения. Было составлено по 5 сетов каждой категории, где использовались: все изображения, каждое 3-е изображение, каждое 5-е, каждое 10-е и каждое 15-е. Для валидации наборы подразделялись только на продольные, поперечные и все изображения. В таблице 3 приведено количественное распределение снимков для каждого из наборов.

Таблица 3 – Распределение количества снимков при обучении YOLOv5

	Все	Поперечные	Продольные
15	552	204	259
10	704	338	436
5	1407	606	772
3	2343	1008	1288
ALL	7131	3115	3856

Таблица 4 – Распределение количества снимков для валидации YOLOv5

	Все	Поперечные	Продольные
15	105	178	200
10	184	191	114
5	123	136	214
3	136	140	209
ALL	108	178	155

15	510	409	303
10	510	409	303
5	510	409	303
3	510	409	303
ALL	510	409	303

В связи с архитектурными особенностями YOLOv5 (при достижении лучшего показателя валидации, сеть продолжает обучение еще 100 эпох и, если показатель метрики не улучшается, сеть останавливает обучение, и сохраняет лучший результат по валидации) существуют различия в количестве эпох. Распределение по количеству эпох обучения приведено в таблице 5.

Таблица 5 – Распределение количества тренировочных эпох при обучении YOLOv5

	Все	Поперечные	Продольные
15	105	178	200
10	184	191	114
5	123	136	214
3	136	140	209
ALL	108	178	155

Исходя из результатов, продемонстрированных на графиках (рис. 4), количество снимков не влияет на результаты обучения. Различия показателей метрик составляют менее 10%. При этом наилучший результат показали сети при

обучении на поперечных проекциях УЗИ щитовидной железы.

Обучение DeepLabV3.

Для тренировки сетей было составлено 6 наборов снимков. В каждый набор входили и поперечные, и продольные проекции УЗИ. В каждый из 6 наборов входили: все изображения, каждый 2-ой, каждый 4-й, каждый 8-й, каждый 16-й и каждый

32-ой. В таблице 6 приведено количественное распределение снимков для каждого из наборов. Модели обучались на одном количестве эпох и валидировались на одном наборе изображений.

Таблица 6 – Распределение количества снимков при обучении и условное обозначение на графиках рисунка 5

	Количество снимков в наборе	Условное обозначение
32	223	розовый
16	431	голубой
8	870	красный
4	1783	синий
2	3523	желтый
ALL	7131	серый

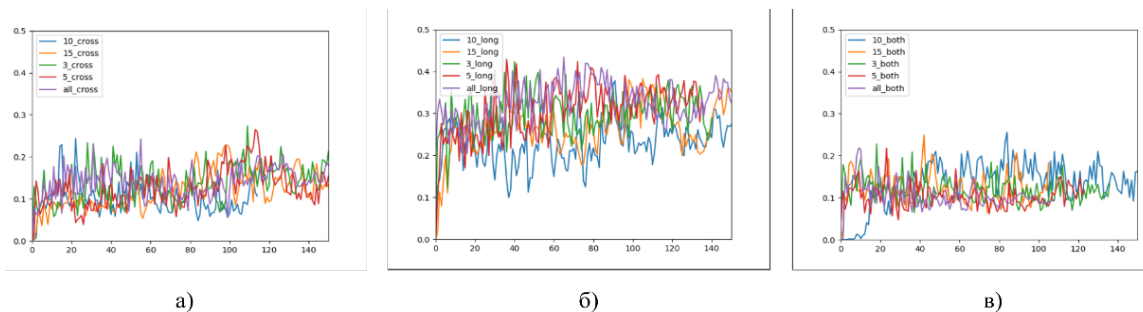


Рис. 4. Показатели метрики mAP-0.5 в процессе обучения моделей: а) модели с обучением на продольных снимках; б) модели с обучением на поперечных снимках; в) модели с обучением на всех снимках.

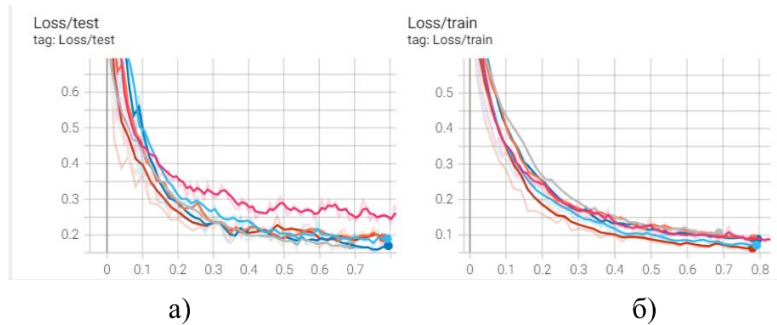


Рис. 5. Показатели метрики Loss в процессе обучения и при тестировании моделей: а) Loss при тестировании моделей; б) Loss при тренировке моделей.

По результатам обучения, представленным на рисунке 5, количество снимков не влияет на результаты обучения. Различия показателей метрик составляют менее 10%. При этом при обучении на каждом 32-ом снимке Loss при тестировании начинает сходиться медленнее. Однако за счет малого количества данных при тренировке модель быстрее переобучается.

VI. ПРОВЕДЕНИЕ ОЦЕНКИ КАЧЕСТВА ИСПОЛЬЗОВАНИЯ ГЛУБОКИХ АРХИТЕКТУР

Полученные в результате обучения модели были протестированы на снимках тестовой выборки. Часть результатов тестирования приведена на рисунках 6 и 7.

Наилучшие показатели при тестировании были получены при использовании моделей, обученных на снимках поперечных проекций

УЗИ. Наилучшие показатели относительно объема используемых данных при обучении были получены в соответствии с результатами

валидации: различия в метриках показателей правильности предсказаний варьируются в 10-15%.

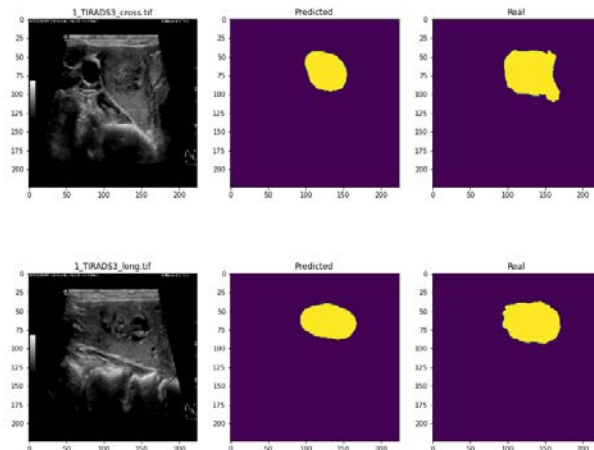


Рис. 6. Результаты тестирования DeepLabV3 для сегментирования моделями, обученными на снимках продольной проекции (верхний ряд) и на снимках поперечной проекции (нижний ряд).

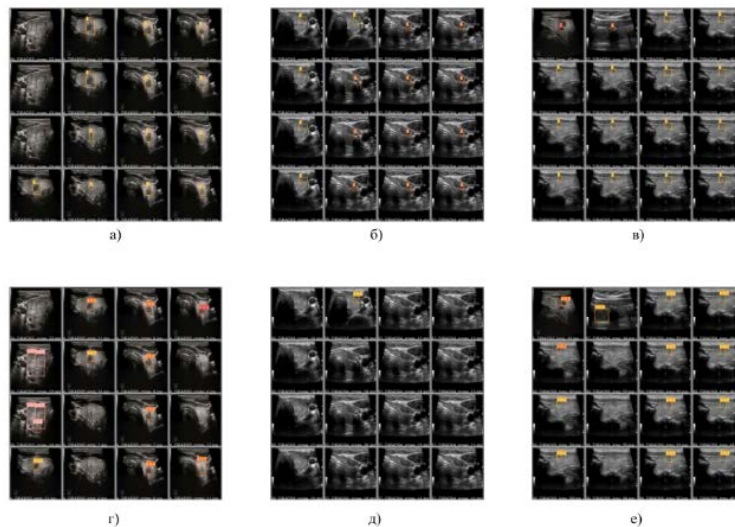


Рис. 7. Результаты тестирования YOLOv5 для обнаружения моделями, обученными на всех снимках (а, г), снимках продольной проекции (б, д) и на снимках поперечной проекции (в, е). Верхний ряд – исходные маски, нижний ряд – предсказанное положение узловых образований

VII. ДИСКУССИЯ

Вопрос объема данных для обучения наиболее актуален в настоящее время. Несмотря на то, что основной задачей при проектировании новых архитектур является оптимизация строения моделей таким образом, чтобы при наименее возможном объеме и качестве данных получать наиболее точные предсказания при валидации и тестировании, многие глубокие архитектуры требуют большого объема уникальных данных. Из проведенных в настоящем исследовании

экспериментов следует, что объем обучающей выборки должен быть увеличен не за счет повторения аналогичных изображений одного пациента, а за счет расширения датасета новыми уникальными экземплярами других пациентов и и/или данными процесса аугментации. Полученные выводы нашли подтверждение или опровержение в других работах на других датасетах.

Например, результаты в исследовании [9] подтверждают выводы, сделанные в ходе данного

исследования. В этой статье рассматривается влияние объема данных на обучение глубокой архитектуры классификации результатов ЭКГ. В статье проводится анализ использования разного количества не аугментированных данных при тренировке сети. Следует заметить, что обучающая выборка состояла лишь из уникальных снимков ЭКГ. В результате описанного эксперимента делается вывод о том, что показатели метрик оценки обучения сети выходят на плато при использовании 45% обучающей выборки. Помимо этого, эксперимент для анализа влияния объема одного экземпляра обучающей выборки демонстрирует, что размер уникального обучающего экземпляра не влияет на результаты обучения. После достижения оптимального для обучения объема показатели метрики выходят на плато.

В статье [11] рассматриваются эксперименты по оценке влияния качества и объема данных для обучения моделей. Здесь рассматриваются глубокие нейронные архитектуры и различные выборки медицинских данных. Проблемы, обнаруженные в результате проведения исследования, подтверждают выводы о необходимости высокого качества изображений и большого количества уникальных данных.

В работе [12] исследуется зависимость точности предсказания нейросетевых архитектур от объема и качества обучающей выборки. Это исследование скорее доказывает, что большой объем данных наиболее важный аспект при формировании гиперпараметров модели. Этот вывод не совсем совпадает с результатами настоящего исследования. Это различие может иметь место из-за различия в датасетах. Там работа ведется и с уникальными и с аугментированными экземплярами данных, – (LVIS-OneShot, PASCAL-5, FSS-1000), тогда как в настоящей работе проводилось исследование влияния на качество нейросетевых моделей однотипных снимков петель небольшого числа пациентов из обучающей выборки.

VII. ЗАКЛЮЧЕНИЕ

В работе исследовалось влияние объема и качества данных для обучения глубоких архитектур YOLOv5 и DeepLabV3. Для проведения анализа использовались кинопетли УЗИ щитовидной железы 166 пациентов, имеющих узловые образования. Подбор пациентов и запись кинопетель узловых

образований проводилась в рамках реализации проекта №22-15-00135 гранта Российского научного фонда.

В ходе исследования был проведен сравнительный анализ результатов обучения модели разными типами и снимков и разными количествами данных. В результате экспериментов было получено 15 моделей архитектуры YOLOv5 и 6 моделей архитектуры DeepLabV3 с различными гиперпараметрами и показателями на обучении и валидации. Показатели метрик демонстрировали разброс лишь в 10%, что не является существенным при исследуемых архитектурах.

Исходя из результатов, полученных при валидации и тестировании моделей, а также при анализе метрик при обучении, было выявлено, что на качество обучения практически не влияет переизбыток однотипных данных. При обучении таких глубоких архитектур необходимо иметь большой объем уникальных и качественных данных.

Подводя итог, можно сделать вывод, что схожие снимки (снимки одной проекции одного пациента) нельзя использовать как независимые данные. Для улучшения показателей качества глубоких архитектур необходимо расширять обучающую выборку за счет увеличения числа снимков разных пациентов.

БИБЛИОГРАФИЯ

1. Сафронов Д.А., Кацер Ю.Д., Зайцев К.С. Поиск аномалий с помощью автоэнкодеров // International Journal of Open Information Technologies, 2022г. Т. 10, Вып. 8 с. 39-45
2. Дюльдин Е.В., Зайцев К.С. Применение глубокого обучения для выявления и классификации DGA доменов // International Journal of Open Information Technologies, 2022г. Т. 10, Вып. 8 с. 3-10
3. Junying Chena, , Haijun Youa, Kai Li A review of thyroid gland segmentation and thyroid nodule segmentation methods for medical ultrasound images // Computer Methods and Programs in Biomedicine, v.185, March 2020, 105329 <https://www.sciencedirect.com/science/article/pii/S0169260719308454>
4. Pengju Deng, Xiaohong Han, Xi Wei, Luchen Chang Automatic classification of thyroid nodules in ultrasound images using a multi-task attention network guided by clinical knowledge // Computer Methods and Programs in Biomedicine, October 2022, 105329 <https://www.sciencedirect.com/science/article/pii/S0010482522008800>
5. Xinyu Zhanga, Vincent CS. Leea, Jia Ronga & etc. Deep convolutional neural networks in thyroid disease detection: A multi-classification comparison by ultrasonography and computed tomography // Computer Methods and Programs in Biomedicine, v. 220 (2022) June 2022, 106823 <https://www.sciencedirect.com/science/article/pii/S016926072200205X>

6. Yasaman Sharifi, Mohamad Amin Bakhshali, Toktam Dehghani & etc. Deep learning on ultrasound images of thyroid nodules // *Biocybernetics and Biomedical Engineering* 2022, v. 41, Iss 2, April-June, p. 636-655
<https://www.sciencedirect.com/science/article/abs/pii/S0208521621000152>
7. Tessler, F. N., Middleton, W. D., & Grant, E. G. (2018). Thyroid Imaging Reporting and Data System (TI-RADS): A User's Guide. *Radiology*, 287(1), 29–36.
<https://doi.org/10.1148/radiol.2017171240>
8. Timo Luddecke, Alexander Ecker; The Role of Data for One-Shot Semantic Segmentation; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2021, pp. 2653-2658
9. Toldo, M.; Maracani, A.; Michieli, U.; Zanuttigh, P. Unsupervised Domain Adaptation in Semantic Segmentation: A Review. *Technologies* 2020, 8, 35.
<https://doi.org/10.3390/technologies8020035>
10. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). *You Only Look Once: Unified, Real-Time Object Detection*. <http://arxiv.org/abs/1506.02640>
11. Ananov, V. V., Skorik, S. N., Shaklein, V. V., Avetisyan, A. A., Teregulov, Y. E., Turdakov, D. Y., Gliner, V., Schuster, A., & Karpulevich, E. A. (2021). Assessment of the impact of non-architectural changes in the predictive model on the quality of ECG classification. *Proceedings of the Institute for System Programming of the RAS*, 33(4), 87–98.
[https://doi.org/10.15514/ispras-2021-33\(4\)-7](https://doi.org/10.15514/ispras-2021-33(4)-7)
12. Luca, A. R., Ursuleanu, T. F., Gheorghe, L., Grigorovici, R., Iancu, S., Hlusneac, M., & Grigorovici, A. (2022). Impact of quality, type and volume of data used by deep learning models in the analysis of medical images. In *Informatics in Medicine Unlocked* (Vol. 29). Elsevier Ltd.
<https://doi.org/10.1016/j.imu.2022.100911>

Dependence of the performance quality of neural networks on the characteristics of training data when working with thyroid ultrasound images

E.A. Ekhlakov, K.V. Tsyguleva, M.E. Dunaev, S.M. Zakharova

Abstract — The purpose of this work is to test the hypothesis that the performance of neural network models for the detection and segmentation of nodular formations in thyroid ultrasound images is practically independent of the number of analyzed images of one patient obtained at the same time. Two deep architectures were used for verification: YOLOv5 when solving the detection problem and DeepLabV3 when solving the segmentation problem. During the experiments, video loops (frame sequences) of thyroid ultrasound were used, containing more than 7000 images of transverse and longitudinal projections of 166 patients. The performance of deep architectures was evaluated both at the stage of their training and validation, and at the stage of their testing. According to the results of the experiments, it was found that an increase in the number of similar images in loops with a constant number of patients under study does not affect the operation of deep architectures.

Keywords – neural networks, machine learning, deep architectures, thyroid ultrasound imaging, gland segmentation, gland detection.

REFERENCES

1. Safronov D.A., Katser Yu.D., Zaytsev K.S. Finding anomalies with autoencoders // International Journal of Open Information Technologies, 2022. T. 10, Issue. 8 p. 39-45.
2. Dyuldin E.V., Zaytsev K.S. Application of deep learning to identify and classify DGA domains // International Journal of Open Information Technologies, 2022. T. 10, Issue. 8 p. 3-10.
3. Junying Chena, , Haijun Youa, Kai Li A review of thyroid gland segmentation and thyroid nodule segmentation methods for medical ultrasound images // Computer Methods and Programs in Biomedicine, v.185, March 2020, 105329 <https://www.sciencedirect.com/science/article/pii/S0169260719308454>
4. Pengju Deng, Xiaohong Han, Xi Wei, Luchen Chang Automatic classification of thyroid nodules in ultrasound images using a multi-task attention network guided by clinical knowledge // Computer Methods and Programs in Biomedicine, October 2022, 105329 <https://www.sciencedirect.com/science/article/pii/S0010482522008800>
5. Xinyu Zhanga, Vincent CS. Leea, Jia Ronga & etc. Deep convolutional neural networks in thyroid disease detection: A multi-classification comparison by ultrasonography and computed tomography // Computer Methods and Programs in Biomedicine, v. 220 (2022) June 2022, 106823 <https://www.sciencedirect.com/science/article/pii/S016926072200205X>
6. Yasaman Sharifi, Mohamad Amin Bakhshali, Toktam Dehghani & etc. Deep learning on ultrasound images of thyroid nodules // Biocybernetics and Biomedical Engineering v. 41, Iss 2, April-June, p. 636-655 <https://www.sciencedirect.com/science/article/abs/pii/S0208521621000152>
7. Tessler, F. N., Middleton, W. D., & Grant, E. G. (2018). Thyroid Imaging Reporting and Data System (TI-RADS): A User's Guide. *Radiology*, 287(1), 29–36. <https://doi.org/10.1148/radiol.2017171240>
8. Timo Luddecke, Alexander Ecker; The Role of Data for One-Shot Semantic Segmentation; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2021, pp. 2653-2658
9. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). *You Only Look Once: Unified, Real-Time Object Detection*. <http://arxiv.org/abs/1506.02640>
10. Ananev, V. V., Skorik, S. N., Shaklein, V. V., Avetisyan, A. A., Teregulov, Y. E., Turdakov, D. Y., Gliner, V., Schuster, A., & Karpulevich, E. A. (2021). Assessment of the impact of non-architectural changes in the predictive model on the quality of ECG classification. *Proceedings of the Institute for System Programming of the RAS*, 33(4), 87–98. [https://doi.org/10.15514/ispras-2021-33\(4\)-7](https://doi.org/10.15514/ispras-2021-33(4)-7).
11. Luca, A. R., Ursuleanu, T. F., Gheorghe, L., Grigorovici, R., Iancu, S., Hlusneac, M., & Grigorovici, A. (2022). Impact of quality, type and volume of data used by deep learning models in the analysis of medical images. In *Informatics in Medicine Unlocked* (Vol. 29). Elsevier Ltd. <https://doi.org/10.1016/j.imu.2022.100911>
12. Chen, L.-C., Papandreou, G., Schroff, F., & Adam, H. (2017). *Rethinking Atrous Convolution for Semantic Image Segmentation*. <http://arxiv.org/abs/1706.05578>