

УДК 004.932

АНАЛИЗ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ДЕТЕКЦИИ СТЕНОЗОВ КОРОНАРНЫХ АРТЕРИЙ

© 2021 г. В. В. Данилов^{а,с,*}, О. М. Гергет^{а,**}, К. Ю. Клышников^{б,***},
А. Ф. Франжи^{с,****}, Е. А. Овчаренко^{б,*****}

^а Томский политехнический университет,
634050 Томск, пр. Ленина, д. 30, Россия

^б Научно-исследовательский институт комплексных проблем сердечно-сосудистых заболеваний,
650000 Кемерово, Сосновый бульвар, д. 6, Россия

^с University of Leeds, EC Stoner Building, Woodhouse Lane,
Leeds, United Kingdom

*E-mail: viacheslav.v.danilov@gmail.com

**E-mail: gerget@tpu.ru

***E-mail: klyshku@kemcardio.ru

****E-mail: A.Frangi@leeds.ac.uk

*****E-mail: ovchea@kemcardio.ru

Поступила в редакцию 18.11.2020 г.

После доработки 25.12.2020 г.

Принята к публикации 12.01.2021 г.

В статье освещается подход на основе технологии машинного обучения, который представляет особый интерес для локализации и определения характеристик как одноочаговых стенозов, так и многососудистых, многоочаговых поражений. В связи со сложностью анализа большого количества данных клиницистом/кардиохирургом, в исследовании большое внимание уделено анализу, обучению и сравнению популярных детекторов для классификации и локализации очагов стеноза на данных коронарной ангиографии. Полный набор данных был собран в НИИ Комплексных проблем сердечно-сосудистых заболеваний на основе исследования коронарографии, среди которых ретроспективно выбраны данные 100 пациентов. Для автоматизированного анализа медицинских данных, в статье подробно рассмотрены 3 модели (SSD MobileNet V1, Faster-RCNN ResNet-50 V1, Faster-RCNN NASNet), которые варьировались по архитектуре, сложности и количеству весов. Приведено сравнение моделей по основным характеристикам эффективности: точность, время обучения и время предсказания. Результаты тестирования показали, что время обучения/предсказания прямо пропорционально сложности модели. Так, наименьшее время предсказания показала модель Faster-RCNN NASNet (среднее время обработки одного изображения составило 880 мс). Что касается точности, то наибольшая точность предсказания была получена моделью Faster-RCNN ResNet-50 V1. Данная модель достигла уровня 0.92 метрики mAP на валидационном наборе данных. С другой стороны, наиболее быстрой оказалась модель SSD MobileNet V1, которая способна выполнять предсказания с частотой предсказания 23 кадра в секунду.

DOI: 10.31857/S0132347421030031

1. ВВЕДЕНИЕ

На сегодняшний день ишемическая болезнь сердца (ИБС) является основной причиной смертности в мире [1]. Наличие атеросклеротических обструкций является наиболее частой причиной ИБС и заключается в уменьшении просвета коронарных артерий бляшкой. Клинически “золотым стандартом” для оценки стеноза коронарных артерий является инвазивная ангиография коронарных артерий (коронарография) с рентгенологическим контрастом. В свою очередь, анализ

данных ангиографии играет важную роль в локализации и интерпретации коронарного стеноза. В зависимости от анализа очагов стеноза и оценки метрики SYNTAX формируется тактика лечения пациента: минимально инвазивная чрессосудистая хирургия или чрезмерно инвазивное вмешательство.

Современное состояние исследований, связанных с локализацией, классификацией и интерпретацией таково, что существующие решения поставленных задач [2–8], несмотря на возможность

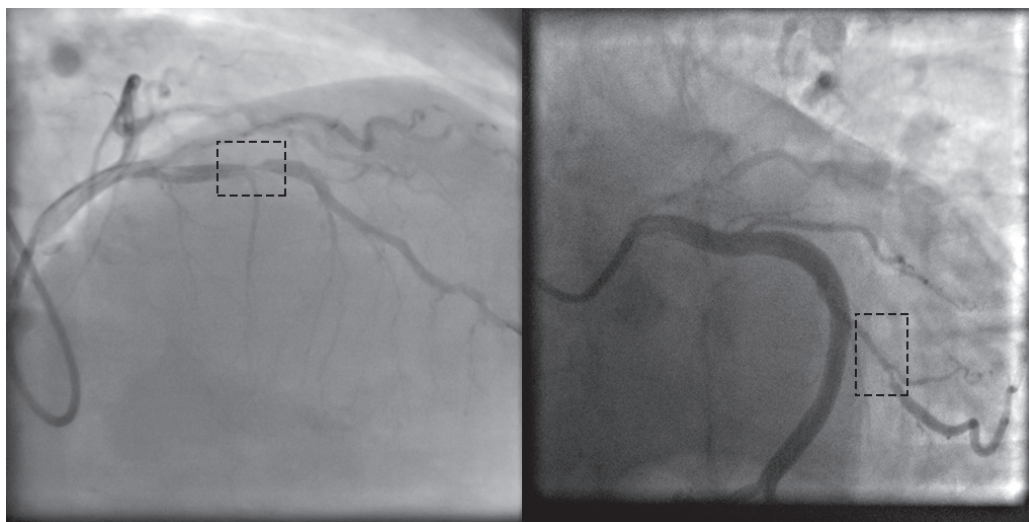


Рис. 1. Пример данных со стенозом для двух пациентов — область интереса обозначена прямоугольником.

получения достаточно высоких результатов по точности локализации (82–95%) и классификации (80–97%), обладают рядом существенных недостатков [9]. Одним из таких недостатков является плохая масштабируемость и не гибкость алгоритмов предварительной обработки, требующей тонкой настройки. Стоит также отметить, что большинство алгоритмов локализации построены по принципу каскадности, который склонен к накоплению ошибки. Исходя из этого, нами предпринята попытка по анализу, обучению и последующему сравнению нейронных сетей, которые способны выполнять оценку стеноза на сырых данных коронарной ангиографии с высокой степенью точности.

2. ИСХОДНЫЕ ДАННЫЕ

Исходные данные представляют собой записи ангиографии коронарных артерий, которые были получены на ангиографических установках “Coroscor” и “Innova” (Siemens) в НИИ Комплексных проблем сердечно-сосудистых заболеваний (г. Кемерово, Россия). Ретроспективно выбраны 100 пациентов, по данным которых было проведено детектирование гемодинамически значимых стенозов. Критерием исключения стало наличие множественных стенозов (двух и более коронарных артерий). В результате каждого ангиографического исследования получен видеоряд прохождения контраста по сосудистой системе сердца. Первичная сортировка данных осуществлялась с участием эксперта-кардиолога, который выбрал информативные изображения, содержащие только этап прохождения контраста и визуализацию

стенозированного участка. Таким образом был выделен набор изображений в оттенках серого (1 канал) с разрешением от 512×512 до 1000×1000 пикселей. Суммарно выборка составила 8325 изображений, из которых 7492 (90%) изображения использовались в качестве обучающей выборки, а 833 (10%) изображения — в качестве валидационной выборки. Для разметки данных была использована бесплатная версия SaaS (Software as a Service) решения “LabelBox”. Примеры размеченных входных данных для двух пациентов отражены на рис. 1.

Для анализа исходного набора данных нами проведена оценка размера области стеноза (площадь ограничивающего прямоугольника). Аналогично набору данных Common Objects in Context (COCO), размеры прямоугольников были разделены на 3 класса: “small” (площадь региона менее 32^2 пикселей), “medium” (площадь региона более 32^2 пикселей, но менее 96^2 пикселей) и “large” (площадь региона более 96^2 пикселей). В результате в исходном наборе данных общее число “small” объектов составило 2509 (30%), “medium” объектов — 5704 (69%) и “large” объектов — 113 (1%). Исходя из данного распределения видна несбалансированность исследуемого набора данных. Потенциально распознавание крупных областей стеноза может быть выполнено с меньшей точностью, чем маленьких и средних областей.

С целью решения задачи локализации в рамках исследования выполнена оценка распределения координат стенозов на изображениях исходного набора данных. Для этого оценивались нормализованные координаты центров ограничивающих пря-

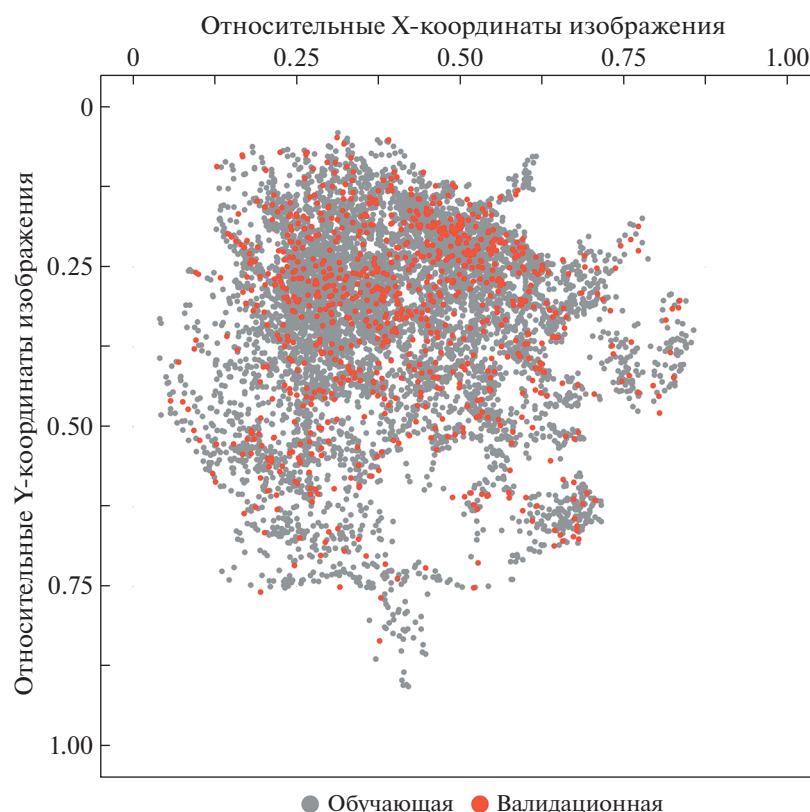


Рис. 2. Карта плотности распределения очагов стеноза для обучающей и валидационной выборок. В качестве координат выбраны условные единицы, характеризующие положение точек, где точка (0; 0) соответствует левому верхнему углу изображения (ангиографии), точка (1; 1) — правому нижнему углу.

моугольников областей стеноза. На основе такой оценки получена карта распределения координат центров очагов стеноза, которая отражена ниже на рис. 2. Опираясь на результаты распределения координат, видно, что на карте наблюдаются 2 центра с относительными координатами (0.50; 0.20) и (0.27; 0.27), где располагается наибольшее количество очагов стеноза. Отметим, что координаты центров очагов распределены равномерно, а явные статистические выбросы при этом отсутствуют.

3. МЕТОДЫ

3.1. Описание моделей

Для решения задачи локализации стеноза на основе данных ангиографии коронарных артерий использовался подход машинного обучения, который хорошо зарекомендовал себя в сфере компьютерного зрения и обработки изображений. Для решения данной задачи использовались детекторы SSD [10] и Faster-RCNN [11] из репозитория Tensorflow Detection Model Zoo [12] на основе таких моделей как MobileNet [13, 14], ResNet

Таблица 1. Краткая характеристика используемых моделей

Модель	Время предсказания, мс	mAP@[0.5:0.95]	Количество весов, млн.	Размер модели, Мб.
SSD MobileNet V1	56	32	4.2	44
Faster-RCNN ResNet-50 V1	89	30	25.6	114
Faster-RCNN NASNet	540	—	88.9	416

Таблица 2. Используемые опции при обучении моделей

Модель	Входной размер	Аугментация	Размер батча	Тип LR	LR
SSD MobileNet V1	$640 \times 640 \times 3$	Random horizontal flip Random crop image	4	Cosine decay with warm up	0.04
Faster-RCNN ResNet-50 V1	$600 \times 600 \times 3$	Random horizontal flip	2	Constant LR	0.0003
Faster-RCNN NASNet	$1200 \times 1200 \times 3$	Random horizontal flip	1	Constant LR	0.0003

[15, 16] и NASNet [17, 18]. Всего исследовано 3 модели, которые варьировались по архитектуре, сложности и количеству весов. В качестве модели, использующей наименьшее количество весов, был выбран детектор SSD MobileNet V1, который способен обрабатывать данные в реальном формате времени. Наиболее сложной моделью из исследуемых является Faster-RCNN NASNet с более 80 млн весов. Выше в табл. 1 отражена краткая характеристика моделей, которые использовались в данном исследовании. Характеристика, а в особенности метрика mAP, получена на основе обучения нейронных сетей на наборе данных Common Objects in Context (COCO).

3.2. Обучение моделей

При обучении нейронных сетей их базовая конфигурация унаследована от конфигурации сетей, обученных на наборе данных COCO. Для однозначного сравнения исследуемых моделей общее число шагов обучения составило 100. Для лучшего восприятия в работе используется сокращенное обозначение количества шагов обучения, т.е. 100 шагов означает 100000 итераций обуче-

ния. В качестве локализационной функции потерь использовался Weighted Smooth L1 loss, а классификационной функции потерь – Weighted Focal Loss [19]. Важным будет отметить, что модель на основе SSD детектора обучалась с использованием техники Cosine decay with warm up. При использовании такой техники скорость обучения LR постепенно снижалась в зависимости от шага обучения. В качестве оборудования для обучения нами были использованы инстансы P2 (Nvidia Tesla K80 12 Gb, 1.87 TFLOPS) и P3 (Nvidia Tesla V100 16 Gb, 7.8 TFLOPS) от Amazon Web Services. Выше в табл. 2 приведены основные характеристики обучения моделей.

В процессе обучения нейронных сетей проведена оценка динамики точности на валидационном наборе данных. В качестве оценки точности выбрана метрика mean Average Precision (mAP) с использованием порогового значения Intersection over Union = 0.5 (mAP@0.5). Ниже на рис. 3 отражена сглаженная динамика изменения метрики mAP на основе валидационной выборки при обучении моделей. Видно, что практически все мо-

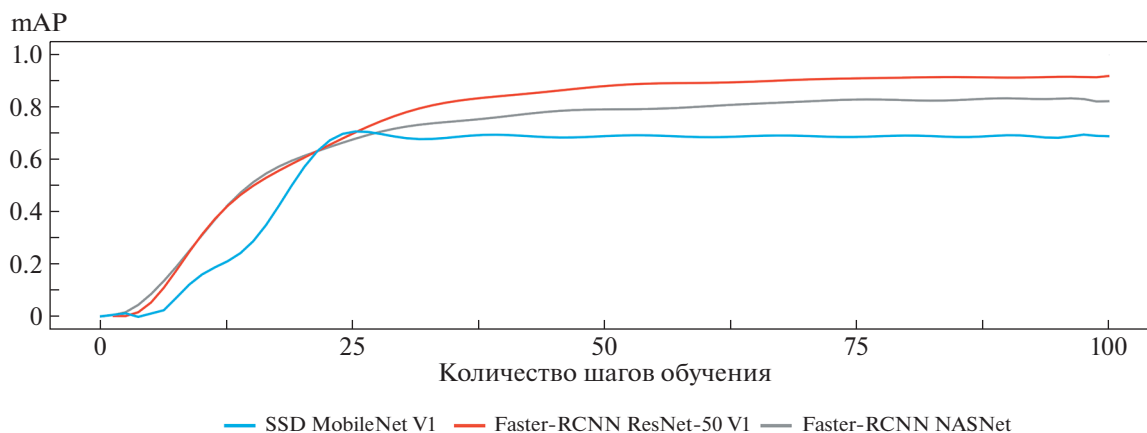
**Рис. 3.** Динамика изменения метрики mAP при обучении исследуемых моделей.

Таблица 3. Сравнение результатов исследуемых моделей

Модель	Количество весов, млн		Время обучения, часы		Время предсказания, мс		mAP@0.5	
	Абс.	Отн.	Абс.	Отн.	Абс.	Отн.	Абс.	Отн.
SSD MobileNet V1	4.2	1.0×	16	1.0×	43	1.0×	0.70	1.00×
Faster-RCNN ResNet-50 V1	25.6	6.0×	28	1.8×	98	2.3×	0.92	1.33×
Faster-RCNN NASNet	88.9	21.0×	147	9.5×	880	20.4×	0.84	1.22×

дели сходятся к некоторому значению асимптотической точности mAP.

4. РЕЗУЛЬТАТЫ

Сравнительный анализ полученных результатов на трех исследованных нейросетевых моделях приведен в табл. 3. Для относительного сравнения используются значения метрик для модели SSD MobileNet V1, а модели Faster-RCNN ResNet-50 V1 и Faster-RCNN NASNet с ней сравниваются. Ниже на рис. 4 и 5 отражены несколько основных метрик эффективности моделей, а именно метрика точности mAP, время обучения и время предсказания. Графики отсортированы таким образом, что на первом месте находятся модели, имеющие наилучшее значение по рассматриваемой метрике.

Для оценки времени предсказания использовался инстанс P3 (Nvidia Tesla V100 16 Gb, 7.8 TFLOPS) от Amazon Web Services. Как показало тестирование, скорость предсказания напрямую зависит от сложности моделей и количества весов. Так, наиболее медленной для предсказания оказалась модель Faster-RCNN NASNet. Среднее время обработки одного изображения для данной модели составляет 880 миллисекунд. С другой стороны, наиболее быстрой из рассматриваемых оказалась модель, базирующаяся на архитектуре MobileNet. Время предсказания одного изображения данной моделью составляет 43 миллисекунды, что говорит о возможности ее использования для предсказания расположения очага стеноза в реальном масштабе времени.

Исходя из полученных результатов видно, что наиболее точной моделью оказалась Faster-

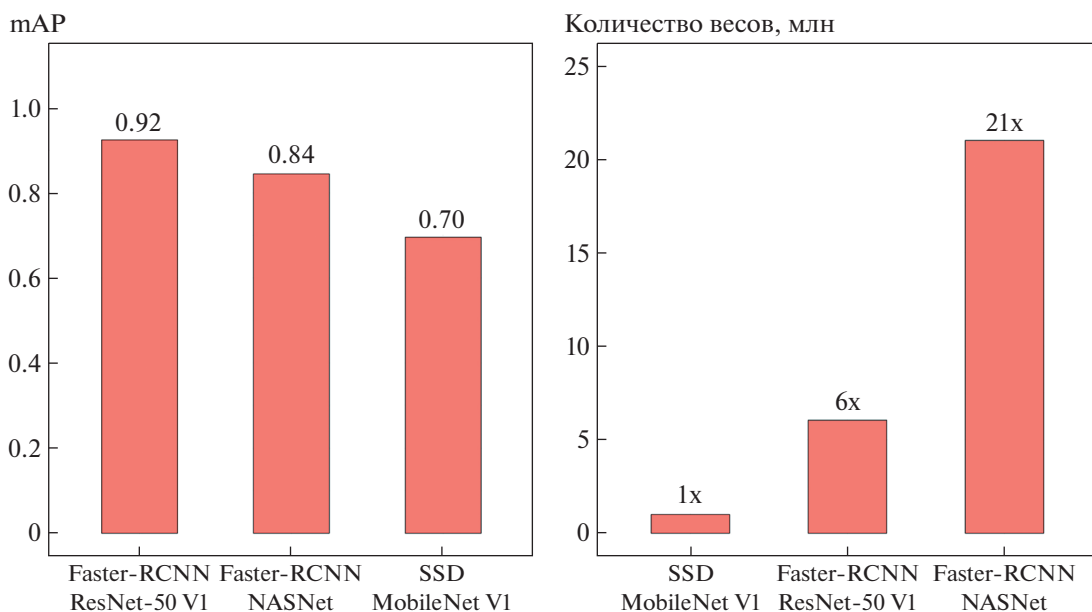


Рис. 4. Сравнение метрики mAP и количества весов исследуемых моделей нейросетей, наглядно визуализирующее различия архитектур.

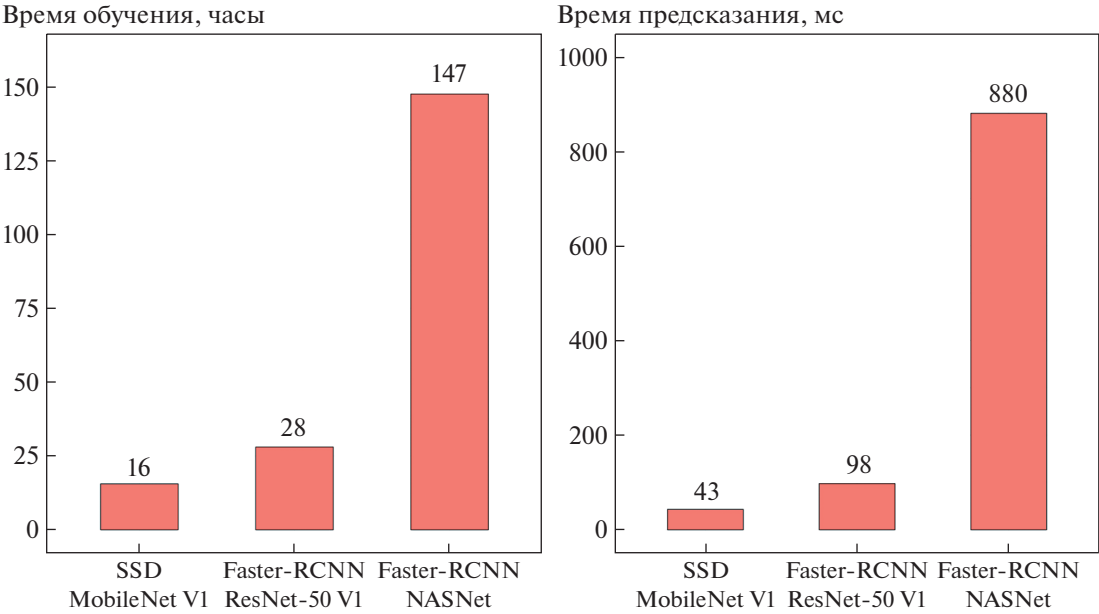


Рис. 5. Сравнение времени обучения и времени предсказания исследуемых моделей нейросетей, наглядно визуализирующее различия архитектур.

RCNN ResNet-50 V1. Средняя точность предсказания данной модели на валидационной выборке составила 0.92, а скорость предсказания – 98 миллисекунд/изображение (≈ 10 кадров в секунду). Наиболее быстрая и относительно легковесная модель SSD MobileNet V1 имеет среднюю точность предсказания на валидационной выборке равной 0.70, при этом ее средняя скорость предсказания составляет 43 миллисекунды/изображение (≈ 23 кадра в секунду). Стоит отметить модель Faster-RCNN NASNet, которая обладает более чем 3-х кратным преимуществом по количеству весов по сравнению с моделью Faster-RCNN ResNet-50 V1. Однако точность Faster-RCNN ResNet-50 V1 на 12% выше модели Faster-RCNN NASNet. В связи с этим, в качестве оптимального решения стоит выделить

именно модель Faster-RCNN ResNet-50 V1, которая способна достаточно быстро обрабатывать данные с высокой точностью.

4.1. Тестирование моделей

В данном пункте продемонстрирована работа нейронных сетей на данных двух пациентов, для которых известна эталонная разметка, а их данные не использовались при обучении моделей. Примеры этих изображений с отмеченной областью стеноза отражены на рис. 6. Для тестирования использовались модели, которые давали наилучшие значения функции потерь и метрики mAP. Значения шагов обучения, на которых модели имеют оптимальные веса отражены ниже в табл. 4.

При сравнении полученных результатов на рис. 6 видно, что модели относительно точно выполняют локализацию очагов стеноза. Однако, существует ряд ложных срабатываний, а именно наличие ложноположительных случаев при тестировании модели Faster-RCNN NASNet. В обоих случаях данная модель дополнительно к размеченному эталонному очагу стеноза, с вероятностью более 90%, локализует дополнительные ложные очаги, которые находятся в правой коронарной артерии и передней нисходящей артерии в двух разных сегментах. С другой стороны, модель SSD MobileNet V1 чаще других ошибочно предсказывает отсутствие очага стеноза т.е. данная модель имеет более

Таблица 4. Значения шагов обучения, на которых модели имеют оптимальные веса

Модель	Оптимальный шаг обучения
SSD MobileNet V1	24
Faster-RCNN ResNet-50 V1	84
Faster-RCNN NASNet	95

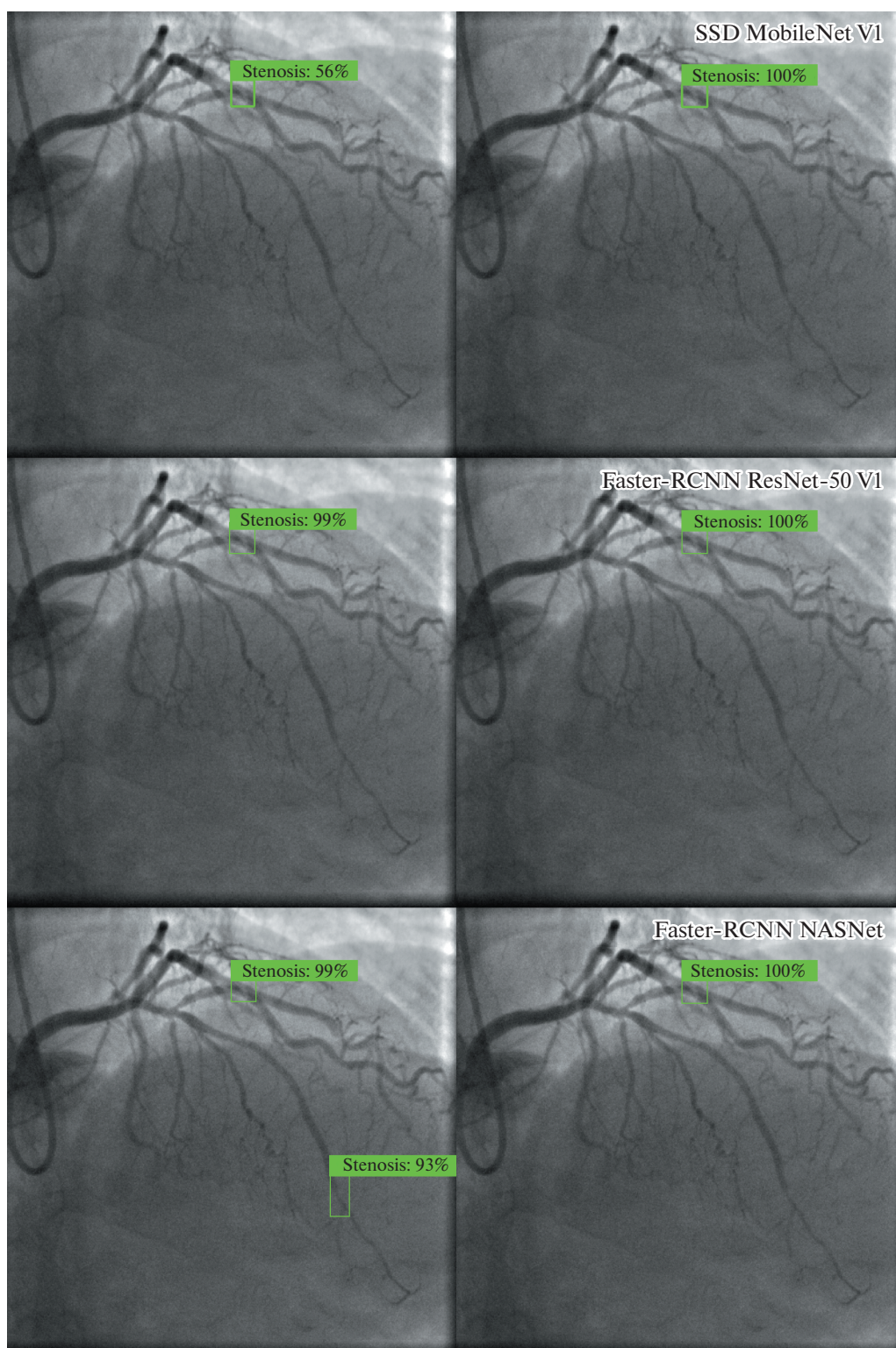


Рис. 6. Пример предсказания новых данных на двух соседних кадрах ангиографии для трех исследованных нейросетей.

высокую ложноотрицательную ошибку. Дополнительно стоит отметить эффективность работы детектора на базе архитектуры ResNet, а именно

Faster-RCNN ResNet-50 V1. Среднее значение индекса Дайса на тестовых данных для этой модели составило 0.85.

5. ЗАКЛЮЧЕНИЕ

В рамках рассмотренного исследования с целью выявления одноочаговых стенозов артерий проведено обучение и сравнительный анализ 3 популярных и перспективных детекторов, которые базировались на разных архитектурах нейронных сетей (MobileNet, ResNet-50, NASNet). Так, модель Faster-RCNN ResNet-50 V1 обладает оптимальным соотношением точность/время предсказания. Средняя точность предсказания (метрика mAP@0.5) данной модели на валидационной выборке составила 0.92, а скорость предсказания 10 кадров в секунду. Наиболее быстрая и относительно легковесная модель SSD MobileNet V1 имеет среднюю точность предсказания на валидационной выборке равной 0.70, при этом ее средняя скорость предсказания составляет 23 кадра в секунду. Важным будет отметить, что предложенный в работе подход представляет особый интерес для локализации и определения характеристик многососудистых, многоочаговых поражений. При таком подходе существует возможность не только детектировать очаги стеноза, но и оценивать ряд дополнительных характеристик стенозированного участка, а именно протяженность, диаметр, наличие боковых ветвей и параллельного шунта.

6. ФИНАНСИРОВАНИЕ

Работа по сбору исходных данных, предварительной обработке и созданию подхода для локализации очагов стеноза, основанного на машинном обучении, выполнена за счет гранта Российского научного фонда проект № 18-75-10061 “Исследование и реализация концепции роботизированного малоинвазивного протезирования клапана аорты”. Обучение разработанных моделей с помощью сервиса Amazon Web Services финансировалось в рамках государственного задания “Наука” № FFSWW-2020-0014 “Разработка научных основ технологии роботизированной мультипараметрической томографии на основе методов обработки больших данных и машинного обучения для исследования перспективных композиционных материалов”. Выбор основных метрик эффективности моделей и их анализ выполнен при поддержке гранта РФФИ 19-07-00351 “Методы и интеллектуальные технологии научного обоснования стратегических решений по цифровой трансформации”.

СПИСОК ЛИТЕРАТУРЫ

1. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* (London, England). 2018. V. 392. № 10159. P. 1736–1788.
2. Toledano M., Lindenbaum M., Lessick J., Dragu R., Ghersin E., Engel A. Learning to Detect Coronary Artery Stenosis from Multi-Detector CT imaging. Haifa, 2010. 1–23 p.
3. Kang D., Dey D., Slomka P.J., Arsanjani R., Nakazato R., Ko H., Berman D.S., Li D., Kuo C.-C.J. Structured learning algorithm for detection of nonobstructive and obstructive coronary plaque lesions from computed tomography angiography // *J. Med. Imaging. SPIE-Intl. Soc. Optical Eng.* 2015. V. 2. № 1. P. 014003.
4. Kang D., Slomka P.J., Nakazato R., Arsanjani R., Cheng V.Y., Min J.K., Li D., Berman D.S., Jay Kuo C.C., Dey D. Automated knowledge-based detection of nonobstructive and obstructive arterial lesions from coronary CT angiography // *Med. Phys.* John Wiley and Sons Ltd. 2013. V. 40. № 4. P. 041912.
5. Chi Y., Huang W., Zhou J., Toe K.K., Zhang J.-M., Wong P., Lim S.T., Tan R.S., Zhong L. Stenosis detection and quantification on cardiac CTCA using panoramic MIP of coronary arteries / 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE., 2017. P. 4191–4194.
6. Antczak K., Liberadzki Ł. Stenosis Detection with Deep Convolutional Neural Networks. *MATEC Web Conf.* / Ed. Mastorakis N., Mladenov V., Bulucea A. 2018. V. 210. P. 04001.
7. Zreik M., Van Hamersvelt R.W., Wolterink J.M., Leiner T., Viergever M.A., Išgum I. A Recurrent CNN for Automatic Detection and Classification of Coronary Artery Plaque and Stenosis in Coronary CT Angiography // *IEEE Trans. Med. Imaging.* Institute of Electrical and Electronics Engineers Inc. 2019. V. 38. № 7. P. 1588–1598.
8. de Vos B.D., Wolterink J.M., Leiner T., de Jong P.A., Lessmann N., Išgum I. Direct Automatic Coronary Calcium Scoring in Cardiac and Chest CT // *IEEE Trans. Med. Imaging. NLM (Medline).* 2019. V. 38. № 9. P. 2127–2138.
9. Danilov V.V., Gerget O.M., Skirnevskii I.P., Manakov R.A., Kolpashchikov D.Y. Segmentation Based on Propagation of Dynamically Changing Superpixels // *Program. Comput. Softw.* 2020. V. 46. № 3. P. 191–202.
10. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A.C. SSD: Single Shot MultiBox Detector / *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, 2016. V. 9905 LNCS. P. 21–37.
11. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // *IEEE Trans. Pattern Anal. Mach. Intell.* IEEE Computer Society. 2017. V. 39. № 6. P. 1137–1149.
12. Huang J., Rathod V., Sun C., Zhu M., Korattikara A., Fathi A., Fischer I., Wojna Z., Song Y., Guadarrama S., Murphy K. Speed/accuracy trade-offs for modern convolutional object detectors / *Proceedings – 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017.* Institute of Electrical and Electronics Engineers Inc., 2017. V. 2017-January. P. 3296–3305.

13. *Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L.-C.* MobileNetV2: Inverted Residuals and Linear Bottlenecks / Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. IEEE Computer Society. 2018. P. 4510–4520.
14. *Howard A.G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M., Adam H.* MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. 2017.
15. *He K., Zhang X., Ren S., Sun J.* Deep residual learning for image recognition / Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society. 2016. V. 2016-Decem. P. 770–778.
16. *He K., Zhang X., Ren S., Sun J.* Identity Mappings in Deep Residual Networks / Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). Springer Verlag, 2016. V. 9908 LNCS. P. 630–645.
17. *Zoph B., Le Q.V.* Neural Architecture Search with Reinforcement Learning / 5th Int. Conf. Learn. Represent. ICLR 2017 – Conf. Track Proc. International Conference on Learning Representations, ICLR., 2016.
18. *Zoph B., Vasudevan V., Shlens J., Le Q. V.* Learning Transferable Architectures for Scalable Image Recognition / 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE., 2018. P. 8697–8710.
19. *Lin T.-Y., Goyal P., Girshick R., He K., Dollar P.* Focal Loss for Dense Object Detection / 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, 2017. V. 42. № 2. P. 2999–3007.