



OPEN

Real-time coronary artery stenosis detection based on modern neural networks

Viacheslav V. Danilov^{1,3✉}, Kirill Yu. Klyshnikov², Olga M. Gerget¹, Anton G. Kutikhin², Vladimir I. Ganyukov², Alejandro F. Frangi³ & Evgeny A. Ovcharenko²

Invasive coronary angiography remains the gold standard for diagnosing coronary artery disease, which may be complicated by both, patient-specific anatomy and image quality. Deep learning techniques aimed at detecting coronary artery stenoses may facilitate the diagnosis. However, previous studies have failed to achieve superior accuracy and performance for real-time labeling. Our study is aimed at confirming the feasibility of real-time coronary artery stenosis detection using deep learning methods. To reach this goal we trained and tested eight promising detectors based on different neural network architectures (MobileNet, ResNet-50, ResNet-101, Inception ResNet, NASNet) using clinical angiography data of 100 patients. Three neural networks have demonstrated superior results. The network based on Faster-RCNN Inception ResNet V2 is the most accurate and it achieved the mean Average Precision of 0.95, F1-score 0.96 and the slowest prediction rate of 3 fps on the validation subset. The relatively lightweight SSD MobileNet V2 network proved itself as the fastest one with a low mAP of 0.83, F1-score of 0.80 and a mean prediction rate of 38 fps. The model based on RFCN ResNet-101 V2 has demonstrated an optimal accuracy-to-speed ratio. Its mAP makes up 0.94, F1-score 0.96 while the prediction speed is 10 fps. The resultant performance-accuracy balance of the modern neural networks has confirmed the feasibility of real-time coronary artery stenosis detection supporting the decision-making process of the Heart Team interpreting coronary angiography findings.

Coronary artery disease (CAD) is the leading cause of death worldwide¹, affecting over 120 million people². The main cause of CAD is atherosclerotic plaque accumulation³ in the epicardial arteries leading to a mismatch between myocardial oxygen supply and myocardial oxygen demand and commonly resulting in ischemia. Chest pain is the most likely symptom that occurs during physical and/or emotional stress, relieved promptly with rest or by taking nitroglycerin. This process can be modified by lifestyle adjustments, pharmacological therapies, and invasive interventions designed to achieve disease stabilization or regression⁴. Despite novel imaging modalities (e.g. coronary CT angiography) have been developed, invasive coronary angiography is the preferred diagnostic tool to assess the extent and severity of complex coronary artery disease according to the 2019 guidelines of the European Society of Cardiology^{5,6}. Multivessel coronary artery disease affecting two or more coronary arteries requires interpretive expertise on the assessment of multiple parameters (the number of affected major coronary arteries, the location of lesions, the severity of stenosis, the length of the stenotic segment, tortuosity, etc.) during an intervention. The process of interpreting complex coronary vasculature, image noise, low contrast vessels, and non-uniform illumination is time-consuming⁷, thereby posing certain challenges to the operator. Real-time automatic CAD detection and labeling may overcome the abovementioned difficulties by supporting the decision-making process.

A number of approaches for automatic or semi-automatic assessment of coronary artery diseases have been proposed by different research groups. These approaches follow the general scheme: (1) coronary artery tree extraction, (2) calculation of geometric dimensioning, and (3) analysis of the stenotic segment. The key stage that determines the speed and accuracy of such algorithms is based on the coronary artery tree extraction using the centerline extraction^{8,9}; the graph-based method^{10–12}; superpixel mapping^{13,14}; and machine/deep learning^{15–17}. The last, being a powerful tool for computer vision and image classification, has shown great promise in CAD detection due to their performance, tuning flexibility, and optimization. The ultimate purpose that CNN developers and users are trying to meet is to strike an optimal balance between accuracy and speed, the so-called speed/

¹Tomsk Polytechnic University, Tomsk, Russia. ²Research Institute for Complex Issues of Cardiovascular Diseases, Kemerovo, Russia. ³University of Leeds, Leeds, UK. ✉email: viacheslav.v.danilov@gmail.com

Parameter	Value
Total number of patients	100
Mean age \pm SD, years	60.3 \pm 13.8
Men, n (%)	68 (68%)
Women, n (%)	32 (32%)
Body mass index (kg/m ²)	21.6 \pm 5.1
Diagnosis	CAD
Class I NYHA	5 (5%)
Class II NYHA	84 (84%)
Class III NYHA	11 (11%)
Comorbidities	
Arterial hypertension	53 (53%)
Diabetes mellitus	14 (14%)
Chronic heart failure, classes 1–2	36 (36%)
Coronary artery stenosis > 70% (n, %)	100 (100%)

Table 1. Clinical and demographic data of the study population.

accuracy trade-off¹⁸. While some CNNs with high performance and optimal accuracy suitable for real-time segmentation can be used on mobile devices and low-end PCs, others with low performance are highly efficient for object detection (precision, recall, F1-score, mAP). Depending upon the task complexity and scope, this balance may vary and be achieved using the proper CNN architecture. The speed/accuracy trade-off for CAD detection should be adjusted to both, elective and urgent diagnosis. On the one hand, neural networks used for determining the severity of atherosclerotic lesions should possess superior detection rate as their decision-making ability will specify the selection of treatment strategies, including life-saving procedures. This situation is typical for stable patients undergoing elective coronary angiography. Therefore, heavy-weight CNNs requiring time to process angiographic data accurately can be applied. On the other, CNNs should ensure the highest performance of real-time image processing for urgent patients who do not have time for prolonged preoperative management and should undergo percutaneous coronary intervention (PCI) immediately following the diagnostic catheterization (ad-hoc PCI)^{19,20}.

Albeit several CNN-based approaches focused on achieving optimal accuracy for CAD detection with the Dice Similarity Coefficient of more than 0.75^{12,13} and/or the Sensitivity metric of more than 0.70²¹ have been proposed, their speed remains disregarded. Image processing time is an important indicator for the applied use of these methods that can reach 1.1–11.87 s¹⁰, 20 s^{10,13}, and over 60 s⁹. However, this time is unacceptable for real-time CAD detection with the processing rate of 7.5–15 fps instead of the required 0.13–0.07 s per frame^{22,23}. Slow data processing does not allow providing real-time support for the operator during the procedure and may be performed after diagnosis and data collection. Some researchers try to improve the performance of these algorithms by segmenting only large vessels of the coronary bed¹⁸. This approach allows achieving the inference time of 0.04 frames per second, but it does not take into account stenotic lesions in small branches. Another approach using convolutional neural networks to speed up the algorithm includes the extraction of individual regions of interest with stenotic sites without the entire coronary artery tree. A similar principle has been reported by Cong et al.¹⁹ describing the Inception V3 neural network and Hong et al.²⁰ describing the M-net (improved version of U-net).

Our study presents a detailed analysis of available neural network architectures and their potential in terms of accuracy and performance to detect single-vessel disease. This approach is aimed at selecting the most efficient CNN architecture and further exploring the ways of its modification and optimization to ensure superior real-time classification potential for detecting multivessel coronary artery stenosis.

To summarize, our main contributions are as follows:

- A comparative analysis of the speed/accuracy trade-off for detecting single stenoses of the coronary arteries of specific state-of-the-art CNN architectures (N = 8).
- The use of RFCN ResNet-101 V2 as is without any modification allows achieving promising real-time performance (10 fps) without a big loss in accuracy.
- The benefits of CNNs reported in our study may be leveraged for the development of software aimed at optimizing and facilitating invasive angiography.

Source data

Initial angiographic imaging series of one hundred patients who underwent coronary angiography using Coroscop (Siemens) and Innova (GE Healthcare) at the Research Institute for Complex Problems of Cardiovascular Diseases (Kemerovo, Russia) were retrospectively enrolled in the study (Table 1). All patients had angiographically and/or functionally confirmed one-vessel coronary artery disease ($\geq 70\%$ diameter stenosis (by QCA (quantitative coronary analysis) or 50–69% with FFR (fractional flow reserve) ≤ 0.80 or stress echocardiography evidence of regional ischemia). Significant coronary stenosis for the purpose of our study was defined according to 2017 US appropriate use criteria for coronary revascularization in patients with stable ischemic heart disease²¹.

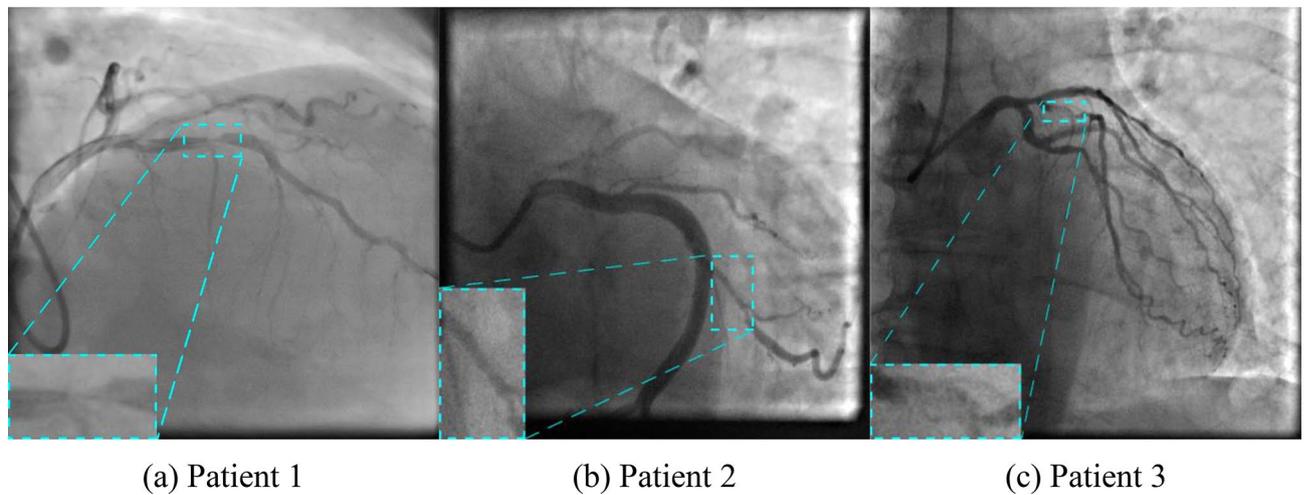


Figure 1. Data labelling of the source images with the callouts of the detected stenotic lesions.

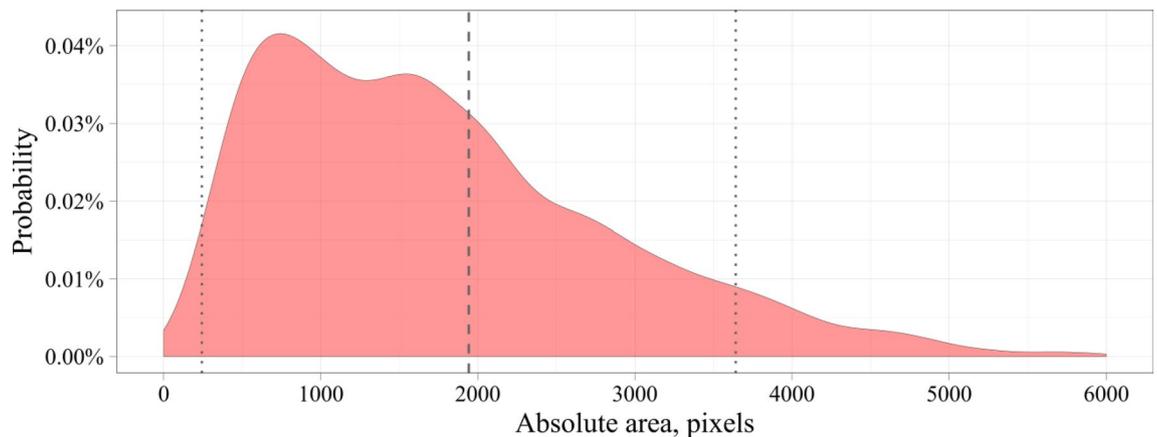


Figure 2. Distribution of the absolute stenotic area in the input dataset.

The study design was approved by the Local Ethics Committee of the Research Institute for Complex Issues of Cardiovascular Diseases (approval letter No. 112 issued on May 11, 2018). All participants provided written informed consent to participate in the study. Coronary angiography was performed by the single operator according to the indications and recommendations stated in the 2018 ESC/EACTS Guidelines on myocardial revascularization. The presence or absence of coronary stenosis was confirmed by the same operator using angiography imaging series according to the 2018 ESC/EACTS Guidelines on myocardial revascularization.

Angiographic images of the radiopaque overlaid coronary arteries with stenotic segments were selected and converted into separate images. An interventional cardiologist rejected non-informative images and selected only those containing contrast passage through a stenotic vessel. A total of 8325 grayscale images (100 patients) of 512×512 to 1000×1000 pixels were included for further study. Of them, 6660 (80%), 833 (10%), and 832 (10%) images were used for training, validation, and testing respectively. In order to correctly estimate model performance, we did not randomly shuffle all 8325 images and then form data subsets. We first randomly choose patient series for the training, validation, and testing subsets in an 80:10:10 ratio, and then form those subsets. Such data split allows us to know that the validation and testing are done on the independent subsets of images and avoid bias in performance metrics. Since the training process is quite time-consuming, we excluded the usage of cross-validation for the models. Data were labeled using the [LabelBox](#), a free version of SaaS (Software as a Service). It allows joint data labeling and subsequent validation by several specialists. Typical data labeling of the source images is shown in Fig. 1.

To analyse the source dataset, we estimated the size of the stenotic region computing the area of the bounding box. Similarly to the Common Objects in Context (COCO) dataset, we divided objects by their area into three types: small (area $< 32^2$), medium ($32^2 \leq \text{area} \leq 96^2$), and large (area $> 96^2$). 2509 small objects (30%), 5704 medium objects (69%), and 113 large objects (1%) were obtained in the input data. Since our data were unbalanced, we suppose that image analysis may be poorer on larger objects than on small and medium ones.

Figures 2 and 3 show the distributions of the absolute and relative stenotic areas. To generate the distribution of the absolute area, we estimated the absolute values of the bounding box stenotic areas in pixels. To generate the distribution of the relative area, we estimated the value of the area of the bounding box relative to the area of

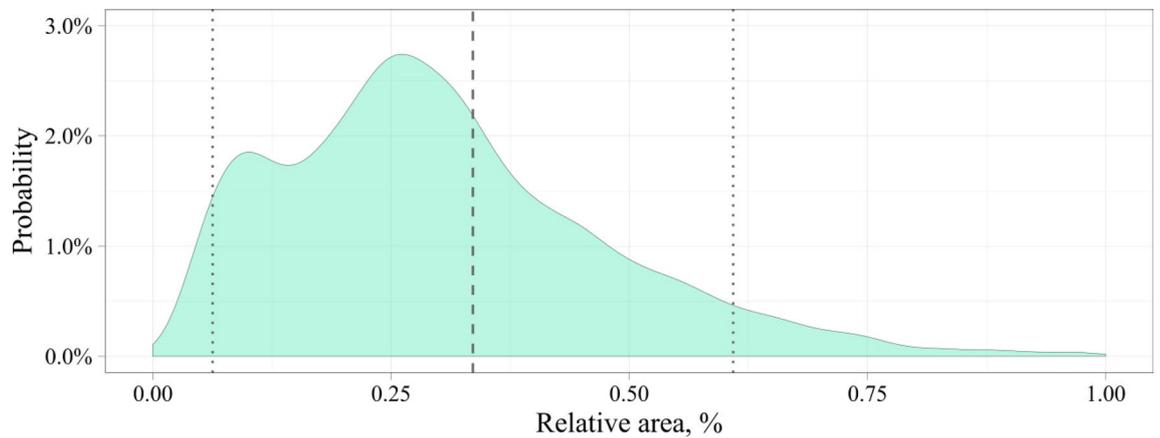


Figure 3. Distribution of the relative stenotic area in the input dataset.

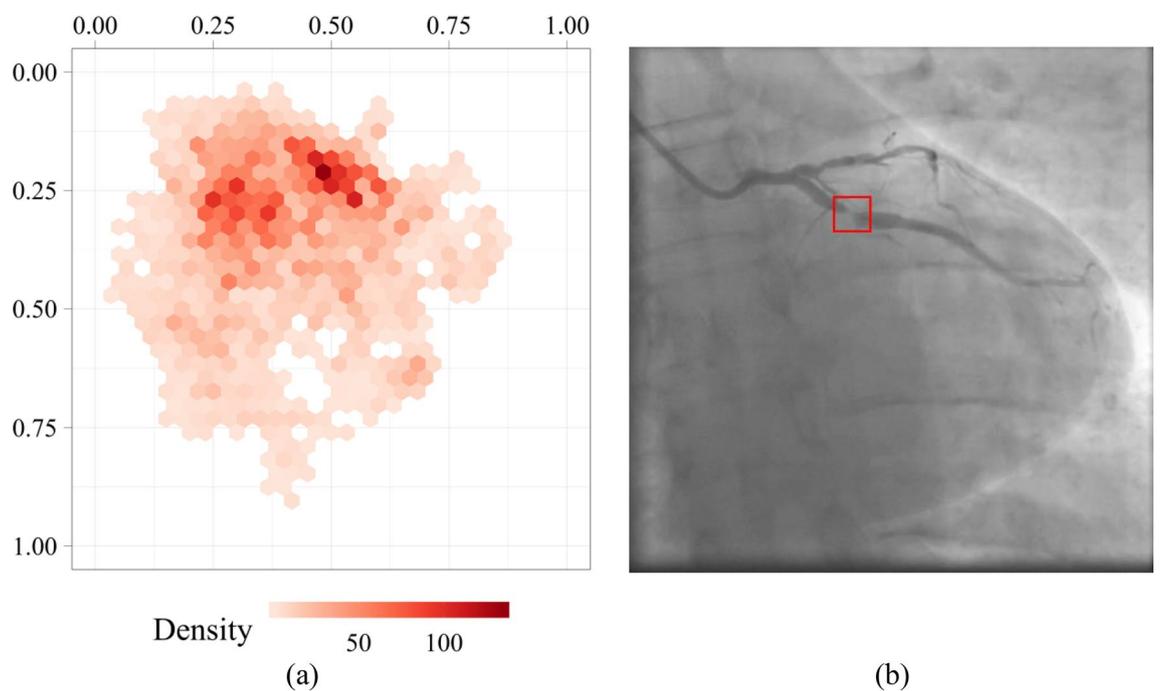


Figure 4. Distribution mapping of the stenosis density over the dataset (a) and an example of an angiography image with the labeled stenotic area (b).

the entire image in percentages. The dashed lines represent the mean values and standard deviations of the area. Based on the input data, the absolute stenotic area was 1942 ± 1699 pixels (Fig. 2). Since the size of the images from the input dataset varied within a certain range of values, we calculated the relative stenotic area. We selected images with normalized X and Y coordinates in the range of values [0; 1]. As a result, the relative stenotic area was $0.34 \pm 0.27\%$ (Fig. 3). As seen, the stenotic area is quite small compared to the area of the whole image that may confuse some detectors typically applied to detect objects in an unconstrained environment.

To determine the location of stenosis accurately, we evaluated the distribution of the stenosis coordinates along the vessel in the input images. We estimated the normalized coordinates of the center point of the bounding box around the stenotic lesion. Based on this assessment, a distribution map of the coordinates of the stenosis centers was generated and is shown in Fig. 4. Each hexagon on this map reflects a number of the stenosis centers of the bounding box around the stenotic lesion. Distributing the coordinates highlights two centres with relative coordinates (0.50; 0.20) and (0.27; 0.27) along the stenotic segment. The coordinates of the stenosis centers are evenly distributed without explicit outliers. The latter should have a positive effect on training regressors based on neural networks that predict the coordinates of the bounding boxes.

Model	Inference time, ms	mAP@ [0.5:0.95]	Weights, mln	Model size, Mb
SSD MobileNet V1	56	32	4.2	44
SSD MobileNet V2	31	22	6.1	19
SSD ResNet-50 V1	76	35	25.6	127
Faster-RCNN ResNet-50 V1	89	30	25.6	114
RFCN ResNet-101 V2	92	30	44.7	199
Faster-RCNN ResNet-101 V2	106	32	44.7	190
Faster-RCNN Inception ResNet V2	620	37	55.9	241
Faster-RCNN NASNet	540	–	88.9	416

Table 2. Brief characteristics of the use.

Model	Input size	Augmentation	Batch size	Type of LR	LR
SSD MobileNet V1	640 × 640 × 3	Random horizontal flip. Random crop image	4	Cosine decay with warm up	0.04
SSD MobileNet V2	300 × 300 × 3	Random horizontal flip. SSD random crop	4	Exponential decay	0.004
SSD ResNet-50 V1	640 × 640 × 3	Random horizontal flip. Random crop image	2	Cosine decay with warm up	0.04
Faster-RCNN ResNet-50 V1	600 × 600 × 3	Random horizontal flip	2	Constant LR	0.0003
RFCN ResNet-101 V2	600 × 600 × 3	Random horizontal flip	1	Constant LR	0.0003
Faster-RCNN ResNet-101 V2	600 × 600 × 3	Random horizontal flip	1	Constant LR	0.0003
Faster-RCNN Inception ResNet V2	600 × 600 × 3	Random horizontal flip	1	Constant LR	0.0003
Faster-RCNN NASNet	1200 × 1200 × 3	Random horizontal flip	1	Constant LR	0.0003

Table 3. Model training settings.

Methods

Models description. We applied machine learning algorithms to detect coronary artery stenosis on the coronary angiography imaging series (see “Source Data” section). We examined eight models with various architectures, network complexity, and a number of weights: SSD²², Faster-RCNN²³, and RFCN²⁴ object detectors from the Tensorflow Detection Model Zoo²⁵ based on MobileNet^{26,27}, ResNet^{28,29}, Inception ResNet³⁰ and NASNet^{31,32}. The lightweight SSD MobileNet V1 and SSD MobileNet V2 enable real-time data processing. While Faster-RCNN Inception ResNet V2 and Faster-RCNN NASNet, with over 50 million weights, were the most complex models selected for the study. Table 2 shows a brief description of the models. Characteristics of neural networks, including mAP, are reported based on their training on the COCO dataset.

Model training. When training neural network models, their base configuration is similar to that used to train on the COCO 2017 dataset. For the unambiguous comparison of the selected models, the total number of training steps was set to 100 equal to 100'000 iterations of learning. Regarding the loss functions, the weighted Smooth L1 loss (see equation 3 in³³) was the localization loss, and the Weighted Focal Loss was the classification loss³⁴. The SSD-based models were trained using the cosine decay with the warm-up and exponential decay. When using these techniques, the learning rate gradually decreased depending on the learning step. It is also worth noting a distinctive feature of the SSD MobileNet V2 neural network, which is the use of the Hard Example Mining technique^{22,35}. It allows getting additional samples of the negative class and then learns from them. Using additional samples often improves the accuracy of the stenosis location.

To train the abovementioned networks, we used models pre-trained on the COCO 2017 dataset. Using Amazon SageMaker, we tuned given models and found their best versions through a series of training jobs run on the collected dataset. Having performed hyperparameter tuning based on Bayesian optimization strategy, a set of hyperparameter values for the best performing models was found, as measured by a validation mAP. Since the network architectures significantly vary and include many parameters, we summarize the main characteristics of the training in Table 3. To train models, we used P2 (Nvidia Tesla K80 12 Gb, 1.87 TFLOPS) and P3 instances (Nvidia Tesla V100 16 Gb, 7.8 TFLOPS) from Amazon Web Services. We also divided the models into 4 groups according to their complexity for further comparison.

Serial changes in accuracy were obtained on the validation set during the training process. Two evaluation metrics were used to compare the performance of the selected neural networks. Precision, Recall, and F1-score were used to compare the classifiers and the mAP metric was used to judge object localization³⁶. For mAP a predefined threshold value for Intersection over Union equal to 0.5 was used.

Figure 5 shows smooth changes in the mAP on the validation set during the training process. All models converge to a specific value of the asymptotic accuracy. SSD ResNet-50 V1 could achieve higher quality with longer training, but this would require more steps.

Results

Comparative assessment. Table 4 presents the results of the comparative study of the neural networks. In addition to the absolute values of the metrics, the relative values are also reported. The metrics of SSD MobileNet V1 were used as a benchmark to compare with other models. Color scale formatting reflects the distribution of models by their accuracy, training and inference times, and a number of weights, where deep blue shows the best value, and white—the worst. Figures 6, 7 and 8 show three basic metrics, the inference time, mAP, F1-score for the prediction of the stenotic lesion bounding box on an image.

The inference time was estimated using the P3 instance (Nvidia Tesla V100 16 Gb, 7.8 TFLOPS) of Amazon Web Services. We concluded that the inference time directly depended on the complexity of the model and the total number of its weights. Thus, Faster-RCNN Inception ResNet v2 and Faster-RCNN NASNet were the slowest in predictions. Their mean processing times per one image were 363 and 880 ms, respectively. While testing the lightweight models based on the MobileNet backbone, we found that MobileNet V2 with a larger number of weights (6.1 mln) demonstrated superior inference time than Mobile Net V1 (4.2 mln). In general, MobileNet V2 had the most superior inference time than other models. Thus, it may be used for predicting the location of stenosis in real-time.

In terms of the mAP metric and F1-score, Faster-RCNN Inception ResNet V2 was the most accurate model. The mean Average Precision of this model on the validation set was 0.95, F1-score 0.96 with the inference time of 363 ms/image (≈ 3 frames per second). The fastest and relatively lightweight SSD MobileNet V2 had the mean Average Precision of 0.83, F1-score 0.80 with an inference time of 26 ms/image (≈ 38 frames per second). Based on the obtained results, we concluded that RFCN ResNet-101 V2 is an optimal one to solve the set tasks. The mAP of this model is 0.94, F1-score 0.96 and the inference time is 99 ms/image (≈ 10 frames per second). In terms of both classification (F1 score) and localization (mAP) metrics, Faster-RCNN ResNet-101 V2, RFCN ResNet-101 V2, and Faster-RCNN Inception ResNet V2 remain the most effective models for the task of stenosis detection. Additional performance metrics, such as precision and recall, are reflected in Online Appendices F and G.

Model testing. The capabilities of the selected neural networks are presented using the data of three patients with the referenced labeling (Fig. 9a–c). Detailed visualization for predictions is presented in Online Appendices H–J. The models with the best values of the loss function and mAP were used for testing.

Table 5 reports the best steps with the model optimal weights. Such localization metrics as Intersection over Union (IoU) and Dice Similarity Coefficient (DSC) were also computed and shown.

Almost all models may accurately detect the location of stenosis. However, we faced several false positives while testing the Faster-RCNN NASNet model. In all three cases, this model detected the location of false stenotic segments with a probability of more than 90% in the right coronary artery (Fig. 9d) and the anterior descending artery (Fig. 9e, f) besides the reference stenotic region. SSD MobileNet V1 and SSD ResNet-50 V1 models failed to detect the location of stenosis in patient 1. SSD MobileNet V2 model demonstrated one of the best results in predicting the location of stenosis (Fig. 10). Despite the DSC metric of 0.65 in patient 3, it had the highest DSC metric in patients 1 and 2 (0.93 and 0.98, respectively). Additionally, the detectors based on the ResNet architecture, Faster-RCNN ResNet-50 V1 and Faster-RCNN ResNet-101 V2, should be noted. The average DSC metric on the test data was 0.85 and 0.84, respectively.

Discussion

The ultimate goal of our study is to develop a novel stenosis detection algorithm for patients with multivessel CAD, as they represent the most difficult group for diagnosis and interpretation. We believe that automatic detection and grading of multivessel CAD may facilitate the operator work by minimizing the risk of misinterpretation and accelerate the decision-making regarding the proper treatment strategy. To date, the accuracy and certainty of interpreting coronary angiograms fully rely on the operator who needs to identify the location of the stenosis and describe individual coronary vasculature, including the diameter of the affected vessels, the length of the stenotic segments, the presence of any lateral branches, any shunts, tortuosity, etc.³⁷. We have successfully tested our algorithm for detecting single-vessel CAD to assess its potential for the key task. Real-time detection of multivessel disease and its automatic grading is a more complex and multicomponent task. According to the obtained results, we concluded that the current version of our algorithm fully corresponds to the following key criteria—sufficient processing speed and detection accuracy.

Image processing speed. From the technical point of view, the speed of the algorithm for real-time detecting coronary artery stenosis and grading its severity is one of the key parameters empowering accurate CAD diagnosis and treatment. Coronary angiography is an invasive procedure that is associated with radiologic exposure, obviating repeated contrast injections, and limiting interventional cardiologists in their manipulations. In this respect, the ability to perform real-time detection of the stenotic lesions and their simultaneous grading in the cath-lab significantly increases the diagnostic efficiency (e.g. if the algorithm is sufficiently accurate, the operator may refuse additional contrast injection and proceed with stenting). Algorithms that generate predictions slowly (inference time of 600–800 s per angiography projection) are limited in use. They should be used separately, after coronary angiography, and may serve for off-line research descriptive tasks. Since the prolonged door-to-balloon time significantly affects the patient's outcome³⁸ and is directly associated with mortality³⁹, the

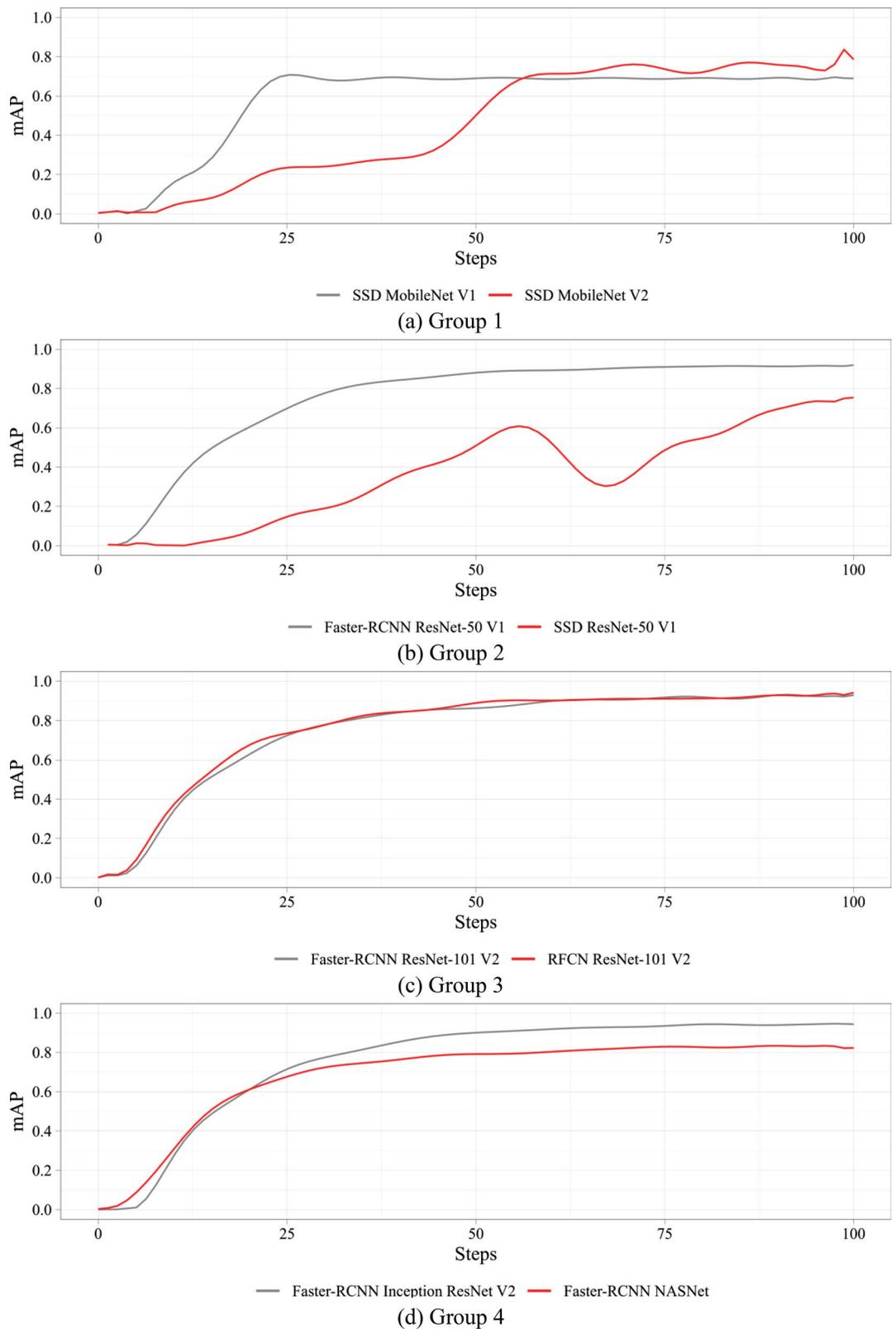


Figure 5. Dynamics of the mAP metric over the training process.

minimization of time spend on diagnosis will facilitate the decision-making process, especially for severe cases (e.g. myocardial infarction).

Model	Weights, mln		Training time, hours		Inference time, ms		F1-score		mAP@0.5	
	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.
SSD MobileNet V1	4.2	1.0×	16	1.0×	43	1.0×	0.72	1.0×	0.69	1.00×
SSD MobileNet V2	6.1	1.4×	20	1.3×	26	0.6×	0.80	1.10×	0.83	1.20×
SSD ResNet-50 V1	25.6	6.0×	47	3.0×	61	1.4×	0.73	1.01×	0.76	1.09×
Faster-RCNN ResNet-50 V1	25.6	6.0×	28	1.8×	98	2.3×	0.88	1.21×	0.92	1.33×
RFCN ResNet-101 V2	44.7	10.5×	55	3.6×	99	2.3×	0.96	1.32×	0.94	1.36×
Faster-RCNN ResNet-101 V2	44.7	10.5×	55	3.5×	118	2.7×	0.96	1.32×	0.94	1.35×
Faster-RCNN Inception ResNet V2	55.9	13.2×	93	6.0×	363	8.4×	0.94	1.30×	0.95	1.38×
Faster-RCNN NASNet	88.9	21.0×	147	9.5×	880	20.4×	0.82	1.13×	×	1.22×

Table 4. Comparative study of the selected models.

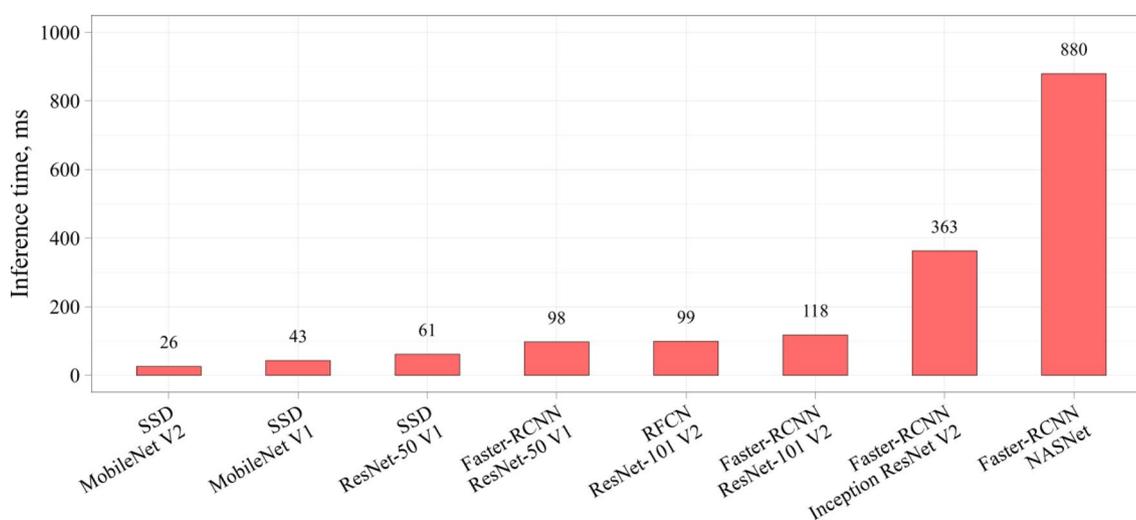


Figure 6. The inference time of the selected neural network models.

The existing research teams mainly focus on the accuracy of the algorithms rather than their speed. Most of them do not fit for routine medical image processing. Some of the recently reported image processing algorithms are generally perceived as slow with a high “cost” of frame analysis: Fang et al. reported the inference time varying from 1.1 to 11.87 s¹⁰; M’Hiri et al.—20 s¹³; and Wan et al.—63.3 s. to build the skeleton of the artery and 70.9 s. for the subsequent processing cycle⁹. Other studies have demonstrated a faster data analysis, spending almost 1.8 s per artery¹⁷, and 32 ± 21 s per each stenotic segment²⁰. However, these algorithms use computed tomography imaging series, which are commonly obtained during routine preoperative management but not urgently. Therefore, they are spending much more time on the descriptive analysis, empowering the decision-making process. Yang et al. have recently reported the use of convolutional neural networks for segmenting

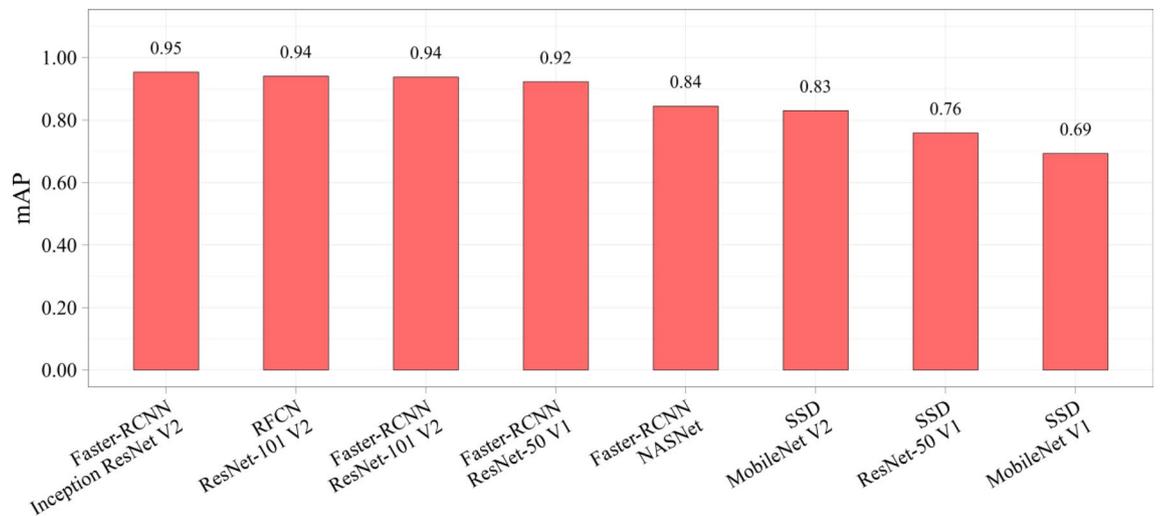


Figure 7. The mAP metric of the selected neural network models.

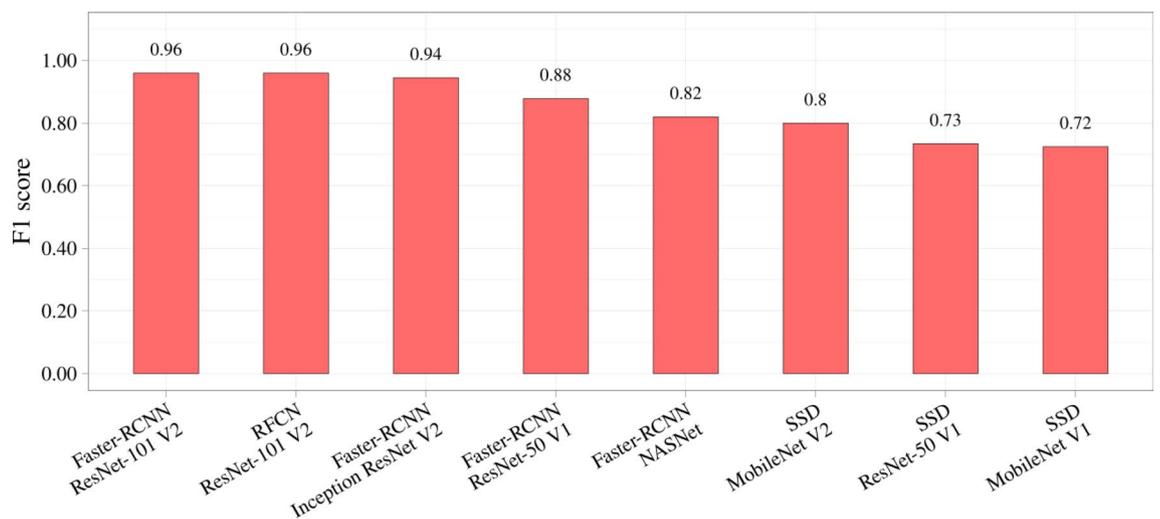


Figure 8. The F1 score of the selected neural network models.

major coronary arteries¹⁸. The algorithm spends 60 ms per angiogram, but it does not predict stenotic lesions of other small vessels.

There are no strict requirements for the processing speed of the angiography imaging series. It depends mainly on individual application settings. Thus, algorithms developed to support diagnostic angiography, performed with the aim of subsequent emergent blood flow restoration, should correspond to the following requirements: input video frame rate of 7.5–15 frames per second^{40,41}, the duration of the procedure less than 25 min, and individual preferences of the operator³⁶. We concluded that neural network architectures with an inference time of less than 66 ms are suitable for this task (Table 4. SSD MobileNet V1, SSD MobileNet V2, and SSD ResNet-50 V1), as they process at least 15 frames per second. However, their performance was assessed on a relatively simple case requiring detecting the location of stenosis without calculating its quantitative parameters. Thus, we expect that a detailed analysis of multivessel CAD may require a much longer time. Neural network models with the inference time of 98–118 ms per frame (Table 4. Faster-RCNN ResNet-50 V1, RFCN ResNet-101 V2, and Faster-RCNN ResNet-101 V2) may be assigned to the “grey zone”, processing 8–10 frames per second. Their resultant performance is insufficient, but they can be used in the cath-lab with the detection lag. The heavyweight models with the inference time of over 360 ms per frame (Table 4. Faster-RCNN Inception ResNet V2 and Faster-RCNN NASNet), do not fully correspond to the needs of the real-time angiography analysis, as they will fail to provide adequate productivity in complex cases.

CNN performance correlates with the complexity of their architectures. The number of weights is the foremost parameter responsible for the inference time. An increase in the number of weights has resulted in improved inference time (Table 4). Therefore, a number of CNN developers (e.g. GoogLeNet, ResNet, MobileNetV2) aim at minimizing the number of weights and size of neural networks for real-time applications, compacting them, and reducing the requirements for hardware performance^{45,46}. Different approaches to these modifications have been reported, including neural network compression accelerating the inference time: tensor decomposition,

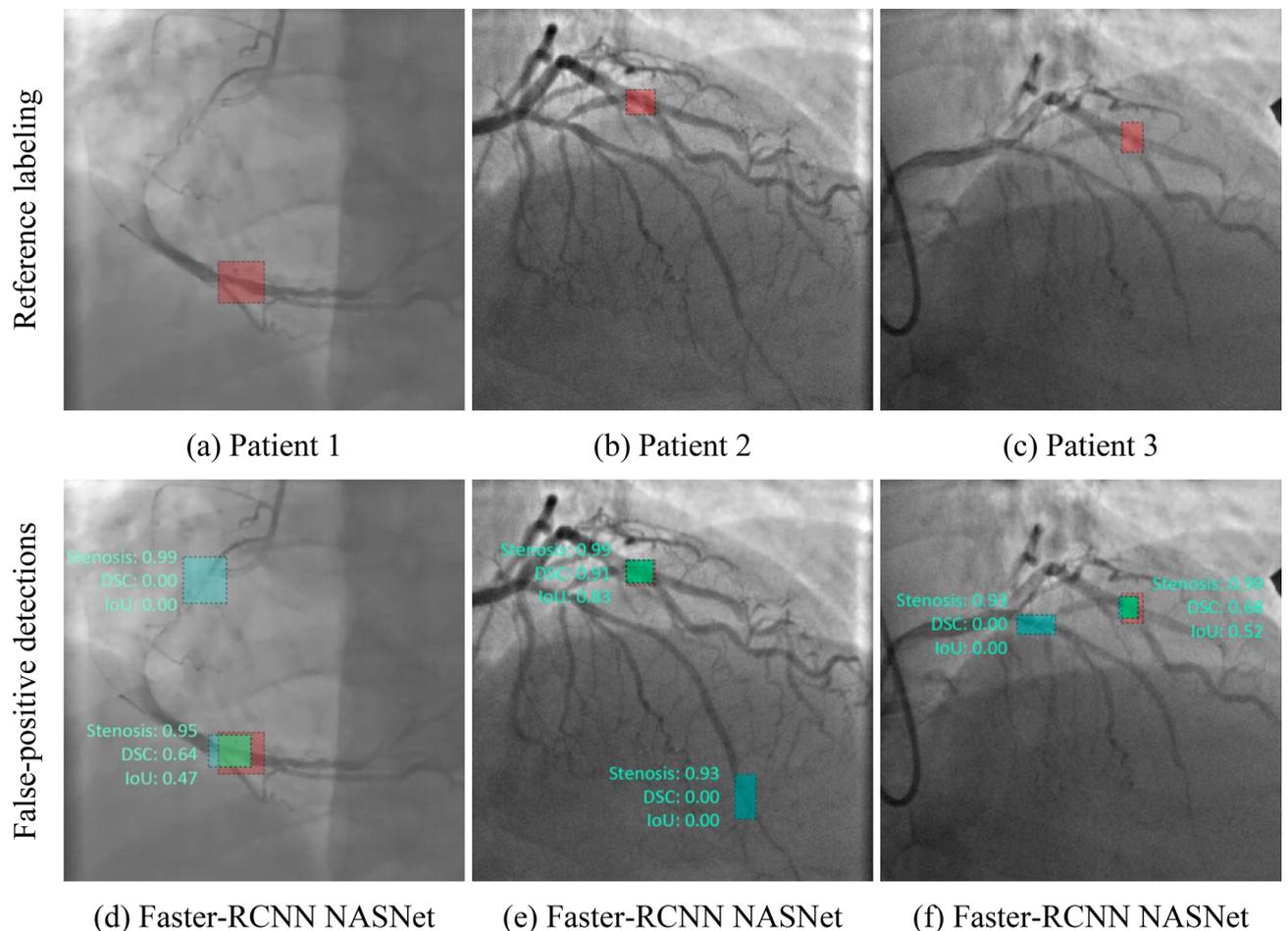


Figure 9. Example of the false positive predictions obtained using Faster-RCNN NASNet neural network for three test patients with the referenced labelling.

Model	Best step
SSD MobileNet V1	24
SSD MobileNet V2	99
SSD ResNet-50 V1	100
Faster-RCNN ResNet-50 V1	84
RFCN ResNet-101 V2	97
Faster-RCNN ResNet-101 V2	94
Faster-RCNN Inception ResNet V2	83
Faster-RCNN NASNet	95

Table 5. Best steps with optimal model weights.

quantization⁴⁷, pruning⁴⁸, teacher-student approaches⁴⁹, specific layer pruning and fusions⁵⁰, using many fewer proposals than is usual for Faster R-CNN¹⁸, Low-rank decomposition⁵¹.

Accuracy. Detection accuracy is another important parameter indicative to the quality of the algorithm, particularly for borderline cases, when the treatment strategy is not clearly defined and false positives may mislead the Heart Team to choose a more invasive treatment option³⁸. Therefore, it seems necessary to discuss these two cases separately—false positives and false negatives in the detection of stenosis. A false positive is an error in data reporting when an algorithm detects incorrectly the presence of stenosis. It may result in choosing coronary artery bypass grafting (CABG) rather than PCI since the operator relies on the misinterpreted data regarding the multiple stenotic lesions that increase individual SYNTAX Score^{38,42}. Thus, we should take seriously false positives produced by the Faster-RCNN NASNet network, that misinterpreted the clinical states of three control patients (Fig. 9d–f). Alternatively, a false negative is an error in data reporting when an algorithm reports the

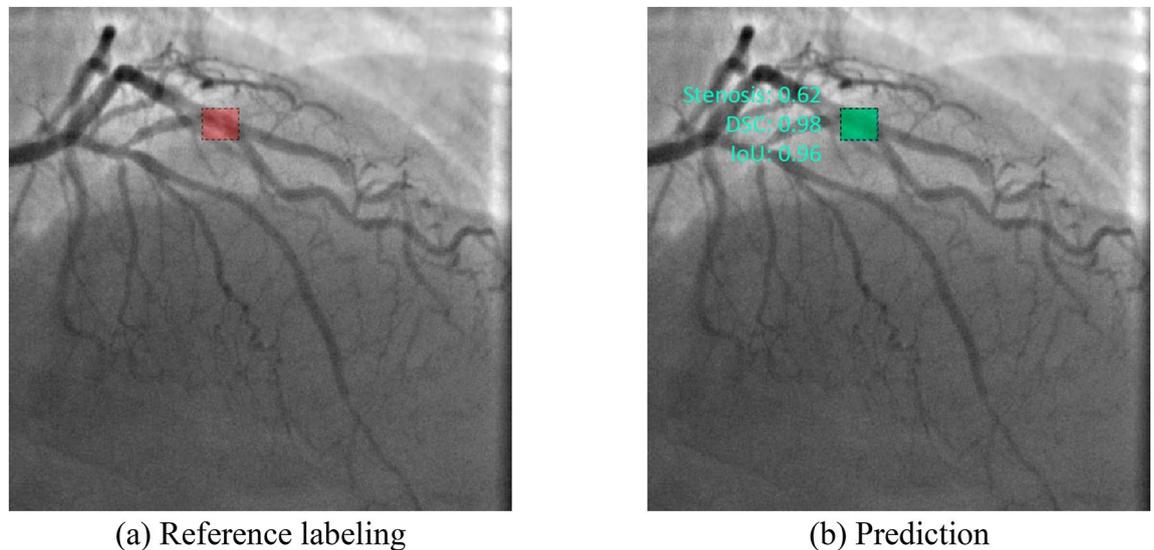


Figure 10. Example of the best prediction compared to reference labeling: data of patient 2 processed with SSD MobileNet V2 network.

absence of the existing stenosis. However, false negatives are less serious than false positives, as they can be leveled out during stenting by repeated contrast injection that will visualize the missed stenosis. This type of error was encountered for the two selected neural networks, the lightweight SSD MobileNet V1 and SSD ResNet-50 V1. Both these models showed the worst mAP of 0.69 and 0.76; F1-score of 0.72 and 0.73, respectively. Since these neural networks have demonstrated the worst mAP and F1-score, they are considered to be unpromising candidates for further optimization. Other models with an mAP of 0.94–0.95 and F1-score > 0.9 (Table 4) have room for further acceleration to detect multivessel CAD.

Resultant values of the classification and localization metric parameters are generally consistent with the recently published studies. Fang et al. reported an F1-score of 0.81–0.89¹⁰. Similar results were shown by Wan et al.⁹ and Zheik et al.¹⁷ equal to 0.83–0.94 and 0.75–0.88, respectively. While Yang et al. demonstrated the range of F1-score from 0.64 to 0.94²⁴. Faster-RCNN InceptionResNet-v2 has been reported as the most accurate (F1-score up to 0.94) in a similar study focusing on exploring the performance of CNN architectures for detecting large arteries²⁴. In our study, F1-score ranged from 0.72 to 0.96. The direct comparison of mAP values with those obtained in other studies is complicated by the different underlying performance metrics, as the Dice coefficient was reported. Therefore, we computed the Dice Similarity Coefficient that varied from 0.64 to 0.93 on the validation set and found that our data are in line with the previously reported studies: the Sensitivity metric varying from 0.59 to 0.72 in¹⁹, the Dice Similarity Coefficient of 0.75 in¹³ and 0.74 to 0.79 in¹².

We found that RFCN ResNet-101 V2 neural network provides the best speed/accuracy trade-off. In addition, the task for real-time CAD detection may be progressed through its modification and hardware upgrade^{18,47–51}. This balance may be achieved for other high-speed CNNs (SSD MobileNet V2) by improving their accuracy. Both, the accuracy and the number of errors, may potentially be improved using traditional approaches, including an increase of the training set size and its heterogeneity in addition to the use of more scalable and efficient neural network architectures (e.g. EfficientDet or CenterNet detectors^{43,44}).

Conclusion

The imbalance between accuracy and computer performance has been previously limited to the introduction of an automatic CAD detection algorithm in clinical practice. We have demonstrated that the development of hardware performance and appearance of the recent neural network architectures may significantly reduce the labor-intensive process during conventional invasive coronary angiography. We trained eight promising detectors based on different neural network architectures (MobileNet, ResNet-50, ResNet-101, Inception ResNet, NAS-Net) to detect the location of stenotic lesions using angiography imaging series and assessed their performance. Out of them, three neural networks have demonstrated superior results. Faster-RCNN Inception ResNet V2 is the most accurate to detect single-vessel disease. It demonstrates the mean Average Precision of 0.954, and the prediction rate of 363 ms per image (\approx 3 frames per second) on the validation set. The relatively lightweight SSD MobileNet V2 model is the fastest with an mAP of 0.830 and a mean prediction rate of 26 ms per image (\approx 38 frames per second). RFCN ResNet-101 V2 has demonstrated an optimal accuracy-to-speed ratio. Its mAP is 0.94, and the prediction speed is 99 ms per image (\approx 10 frames per second). The resultant performance-accuracy balance using the described neural networks has confirmed the feasibility of real-time CAD tracking supporting the decision-making process of the Heart Team. Real-time automatic labeling has opened new horizons for the diagnosis and treatment of complex coronary artery disease.

Received: 17 December 2020; Accepted: 24 March 2021

Published online: 07 April 2021

References

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)* **392**(10159), 1736–1788 (2018).
2. Virani, S. S. *et al.* Heart disease and stroke statistics-2020 update: a report from the American heart association. *Circulation* **141**(9), e139–e596 (2020).
3. Jensen, R. V., Hjortbak, M. V. & Bøtker, H. E. Ischemic Heart Disease: An Update. *Semin. Nucl. Med.* **50**, 195–207 (2020).
4. Knuuti, J. *et al.* 2019 ESC guidelines for the diagnosis and management of chronic coronary syndromes. *Eur. Heart J.* **41**(3), 407–477 (2020).
5. Saraste, A. & Knuuti, J. ESC 2019 guidelines for the diagnosis and management of chronic coronary syndromes: recommendations for cardiovascular imaging. *Herz. Springer Medizin* **45**(5), 409–420 (2020).
6. Collet, C. *et al.* Coronary computed tomography angiography for heart team decision-making in multivessel coronary artery disease. *Eur. Heart J.* **39**(41), 3689–3698 (2018).
7. Janssen, J. P. *et al.* New approaches for the assessment of vessel sizes in quantitative (cardio-)vascular X-ray analysis. *Int. J. Cardiovasc. Imag.* **26**(3), 259–271 (2010).
8. Gao, Y. & Sundar, H. Coronary arteries motion modeling on 2D X-ray images. In *Medical Imaging 2012: Image-Guided Procedures, Robotic Interventions, and Modeling* (eds Holmes, D. R., III & Wong, K. H.) 83161A (SPIE, 2012).
9. Wan, T. *et al.* Automated identification and grading of coronary artery stenoses with X-ray angiography. *Comput. Methods Progr. Biomed.* **167**, 13–22 (2018).
10. Fang, H. *et al.* Greedy soft matching for vascular tracking of coronary angiographic image sequences. *IEEE Trans. Circuits Syst. Video Technol.* **30**, 1466–1480 (2020).
11. M'Hiri F. *et al.* Vesselwalker: coronary arteries segmentation using random walks and Hessian-based vesselness filter. In *Proceedings: International Symposium on Biomedical Imaging*, 918–921 (2013).
12. M'Hiri F. *et al.* Hierarchical segmentation and tracking of coronary arteries in 2D X-ray Angiography sequences. In *Proceedings: International Conference on Image Processing, ICIP. IEEE Computer Society*, vol. 2015-December, 1707–1711 (2015).
13. M'Hiri, F. *et al.* A graph-based approach for spatio-temporal segmentation of coronary arteries in X-ray angiographic sequences. *Comput. Biol. Med.* **79**, 45–58 (2016).
14. Ren, X. & Malik J. Learning a classification model for segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*. Institute of Electrical and Electronics Engineers Inc., vol. 1, 10–17 (2003).
15. Jo, K. *et al.* Segmentation of the main vessel of the left anterior descending artery using selective feature mapping in coronary angiography. *IEEE Access* **7**, 919–930 (2019).
16. Nasr-Esfahani, E. *et al.* Segmentation of vessels in angiograms using convolutional neural networks. *Biomed. Signal Process. Control* **40**, 240–251 (2018).
17. Zreik, M. *et al.* A recurrent CNN for automatic detection and classification of coronary artery plaque and stenosis in coronary CT angiography. *IEEE Trans. Med. Imag.* **38**(7), 1588–1598 (2019).
18. Yang, S. *et al.* Deep learning segmentation of major vessels in X-ray coronary angiography. *Sci. Rep. Nat. Res.* **9**(1), 1–11 (2019).
19. Cong, C. *et al.* automated stenosis detection and classification in x-ray angiography using deep neural network. In *Proceedings: 2019 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2019*, (2019).
20. Hong, Y. *et al.* Deep learning-based stenosis quantification from coronary CT angiography. In *Proceedings of SPIE-the International Society for Optical Engineering*. vol. 10949, p. 88 (2019).
21. Patel, M. R. *et al.* ACC/AATS/AHA/ASE/ASNC/SCAI/SCCT/STS 2017 appropriate use criteria for coronary revascularization in patients with stable ischemic heart disease. *J. Am. Coll. Cardiol.* **69**(17), 2212–2241 (2017).
22. Liu, W. *et al.* SSD: single shot multibox detector. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, vol. 9905 LNCS, 21–37 (2016).
23. Ren, S. *et al.* Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(6), 1137–1149 (2017).
24. Dai, J. *et al.* R-FCN: Object detection via region-based fully convolutional networks. In *Advances in Neural Information Processing Systems. Neural Information Processing Systems Foundation*, 379–387 (2016).
25. Huang, J. *et al.* Speed/accuracy trade-offs for modern convolutional object detectors. In *Proceedings: 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. Institute of Electrical and Electronics Engineers Inc.*, vol. 2017-Janua., 3296–3305 (2017).
26. 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.* 4510–4520 (2018).
27. Sandler, M. *et al.* MobileNetV2: inverted residuals and linear bottlenecks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE*, 4510–4520 (2018).
28. He, K. *et al.* Identity mappings in deep residual networks. In *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, vol. 9908 LNCS. 630–645 (2016).
29. He, K. *et al.* Deep residual learning for image recognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society*, vol. 2016-Decem. 770–778 (2016).
30. Szegedy, C. *et al.* Inception-v4, inception-ResNet and the impact of residual connections on learning. In *31st AAAI Conference on Artificial Intelligence, AAAI 2017. AAAI press*, 4278–4284 (2017).
31. Zoph, B. & Le Q. V. Neural architecture search with reinforcement learning. In *5th International Conference on Learning Representations. ICLR 2017: Conference Track Proceedings*, (2016).
32. Zoph B. *et al.* Learning transferable architectures for scalable image recognition. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE*, 8697–8710 (2018).
33. Girshick R. Fast R-CNN. In *2015 IEEE International Conference on Computer Vision (ICCV). IEEE*, 1440–1448 (2015).
34. Lin, T.-Y. *et al.* Focal loss for dense object detection. In *2017 IEEE International Conference on Computer Vision (ICCV). IEEE*, vol. 42, (2), 2999–3007 (2017).
35. Shrivastava, A., Gupta, A. & Girshick R. Training region-based object detectors with online hard example mining. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society*, vol. 2016-Decem. 761–769 (2016).
36. Ferreira-Neto, A. N. *et al.* Clinical and technical characteristics of coronary angiography and percutaneous coronary interventions performed before and after transcatheter aortic valve replacement with a balloon-expandable valve. *J. Interv. Cardiol.* **2019**, 3579671 (2019).
37. Farooq, V., Brugaletta, S. & Serruys, P. W. The SYNTAX score and SYNTAX-based clinical risk scores. *Semin. Thorac. Cardiovasc. Surg. USA.* **23**(2), 99–105 (2011).

38. Sousa-va, M. *et al.* 2018 ESC/EACTS guidelines on myocardial revascularization. *Eur. J. Cardio-Thoracic Surg. Off. J. Eur. Assoc. Cardio-Thoracic Surg.* **55**(1), 4–90 (2019).
39. De Luca, G. *et al.* Time delay to treatment and mortality in primary angioplasty for acute myocardial infarction: every minute of delay counts. *Circulation* **109**, 1223–1225 (2004).
40. Abdelaal, E. *et al.* Effectiveness of low rate fluoroscopy at reducing operator and patient radiation dose during transradial coronary angiography and interventions. *JACC Cardiovasc. Interv.* **7**(5), 567–574 (2014).
41. Badawy, M. K. *et al.* Feasibility of using ultra-low pulse rate fluoroscopy during routine diagnostic coronary angiography. *J. Med. Radiat. Sci.* **65**(4), 252–258 (2018).
42. Capodanno, D. *et al.* Usefulness of SYNTAX score to select patients with left main coronary artery disease to be treated with coronary artery bypass graft. *JACC. Cardiovasc. Interv. USA* **2**(8), 731–738 (2009).
43. Tan, M., Pang, R. & Le, Q. V. EfficientDet: Scalable and efficient object detection. [arXiv:1911.09070](https://arxiv.org/abs/1911.09070) (2019).
44. Duan, K. *et al.* CenterNet: Keypoint Triplets for Object Detection. [arXiv:1904.08189](https://arxiv.org/abs/1904.08189) (2019).
45. Huang, Y., Qiu, C., Wang, X., Wang, S. & Yuan, K. A compact convolutional neural network for surface defect inspection. *Sensors* **2020** **20**, 1974 (2020).
46. Yu, D. *et al.* An efficient and lightweight convolutional neural network for remote sensing image scene classification. *Sensors* **20**, 1999 (2020).
47. Xu, T.-B., Yang, P., Zhang, X.-Y. & Liu, C.-L. Margin-aware binarized weight networks for image classification. In: *Image and graphics. ICGI 2017. Lecture notes in computer science*, vol 10666. (eds Zhao Y., Kong X. & Taubman D.) 590–601. https://doi.org/10.1007/978-3-319-71607-7_52 (Springer, Cham, 2017).
48. Molchanov, P., Tyree, S., Karras, T., Aila, T. & Kautz, J. Pruning convolutional neural networks for resource efficient inference. In *5th International conference on learning representations, ICLR 2017 - Conference Track Proceedings* (2017).
49. Lebedev, V. & Lempitsky, V. Speeding-up convolutional neural networks: A survey. *Bull. Polish Acad. Sci. Tech. Sci.* **66**(6), 799–810 (2018).
50. Borkovkina, S., Camino, A., Janpongsri, W., Sarunic, M. V. & Jian, Y. Real-time retinal layer segmentation of OCT volumes with GPU accelerated inferencing using a compressed, low-latency neural network. *Biomed. Opt. Express* **11**, 3968–3984 (2020).
51. Tai, C., Xiao, T., Zhang, Y., Wang, X. & Weinan, E. Convolutional neural networks with low-rank regularization. In *4th International conference on learning representations, ICLR 2016 - Conference Track Proceedings* (2016).

Acknowledgements

Data mining, data pre-processing, and development of the ML-based approach to detect stenosis were supported by a grant from the Russian Science Foundation, Project No. 18-75-10061 “Research and implementation of the concept of robotic minimally invasive prosthetics of the aortic valve”. The training of the developed models using Amazon Web Services was funded by the Ministry of Science and Higher Education, Project No. FFSWW-2020-0014 “Development of the technology for robotic multiparametric tomography based on big data processing and machine learning methods for studying promising composite materials”.

Author contributions

V.D., K.K., and E.O. conceived the idea of the study. V.D., O.G., and A.F. contributed to the methodology used in the field of stenosis detection; K.K. and E.O. collected the data. V.D. prepared the software and algorithms for data analysis. V.D., A.K., V.G., and E.O. performed an analysis of the current state of research in the field of detection and quantitative assessment of stenosis; V.D. and A.F. trained deep learning models using the Multi-X platform. V.D. and K.K. wrote the manuscript. A.K., V.G., O.G., A.F., and E.O. reviewed and edited the manuscript. K.K. and E.O. contributed critical discussions and revisions of the manuscript. A.F. and E.O. were supervising and administering the project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-87174-2>.

Correspondence and requests for materials should be addressed to V.V.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021