



OPEN

## AngioNet: a convolutional neural network for vessel segmentation in X-ray angiography

Kritika Iyer<sup>1</sup>, Cyrus P. Najarian<sup>1</sup>, Aya A. Fattah<sup>1</sup>, Christopher J. Arthurs<sup>2</sup>, S. M. Reza Soroushmehr<sup>1</sup>, Vijayakumar Subban<sup>3</sup>, Mulasari A. Sankardas<sup>3</sup>, Raj R. Nadakuditi<sup>1</sup>, Brahmajee K. Nallamothu<sup>1</sup> & C. Alberto Figueroa<sup>1</sup>✉

Coronary Artery Disease (CAD) is commonly diagnosed using X-ray angiography, in which images are taken as radio-opaque dye is flushed through the coronary vessels to visualize the severity of vessel narrowing, or stenosis. Cardiologists typically use visual estimation to approximate the percent diameter reduction of the stenosis, and this directs therapies like stent placement. A fully automatic method to segment the vessels would eliminate potential subjectivity and provide a quantitative and systematic measurement of diameter reduction. Here, we have designed a convolutional neural network, AngioNet, for vessel segmentation in X-ray angiography images. The main innovation in this network is the introduction of an Angiographic Processing Network (APN) which significantly improves segmentation performance on multiple network backbones, with the best performance using Deeplabv3+ (Dice score 0.864, pixel accuracy 0.983, sensitivity 0.918, specificity 0.987). The purpose of the APN is to create an end-to-end pipeline for image pre-processing and segmentation, learning the best possible pre-processing filters to improve segmentation. We have also demonstrated the interchangeability of our network in measuring vessel diameter with Quantitative Coronary Angiography. Our results indicate that AngioNet is a powerful tool for automatic angiographic vessel segmentation that could facilitate systematic anatomical assessment of coronary stenosis in the clinical workflow.

Coronary Artery Disease (CAD) affects over 20 million adults in the United States and accounts for nearly one-third of adult deaths in western countries<sup>1-3</sup>. The annual cost to the United States healthcare system for the first year of treatment is \$5.54 billion<sup>4</sup>. The disease is characterized by the buildup of plaque in the coronary arteries<sup>5,6</sup>, which causes a narrowing of the blood vessel known as stenosis.

CAD is most commonly diagnosed using X-ray angiography (XRA)<sup>7</sup>, whereby a catheter is inserted into the patient and a sequence of X-ray images are taken as radio-opaque dye is flushed into the coronary arteries. Cardiologists typically approximate stenosis severity via visual inspection of the XRA images, estimating the percent reduction in diameter or cross-sectional area. If the area reduction is believed to be greater than 70%, a revascularization procedure, such as stent placement or coronary artery bypass grafting surgery, may be performed to treat the stenosis<sup>8,9</sup>.

Quantitative Coronary Angiography, or QCA, is a diagnostic tool used in conjunction with XRA to more accurately determine stenosis severity<sup>10,11</sup>. QCA is an accepted standard for assessment of coronary artery dimensions and uses semi-automatic edge-detection algorithms to quantify the change in vessel diameter. The QCA software then reports the diameter at user-specified locations as well as the percentage diameter reduction at the stenosis<sup>12</sup>. Although QCA is more quantitative than visual inspection alone, it requires substantial human input to identify the stenosis and to manually correct the vessel boundaries before calculating the stenosis percentage. This has led to QCA largely being used in the setting of clinical studies with limited impact on patient care. A fully automatic angiographic segmentation algorithm would speed up the process of determining stenosis severity, eliminate the need for subjective manual corrections, and potentially lead to broader utilization in clinical workflows.

Fully-automated angiographic segmentation is particularly challenging due to the poor signal-to-noise ratio and overlapping structures such as the catheter and the patient's spine and rib cage<sup>13</sup>. Several filter-based or region-growing approaches<sup>13-22</sup> have been developed for angiographic segmentation. The principal limitation

<sup>1</sup>University of Michigan, 500 S State St, Ann Arbor, MI 48109, USA. <sup>2</sup>King's College London, Strand, London, UK. <sup>3</sup>Madras Medical Mission, Chennai, Tamil Nadu, India. ✉email: [figueroa@med.umich.edu](mailto:figueroa@med.umich.edu)

of these methods is that they cannot separate overlapping objects such as catheters and bony structures from the vessels, requiring manual correction which can be time-consuming and subjective. To address these limitations, some have turned to convolutional neural networks (CNNs) for angiographic segmentation.

CNNs have been used for segmentation in numerous applications<sup>23–28</sup>. Many CNNs have been designed specifically for angiographic segmentation<sup>29–34</sup>, including several based on U-Net<sup>35</sup>. U-Net is a CNN designed for biomedical segmentation and has been widely adopted in other fields due to its relatively simple architecture and high accuracy on binary segmentation problems<sup>36–38</sup>. The main advantage of this network is that it can be trained on small datasets of hundreds of images due to its simple architecture, making it well-suited for medical imaging applications<sup>35</sup>. Yang et al.<sup>29</sup> developed a CNN based on U-Net to segment the major branches of the coronary arteries. Despite its high segmentation accuracy, this network was only developed for single vessel segmentation. Multichannel Segmentation Network with Aligned input (MSN-A)<sup>30</sup>, is another CNN based on U-Net designed to segment the entire coronary tree. The inputs to MSN-A are the angiographic image and a co-registered “mask” image taken before the dye was injected into the vessel. The main drawback of this network is that the multi-input strategy requires the entire angiographic sequence to be acquired with minimal table motion, whereas standard clinical practice involves moving the patient table to follow the flow of dye within the vessels. Nasr-Esfahani et al.<sup>31</sup> developed their own CNN architecture for angiographic segmentation, combining contrast enhancement, edge detection, and feature extraction. Shin et al.<sup>32</sup> combined a feature-extraction convolutional network with a graph convolutional network and inference CNN to create Vessel Graph Network (VGN) and improve segmentation performance by learning the tree structure of the vessels.

To address these shortcomings, we have developed a new CNN for angiographic segmentation: *AngioNet*, which combines an Angiographic Processing Network (APN) with a semantic segmentation network. The APN was trained to address several of the challenges specific to angiographic segmentation, including low contrast images and overlapping bony structures. *AngioNet* uses Deeplabv3+<sup>39,40</sup> as its backbone semantic segmentation network instead of U-Net or other fully convolutional networks (FCNs), which are more commonly used for medical segmentation. In this paper, we explored the specific benefits of the APN—and the importance of using Deeplabv3+ as the backbone—by comparing segmentation accuracy in Deeplabv3+, U-Net, and Attention U-Net<sup>36</sup>, trained both with and without the APN. Deeplabv3+ was chosen as the main backbone network since its deep architecture has greater expressive power<sup>41–43</sup> compared to the FCNs typically used for medical segmentation. Its ability to approximate more complex functions is what enables *AngioNet* perform well in challenging cases. We chose to investigate the effect of the APN on U-Net due to its extensive use in medical image segmentation. Attention-based networks have been shown to improve segmentation performance compared to pure CNNs by suppressing irrelevant features and learning feature interdependence<sup>36,44</sup>; thus, Attention U-Net was chosen as an appropriate network backbone for comparison. Lastly, we performed clinical validation of segmentation accuracy by comparing *AngioNet*-derived vessel diameter against QCA-derived diameter.

The main contribution of this work is the introduction of an APN to various segmentation backbone networks, creating an end-to-end pipeline encompassing image pre-processing and segmentation for angiographic images of coronary arteries. The APN was shown to improve segmentation accuracy in all three tested backbone networks. Furthermore, networks containing the APN were better able to preserve the topology of the coronary tree compared to the backbone networks alone. Another contribution of our work is the comparison against a clinically validated standard for measuring vessel diameter, QCA. This validation study provides more insight into *AngioNet*'s ability to accurately detect the vessel boundaries than traditional segmentation evaluation metrics. Our statistical analyses demonstrate that *AngioNet* and QCA are interchangeable methods of determining vessel diameter.

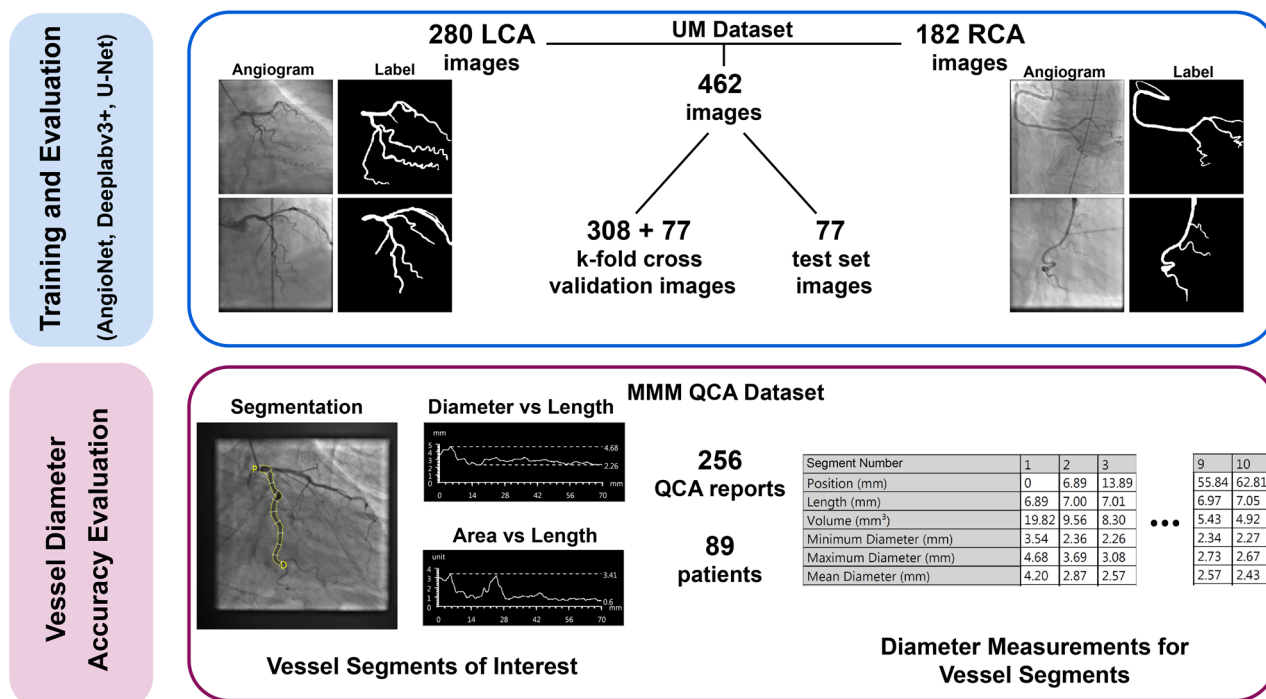
## Methods

**Datasets.** Figure 1 summarizes the two patient datasets used in this work for neural network training and evaluation of performance. All data were collected in compliance with ethical guidelines.

**UM dataset.** This dataset was composed of 462 de-identified angiograms acquired using a Siemens Artis Q Angiography system at the University of Michigan (UM) Hospital. The study protocol to access this data (HUM00084689) was reviewed by the Institutional Review Boards of the University of Michigan Medical School (IRBMED). Since the data was collected retrospectively, IRBMED approved use without requiring informed consent. This dataset was composed of patients who were referred for a diagnostic coronary angiography procedure at the UM Hospital in 2017. Patients with pacemakers, implantable defibrillators, or prior bypass grafts were excluded, as these prior procedures introduce artifacts and additional vascular conduits. Furthermore, patients with diffuse stenosis were excluded as this is less common in arteries suitable for revascularization. In our sample of 161 patients, 14 had severe stenosis ( $\leq 80\%$  diameter reduction) and the remaining had mild to moderate stenosis. The dataset was composed of 280 images of the left coronary artery (LCA) and 182 images of the right coronary artery (RCA).

The data were equally split by patient into a fivefold cross-validation set and test set to avoid having images from the same patient in both the training and test sets. Labels for all images were manually annotated to include vessels with a diameter greater than 1.5 mm (4 pixels) at their origin. Labels were created by selecting the vessels using Adobe Photoshop's Magic Wand tool followed by manual refinement of the vessel boundaries, and were reviewed by a board-certified cardiologist. The fivefold cross-validation portion of the dataset was used for neural network training and hyperparameter optimization, whereas the test set was used to evaluate segmentation accuracy.

There is a great number of artifacts in XRA images, including borders from X-ray filters, rotation of the image frame, varying levels of contrast, and magnification during image acquisition. Data augmentation of the UM



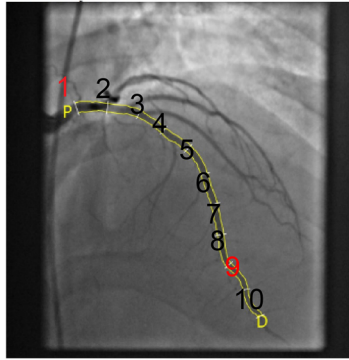
**Figure 1.** Diagram of datasets for CNN training and evaluation. AngioNet's performance was compared against state-of-the-art neural networks, all trained on the UM Dataset. The MMM QCA dataset was used to quantify segmentation diameter accuracy by comparing AngioNet's results against the diameters reported in QCA. Left coronary artery (LCA); right coronary artery (RCA); Madras Medical Mission (MMM).

dataset was employed to account for this variability. Horizontal and vertical flips of the images were included to make the network segmentation invariant to image orientation. Random zoom up to 20%, rotation up to 10%, and shear up to 5% were used to account for variation in magnification and imaging angles. When zooming out, shearing, or rotating the image, a constant black fill was used to mimic images acquired using physical X-ray filters. The combination of the above data augmentations created a training dataset of over half a million images to improve network generalizability. Data augmentation was not applied to the test set. The augmented UM dataset was used for neural network training, and the test set was used to compare segmentation accuracy.

**MMM QCA dataset.** The percent change in vessel diameter at the region of stenosis is a key determinant of whether a patient requires an intervention or not; therefore, the accuracy of AngioNet's segmented vessel diameters was assessed in addition to its overall segmentation accuracy. Although the main result of a QCA report is the overall percent change in vessel diameter, these reports also contain measurements of maximum, minimum, and mean diameter in 10 equal segments of the vessel of interest (Fig. 1). These diameter measurements in the MMM QCA Dataset were used to evaluate the discrepancies between QCA and AngioNet.

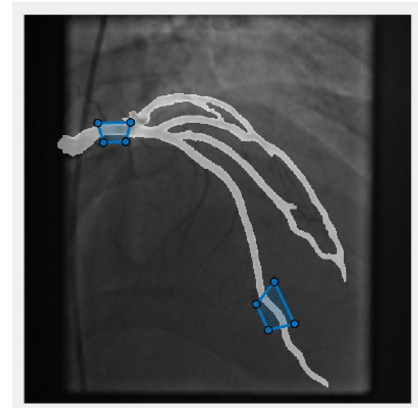
The Madras Medical Mission (MMM) QCA dataset contained independently generated three-vessel QCA reports of 89 patients, encompassing 223 vessels in both the LCA and RCA. All patients presented with mild to moderate stenosis. The data were acquired from the Indian Cardiovascular Core Laboratory (ICRF) at the MMM, which serves as a core laboratory with experience in clinical trials and other studies and has expertise in QCA. The data provided by the MMM ICRF Cardiovascular Core Laboratory includes independent and detailed analysis of quantitative angiographic parameters (minimum lesion diameter, percent diameter stenosis, etc.) as per American College of Cardiology/American Heart Association standards, through established QCA software (CAAS-5.10.2, Pie Medical Corp). The study protocol for this data (Computer-Assisted Diagnosis of Coronary Angiography) was approved by the Institutional Ethics Committee of the Madras Medical Mission. This data was obtained using an unfunded Materials Transfer Agreement between UM and MMM. Since the data is completely anonymized and cannot be re-identified, it does not qualify as human subjects research according to OHRP guidelines.

To validate the accuracy of AngioNet's segmented vessel diameters, a MATLAB script was employed for user specification of the same vessel regions as those in the QCA report. Two regions from the QCA report were sampled in each angiogram. The first was the most proximal region, containing the maximum vessel diameter, and the second was the region of stenosis (given in the QCA report), if present. If no stenosis was reported, the region containing minimum diameter was selected. A skeletonization algorithm<sup>45</sup> was used to identify the centerline and radius map of the selected vessel region. Using the output of the skeletonization algorithm, the script reported the maximum and minimum diameters at the selected regions and compared them against the diameters in the QCA report. Maximum and minimum vessel diameter were chosen rather than the diameters on either side of the stenosis since the purpose of using the QCA reports was to systematically assess overall

**a****QCA Report Diameter**

	1	2	3	4	5	6	7	8	9	10
Position (mm)	0	10.55	21.09	31.76	42.25	52.93	63.47	74.07	84.69	95.27
Length (mm)	10.55	10.54	10.67	10.49	10.67	10.55	10.60	10.62	10.58	10.62
Volume (mm <sup>3</sup> )	23.54	8.98	7.52	7.85	7.12	5.41	5.09	4.74	3.98	3.06
Minimum diameter (mm)	2.72	2.43	1.98	2.01	2.11	2.05	1.95	1.86	1.64	1.39
Maximum diameter (mm)	3.60	2.92	2.59	2.57	2.41	2.35	2.32	2.15	1.97	1.85
Mean diameter (mm)	3.15	2.73	2.16	2.39	2.27	2.19	2.13	1.98	1.82	1.62

Proximal Region max diameter: 3.60 mm  
Distal Region min diameter: 1.64 mm

**b****Segmentation Diameter**

Proximal Region max diameter: 3.74 mm  
Distal Region min diameter: 1.72 mm

**Figure 2.** (a) Annotated QCA report, along with the corresponding diameters in the report table. Highlighted values correspond to maximum (proximal) and minimum (distal) diameters (segments 1 and 9, respectively). (b) Schematic of how a MATLAB script was used to delineate regions in the neural network segmentation corresponding to the regions measured in the QCA report, along with the computed proximal and distal diameters.

vessel diameter accuracy, not the percent diameter reduction. A diagram of the comparison between QCA and AngioNet diameters is shown in Fig. 2.

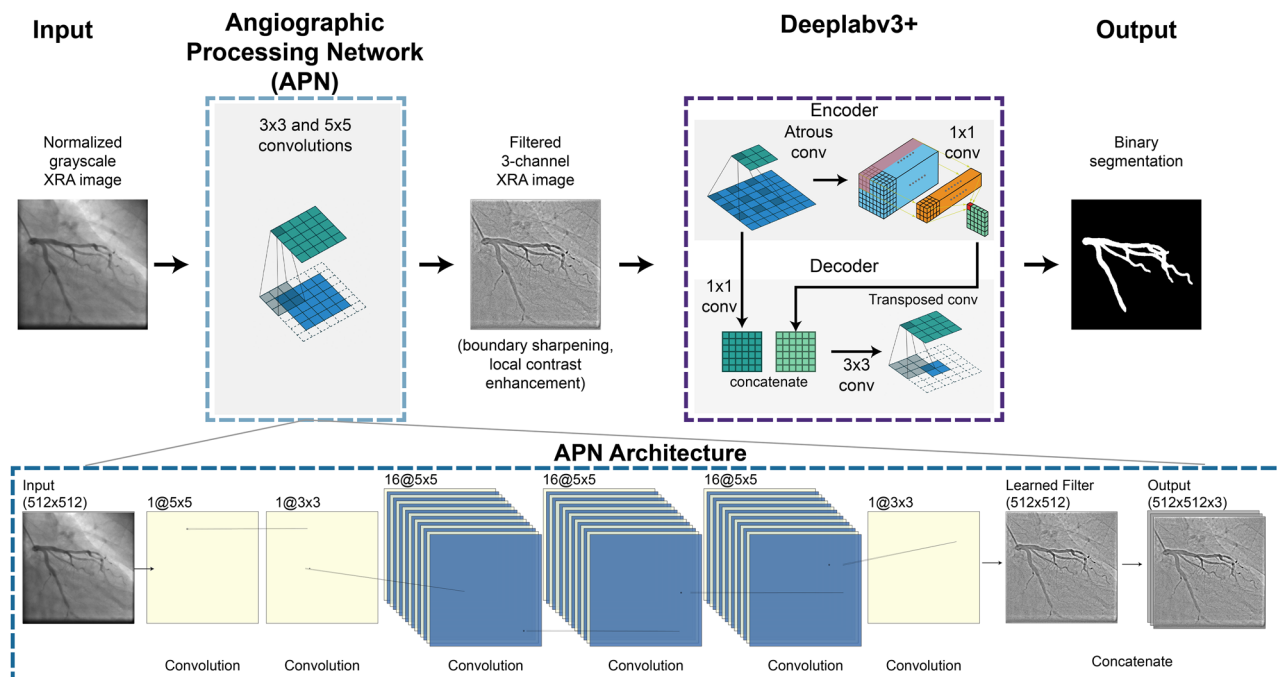
**CNN design and training.** *Design.* AngioNet was created by combining Deeplab v3+ and an Angiographic Processing convolutional neural network (APN). A diagram of the network architecture is given in Fig. 3. Each component of AngioNet, the APN and Deeplabv3+, was trained separately before fine-tuning the entire network.

The purpose of the APN was to address some of the challenges specific to angiographic segmentation, namely poor contrast and the lack of clear vessel boundaries. The intuition behind the APN was that the best possible pre-processing filter and its parameters are unknown; we hypothesized that learning the best possible filter would lead to higher accuracy than manually sampling several filters. The APN was initially trained to mimic a combination of standard image processing filters instead of initializing with random weights, since it would later be fine-tuned with a pre-trained backbone network. A combination of unsharp mask filters was chosen as these can improve boundary sharpness and local contrast at the edges of the coronary vessels, making the segmentation task easier. Starting with an initialization of unsharp masking, the fine-tuning process was used to learn a new filter that was best suited for angiographic segmentation. The single-channel filtered image from the APN was then copied and concatenated to form a 3-channel image, as this was the expected input of Deeplabv3+, a network typically used to segment RGB camera images (see Concatenation and Output in the APN Architecture, Fig. 3).

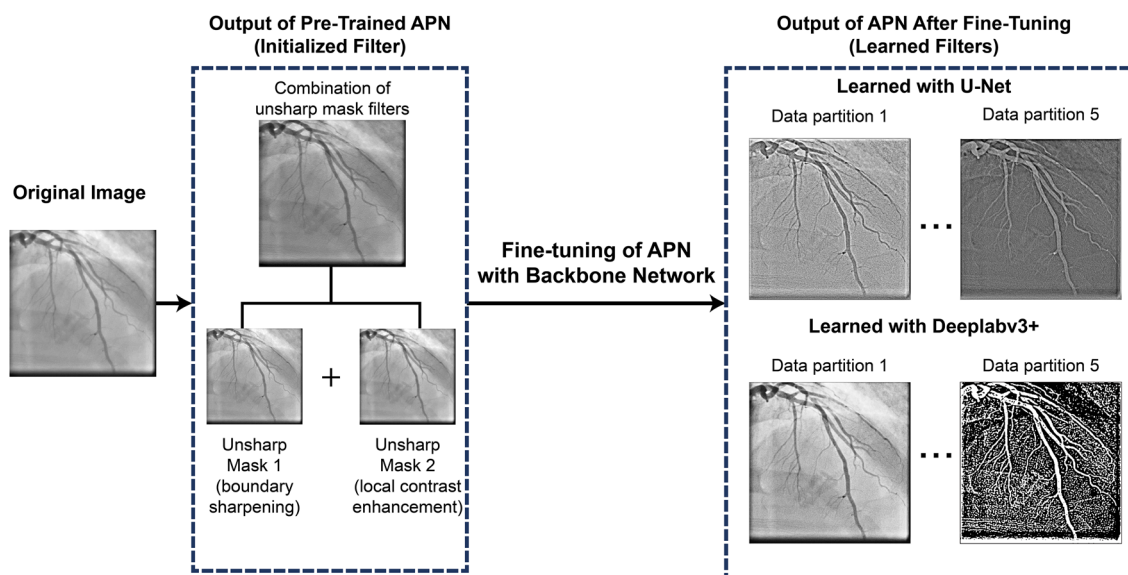
*Training.* The Deeplabv3+ CNN architecture was cloned from the official Tensorflow Deeplab GitHub repository, maintained by Liang-Chieh Chen and co-authors<sup>39</sup>. The network was initialized with pre-trained weights from their repository, as recommended by the authors for training on a new dataset. The input to this network were normalized angiographic coronary images, and the output was a binary segmentation. Training was conducted using four NVIDIA Tesla K80 GPUs on the American Heart Association Precision Medicine Platform (<https://precision.heart.org/>), hosted by Amazon Web Services. Hyperparameters such as batch size, learning rate, learning rate optimizer, and regularization were tuned. We observed that training with larger batch size led to better generalization to new data. A batch size of 16 was used as this was the largest batch size we could fit into memory using four GPUs. The Adam optimizer was chosen to adaptively adjust the learning rate, and L2 regularization was used to reduce the chance of over-fitting. The vessel pixels account for 15–19% of the total pixels in any given angiography image. Due to this class imbalance, it was important to encourage classification of vessel pixels over background using weighted cross-entropy loss<sup>46</sup>.

The APN was initially trained to mimic the output of several unsharp mask filters applied in series (parameters: radius = 60, amount = 0.2 and radius = 2, amount = 1) as seen in Fig. 4. This ensured the APN architecture was complex enough to learn the equivalent of multiple filters with sizes up to 121 × 121 using only 3 × 3 and





**Figure 3.** AngioNet Architecture Diagram. AngioNet is composed of an Angiographic Processing Network (APN) in tandem with Deeplabv3+. The APN is designed to improve local contrast and vessel boundary sharpness. The output of the APN, a single channel filtered image, is copied and concatenated to form a 3-channel image which is input into the backbone network.



**Figure 4.** Examples of learned filters when the APN is fine-tuned together with Deeplabv3+ and U-Net. The APN was initialized with the combination of unsharp mask filters shown above, and learned new filters to aid segmentation. Each example image is the output of the APN after training with different data partitions.

$5 \times 5$  convolutions. The number of  $3 \times 3$  versus  $5 \times 5$  convolutions as well as the network width and depth were adjusted until the APN could reproduce the results of the serial unsharp mask filters. The inputs to the APN were the normalized single-channel images from the augmented UM Dataset, whereas the output was a filtered version of the image, copied to form a 3-channel image. The ground truth images for the APN were generated by applying several unsharp mask filters with various parameters to each normalized clinical image. The APN was composed of several  $3 \times 3$  and  $5 \times 5$  convolutional layers (Fig. 3) and was trained to mimic the unsharp mask filters by minimizing the Mean Squared Error (MSE) loss between the prediction and ground truth images (MSE on the order of  $1e-4$ ). The APN design and training were carried out using TensorFlow 2.0, integrated with Keras<sup>47,48</sup>.

Once the APN and Deeplabv3+ networks were individually trained, the two CNNs were combined to form AngioNet using the Keras functional Model API<sup>49</sup>. The resulting network was trained with a low learning rate to fine-tune the combined model. Since neither the APN nor Deeplabv3+ were frozen during fine-tuning, both were able to adjust their weights to better complement each other: the APN learned a better filter than its unsharp mask initialization, and Deeplabv3+ learned the weights that could most accurately segment the vessel from the filtered image that was output by the APN. Network hyperparameters were once again tuned. The same process of pre-training, combining models, and fine-tuning was carried out with the APN and each of the other network backbones, U-Net and Attention U-Net, to determine how much the backbone network contributes to segmentation performance. U-Net and Attention U-Net were not initialized with pre-trained weights as our dataset was adequately large to train these networks from random initialization.

During all phases of training, batch normalization layers were frozen at their pre-trained values as we did not have a large enough dataset to retrain these layers. Furthermore, all hyperparameter optimization was performed on the fivefold cross validation holdout set and accuracy was measured on the test set.

## Results

**Learned filters using the angiographic processing network.** Figure 4 contains examples of the filters that the APN learned when it was trained with Deeplabv3+ or U-Net, respectively. The images represent the output of the APN, and thus the input to the backbone network. Although the APN was initialized with the combination of unsharp mask filters shown in Fig. 4, the network learns different filters that perform a combination of contrast-enhancement and boundary sharpening. The examples given are the results of training with different data partitions during k-fold cross-validation. The large variations in the learned filters come from an inherent property of neural network training; since minimization of the neural network's loss function is a non-convex optimization problem<sup>50</sup>, there are many combinations of network weights which will lead to similar values of the loss function, and consequently, similar overall accuracy. The effect of these varied learned filters on segmentation accuracy is described in the next section.

**Segmentation accuracy metrics.** Segmentation accuracy was measured using the Dice score, given by

$$Dice = \frac{2|Y \cap \hat{Y}|}{|Y| + |\hat{Y}|}. \quad (1)$$

Here,  $Y$  is the label image and  $\hat{Y}$  is the neural network prediction, each of which is a binary image where vessel pixels have a value of 1 and background pixels have a value of 0.  $|Y|$  denotes the number of vessel pixels (1s) in image  $Y$ , and  $\cap$  represents a pixel-wise logical AND operation. Alternatively, the Dice score can be defined in terms of the true positives (TP), false positives (FP) and false negatives (FN) of the neural network prediction with respect to the label image, and is then given by

$$Dice = \frac{2TP}{2TP + FP + FN} \quad (2)$$

In addition to the standard Dice score, centerline Dice or clDice was also measured. Rather than measuring pixel-wise accuracy, clDice measures connectivity of tubular structures and can be used to determine how well the predicted image maintains the topology of the vessel tree in the label image<sup>51</sup>. clDice is measured by finding the centerlines of the prediction and label images,  $cl_{\hat{Y}}$  and  $cl_Y$ . The proportion of  $cl_{\hat{Y}}$  which lies in the label  $Y$ ,  $cl_{\hat{Y}} \cap Y$ , and the proportion of  $cl_Y$  which lies in the prediction  $\hat{Y}$ ,  $cl_Y \cap \hat{Y}$ , are computed as analogs for precision and recall. clDice is then given by

$$clDice = 2 \times \frac{cl_{\hat{Y}} \cap Y \times cl_Y \cap \hat{Y}}{cl_{\hat{Y}} \cap Y + cl_Y \cap \hat{Y}}. \quad (3)$$

We also report the Area under the Receiver-Operator Curve (AUC), which measures the ability of the network to separate classes, in this case, vessel and background pixels. An AUC of 0.5 indicates a model that is no better than random chance, whereas an AUC of 1 indicates a model that can perfectly discriminate between classes. Finally, we also report the pixel accuracy of the binary segmentation, defined as

$$Pixel Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (4)$$

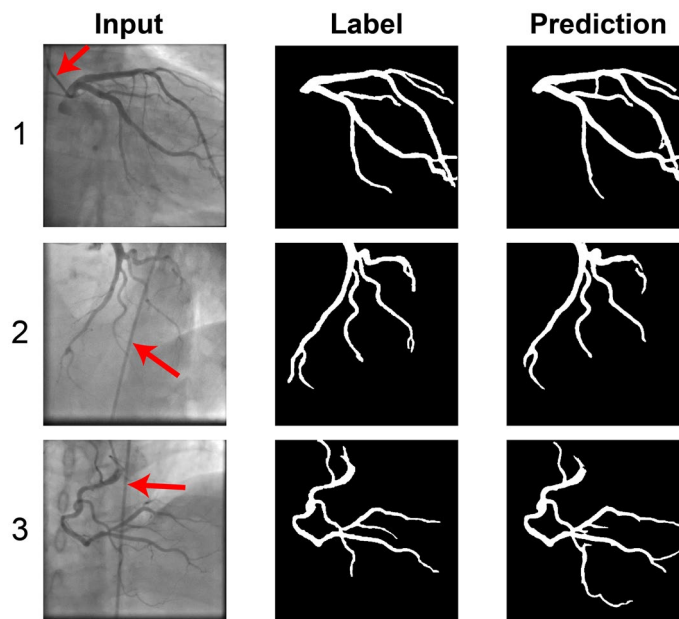
## Comparison of AngioNet versus current state-of-the-art semantic segmentation neural networks.

The accuracy of AngioNet was validated using a fivefold cross-validation study, in which the neural network was trained on 4 out of the 5 UM Dataset training partitions at a time, with the fifth partition reserved for validation and hyperparameter optimization (hold-out set). This process was repeated five times, holding out a different partition each time. The accuracy of the resulting five trained networks was measured on the sixth partition, the test set, which was never used for training. The mean k-fold accuracy and the accuracy when trained on all five training data partitions are summarized in Table 1. The network performs well on both LCA and RCA input images (Fig. 5) and does not segment the catheter or other imaging artifacts despite uneven brightness, overlapping structures, and varying contrast.

The Dice score distributions on the test set for AngioNet, Deeplabv3+, APN + U-Net, U-Net, APN + Attention U-Net, and Attention U-Net are shown in Fig. 6a. All networks were trained using the UM Dataset. AngioNet has the highest mean Dice score on the test set (0.864) when trained on all five partitions of the training data,

	Dice score	Sensitivity	Specificity	AUC	Pixel accuracy
k-fold test	0.856 ± 0.004	0.913 ± 0.013	0.987 ± 0.001	0.991 ± 0.002	0.982 ± 0.004
k-fold holdout	0.857 ± 0.012	0.909 ± 0.012	0.987 ± 0.001	0.990 ± 0.002	0.980 ± 0.003
All data	0.864	0.918	0.987	0.991	0.983

**Table 1.** Accuracy of AngioNet using K-fold cross validation.

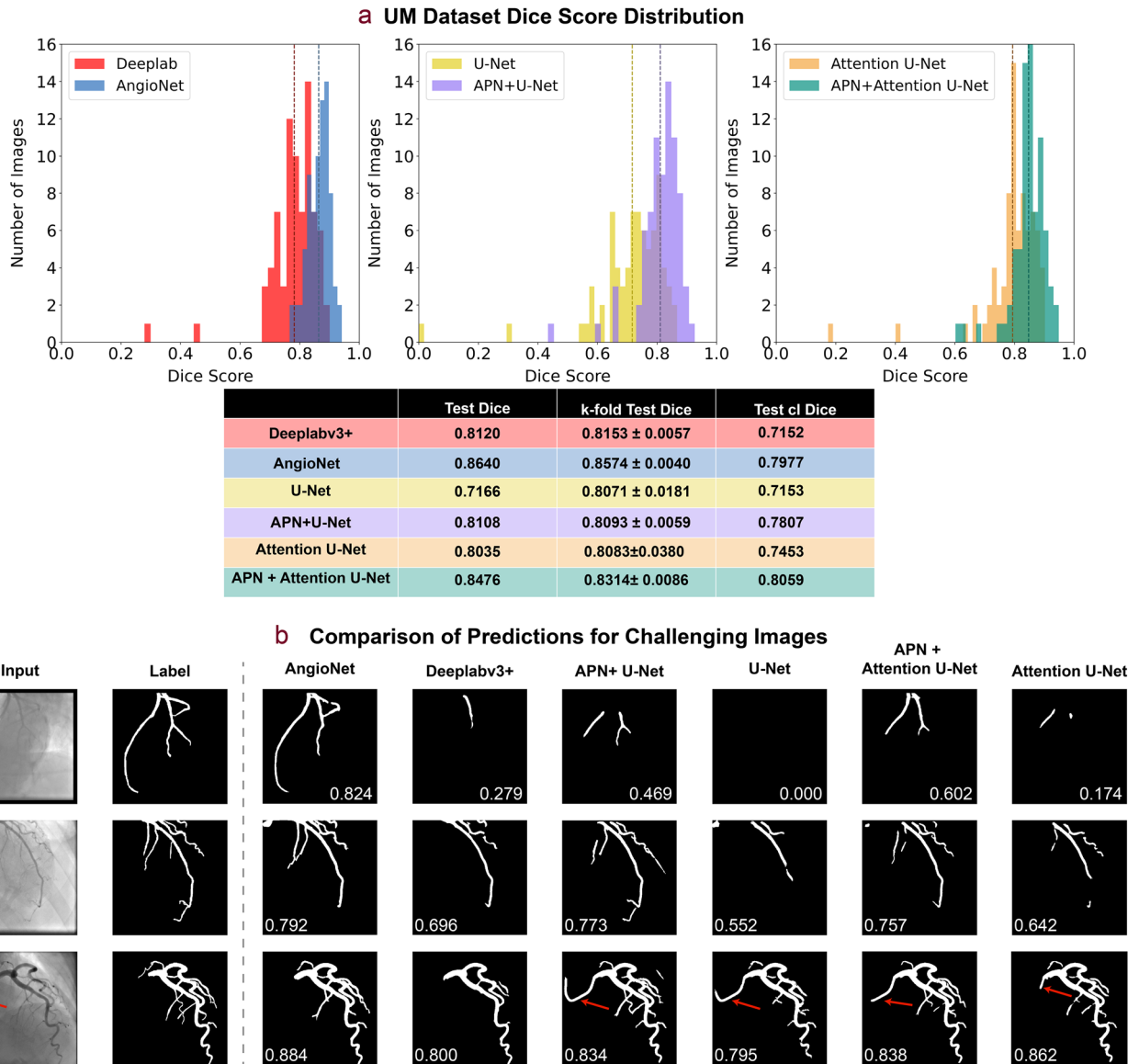


**Figure 5.** Examples of AngioNet segmentation on left coronary tree, taken at two different angles (1,2), and right coronary tree (3). AngioNet does not segment the catheter (red arrows), despite its similar diameter and pixel intensity as the vessels (2,3). It also ignores bony structures such as the spine in (3) and ribs in (1).

compared to 0.815 for Deeplabv3+ alone, 0.811 for APN + U-Net, 0.717 for U-Net alone, 0.848 for APN + Attention U-Net, and 0.804 for Attention U-Net alone. On average, AngioNet has a 10% higher Dice score per image than Deeplabv3+ alone. APN + U-Net has a 14% higher Dice score than U-Net alone, while APN + Attention U-Net has a 10% higher dice score than Attention U-Net alone. A one-tailed paired Student's t-test ( $n=77$ ) was performed to determine if the addition of the APN significantly improved the Dice score for all network backbones. The p-value between AngioNet and Deeplabv3+ was  $5.76e-10$ , the p-value between APN + U-Net and U-Net was  $2.63e-16$ , and the final p-value was  $3.61e-11$  between APN + Attention U-Net and Attention U-Net. All p-values were much less than the statistical significance threshold of 0.05, therefore we can conclude that there are statistically significant differences between the Dice score distributions with and without the APN. Furthermore, all three network backbones exhibit outliers with Dice score lower than 0.5, but adding the APN eliminates these outliers in all networks.

Although the Deeplabv3+ backbone had the highest Dice scores, the Attention U-Net backbone had the highest cIDice scores. APN + Attention U-Net and Attention U-Net had cIDice scores of 0.806 and 0.745 respectively, while AngioNet and Deeplabv3+ had cIDice scores of 0.798 and 0.715. U-Net had the lowest cIDice score of 0.715, and APN + U-Net had a cIDice score of 0.781. In all three cases, the addition of the APN improved the cIDice score by at least 5 points.

Compared to the other networks, AngioNet performs the best on the challenging cases shown in Fig. 6b. The first row of Fig. 6b shows segmentation performance on a low contrast angiography image. AngioNet can segment the major coronary vessels in this image, whereas Deeplabv3+ and Attention U-Net only segment one vessel and U-Net is unable to identify any vessels at all. APN + U-Net and APN + Attention U-Net can segment more vessels than either backbone alone, indicating once again that the addition of the APN improves segmentation performance on these low contrast images. In the second row, the networks including the APN can segment fainter and smaller diameter vessels than the backbone networks. Finally, we observe that AngioNet and Deeplabv3+ did not segment the catheter in the third column, although it is of similar diameter and curvature as the coronary vessels. Conversely, the U-Net and Attention U-Net based networks included part of the catheter in their segmentations. Overall, AngioNet segmented the catheter in 2.6% of the images, where the catheter curved across the image and overlapped with the vessel. In contrast, Deeplabv3+ segmented the catheter in 6.4% of images. Both networks performed better than U-Net and APN + U-Net, which segmented the catheter in 19.5% and 9.1% of images respectively. The Deeplabv3+ backbone networks also outperformed the Attention U-Net



**Figure 6.** Summary of APN performance. All results are derived from the networks trained on all five partitions of the UM training set, unless otherwise noted as a k-fold result. **(a)** Comparison of Dice score distribution on test set. AngioNet has the highest average Dice score, with scores ranging from 0.737 to 0.946. Adding the APN improves the Dice scores of all backbone networks (Deeplabv3+, U-Net, and Attention U-Net). Dashed lines correspond to the Test Dice in the table below. cIDice is also summarized in the table, with the Attention U-Net backbone demonstrating highest topology preservation. **(b)** Segmentation comparison on challenging images with low contrast, faint vessels, and a curved catheter. AngioNet can segment more vessels in these images without segmenting the catheter (red arrows).

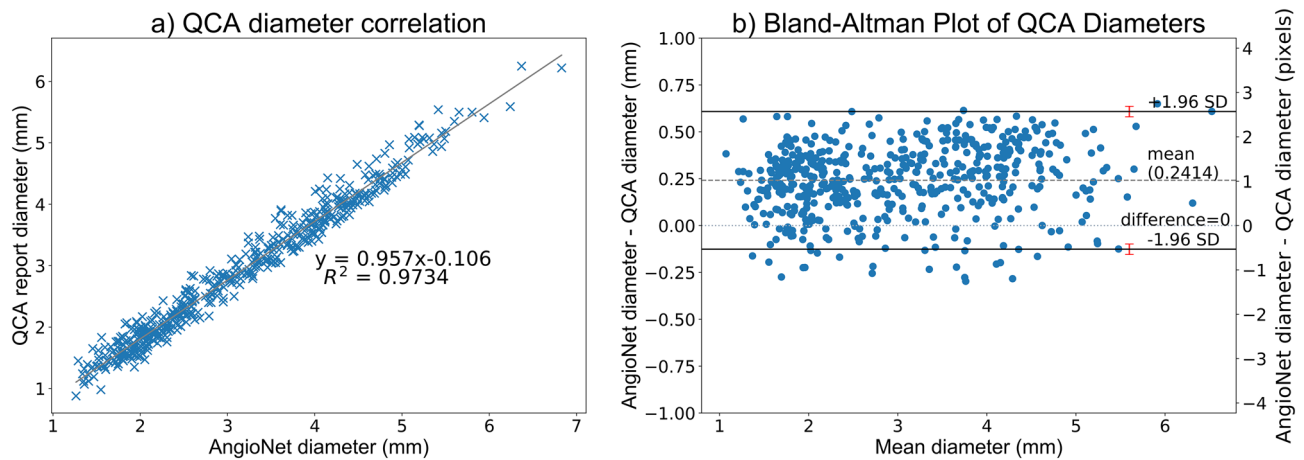
networks, as Attention U-Net segmented the catheter in 11.7% of images and APN + Attention U-Net included in the catheter in 9.1% of images.

**Evaluation of vessel diameter accuracy versus QCA.** Evaluation of vessel diameter accuracy was done using the MMM QCA dataset. Maximum and minimum vessel diameter were compared in 255 vessels including both the RCA and LCA. On average, the absolute error in vessel diameter between the AngioNet segmentation and QCA report was 0.272 mm or 1.15 pixels.

The linear regression plot in Fig. 7A shows that vessel diameter estimates of both methods ( $n = 255$ ) are linearly proportional and tightly clustered around the line of best fit,  $y = 0.957x - 0.106$ , Pearson's correlation coefficient,  $r = 0.9866$ . The standardized difference<sup>52</sup>, also known as Cohen's effect size<sup>53</sup>, was used to determine the difference in means between the diameter distributions of AngioNet and QCA. The standardized difference can determine significant differences between two groups in clinical studies<sup>52</sup>. The standardized difference between the AngioNet and QCA diameter distributions is 0.215, suggesting small differences between the two method.

Figure 7B is a Bland–Altman plot demonstrating the interchangeability of the QCA and AngioNet-derived diameters. The mean difference between both measures,  $\bar{d}$ , is 0.2414. The magnitude of the diameter difference





**Figure 7.** (A) Correlation plot of QCA and AngioNet derived vessel diameters. (B) The Bland–Altman plot demonstrates that AngioNet’s segmentation and QCA are interchangeable methods to determine vessel diameter since more than 95% of points lie within the limits of agreement. The red error bars represent the 95% confidence interval containing the limits of agreement. The mean difference in diameter between methods is 0.24 mm or 1.1 pixels.

remains relatively constant for all mean diameter values, indicating that there is a systematic error and not a proportional error between the two measurements. The limits of agreement are defined as  $\bar{d} \pm 1.96 \text{ SD}$ , where SD is the standard deviation of the diameter differences. For both measurements to be considered interchangeable, 95% of data points must lie between these limits of agreement. In this plot, 96% of the 255 data points are within  $\bar{d} \pm 1.96 \text{ SD}$ . When including the 95% confidence interval of the limits of agreement as recommended by Bland and Altman<sup>54</sup>, 97% of data points are within the range.

## Discussion

In the following sections, we will demonstrate that AngioNet is comparable to state-of-the-art methods for angiographic segmentation. The key findings and clinical implications of our study are also described below.

**Learned filters using the angiographic processing network.** As seen in Fig. 4, the APN learns many different preprocessing filters that improve segmentation performance based on the data partition used for training. Despite the variation in the learned filters, all learned filters exhibit both boundary sharpening and local contrast enhancement. Moreover, the overall segmentation accuracy of the various learned filters remains relatively constant as indicated by the small standard deviation from k-fold cross validation (0.004 for AngioNet, 0.006 for APN + U-Net, and 0.009 for APN + Attention U-Net, see table in Fig. 6a). This demonstrates that the different combinations of learned weights all achieved similar local minima of the loss function, leading to similar Dice scores. Furthermore, the addition of the APN to U-Net and Attention U-Net led to a lower k-fold Dice score standard deviation (0.006 for APN + U-Net compared to 0.018 for U-Net, and 0.009 for APN + Attention U-Net compared to 0.038 for Attention U-Net). This could indicate that the networks incorporating the APN are more robust than the backbone networks alone.

To better understand the effect of using the APN for image preprocessing compared to selecting a particular pre-processing filter, we performed a follow-up experiment where Deeplabv3+ was trained on images pre-processed with the series of unsharp mask filters used to initialize the APN. These unsharp mask filters yielded the highest accuracy of all pre-processing filters we tested, including CLAHE for improved contrast, SVD-based background subtraction, and Gaussian blur for denoising. The average Dice score on the test set for unsharp Deeplabv3+ was 0.833, while the average Dice score for AngioNet was 0.864. A one-tailed paired Student’s t-test ( $n = 77$ ) was used to determine if there were any significant differences between the two Dice score distributions. The p-value of this test was  $1.41 \times 10^{-11}$ , which is much less than the threshold of 0.05. We therefore conclude that the APN’s learned filter has a significant impact on overall segmentation accuracy compared to standard pre-processing filters. This suggests that the learned preprocessing filter implemented in this work is superior to manually selecting a particular contrast enhancement or boundary sharpening filter for preprocessing.

Our pre-training strategy for the APN and backbone networks allows the backbone network to gradually adjust its learned weights to better suit the filter learned by APN. Had the APN been randomly initialized, the filtered image may not be of high quality and could cause the quality of the backbone weights to degrade. Similarly, if only the APN was pre-trained and the backbone was randomly initialized, the much larger learning rate required to train the backbone network would cause the APN weights to degrade. Finally, initializing both the APN and backbone networks randomly can lead to the network getting trapped in local minima since the input to the backbone network may be of poor quality.

Another advantage of the pre-training and fine-tuning strategy is that it allows the user to design for the type of filter that is learned by the APN. By pre-training the APN to mimic a series of unsharp mask filters and applying a small learning rate during fine-tuning, we are able to prime the network to learn a filter that performs

similar functions of boundary sharpening and local contrast enhancement. Indeed, despite the large variation in learned filters as seen in Fig. 4, all the learned filters perform some variation of boundary sharpening and contrast enhancement at the edges of the vessel. This concept could be expanded to take advantage of all three channels of the filtered image that is input into the backbone network. Rather than training the APN to learn a single filter and concatenating it to form a three-channel image, we can pre-train the APN to mimic three different filters for various purposes such as edge detection, contrast enhancement (CLAHE), vesselness enhancement (Frangi filter), and texture analysis (Gabor filter). Learning several new filters for these purposes through the APN may provide the backbone network with richer information with which to segment the vessels.

**Comparison of AngioNet to current state-of-the-art semantic segmentation neural networks.** Several aspects of AngioNet's design contribute to its enhanced segmentation performance compared to existing state of the art networks. The APN successfully improves segmentation performance on low contrast images compared to previous state-of-the-art semantic segmentation networks (Fig. 6a). The APN also enhances performance on smaller vessels, which have lower contrast than larger vessels because they contain less radio-opaque dye. Without the APN, Deeplabv3+, U-Net, and Attention U-Net are not equipped to identify these faint vessels and underpredict the presence of small coronary branches. As seen in Fig. 6a,b, both Deeplabv3+ and AngioNet perform better than U-Net on angiographic segmentation. The addition of the APN to U-Net significantly increases the mean Dice score, facilitates segmentation of more vessels compared to U-Net alone, and greatly reduces the proportion of segmentation that include the catheter; yet APN + U-Net has some of the same drawbacks of U-Net such as disconnected vessels and more instances of the catheter being segmented compared to Deeplabv3+ and AngioNet. Although the Attention U-Net backbone and APN + Attention U-Net outperform U-Net and APN + U-Net, both Attention U-Net networks are still more susceptible to including the catheter (in 19.5% and 9.5% of test images) than Deeplabv3+ (6.4%) and AngioNet (2.6%).

The main benefit of the attention gates in the Attention U-Net backbone seems to be in suppressing background artifacts and preserving the connectivity of the vessels, which can be explained by the global nature of attention-based feature maps. This is demonstrated by Attention U-Net and APN + Attention U-Net having the highest cDice scores compared to the other network backbones. Deeplabv3+ and AngioNet still demonstrate higher mean Dice scores and lower standard deviations than the attention networks; the discrepancy between Dice and cDice scores could be because Deeplabv3+ and AngioNet are able to segment more vessels than the attention-based networks, but the branches are not necessarily connected. While vessel connectivity is an important feature of coronary segmentation, the drawback of catheter segmentation outweighs this benefit of the attention-based networks. It is difficult to avoid segmenting the catheter using filter-based methods<sup>13–22</sup>, thus the main advantage of deep learning methods is their ability to avoid the catheter. AngioNet excels at this task compared to the other networks. Although U-Net and Attention U-Net have demonstrated great success in other binary segmentation applications<sup>35–37</sup>, the presence of catheters and bony structures with similar dimensions and pixel intensity as the vessels of interest make this a particularly challenging segmentation task. Deeplabv3+ and AngioNet have a deeper, more complex architecture, which allows these networks to learn more features with which to identify the vessels in each image<sup>41–43</sup>.

The effective receptive field size U-Net and Attention U-Net is  $64 \times 64$  pixels whereas that of Deeplabv3+ is  $128 \times 128$  pixels<sup>55</sup>. A larger receptive field is associated with better pixel localization and segmentation accuracy, as well as classification of larger scale objects in an image<sup>56,57</sup>. Deeplabv3+'s larger receptive field may explain why Deeplabv3+ and AngioNet are more successful in avoiding segmentation of the catheter, an object typically larger than U-Net or Attention U-Net's  $64 \times 64$  pixel receptive field. The larger receptive field may also explain why Deeplabv3+ and AngioNet are better able to preserve the continuity of the coronary vessel tree and produce fewer broken or disconnected vessels than U-Net and APN + U-Net. The attention gates in Attention U-Net perform a similar function in terms of preserving connectivity, however, the attention mechanism is not able to suppress the features that define catheter. Thus, Deeplabv3+ was an appropriate choice of network backbone for AngioNet.

AngioNet's strengths compared to previous networks include ignoring overlapping structures when segmenting the coronary vessels, smaller sensitivity to noise, and the ability to segment low contrast images. The ability to avoid overlapping bony structures or the catheter is especially important as this eliminates the need for manual correction of the vessel boundary, which is a major advantage over mechanistic segmentation approaches.

AngioNet's greatest limitation is that it overpredicts the vessel boundary in cases of severe (> 85%) stenosis. The network performs well on mild and moderate stenoses, but it has learned to smooth the vessel boundary when the diameter sharply decreases to a single pixel. This is likely due to the low number of training examples containing severe stenosis: only 14 out of the 462 images in the entire UM Dataset contained severe stenosis, and two of these were in the test set. This drawback can be addressed by increasing the training data to encompass more examples of severe stenosis.

**Evaluation of vessel diameter accuracy.** A significant clinical implication of our findings was the comparison between AngioNet and QCA. In Fig. 7A, we observe that QCA and AngioNet results are clustered around the line of best fit,  $y = 0.957x - 0.106$ . Given that the slope of the line of best fit is nearly 1, the intercept is close to 0, and the Pearson's coefficient  $r$  is 0.9866, the line of best fit indicates strong agreement between these two methods of determining vessel diameter. The  $R^2$  coefficient for the linear regression model implies that 97.34% of the variance in the data can be explained by the line of best fit.

The standardized difference, or effect size, is a measure of how many pooled standard deviations separate the means of two distributions<sup>52</sup>. According to Cohen, an effect size of 0.2 is considered a small difference between both groups, 0.5 is a medium difference, and 0.8 is a large difference<sup>53</sup>. Given that the effect size between the QCA and AngioNet diameter distributions was 0.215 (91.5% overlap between the two distributions), we can conclude

that the difference between QCA and AngioNet diameters are small. Furthermore, since the standardized difference indicated no large difference between QCA and AngioNet diameter estimations, these results suggest that both methods can be used interchangeably from a clinical perspective for the dataset examined.

The Bland–Altman plot in Fig. 7B shows that the mean difference between QCA and AngioNet diameters is approximately 1.1 pixels. AngioNet under-predicts the vessel boundary by no more than 1 pixel and over-predicts by no more than 2.5 pixels. To put these values in context, the inter-operator variability for annotating the vessel boundary is  $0.18 \pm 0.24$  mm or slightly above 1 pixel according to a study by Hernandez-Vela et al.<sup>58</sup>. The 95% confidence intervals of the limits of agreement were taken into consideration when determining how many data points lie between the limits of agreement as recommended by Bland and Altman<sup>54</sup>. 97% of the data points lie within the range, which is greater than the 95% threshold. Given these results, and that the standardized difference test which produced no significant difference between the methods, one can conclude that QCA and AngioNet are interchangeable methods to determine vessel diameter. Given AngioNet's fully automated nature, the workload required for generating QCA due to human input could be substantially reduced. Although our direct comparison of AngioNet-derived diameters with QCA-derived diameters required user interaction, future work will focus on developing an automated algorithm for stenosis detection and measurement based on the outputs of AngioNet's segmentation.

## Conclusions

In conclusion, AngioNet was designed to address the shortcomings of current state-of-the-art neural networks for X-ray angiographic segmentation. The APN was found to be a critical component to improve detection and segmentation of the coronary vessels, leading to 14%, 10%, and 10% improved Dice score compared to U-Net, Attention U-Net, or Deeplabv3+ alone. AngioNet demonstrated better Dice scores than all other networks, particularly on images with poor contrast or many small vessels. Although APN + Attention U-Net demonstrated the highest cIDice score (0.806), the connectivity of AngioNet was comparable (0.798) and AngioNet was better able to avoid segmenting the catheter and other imaging artifacts. Furthermore, our statistical analysis of the vessel diameters determined by AngioNet and traditional QCA demonstrated that the two methods may be interchangeable which could have large implications for clinical workflows. Future work to improve performance will focus on increasing accuracy on severe stenosis cases and automating stenosis measurement. We also aim to increase the versatility of the APN by pre-training and learning several filters for edge detection, texture analysis, or vesselness enhancement in addition to our current contrast-enhancing implementation. Combining these learned filters into a multi-channel image may improve the semantic segmentation performance of AngioNet.

## Data availability

Since the datasets used in this work contain patient data, these cannot be made generally available to the public due to privacy concerns.

## Code availability

The code for the AngioNet architecture and examples of synthetic angiograms are available at <https://github.com/kritiyer/AngioNet>. This code is licensed under a Polyform Noncommercial license.

Received: 7 April 2021; Accepted: 11 August 2021

Published online: 10 September 2021

## References

- Sanchis-Gomar, F., Perez-Quilis, C., Leischik, R. & Lucia, A. Epidemiology of coronary heart disease and acute coronary syndrome. *Ann. Transl. Med.* **4**, 256 (2016).
- Townsend, N. et al. Cardiovascular disease in Europe: Epidemiological update 2016. *Eur. Heart J.* **37**, 3232–3245 (2016).
- Go, A. S. et al. Heart disease and stroke statistics—2013 update: A report from the American Heart Association. *Circulation* **127**, e6–e245 (2013).
- Russell, M. W., Huse, D. M., Drowns, S., Hamel, E. C. & Hartz, S. C. Direct medical costs of coronary artery disease in the United States. *Am. J. Cardiol.* **81**, 1110–1115 (1998).
- Nichols, W. W., O'Rourke, M. F., Vlachopoulos, C. & McDonald, D. A. *McDonald's Blood Flow in Arteries: Theoretical, Experimental and Clinical Principles* (Hodder Arnold, 2011).
- Feigl, E. O. Coronary physiology. *Physiol. Rev.* **63**, 1–205 (1983).
- Nieman, K. et al. Usefulness of multislice computed tomography for detecting obstructive coronary artery disease. *Am. J. Cardiol.* **89**, 913–918 (2002).
- Morice, M.-C. et al. A randomized comparison of a sirolimus-eluting stent with a standard stent for coronary revascularization. *N. Engl. J. Med.* **346**, 1773–1780 (2002).
- Stadius, M. L. & Alderman, E. L. Editorial: Coronary artery revascularization critical need for, and consequences of, objective angiographic assessment of lesion severity. *Circulation* **82**, 2231–2234 (1982).
- Klein, A. K., Lee, F. & Amini, A. A. Quantitative coronary angiography with deformable spline models. *IEEE Trans. Med. Imaging* **16**, 468–482 (1997).
- Reiber, J. H. C. An overview of coronary quantitation techniques as of 1989. In *Quantitative Coronary Arteriography* (eds Reiber, J. H. C. & Serruys, P. W.) 55–132 (Kluwer Academic Publishers, 1997).
- Mancini, G. B. et al. Automated quantitative coronary arteriography: Morphologic and physiologic validation in vivo of a rapid digital angiographic method. *Circulation* **75**, 452–460 (1987).
- Lin, C. Y. & Ching, Y. T. Extraction of coronary arterial tree using cine X-ray angiograms. *Biomed. Eng. Appl. Basis Commun.* **17**, 111–120 (2005).
- Herrington, D. M., Siebes, M., Sokol, D. K., Siu, C. O. & Walford, G. D. Variability in measures of coronary lumen dimensions using quantitative coronary angiography. *J. Am. Coll. Cardiol.* **22**, 1068–1074 (1993).
- Canero, C. & Radeva, P. Vesselness enhancement diffusion. *Pattern Recogn. Lett.* **24**, 3141–3151 (2003).

16. Frangi, A. F., Niessen, W. J., Vincken, K. L. & Viergever, M. A. Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI'98* (eds Wells, W. M. *et al.*) (Springer, 1998).
17. Shang, Y. *et al.* Vascular active contour for vessel tree segmentation. *IEEE Trans. Biomed. Eng.* **58**, 1023–1032 (2011).
18. Xia, S. *et al.* Vessel segmentation of X-ray coronary angiographic image sequence. *IEEE Trans. Biomed. Eng.* **67**, 1338–1348 (2020).
19. Pappas, T. & Lim, J. S. A new method for estimation of coronary artery dimensions in angiograms. *IEEE Trans. Acoust. Speech Signal Process.* **36**, 1501–1513 (1988).
20. M'Hiri, F., Duong, L., Desrosiers, C. & Chieriet, M. Vesselwalker: coronary arteries segmentation using random walks and Hessian-based vesselness filter. In *Proceedings—International Symposium on Biomedical Imaging* 918–921 (2013) <https://doi.org/10.1109/ISBI.2013.6556625>.
21. Lara, D. S. D., Faria, A. W. C., Araújo, A. D. A. & Menotti, D. A novel hybrid method for the segmentation of the coronary artery tree in 2D angiograms. *Int. J. Comput. Sci. Inf. Technol.* **5**, 45 (2013).
22. Chen, S. J. & Carroll, J. D. 3-D reconstruction of coronary arterial tree to optimize angiographic visualization. *IEEE Trans. Med. Imaging* **19**, 318–336 (2000).
23. Rawat, W. & Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* **29**, 2352–2449 (2017).
24. Badrinarayanan, V., Kendall, A. & Cipolla, R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 2481–2495 (2017).
25. Khanmohammadi, M., Egan, K., Sæland, C., Eftestøl, T. & Larsen, A. I. Automatic estimation of coronary blood flow velocity step 1 for developing a tool to diagnose patients with micro-vascular angina pectoris. *Front. Cardiovasc. Med.* **6**, 1 (2019).
26. Pohlen, T., Hermans, A., Mathias, M. & Leibe, B. Full-resolution residual networks for semantic segmentation in street scenes. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017).
27. Zhao, H., Shi, J., Qi, X., Wang, X. & Jia, J. Pyramid scene parsing network. In *2017 IEEE Conf. Comput. Vis. Pattern Recognit. CVPR* 6230–6239 (2017) <https://doi.org/10.1109/CVPR.2017.660>.
28. Paszke, A., Chaurasia, A., Kim, S. & Culurciello, E. ENet: A deep neural network architecture for real-time semantic segmentation. Preprint at <https://arxiv.org/abs/1606.02147> (2016).
29. Yang, S. *et al.* Deep learning segmentation of major vessels in X-ray coronary angiography. *Sci. Rep.* **9**, 1–11 (2019).
30. Fan, J. *et al.* Multichannel fully convolutional network for coronary artery segmentation in X-ray angiograms. *IEEE Access* **6**, 44635–44643 (2018).
31. Nasr-Esfahani, E. *et al.* Segmentation of vessels in angiograms using convolutional neural networks. *Biomed. Signal Process. Control* **40**, 240–251 (2018).
32. Shin, S. Y., Lee, S., Yun, I. D. & Lee, K. M. Deep vessel segmentation by learning graphical connectivity. Preprint at <https://arxiv.org/abs/1806.02279> (2018).
33. Cervantes-Sanchez, F., Cruz-Aceves, I., Hernandez-Aguirre, A., Hernandez-Gonzalez, M. A. & Solorio-Meza, S. E. Automatic segmentation of coronary arteries in X-ray angiograms using multiscale analysis and artificial neural networks. *Appl. Sci.* **9**, 5507 (2019).
34. Zhu, X., Cheng, Z., Wang, S., Chen, X. & Lu, G. Coronary angiography image segmentation based on PSPNet. *Comput. Methods Programs Biomed.* <https://doi.org/10.1016/j.cmpb.2020.105897> (2020).
35. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *MICCAI 2015: Medical Image Computing and Computer-Assisted Intervention* 234–241 (2015).
36. Oktay, O. *et al.* Attention U-Net: learning where to look for the pancreas. Preprint at <https://arxiv.org/abs/1804.03999> (2018).
37. Akeret, J., Chang, C., Lucchi, A. & Refregier, A. Radio frequency interference mitigation using deep convolutional neural networks. *Astron. Comput.* **18**, 35–39 (2017).
38. Igloukov, V. & Shvets, A. TeraNet: U-Net with VGG11 encoder pre-trained on imageNet for image segmentation. Preprint at <https://arxiv.org/abs/1801.05746> (2018).
39. Chen, L.-C., Papandreou, G., Schroff, F. & Adam, H. Rethinking atrous convolution for semantic image segmentation. Preprint at <https://arxiv.org/abs/1706.05587> (2017).
40. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F. & Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *ECCV 2018* (2018).
41. Raghu, M., Poole, B., Kleinberg, J., Ganguli, S. & Dickstein, J. S. On the expressive power of deep neural networks. In *34th Int. Conf. Mach. Learn. ICML 2017*, Vol. 6, 4351–4374 (2017).
42. Lu, Z., Pu, H., Wang, F., Hu, Z. & Wang, L. The expressive power of neural networks: a view from the width. In *31st Conference on Neural Information Processing Systems* (2017).
43. Eldan, R. & Shamir, O. The power of depth for feedforward neural networks. *J. Mach. Learn. Res.* **49**, 907–940 (2016).
44. Fu, J., Liu, J., Tian, H. & Li, Y. Dual Attention Network for Scene Segmentation. Preprint at <https://arxiv.org/abs/1809.02983> (2018).
45. Telea, A. & Van Wijk, J. J. An augmented fast marching method for computing skeletons and centerlines. In *EUROGRAPHICS—IEEE TCVG Symposium on Visualization* (eds Ebert, D. *et al.*) (2002).
46. Kingman, J. F. C. & Kullback, S. Information theory and statistics. *Math. Gaz.* **54**, 90 (1970).
47. Abadi, M. *et al.* *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. (2015).
48. Chollet, F. & others. Keras. <https://keras.io> (2015).
49. Zakirov, E. keras-deeplab-v3-plus. <https://github.com/bonlime/keras-deeplab-v3-plus> (2019).
50. Allen-Zhu, Z. & Hazan, E. Variance reduction for faster non-convex optimization. In *Convex Optimization*, Vol. 48, 9 (JMLR: W&CP, 2016).
51. Paetzold, J. C. *et al.* cDice—A novel connectivity-preserving loss function for vessel segmentation. Preprint at <https://arxiv.org/abs/2003.07311v6> (2020).
52. Austin, P. C. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat. Med.* **28**, 3083–3107 (2009).
53. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences* (L. Erlbaum Associates, 1988).
54. Bland, J. M. & Altman, D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **327**, 307–310 (1986).
55. fornaxai. fornaxai/receptivefield: Gradient based receptive field estimation for Convolutional Neural Networks. <https://github.com/fornaxai/receptivefield> (2018).
56. Liu, Y., Yu, J. & Han, Y. Understanding the effective receptive field in semantic image segmentation. *Multimed. Tools Appl.* **77**, 22159–22171 (2018).
57. Wang, P. *et al.* Understanding convolution for semantic segmentation. In *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018* 1451–1460 (Institute of Electrical and Electronics Engineers Inc., 2018).
58. Hernandez-Vela, A. *et al.* Accurate coronary centerline extraction, caliber estimation, and catheter detection in angiographies. *IEEE Trans. Inf. Technol. Biomed.* **16**, 1332–1340 (2012).



## Acknowledgements

The authors would like to thank the American Heart Association Precision Medicine Platform (<https://precision.heart.org/>), which was used for data analysis and neural network training on a cluster of Tesla K80 GPUs. Neural network prototyping and initial code development were carried out on a Titan XP GPU granted to our lab by the NVIDIA GPU grant before deploying the model onto the Precision Medicine Platform. Funding for this work was provided by a grant from the American Heart Association [19A1ML34910010]. KI was funded by the National Science Foundation Graduate Research Fellowship Program [DGE1841052] and Rackham Merit Fellowship, and CAF was supported by the Edward B. Diethrich Professorship. CJA was funded by a King's Prize Research Fellowship via the Wellcome Trust Institutional Strategic Support Fund grant to King's College London [204823/Z/16/Z]. R.R. Nadakuditi was funded by ARO grant W911NF-15-1-0479.

## Author contributions

K.I. and C.P.N. performed neural network design and training with advice from RNR. K.I., C.P.N., and A.A.F. created the labels for the training datasets. C.J.A. and S.M.R.S. provided advice on neural network training and image processing techniques. M.A.S. and V.S. provided the MMM QCA dataset. C.A.F. and B.K.N. are the co-investigators of this work and provided guidance on network training and clinical validation metrics. All authors reviewed the written manuscript.

## Competing interests

CAF and BKN are founders of AngioInsight, Inc. AngioInsight is a startup company which uses machine learning and signal processing to assist physicians with the interpretation of angiograms. AngioInsight did not sponsor this work. This work is a part of a patent which has been filed by the University of Michigan. All other authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to C.A.F.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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