

Article

A Preprocessing Method for Coronary Artery Stenosis Detection Based on Deep Learning

Yanjun Li ¹, Takaaki Yoshimura ^{2,3,4,5}, Yuto Horima ⁶ and Hiroyuki Sugimori ^{4,5,7,*}

- ¹ Graduate School of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan; yanjun.li.g9@elms.hokudai.ac.jp
- ² Department of Health Sciences and Technology, Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan; takaaki.ysm@med.hokudai.ac.jp
- ³ Department of Medical Physics, Hokkaido University Hospital, Sapporo 060-8648, Japan
- ⁴ Global Center for Biomedical Science and Engineering, Faculty of Medicine, Hokkaido University, Sapporo 060-8648, Japan
- ⁵ Clinical AI Human Resources Development Program, Faculty of Medicine, Hokkaido University, Sapporo 060-8648, Japan
- ⁶ Department of Central Radiology, JR Sapporo Hospital, Sapporo 060-0033, Japan; marimokoppe09@gmail.com
- ⁷ Department of Biomedical Science and Engineering, Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan
- * Correspondence: sugimori@hs.hokudai.ac.jp; Tel.: +81-11-706-3410

Abstract: The detection of coronary artery stenosis is one of the most important indicators for the diagnosis of coronary artery disease. However, stenosis in branch vessels is often difficult to detect using computer-aided systems and even radiologists because of several factors, such as imaging angle and contrast agent inhomogeneity. Traditional coronary artery stenosis localization algorithms often only detect aortic stenosis and ignore branch vessels that may also cause major health threats. Therefore, improving the localization of branch vessel stenosis in coronary angiographic images is a potential development property. In this study, we propose a preprocessing approach that combines vessel enhancement and image fusion as a prerequisite for deep learning. The sensitivity of the neural network to stenosis features is improved by enhancing the blurry features in coronary angiographic images. By validating five neural networks, such as YOLOv4 and R-FCN-Inceptionresnetv2, our proposed method can improve the performance of deep learning network applications on the images from six common imaging angles. The results showed that the proposed method is suitable as a preprocessing method for coronary angiographic image processing based on deep learning and can be used to amend the recognition ability of the deep model for fine vessel stenosis.

Keywords: coronary angiography; deep learning; image enhancement; image fusion



Citation: Li, Y.; Yoshimura, T.; Horima, Y.; Sugimori, H. A Preprocessing Method for Coronary Artery Stenosis Detection Based on Deep Learning. *Algorithms* **2024**, *17*, 119. <https://doi.org/10.3390/a17030119>

Academic Editors: Laura Antonelli and Lucia Maddalena

Received: 9 February 2024

Revised: 5 March 2024

Accepted: 11 March 2024

Published: 13 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Coronary heart disease, in which lifestyle and environmental factors are risk factors, is one of the leading cardiovascular diseases affecting the global population [1]. Coronary angiography has been used clinically for decades as the “gold standard” for the diagnosis of coronary heart disease [2]. Invasive coronary angiography (ICA) is an important method for arterial stenosis detection and is widely used in clinical treatment. Typically, in ICAs, a radiopaque contrast agent is injected into the coronary arteries, and then the arteries are viewed with the use of consecutive X-rays. In clinical practice, the diagnosis decision usually considers the percent stenosis as one of the judgment factors of severity, which is determined by a physician’s visual assessment. For example, 50% stenosis is generally considered the minimum criterion for coronary revascularization, and 70% stenosis is considered an indicator for identifying clinical lesions. However, this method has some limitations, such as reliance on professionally trained radiologists and judgment errors due to fatigue and other reasons. On the other hand, limitations in imaging modalities, such

as very low contrast caused by a limited radiation dose or illumination and imaging blur caused by a limited number of viewing angles and diluted contrast agents, also make the accuracy of the manual determination of coronary stenosis challenging [3].

These unsolved problems have been developed with the outstanding contributions of deep learning in the field of image processing. Recently, the deep learning methods represented by convolutional neural networks (CNNs) have made it possible to locate stenosis and calculate the stenosis rate more accurately, stably, and reliably. Within the existing body of literature, diligent scholars have undertaken extensive research endeavors encompassing various facets of this complex domain. Some have delved into the intricacies of blood vessel extraction [4], while others have harnessed the power of deep learning methods to precisely locate stenotic lesions [5]. Concurrently, certain researchers have proposed novel mathematical approaches to calculate stenosis rates [6]. Although these pioneering studies have significantly contributed to the field, they often exhibit a predominant focus on stenotic areas within the aorta, inadvertently overlooking the critical implications of stenosis within branch vessels. This oversight can be attributed to challenges stemming from overlapping vessels, inadequate contrast agent distribution, and suboptimal imaging angles, which frequently hinder the comprehensive visualization of branch vessels. Consequently, the detection of stenosis within complete coronary angiography presents an intriguing and formidable research frontier, characterized by abundant opportunities and intricate challenges that await exploration. Addressing this critical aspect holds the potential to substantially augment our understanding and capabilities in diagnosing and managing stenotic conditions within the intricate network of coronary arteries.

To address the formidable challenges posed by the intricate network of coronary arteries, characterized by numerous branches and complex geometries, an efficient approach involves leveraging Hessian matrix multiscale filtering for blood vessel enhancement [7]. This method capitalizes on the inherent correlations among the eigenvalues of the Hessian matrix to identify and extract the tubular structural characteristics of blood vessels within medical images. Notably, the Hessian-based Frangi vesselness (HFV) filter has gained widespread prominence for vascular enhancement and quantification, as evidenced by its utilization in various studies [8,9]. In addition to standalone approaches, hybrid enhancement techniques have also emerged, such as the amalgamation of region growth with the HFV filter [10], aimed at further refining image backgrounds and suppressing noise.

In the realm of convolutional neural network-based methods, significant strides have been made in various aspects of stenosis localization and detection. For instance, a fundamental CNN architecture, with slight modifications within its layers, was employed for the detection of stenotic lesions in angiographic images in 2018 [11]. Experimental outcomes revealed that networks trained on artificial data and subsequently fine-tuned with clinical datasets achieved a remarkable accuracy of approximately 90%, surpassing the performance of traditional feedforward neural networks. Transfer learning has also been explored, where pretrained CNNs are fine-tuned using artificial data. It was observed that a pretrained CNN, with only the top three blocks serving as feature extractors, exhibited promising performance in stenosis detection [12].

Within the domain of blood vessel segmentation, researchers have contributed a spectrum of model- and tracking-based approaches in recent years. These include techniques like multiscale Gabor filtering [13], Hessian filtering with flux flow measurement [14], and active contour-based segmentation methods [15]. Furthermore, deep learning-based approaches, such as the utilization of long short-term memory recurrent neural networks [16], have yielded notable performance improvements, contributing significantly to the advancement of this field. These multifaceted methodologies collectively constitute a rich tapestry of techniques aimed at enhancing our understanding and capabilities in the realm of medical image analysis, particularly in the context of coronary artery stenosis detection and blood vessel segmentation.

In this study, we propose a preprocessing method for coronary artery image enhancement that combines Hessian matrix multiscale filtering and image feature fusion. In this

method, we used the image fusion method to retain as many features as possible in the image based on image enhancement. These features can help strengthen the algorithm's ability to learn and detect small or distal vessels. HFV filtering was used to enhance the blood vessels in angiographic images, and an image fusion step was performed on the enhanced images to accentuate the blood vessels. Then, we employed various deep learning networks for model training and compared their performances. In comparisons, we not only evaluated the performance of different deep learning networks under identical conditions but also compared the performance of images with and without preprocessing across these networks. The experimental results showed that the proposed method can obtain satisfactory object detection results.

2. Materials and Methods

2.1. Materials

This research employed a dataset comprising data extracted from a cohort of 20 patients who had undergone coronary angiography procedures at Hokkaido University Hospital. Notably, all the angiographic images and video streams derived from these patients underwent meticulous and precise diagnosis by experienced radiologists, ensuring the accuracy and reliability of the clinical data underpinning our study. To facilitate subsequent image analysis, the continuous video streams of each patient were discretized into individual images, with a consistent frame rate of 15 frames per second maintained throughout this process. For the purpose of our experiment, a total of 250 contrast images were obtained and classified manually by right coronary artery (RCA) and left coronary artery (LCA) (each artery category encompassed three distinct imaging angles, namely the left anterior oblique (LAO), right anterior oblique (RAO), and the ZERO angle).

2.2. HFV Filter

Enhancement filters are crucial for boosting vessel segmentation efficiency in medical imaging. Commonly used techniques focus on differential information, such as the second derivative or the Hessian matrix, to highlight vessel structures in images. These methods utilize differential analysis to enhance the visibility of blood vessels, making them essential for medical image segmentation and analysis.

This method is based on the eigenvalues analysis of the Hessian matrix ($H(x, y)$) to determine the presence of vessel-like structures in the image. Denote the source image as $I(x, y)$, and then the Hessian matrix is made continuous by convolving it with a two-dimensional (2D) Gaussian filter for images. Equation (1) presents the convolution of the Gaussian filter with the Hessian of image $I(x, y)$ using convolution.

$$H(x, y) \approx G * \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 G}{\partial x^2} & \frac{\partial^2 G}{\partial x \partial y} \\ \frac{\partial^2 G}{\partial y \partial x} & \frac{\partial^2 G}{\partial y^2} \end{bmatrix} * I(x, y) \quad (1)$$

The analysis of the sign and magnitude of the Hessian eigenvalues can be used to improve the local image structure. Thus, the eigenvalues and eigenvectors of $H(x, y)$ are calculated to obtain information about the contrast and direction at each point of the image. The eigen decomposition generates 2 eigenvalues, which are (λ_1, λ_2) , and the corresponding 2 eigenvectors (\bar{u}_1, \bar{u}_2) . These parameters are analyzed to distinguish the blob-like, tubular, or plate-like structure in the image. Because we are interested in the blood vessel-like structures in the image, these structures are represented by the relation of $|\lambda_1| \approx 0$, $|\lambda_1| \ll |\lambda_2|$, and \bar{u}_1 points toward the direction of the vessel. In addition, λ_2 is negative if the vessels appear as glowing tubular structures in dimly lit areas. This signature information is used as a consistency check to exclude other structures existing in the image. The Frangi vesselness function is defined as follows:

$$F(v) = \begin{cases} 0 & \text{if } \lambda_2 > 0, \\ e^{-\frac{R_B^2}{2\beta^2}} \left(1 - e^{-\frac{s^2}{2c^2}}\right) & \text{otherwise,} \end{cases} \quad (2)$$

where

$$R_B = \frac{|\lambda_1|}{|\lambda_2|}, \quad S = \sqrt{\lambda_1^2 + \lambda_2^2} \quad (3)$$

The parameter R_B is the measure of eccentricity of the object in a 2D image. S is the measure of structureness and is used to obtain the background area. The contrast in the background is lowest because of the absence of any structure. S is also low for the background and high for the region with structures. β and c are thresholds to control the sensitivity of the line filter with respect to the parameters R_B and S . The value of β can be set at 0.5, whereas the value of c depends on the gray level intensity of the vessels. This filter is applied to the image to obtain the response value for each pixel. The response values are higher for the pixels belonging to the vessels and vice versa.

2.3. Image Fusion

Medical image fusion represents a critical process involving the integration of multiple images derived from either a single imaging modality or multiple distinct modalities within the medical field. The overarching objective of medical image fusion is to enhance the overall image quality while retaining and accentuating specific features pertinent to the clinical context. This enhancement serves to bolster the clinical applicability of these fused images in the realm of medical diagnosis and evaluation. The domain of medical image fusion spans across a diverse spectrum of scientific disciplines, encompassing image processing, computer vision, pattern recognition, machine learning, and artificial intelligence. The versatile applications of medical image fusion extend to various clinical scenarios, where it equips physicians with the ability to gain a comprehensive understanding of pathological lesions through the amalgamation of medical images obtained from disparate modalities. This transformative approach empowers healthcare professionals with a richer and more informative visualization of medical data, ultimately contributing to more accurate diagnoses and informed treatment decisions.

In the context of medical image fusion methodologies based on the intensity-hue-saturation (IHS) model, a fundamental transformation known as the IHS transform plays a pivotal role in the process. Initially, the input medical image, typically represented in the red-green-blue (RGB) color channels, undergoes a conversion known as the RGB-IHS transform. This transformation translates the image into the IHS color space, comprising intensity, hue, and saturation components, effectively creating a matrix representation of the image in this alternative color space (IHS-RGB). Subsequently, the fused medical image is reconstructed by employing the inverse IHS transform (IHS-RGB). In this fusion approach, two medical images derived from different imaging modalities are initially converted into the IHS color space through the RGB-IHS transformation. Following this, the final fused image is synthesized by reversing this process using the IHS-RGB transformation:

$$I = \frac{R + B + G}{3} \quad (4)$$

$$\begin{cases} H = \frac{G-B}{3I-3B}, S = \frac{I-B}{I} & \text{if } B < R, G, \\ H = \frac{B-R}{3I-3R}, S = \frac{I-R}{I} & \text{if } R < B, G, \\ H = \frac{R-G}{3I-3G}, S = \frac{I-G}{I} & \text{if } G < R, B \end{cases} \quad (5)$$

$$\left\{ \begin{array}{l} R = I(1 + 2S - 3S \times H), \\ G = I(1 - S + 3S \times H), \\ B = I(1 - S) \quad \text{if } B < R, G, \\ R = I(1 - S), \\ G = I(1 + 5S - 3S \times H), \\ B = I(1 - 4S + 3S \times H) \quad \text{if } R < B, G, \\ R = I(1 - 7S + 3S \times H), \\ G = I(1 - S), \\ B = I(1 + 8S - 3S \times H) \quad \text{if } G < R, B \end{array} \right. \quad (6)$$

This comprehensive technique harnesses the benefits of the IHS color space, enabling the integration of distinct modalities into a singular representation that preserves essential features for improved clinical interpretation and analysis.

2.4. You Only Look Once v4

YOLOv4 [17] represents the fourth iteration in the You Only Look Once series, emphasizing real-time object detection and enabling training on a single GPU. Its backbone, often pre-trained on the ImageNet dataset for classification, is crucial for feature identification, which gets refined for detecting objects. The model uses CSPDarknet53, inspired by DenseNet, which aims to mitigate issues like vanishing gradients, enhance feature propagation, foster feature reuse, and minimize parameters. For feature integration, YOLOv4 adopts PANet, although the choice is not extensively justified, indicating potential areas for further investigation given the simultaneous development of NAS-FPN and BiFPN. A Spatial Pyramid Pooling (SPP) block is incorporated post-CSPDarknet53 to expand the receptive field and isolate key features. YOLOv4 maintains the YOLOv3 detection mechanism, utilizing anchor-based steps and multi-scale detection, enabling rapid and accurate object identification, well-suited for real-time scenarios, and setting high benchmarks in accuracy within object detection performance metrics. Currently, the YOLO series of detectors has evolved to YOLOv8. However, due to factors such as the stability of YOLOv4, it is still widely used in object detection tasks in medical image processing [18,19]. In our experiment, the YOLOv4 backbone utilizes CSPDarknet53, which is based on DenseNet, with PANet serving as the neck and YOLOv3 employed as the detection head.

2.5. Faster Region-Convolutional Neural Network

Faster R-CNN [20], standing for “Faster Region-Convolutional Neural Network”, is a cutting-edge object detection framework within the R-CNN family. This network aims to create a comprehensive architecture that not only identifies objects in an image but also pinpoints their exact locations. It integrates the strengths of deep learning, convolutional neural networks (CNNs), and region proposal networks (RPNs), enhancing both the speed and accuracy of the detection process.

RPN is a crucial part of the Faster R-CNN, tasked with identifying potential object regions in images. It leverages predefined anchor boxes of varying sizes and aspect ratios distributed across the image feature maps to pinpoint where objects might be. Each anchor box is assessed for its “objectness” score, indicating the likelihood of containing an object versus background, and adjusted for better alignment with potential objects. The RPN uses a sliding window method with a small convolutional network to evaluate these anchors, producing scores and adjustment values. Through Non-Maximum Suppression (NMS), the RPN filters out overlapping boxes, retaining only the most probable ones as region proposals, ready for further analysis by the Fast R-CNN detector.

The Fast R-CNN detector within the Faster R-CNN architecture plays a pivotal role in detecting objects based on region proposals from the RPN. It first applies RoI pooling to standardize the size of these proposals, then extracts features using the CNN backbone. Subsequently, fully connected layers classify the objects and adjust bounding box

coordinates. The detector employs a multi-task loss function to optimize classification and bounding box regression, enhancing the precision of object detection. Finally, post-processing with non-maximum suppression refines the detection outcomes, eliminating redundant detections and preserving the most accurate ones.

2.6. Region-Based Fully Convolutional Networks

The architecture of region-based fully convolutional networks (R-FCN), as detailed in [21], comprises four key components, each playing a crucial role in the network's functionality. These components are the RPN, the residual network (ResNet), the classification module, and the regression module.

The primary responsibility of the RPN is to extract region proposals, which are candidate regions of interest (RoIs) identified within the input image. These proposals enable subsequent convolutional operations to be performed across the entire image, facilitating the computation of weight layers.

Within the R-FCN structure, the process initiates with input images resized to have a consistent short-side dimension. These images are then subjected to feature extraction through ResNet-101, which consists of five convolutional network blocks. Importantly, the output of the fourth convolutional layer of ResNet-101 is employed as the input for the RPN.

However, due to the high dimensionality inherent in the output of the fifth convolutional layer of ResNet-101, it becomes necessary to downscale the number of channels through the incorporation of additional convolutional layers. This step results in the generation of a 1024-dimensional feature map.

Following this feature extraction phase, two parallel convolutional layers are introduced. These layers are designed for classification and regression tasks, respectively. The generation of position-sensitive score maps with dimensions of $k^2(C + 1)$ and $4k^2$ takes place in preparation for these two parallel tasks. These maps provide the necessary positional information for subsequent operations.

In the final stages of the network, the two distinct dimensional position-sensitive score maps are pooled with the RoIs, which have been extracted earlier by the RPN. This pooling process ultimately yields the classification and regression results, encapsulating the network's ability to identify and localize objects of interest within the input image data.

2.7. Mean Average Precision

Mean average precision (mAP) is a crucial metric for evaluating the effectiveness of machine learning models, especially in object detection. It integrates several sub-metrics to offer a thorough assessment:

- **Confusion Matrix:** A tool that breaks down the model's predictions into true positives, true negatives, false positives, and false negatives, helping in the computation of other key metrics like precision and recall.
- **Intersection over Union (IoU):** This metric quantifies the overlap between the predicted and actual bounding boxes, providing insight into the model's localization accuracy.
- **Precision:** Represents the accuracy of the model's positive predictions, calculated as the ratio of true positives to the total of true positives and false positives.
- **Recall:** Measures the model's ability to identify all relevant cases, computed as the ratio of true positives to the total of true positives and false negatives.

mAP averages the precision values across different classes and/or IoU thresholds, offering a single metric that reflects the model's precision and recall performance across its various categories:

$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} AP_k \quad (7)$$

where n represents the number of classes and AP_k represents the average precision of class k .

3. Experiments and Results

3.1. Experiments

3.1.1. Image Selection and Annotation

Given the variability in contrast agent distribution during the initial and final phases of angiography, which impacts vascular visualization, we meticulously selected 250 representative images from the dataset for model development. Experienced radiologists from Hokkaido University Hospital provided diagnoses for these patients, facilitating precise image annotations, particularly focusing on areas with over 50% stenosis due to their higher risk and clearer visibility.

3.1.2. Data Preparation

The selected images were randomly assigned to training (80%) and validation (20%) sets. To mitigate the limited data volume and enhance model robustness, we applied preprocessing and augmentation techniques, notably rotating the images between -45° and 45° at 5° intervals. This process expanded our dataset to 3800 images, enhancing the diversity and generalizability of our training set.

3.1.3. Model Training and Evaluation

We employed various deep learning architectures to train models on this augmented dataset, assessing their performance in detecting and segmenting vascular structures. Notably, all models evaluated in this study were trained from scratch, without the use of pre-trained weights, to ensure the models' adaptability and performance were solely attributed to the training on our specific dataset. Each model was evaluated under uniform conditions to ensure a fair comparison, focusing on their effectiveness with both preprocessed and original images. The models' performances were quantitatively assessed, considering their ability to accurately identify and delineate areas of stenosis, thereby aiding in the comprehensive analysis of coronary angiography images.

3.2. Results

We separately evaluated and compared the effects of preprocessing with an HFV filter and image fusion with the original image on the performance of the deep learning network for stenosis detection.

Following the completion of image preprocessing, we embarked on the crucial task of annotating the positions of stenosis within the main and branch vessels, specifically targeting stenosis rates exceeding 50%, in accordance with the diagnostic findings provided by medical professionals. For visual clarity and reference, we present Figure 1(a1–a3) within our manuscript, each illustrating a distinct aspect of our analysis. Figure 1(a1) showcases the original, unprocessed image, providing a baseline for our observations. In contrast, Figure 1(a2) displays the image subjected to preprocessing using the HFV filter, revealing the enhancements achieved through this technique. Lastly, Figure 1(a3) presents the image after undergoing image fusion, illustrating the outcome of integrating the preprocessed image with its original counterpart. These visual representations serve as valuable aids in elucidating the effects of our image processing methods and the localization of stenosis positions within the vascular structures under examination.

To establish a reference benchmark for subsequent experiments, we initiated our investigation by evaluating the performance of the original images across various neural network architectures. This comprehensive assessment encompassed a range of state-of-the-art models, including YOLOv4, FasterRCNN-ResNet 101, FasterRCNN-Inceptionresnetv2, R-FCN-ResNet 101, and R-FCN-Inceptionresnetv2. Our objective was to quantify the efficacy of each model by calculating the mAP at varying IoU thresholds within the spectrum of 0.1 to 0.6. Figure 2 serves as a graphical representation of the performance of these diverse neural networks when applied to our dataset, thus providing a visual elucidation of their comparative strengths and weaknesses in the context of stenosis detection.

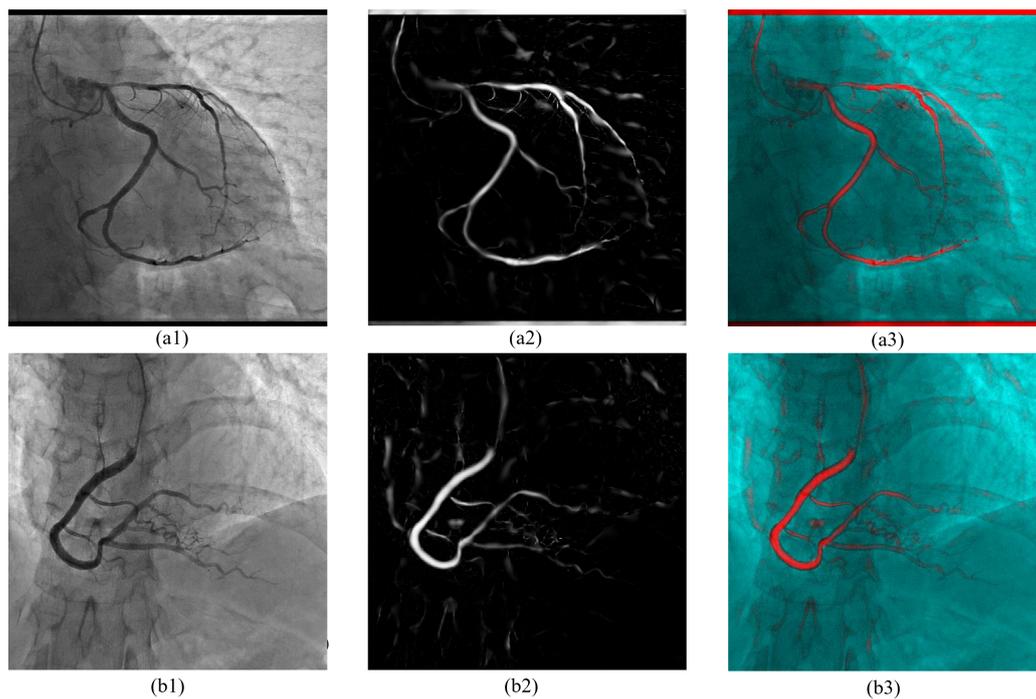


Figure 1. Images of left coronary artery (LCA) and right coronary artery (RCA) at zero imaging angle: (a1,b1) are the original images of LCA and RCA, respectively; (a2,b2) are the images enhanced using Hessian-based Frangi vesselness filter; and (a3,b3) are the images after image fusion.

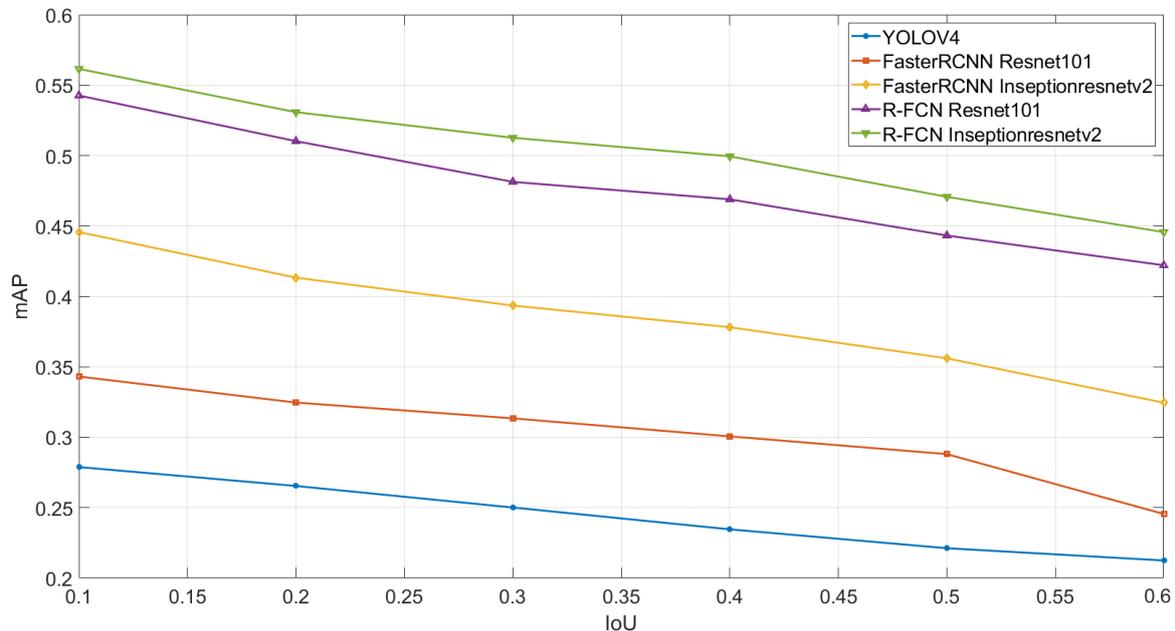


Figure 2. Network performance on the dataset without HFV filtering.

Among these neural network models, YOLOv4 exhibited mAP values of 0.2788 and 0.2125 at IoU thresholds of 0.1 and 0.6, respectively. In contrast, R-FCN-Inceptionresnetv2 demonstrated a more competitive performance, displaying mAP values of 0.5615 and 0.4456 at the corresponding IoU thresholds of 0.1 and 0.6. Notably, it is apparent from Figure 3 that both of these models exhibit limitations in effectively identifying stenotic lesions present within branch vessels, pointing towards a specific challenge in their performance in this context.

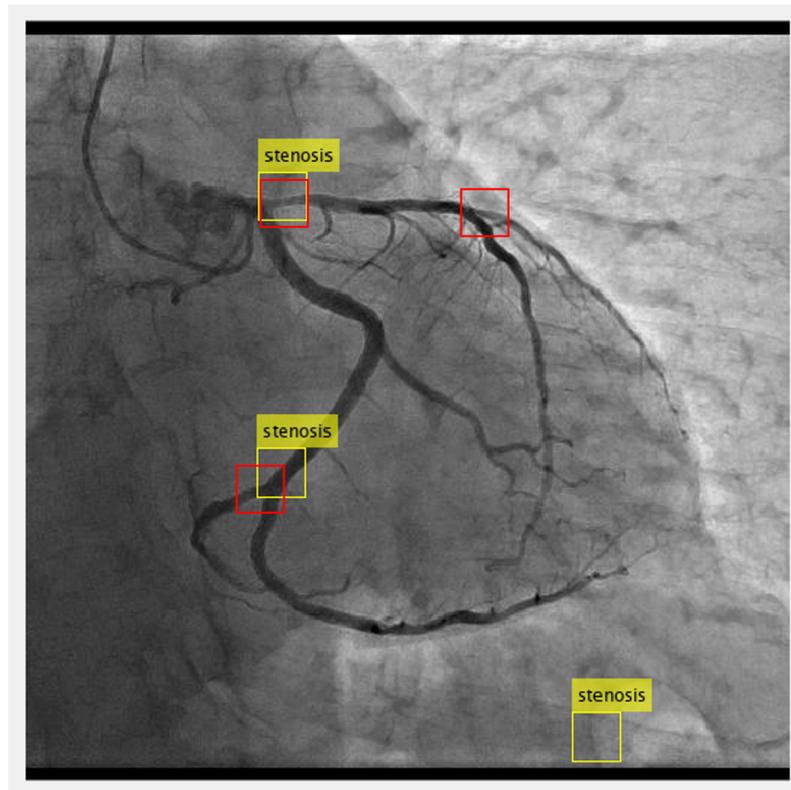


Figure 3. Example of stenosis detection by applying YOLOv4 on the dataset without HFV filter: the yellow bounding boxes respect the stenosis location detected by the trained detector and the red bounding boxes illustrate the annotation.

We compared the preprocessed images with the original images at the next six imaging angles. These images clearly showed that the vascular features in the images, especially the main vessels, were enhanced by applying the HFV filter. At the same time, the noise that appeared in the background was partially eliminated. For example, Figure 1(a2) shows the improved appearance of the tubular structures. Nevertheless, we also observed some pseudo-vessels that resembled the shape of the branch vessels throughout the image. Although these structures appeared realistic, they were present in areas where there were no vessels or where the vessels were blurred in the original image. A similar effect is shown in Figure 1(b2), where the HFV filter enhances the features while weakening the fine vessels in the background that are blurred due to the low amount of contrast.

In our quest to optimize the enhancement of all vascular features within the images while preserving fine vascular details to the greatest extent possible, we contemplated the utilization of image fusion as a potential solution. Figure 4(b1) through Figure 4(b6) vividly illustrates the outcomes of this image fusion technique applied to the processed images across the six imaging angles. Through image fusion, we effectively retained the salient features present in the original images, thereby ensuring their preservation. Concurrently, the noise and random morphological features inherent in the original images, which had the potential to transform into vascular-like structures due to the HFV filter application, were substantially mitigated.

After determining that the vascular structures in the image were enhanced, we again introduced the neural network applied to the original image and examined its performance for detection under the same conditions. Figure 5 shows the performance of these neural networks on preprocessed images.

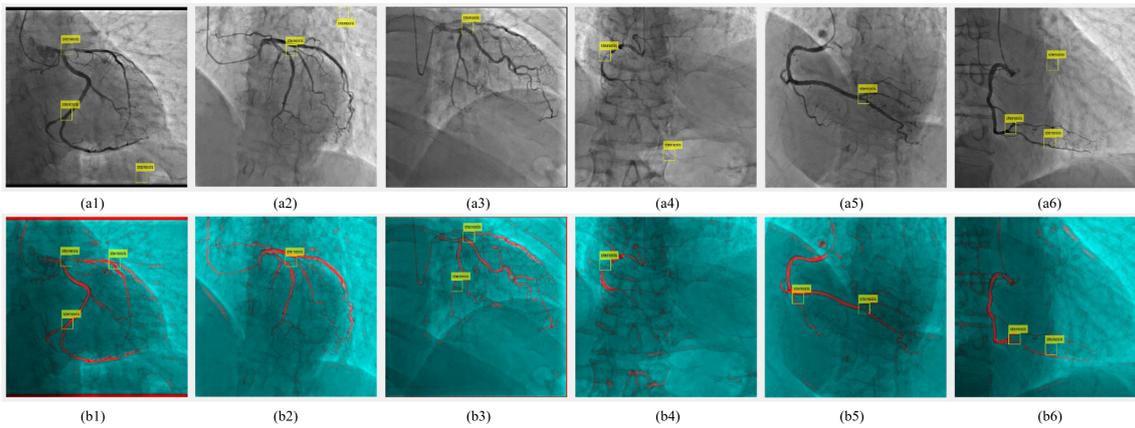


Figure 4. Stenosis detection cases of the model from six camera angles. Among them, (a1–a6) represent the detection results using the R-FCN-Inseptionresnetv2 network under the conditions of left coronary artery (LCA) zero, LCA left anterior oblique (LAO), LCA right anterior oblique (RAO), right coronary artery (RCA) zero, RCA LAO, and RCA RAO, respectively. (b1–b6) corresponds to the stenosis detection situation of (a1–a6) in the case of preprocessing by HFV filter and image fusion.

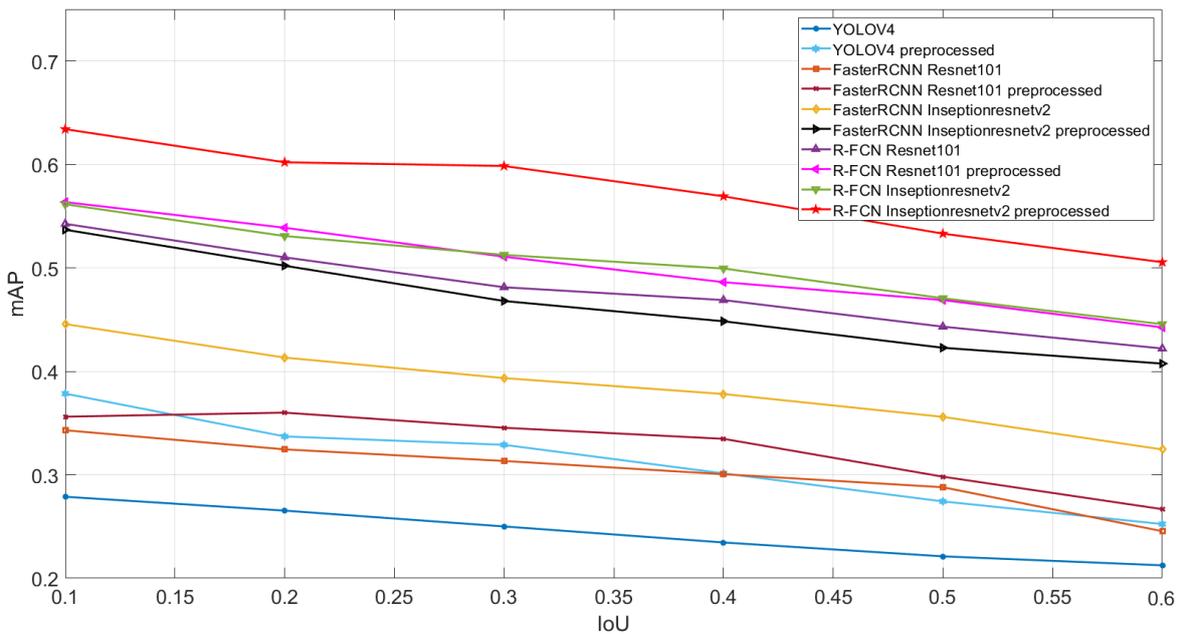


Figure 5. Network performance on original dataset and dataset preprocessed by HFV filter and image fusion.

Similar to the previous results, YOLOv4 still had the lowest accuracy among the five networks, although the accuracy had improved. The mAP values were 0.3785 and 0.2524 at IoU values of 0.1 and 0.6, respectively. The R-FCN-Inseptionresnetv2 network remained the optimal model with mAP values of 0.6341 and 0.5055 at IoU values of 0.1 and 0.6, respectively. By observing the last row of the table, it can be noticed that the preprocessing method has improved the accuracy of the model. For YOLOv4, the detection accuracy was improved by 118–135%. Meanwhile, for R-FCN-Inseptionresnetv2, which was the best network, the detection accuracy was improved by 112–116%.

Nevertheless, it is crucial to address certain challenges posed by specific imaging angles, as exemplified in Figure 1(b2), which pertains to the left anterior oblique (LAO), right anterior oblique (RAO), and ZERO angles within the right coronary artery (RCA).

In these cases, complexities such as vessel overlap arising from the camera angle can significantly impede the model’s ability to accurately detect the precise location of stenosis.

In pursuit of a more intuitive assessment of the potential enhancement conferred by the preprocessing approach and the model’s performance on our dataset, we adopted an image cropping strategy. This strategy involved the extraction of the central portion of the original image, characterized by a concentrated distribution of blood vessels, to form a refined dataset for network training. The primary objective of this approach was twofold: firstly, to mitigate the influence of noise within the dataset and, secondly, to minimize the impact of misclassifications resulting from the uneven distribution of the contrast agent. The outcomes of these experiments are depicted in Figure 6.

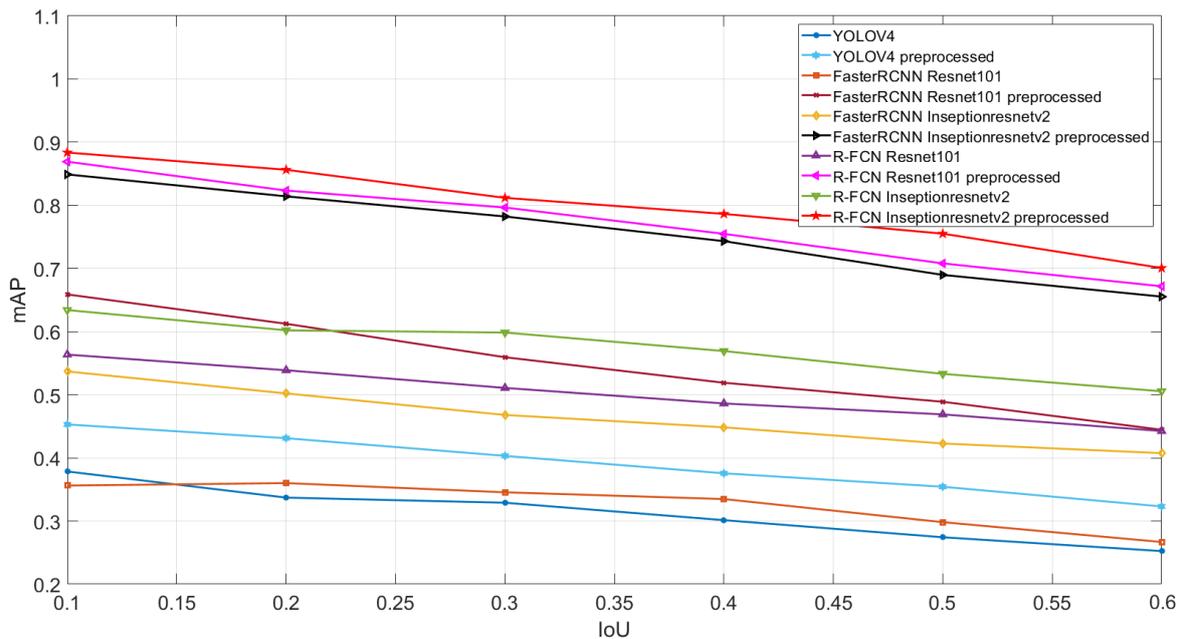


Figure 6. Network performance on cropped original dataset and dataset modified by cropping, HFV filter and image fusion.

In Figure 6, the performance of both YOLOv4 and R-FCN-Inceptionresnet2, with different IoU values, had obviously improved. R-FCN-Inceptionresnet2 obtained the mAP value of 0.7551 in the preprocessed dataset with an IoU value of 0.5. Regarding the magnitude of the improvement, for the cropped images, the preprocessed images showed a 102–111% improvement in the mAP values compared with the original images.

Furthermore, to comprehensively evaluate the performance of our approach from multiple imaging perspectives, we conducted a meticulous analysis for each imaging angle within our test dataset. This rigorous examination involved classifying the test dataset and subsequently assessing the mean average precision (mAP) achieved by the top-performing R-FCN Inceptionresnet2 model, with a stringent IoU (intersection over union) threshold set at 0.5. The tabulated outcomes of these evaluations can be found in Table 1 provides a detailed and insightful view of the model’s performance across a range of imaging angles.

Table 1. Mean average precision of six imaging angles achieved by the top-performing R-FCN Inceptionresnet2 model.

| Angles | LCA ZERO | LCA LAO | LCA RAO | RCA ZERO | RCA LAO | RCA RAO |
|--------|----------|---------|---------|----------|---------|---------|
| mAP | 0.7818 | 0.7462 | 0.7698 | 0.6908 | 0.6522 | 0.6474 |

4. Discussion

In this study, we experimented and examined whether the application of a general HFV filter for vascular contrast enhancement and image fusion would improve the accuracy of deep learning-based stenosis detection in coronary angiographic images. This study included the training and analysis of preprocessed and postprocessed images using five deep learning networks. We demonstrated that image preprocessing based on HFV filters and image fusion can improve the processing efficiency of deep learning networks in complex coronary angiographic images and that this approach does not add much training complexity. It is worth noting that the images obtained by the HFV filter may appear to create false vascular-like structures or fail to enhance the real vessels. This could potentially lead to misclassification of the model during training and detection, resulting in a low mAP index. These findings raise the question of under what conditions the HFV filter can be accurately applied in preprocessing. Therefore, a rigorous evaluation of whether improvements in the visual appearance of images are made at the expense of image accuracy is needed. This is used as a starting point for further processing of the images.

The number of features of interest in an image affects the performance of the trained model. To enable the model to better learn the stenosis features in coronary angiographic images, we introduced image fusion based on the use of HFV filters to enhance the vascular features in the images, especially the stenosis features with respect to fine vessels. Specifically, the image fusion technique complements the problem of feature loss after vessel enhancement and, to some extent, mitigates the problem of edge enhancement of small-sized features caused by the filter. The experimental results also demonstrate that the efficiency of stenosis detection for branching vessels can be enhanced to some extent using our proposed method.

In our research, we sought to investigate the potential impact of the proposed methodology on feature learning. We took care to rigorously assess the degree of improvement in mean average precision (mAP) achieved through our approach, while also considering the possibility of dataset-related issues affecting our conclusions. To address this, we employed a “chopping image” technique to systematically evaluate the dataset’s reliability. Our findings revealed an improvement in the performance of our neural network when applied to preprocessed images compared to their unprocessed counterparts.

Concurrently, our research undertook a meticulous examination of the ramifications of different imaging angles on the precision of stenosis detection within coronary arteries. The empirical findings derived from our experiments unveiled a noteworthy revelation: the accuracy of stenosis detection exhibited conspicuous variations contingent upon the specific imaging angles scrutinized. Specifically, our investigations elucidated that the detection outcomes pertaining to the three angles corresponding to the left coronary artery (LCA) exhibited superior accuracy in comparison to those observed in the context of the right coronary artery (RCA). This intriguing and differentiating pattern in detection accuracy between these two arterial segments delineates a compelling avenue for further scholarly inquiry and investigation. It points to a potentially fertile area for future research endeavors aimed at optimizing the efficiency of coronary artery stenosis detection, offering invaluable options for the advancement of medical diagnostics and treatment strategies within this pivotal domain of healthcare.

Our study is subject to several limitations that warrant discussion. These limitations encompass various facets of our research methodology and dataset, each contributing to a nuanced understanding of the constraints inherent in our investigation. Firstly, the composition of our dataset encompassed common angles of coronary angiography. While this choice was made to reflect clinical practice, it inherently introduced certain challenges. Specifically, some angiographic images, characterized by their complexity or intrinsic difficulty in identifying stenosis, were included. Consequently, our model’s ability to achieve very high mean average precision (mAP) values was constrained in such instances. The presence of these challenging images underscores the practical and clinical variability encountered in real-world scenarios, necessitating further investigation into model robustness

under such conditions. Secondly, the extraction of key frames from coronary angiography video streams, while a necessary step in our approach, came with inherent limitations. These limitations became particularly evident when dealing with branch vessels. Due to the complexities associated with the dynamic nature of coronary blood flow, obtaining a comprehensive set of key frames with branch vessels fully intact proved to be a challenging endeavor. Thirdly, although we conscientiously considered stenosis in branch vessels, many of these cases remained inadequately observed and labeled. This limitation subsequently impacted the model's capacity to effectively recognize and incorporate these branch vessel stenoses into its results. This highlights an area where future research efforts may be directed, with a focus on enhancing the model's ability to accurately detect stenosis in branch vessels within the context of coronary angiography. Fourthly, we employed an older version from the YOLO series, YOLOv4, instead of the latest version, YOLOv8. Our rationale for choosing YOLOv4 was its significant improvements over prior versions and its demonstrated stability and maturity in diverse environments. Although newer versions like YOLOv8 might offer enhanced features, they could be in experimental phases or lack thorough validation. We selected YOLOv4 for its proven reliability, considering it a cautious decision to enhance our study's dependability. The fifth point is that this study does not delve deeply into the optimization of learning parameters in the model training process, which is one of its limitations. The image training in this study is based on existing research and adopts a general approach to similar challenges. The training process, including measures to avoid overtraining, could have enhanced the model's accuracy and generalizability through appropriate parameter optimization. Future research is expected to conduct a more detailed analysis of these aspects, contributing to the improvement of model performance. These limitations underscore the necessity for ongoing research endeavors aimed at addressing these constraints and further refining our methodology. By acknowledging and actively addressing these limitations, we can progress towards a more robust and clinically applicable framework for coronary artery stenosis detection.

5. Conclusions

Overall, the application of the HFV filter significantly improves the visual appearance of the image, fills the image space with blood vessels, and reduces noise. The image fusion approach complements some of the features lost in the vascular feature enhancement stage. The models obtained by deep learning using the fused images outperformed the original images. We believe that future efforts should focus on increasing the diversity of data and maximizing the discrimination of complex coronary angiographic images, for example, by changing the signal processing methods for vessel enhancement and image fusion.

Author Contributions: Conceptualization, H.S. and Y.L.; methodology, H.S. and Y.L.; software, Y.L.; validation, Y.L., T.Y. and Y.H.; formal analysis, Y.L.; investigation, Y.L.; resources, H.S.; data curation, Y.H.; writing—original draft preparation, Y.L.; writing—review and editing, T.Y.; visualization, Y.L.; supervision, H.S.; project administration, H.S.; funding acquisition, H.S. and Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by JST, the establishment of university fellowships towards the creation of science technology innovation, Grant Number JPMJFS2101.

Institutional Review Board Statement: The study involving human participants was reviewed and approved by the Clinical Research Administration Center at Hokkaido University Hospital.

Informed Consent Statement: Informed consent was not obtained for this study, as it was a retrospective study. Accordingly, it is published on the institution's website (https://www.huhp.hokudai.ac.jp/date/rinsho-johokokai/etc_ika/ (accessed on 9 February 2024)) as an opt-out for information disclosure.

Data Availability Statement: The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare that they have no competing interests that could influence the conduct or presentation of the work described in this manuscript.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|--------|---|
| ICA | Invasive coronary angiography |
| CNNs | Convolutional neural networks |
| HFV | Hessian-based frangi vesselness |
| RCA | Right coronary artery |
| LCA | Left coronary artery |
| LAO | Left anterior oblique |
| RAO | Right coronary artery |
| IHS | Intensity-hue-saturation |
| RGB | Red–green–blue |
| R-FCN | Region-based fully convolutional networks |
| RPN | Region proposal network |
| ResNet | Residual network |
| RoIs | Regions of interest |
| mAP | Mean average precision |
| NMS | Non-maximum suppression |
| SPP | Spatial pyramid pooling |
| IoU | Intersection over union |

References

- Malakar, A.K.; Choudhury, D.; Halder, B.; Paul, P.; Uddin, A.; Chakraborty, S. A review on coronary artery disease, its risk factors, and therapeutics. *J. Cell. Physiol.* **2019**, *234*, 16812–16823. [[CrossRef](#)] [[PubMed](#)]
- Husmann, L.; Leschka, S.; Desbiolles, L.; Schepis, T.; Gaemperli, O.; Seifert, B.; Cattin, P.; Frauenfelder, T.; Flohr, T.G.; Marincek, B.; et al. Coronary artery motion and cardiac phases: Dependency on heart rate—Implications for CT image reconstruction. *Radiology* **2007**, *245*, 567–576. [[CrossRef](#)] [[PubMed](#)]
- Zhang, H.; Mu, L.; Hu, S.; Nallamotheu, B.K.; Lansky, A.J.; Xu, B.; Bouras, G.; Cohen, D.J.; Spertus, J.A.; Masoudi, F.A.; et al. Comparison of physician visual assessment with quantitative coronary angiography in assessment of stenosis severity in China. *JAMA Intern. Med.* **2018**, *178*, 239–247. [[CrossRef](#)] [[PubMed](#)]
- Roychowdhury, S.; Koozekanani, D.D.; Parhi, K.K. Blood vessel segmentation of fundus images by major vessel extraction and subimage classification. *IEEE J. Biomed. Health Inform.* **2014**, *19*, 1118–1128.
- Chen, M.; Wang, X.; Hao, G.; Cheng, X.; Ma, C.; Guo, N.; Hu, S.; Tao, Q.; Yao, F.; Hu, C. Diagnostic performance of deep learning-based vascular extraction and stenosis detection technique for coronary artery disease. *Br. J. Radiol.* **2020**, *93*, 20191028. [[CrossRef](#)] [[PubMed](#)]
- Pijls, N.H.; Sels, J.W.E. Functional measurement of coronary stenosis. *J. Am. Coll. Cardiol.* **2012**, *59*, 1045–1057. [[CrossRef](#)] [[PubMed](#)]
- Frangi, A.F.; Niessen, W.J.; Vincken, K.L.; Viergever, M.A. Multiscale vessel enhancement filtering. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Cambridge, MA, USA, 11–13 October 1998; Springer: Berlin/Heidelberg, Germany, 1998; pp. 130–137.
- Bi, R.; Dinish, U.; Goh, C.C.; Imai, T.; Moothanchery, M.; Li, X.; Kim, J.Y.; Jeon, S.; Pu, Y.; Kim, C.; et al. In vivo label-free functional photoacoustic monitoring of ischemic reperfusion. *J. Biophotonics* **2019**, *12*, e201800454. [[CrossRef](#)] [[PubMed](#)]
- Orlova, A.; Sirotkina, M.; Smolina, E.; Elagin, V.; Kovalchuk, A.; Turchin, I.; Subochev, P. Raster-scan optoacoustic angiography of blood vessel development in colon cancer models. *Photoacoustics* **2019**, *13*, 25–32. [[CrossRef](#)] [[PubMed](#)]
- Wang, S.; Li, B.; Zhou, S. A segmentation method of coronary angiograms based on multi-scale filtering and region-growing. In Proceedings of the 2012 IEEE International Conference on Biomedical Engineering and Biotechnology, Macau, China, 28–30 May 2012; pp. 678–681.
- Zreik, M.; Van Hamersvelt, R.W.; Wolterink, J.M.; Leiner, T.; Viergever, M.A.; Išgum, I. A recurrent CNN for automatic detection and classification of coronary artery plaque and stenosis in coronary CT angiography. *IEEE Trans. Med. Imaging* **2018**, *38*, 1588–1598. [[CrossRef](#)] [[PubMed](#)]
- Ovalle-Magallanes, E.; Avina-Cervantes, J.G.; Cruz-Aceves, I.; Ruiz-Pinales, J. Transfer learning for stenosis detection in X-ray coronary angiography. *Mathematics* **2020**, *8*, 1510. [[CrossRef](#)]
- Cruz-Aceves, I.; Oloumi, F.; Rangayyan, R.M.; Aviña-Cervantes, J.G.; Hernandez-Aguirre, A. Automatic segmentation of coronary arteries using Gabor filters and thresholding based on multiobjective optimization. *Biomed. Signal Process. Control* **2016**, *25*, 76–85. [[CrossRef](#)]
- Felfelian, B.; Fazlali, H.R.; Karimi, N.; Soroushmehr, S.M.R.; Samavi, S.; Nallamotheu, B.; Najarian, K. Vessel segmentation in low contrast X-ray angiogram images. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 375–379.

15. Rahman, C.M.A.; Nyeem, H. Active Contour based Segmentation of ROIs in Medical Images. In Proceedings of the 2019 IEEE International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox's Bazar, Bangladesh, 7–9 February 2019; pp. 1–6.
16. Banerjee, R.; Ghose, A.; Mandana, K.M. A hybrid CNN-LSTM architecture for detection of coronary artery disease from ECG. In Proceedings of the 2020 IEEE International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–8.
17. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
18. Kim, D.K.; Kim, B.S.; Kim, Y.J.; Kim, S.; Yoon, D.; Lee, D.K.; Jeong, J.; Jo, Y.H. Development and validation of an artificial intelligence algorithm for detecting vocal cords in video laryngoscopy. *Medicine* **2023**, *102*, e36761. [[CrossRef](#)] [[PubMed](#)]
19. Das, S.; Hasan, O.; Chowdhury, A.; Aslam, S.M.; Minhaz Hossain, S.M. An Automatic Detection of Heart Block from ECG Images Using YOLOv4. In Proceedings of the International Conference on Hybrid Intelligent Systems, Online, 13–15 December 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 981–990.
20. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; Volume 28.
21. Dai, J.; Li, Y.; He, K.; Sun, J. R-fcn: Object detection via region-based fully convolutional networks. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; Volume 29.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.