ImageCAS: A large-scale dataset and benchmark for coronary artery segmentation based on computed tomography angiography images☆

An Zeng [a,1], Chunbiao Wu [a,1], Guisen Lin [h,1], Wen Xie [b,1], Jin Hong [b], Meiping Huang [b], Jian Zhuang [b], Shanshan Bi [c], Dan Pan [d], Najeeb Ullah [e], Kaleem Nawaz Khan [e], Tianchen Wang [f], Yiyu Shi [f], Xiaomeng Li [g], Xiaowei Xu [b,*]

[a] School of Computer Science, Guangdong University of Technology, Guangzhou, China
[b] Guangdong Provincial Key Laboratory of South China Structural Heart Disease, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, 510080, China
[c] Department of Computer Science and Engineering, Missouri University of Science and Technology, Rolla, MO, United States
[d] Department of Computer Science, Guangdong Polytechnic Normal University, Guangzhou, China
[e] Department of Computer Science, University of Engineering and Technology, Mardan, KP, Pakistan
[f] Department of Computer Science and Engineering, University of Notre Dame, Indiana, United States
[g] Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong Special Administrative Region, China
[h] Department of Radiology, Shenzhen Children's Hospital, Shenzhen, China

ARTICLE INFO

ABSTRACT

Cardiovascular disease (CVD) accounts for about half of non-communicable diseases. Vessel stenosis in the coronary artery is considered to be the major risk of CVD. Computed tomography angiography (CTA) is one of the widely used noninvasive imaging modalities in coronary artery diagnosis due to its superior image resolution. Clinically, segmentation of coronary arteries is essential for the diagnosis and quantification of coronary artery disease. Recently, a variety of works have been proposed to address this problem. However, on one hand, most works rely on in-house datasets, and only a few works published their datasets to the public which only contain tens of images. On the other hand, their source code have not been published, and most follow-up works have not made comparison with existing works, which makes it difficult to judge the effectiveness of the methods and hinders the further exploration of this challenging yet critical problem in the community. In this paper, we propose a large-scale dataset for coronary artery segmentation on CTA images. In addition, we have implemented a benchmark in which we have tried our best to implement several typical existing methods. Furthermore, we propose a strong baseline method which combines multi-scale patch fusion and two-stage processing to extract the details of vessels. Comprehensive experiments show that the proposed method achieves better performance than existing works on the proposed large-scale dataset. The benchmark and the dataset are published at https://github.com/XiaoweiXu/ImageCAS-A-Large-Scale-Dataset-and-Benchmark-for-Coronary-Artery-Segmentation-based-on-CT.

1. Introduction

Cardiovascular disease (CVD) is one of the leading health problems around the world nowadays. According to World Health Organization (WHO), 17.9 million deaths due to CVD occurred in 2019, accounting for 32% of all global deaths (Organization et al., 2009). Australian Institute of Health and Welfare (AIHW) reported that CVD was the leading cause of deaths in Australia, representing 42% of all death in 2018 (Zhang, 2010). Among all the CVDs, coronary heart disease is the most common type (Cooper et al., 2000) where the pathophysiology is

mainly attributed to abnormal coronary artery stenosis. Such stenosis often results in decreased myocardial perfusion and hypoxia damage of myocardial cells, and finally leads to myocardial infarction.

In clinical practice, computed tomography angiography (CTA) is widely used for diagnosis and treatment planning of coronary artery diseases since it is noninvasive and able to provides high-resolution 3D imaging (Collet et al., 2018). Once CTA images are obtained, radiologists first locate the coronary arteries and isolate its boundaries. Then, the narrowing part is extracted and quantified for final diagnosis and treatment planning. It is widely acknowledged that, however, such a process based on manual operations are unfavorably time-consuming and error-prone. What is worse, the ever-increasing quantity and variety of medical images (Li et al., 2018c) can even make manual segmentation completely impracticable in terms of cost and reproducibility. Therefore, automatic coronary artery segmentation is highly desirable.

However, this task is very challenging due to multiple reasons. First, the anatomical structure of coronary artery varies significantly from population to population. For example, coronary arteries are usually surrounded by a layer of fat but in some people the arteries are within the heart muscle itself. Second, some CTA images are noisy with artifacts that will cause low quality in segmentation result. Third, the tubular structure of coronary arteries is extremely complex. For example, there is a large number of bifurcations along the arteries and a small proportion of coronary arteries area in the transverse planes (Zhu et al., 2021).

To address the above issues, tens of works have been proposed in the community in the last decade. The approaches used in these works can be divided into two main categories: traditional machine learning (ML) based method, and deep learning (DL) based method. Traditional ML-based method can be further sub-classified into pixel-based method (Doyle et al., 2006; Nguyen et al., 2012; Tabesh et al., 2007; Sirinukunwattana et al., 2015a) and structure-based method (Altunbay et al., 2010; Gunduz-Demir et al., 2010; Fu et al., 2014; Sirinukunwattana et al., 2015b; Lesage et al., 2016). These methods achieve promising results using hand-crafted features and prior knowledge of structures of coronary artery (Zheng et al., 2011; Mohr et al., 2012; Broersen et al., 2012; Shahzad et al., 2013; Chi et al., 2015). However, they suffer considerable degradation when applied to coronaries with serious deformation. Different from ML-based method, the recently proposed DL-based methods require little hand-crafted features or prior knowledge. Works using such an approach has achieved considerable improvement over the first approach (Huang et al., 2018; Shen et al., 2019; Chen et al., 2019; Wolterink et al., 2019; Kong et al., 2020; Gu and Cai, 2021; Zhu et al., 2021; Tian et al., 2021), proving its high effectiveness for coronary artery segmentation.

We perform a detailed analysis of existing works as shown in Table 1 and Table 2, and find that most works fail to make a fair and comprehensive comparison with others. For example, Kong et al. (2020) did not make comparisons with any previous related works but just some popular deep neural networks (DNNs). Shen et al. (2019) just listed the Dice score of related works for comparison which is not fair enough as the two methods were not evaluated using the same dataset. Some works (Chi et al., 2015; Han et al., 2016) even used different sets of evaluation metrics tailored for specific clinical needs, while others (Shen et al., 2019) even used different annotations in which the initial aorta is included, leading to a much higher Dice score. These evaluative biases are mainly due to the lack of a large-scale benchmark dataset available in the public domain. The only two public datasets for coronary artery segmentation (Schaap et al., 2009a; Kirişli et al., 2013) (also shown in Table 2) only contain 8 and 18 images for training, respectively. Besides that, all the existing methods have not released their source code, which has introduced more difficulties for a fair comparison.

In this paper, we propose a large dataset to fairly investigate the effectiveness of automatic coronary artery segmentation methods. This

dataset contains 1000 3D CTA images, which is considerably larger than the existing public datasets. In addition, we also propose a benchmark based on this dataset, in which we not only implement several typical existing methods but also propose a strong baseline method. Experimental results show that our baseline method achieves better performance than all the existing methods. and meanwhile shows a good potential for improvement. The contributions of this work are summarized as follows:

- We have collected a large-scale publicly available dataset containing 1000 patients for coronary artery segmentation available at https://github.com/XiaoweiXu/ImageCAS-A-Large-Scale-Dataset-and-Benchmark-for-Coronary-Artery-Segmentation-based-on-CT. An official data split of the dataset is also provided. We hope this dataset can help to promote related research in the community;
- We have proposed a benchmark in which we implemented a variety of existing methods for coronary artery segmentation. We have also published the benchmark, and hope this could help follow-up works to make fair comparisons;
- We have proposed a strong baseline method which combines multi-scale patch fusion and two-stage processing to extract the details of vessels, and the experimental results show that our method has outperformed existing state-of-the-art methods.

The rest of the paper is organized as follows. Section 2 gives an overview of related works. In Section 3, details of the proposed dataset is presented. Subsequently, Section 4 presents the proposed benchmark including several typical existing methods and our baseline method. The experimental results are presented and discussed in Section 5, and Section 6 concludes the paper.

2. Related work

In this section, we review existing literature on coronary artery segmentation. We first review the related works by the type of approaches used, i.e. traditional ML-based approach and DL-based approach, and then discuss the datasets and benckmarks commonly adopted in these works.

2.1. Traditional ML-based approach

Hand-crafted features and prior knowledge of structures of coronary artery are extensively used in this approach as shown Table 1. Zheng et al. (2011) proposed an ML method to exploit the rich domain-specific knowledge (particularly a set of geometric and image features) embedded in an expert-annotated dataset. Mohr et al. (2012) proposed a level-set based approach for efficient processing. Wang et al. (2012) combines level-sets with an implicit 3D model of the vessels for accurate segmentation. Broersen et al. (2012) adopted a pipeline consisting of three consecutive steps. Shahzad et al. (2013) performed segmentation with the help of extracted centerlines. Lugauer et al. (2014a) used a learning-based boundary detector to enable a robust lumen contour detection via dense ray-casting. Lugauer et al. (2014b) proposed a model-guided segmentation approach based on a Markov random field formulation with convex priors. Chi et al. (2015) integrated coronary artery features of density, local shape and global structure into a learning framework. Lesage et al. (2016) considered vessel segmentation as an iterative tracking process and proposed a new Bayesian tracking algorithm based on particle filters for the delineation of coronary arteries. Han et al. (2016) used an active search method to find branches and seemingly disconnected but actually connected vessel segments. Freiman et al. (2017) used a flow simulation method (Nickisch et al., 2015) in the coronary trees with accounting for partial volume effects (Glover and Pelc, 1980). Gao et al. (2019) located the coronary root through extracting aorta by using circular Hough transform. Du et al. (2021) proposed a new segmentation framework which included noise reduction, candidate region detection, geometric feature extraction, and coronary artery tracking techniques.

Table 1
Existing traditional machine learning based works and their Dice scores (in %) of coronary artery segmentation in the past decade. In $a(b)$, a is the overall quantity and b is the quantity for training. *General* indicates that the method is designed for a variety of applications, while *specific* indicates that the method is designed and optimized for coronary vessel segmentation.

Work	Year	Data quantity	Compared traditional ML methods		Compared deep learning methods		Average dice score	Code available	Data available
			General	Specific	General	Specific			
Zheng et al. (2011)	2011	54(40)	Kroon (2009)	NA	NA	NA	NA	No	No
Mohr et al. (2012)	2012	24[a]	NA	NA	NA	NA	70–73	No	Kirişli et al. (2013)
Wang et al. (2012)	2012	42(18)	NA	NA	NA	NA	68–72	No	Kirişli et al. (2013)
Broersen et al. (2012)	2012	42(18)	NA	NA	NA	NA	66	No	Kirişli et al. (2013)
Shahzad et al. (2013)	2013	48(18)	NA	NA	NA	NA	65	No	Kirişli et al. (2013)
Lugauer et al. (2014a)	2014	48(18)	NA	Mohr et al. (2012), Wang et al. (2012) and Shahzad et al. (2013)	NA	NA	72–74	No	Kirişli et al. (2013)
Lugauer et al. (2014b)	2014	48(18)	NA	Mohr et al. (2012), Lugauer et al. (2014a), Shahzad et al. (2013) and Wang et al. (2012)	NA	NA	75–77	No	Kirişli et al. (2013)
Chi et al. (2015)	2015	10(6)	NA	NA	NA	NA	84	No	No
Lesage et al. (2016)	2016	61(10)	Skare et al. (2003) and Douc et al. (2009)	NA	NA	NA	86.2	No	No
Han et al. (2016)	2016	32(8)	NA	Schaap et al. (2009b)	NA	NA	<84.3	No	Schaap et al. (2009a)
Freiman et al. (2017)	2017	48(18)	NA	Mohr et al. (2012) Lugauer et al. (2014a)	NA	NA	69–74	No	Kirişli et al. (2013)
Gao et al. (2019)	2017	50[a]	NA	NA	NA	NA	93-95[b]	No	No
Du et al. (2021)	2021	100[a]	NA	NA	Han et al. (2014), Yu et al. (2017) and Shen et al. (2019)	NA	82	No	No

[a]Only testing data is involved (training data is not applicable).
[b]The segmentation result includes the initial part of the aorta.

2.2. DL-based approach

Since the rise of deep learning, it has attracted tremendous attention in the related communities. Currently, as shown in Table 2, there are mainly five technique trends including pixel-based segmentation (Moeskops et al., 2016; Kjerland, 2017), direct segmentation (Shen et al., 2019; Lee et al., 2019; Yang et al., 2019; Fu et al., 2020; Lei et al., 2020; Gu et al., 2020; Gu and Cai, 2021; Zhu et al., 2021; Liang et al., 2021; Tian et al., 2021; Cheung et al., 2021; Li et al., 2018a,b; Lin et al., 2022), patch based segmentation (Duan et al., 2018; Chen et al., 2018b; Huang et al., 2018; Chen et al., 2019; Mirunalini et al., 2019; Wang et al., 2021; Pan et al., 2021), tree data based segmentation (Kong et al., 2020), and graph data based segmentation (Wolterink et al., 2019). Some other works like (Wu et al., 2019) performs detailed anatomical labeling of coronary arteries, which is out of the scope of this paper.

Pixel-based segmentation is the pioneer to adopt convolutional neural networks (CNNs) into coronary artery segmentation. Moeskops et al. (2016) used a single convolutional neural network to show the feasibility of deep learning for coronary vessel segmentation. Kjerland (2017) used two neural networks trained on aorta segmentation and coronary segmentation respectively and is able to segment the complete coronary artery tree.

Then, direct segmentation has been popular due to the rise of U-net (Ronneberger et al., 2015). Shen et al. (2019) propose a joint framework based on deep learning and traditional level set method. Lee et al. (2019) adopted template transformer networks where a shape template is deformed to match the underlying structure of interest through an end-to-end trained spatial transformer network. Yang et al. (2019) adopted a discriminative coronary artery tracking method which included two parts: a tracker and a discriminator. Fu et al. (2020) used Mask R-CNN for coronary artery segmentation in which the lung region prior is masked out to avoid the interferences from pulmonary vessels. Lei et al. (2020) integrated deep attention strategy into the fully convolutional network (FCN) model (Long et al., 2015) to highlight the

Table 2

Existing deep learning based works and their Dice scores (in %) of coronary artery segmentation in the past decade. In $a(b)$, a is the overall quantity and b is the quantity for training. *General* indicates that the method is designed for a variety of applications, while *specific* indicates that the method is designed and optimized for coronary vessel segmentation.

Work	Year	Data quantity	Compared traditional ML methods		Compared deep learning methods		Average dice score	Code available	Data available
			General¶	Specific§	General	Specific			
Moeskops et al. (2016)	2016	10(6)	NA	NA	NA	NA	<65[c]	No	No
Kjerland (2017)	2017	42(18)	NA	NA	NA	NA	78	No	Kirişli et al. (2013)
Duan et al. (2018)	2018	50(40)	Frangi et al. (1998)	NA	NA	NA	80	No	No
Chen et al. (2018b)	2018	44(35)	NA	NA	NA	NA	86	No	No
Huang et al. (2018)	2018	52(45)	NA	Mohr et al. (2012) and Shahzad et al. (2013)	NA	Moeskops et al. (2016) and Kjerland (2017)	83[b]	No	Kirişli et al. (2013)
Shen et al. (2019)	2019	70(50)	NA	NA	Milletari et al. (2016)	NA	90[a]	No	No
Chen et al. (2019)	2019	15(11)	NA	NA	NA	NA	80	No	No
Mirunalini et al. (2019)	2019	50(40)	NA	NA	NA	NA	92	No	No
Wolterink et al. (2019)	2019	42(18)	NA	Lugauer et al. (2014b) and Freiman et al. (2017)	NA	NA	73–75	No	Kirişli et al. (2013)
Lee et al. (2019)	2019	548(274)	NA	NA	Ronneberger et al. (2015) and Long et al. (2015)	NA	85	No	No
Yang et al. (2019)	2019	8(7)	Cetin and Unal (2015), Aylward and Bullitt (2002) and Friman et al. (2010)	NA	NA	NA	<99[b]	No	Schaap et al. (2009a)
Fu et al. (2020)	2020	25(20)	NA	NA	NA	NA	90	No	No
Lei et al. (2020)	2020	48(18)	NA	NA	NA	NA	83	No	Kirişli et al. (2013)
Gu et al. (2020)	2020	70(50)	NA	NA	Çiçek et al. (2016), Milletari et al. (2016), Zhang et al. (2019), Long et al. (2015) and Huang et al. (2019)	NA	91	No	No
Kong et al. (2020)	2020	916(733)	NA	NA	Yu et al. (2017)	NA	85	No	No
Wang et al. (2021)	2021	48(18)	NA	NA	NA	Moeskops et al. (2016), Kjerland (2017) and Chen et al. (2019)	91	No	Kirişli et al. (2013)
Gu and Cai (2021)	2021	59(34)	NA	NA	Patravali et al. (2017), Duan et al. (2019), Mortazi et al. (2017), Chen et al. (2018a) and Milletari et al. (2016)	NA	87	No	No

Table 2 (*continued*).

Zhu et al. (2021)	2021	30(24)	NA	NA	Ronneberger et al. (2015), Oktay et al. (2018), Badri-narayanan et al. (2017) and Diakogiannis et al. (2020)	NA	88	Yes	No
Liang et al. (2021)	2021	300(200)	NA	NA	Çiçek et al. (2016)	NA	89	No	No
Tian et al. (2021)	2021	70(50)	NA	NA	Çiçek et al. (2016), Milletari et al. (2016), Zhang et al. (2019), Huang et al. (2019) and Isensee et al. (2021)	NA	94[a]	No	No
Pan et al. (2021)	2021	474(432)	NA	NA	Yu et al. (2016)	NA	94	No	No
Cheung et al. (2021)	2021	69(44)	NA	NA	Zhou et al. (2018)	NA	89	No	No

[a] The segmentation result includes the initial part of the aorta.
[b] Two kinds of datasets including a private dataset and a public dataset are adopted for evaluation, and the comparison is only performed on the public dataset.
[c] The Dice score was estimated from figures.

informative semantic features extracted from CCTA images. Gu et al. (2020) proposed a global feature embedded network for better coronary arteries segmentation. Gu and Cai (2021) proposed a two-stage strategy that retains the advantages of both 2D and 3D CNNs. Zhu et al. (2021) proposed a U-shaped network based on spatio-temporal feature fusion structure which combines features of multiple levels and different receptive fields separately to get more precise boundaries. Liang et al. (2021) proposed an improved U-net which integrated channel attention and spatial attention to distinguish confusing categories and targets with similar appearance features. Tian et al. (2021) combined deep learning and digital image processing algorithms to address the problem of the limited GPU memory. Cheung et al. (2021) proposed a fully automatic two-dimensional U-net model to segment the aorta and coronary arteries.

As direct segmentation usually requires a downsampled image as the input thus cannot obtain the local details well, patch based segmentation becomes popular. Essentially in this method, a rough segmentation on the downsampled image is performed first and then a refinement is adopted in a patch manner to extract the local details. Duan et al. (2018) proposed a context aware 3D FCN for vessel enhancement and segmentation which outperformed conventional Hessian vesselness based approach (Frangi et al., 1998). Chen et al. (2018b) adopted a paired multi-scale 3D CNN to identify which voxels belong to the vessel lumen. Huang et al. (2018) transformed CTA image into small patches and then sent them to a 3D U-net (Çiçek et al., 2016) for processing. Chen et al. (2019) incorporate vesselness maps into the input of 3D U-Net (Çiçek et al., 2016) to highlight the tubular structure of coronary arteries. Mirunalini et al. (2019) combined CNN and recurrent neural networks to identify the presence of coronary arteries in 2D slices. Wang et al. (2021) incorporated the voxel and point cloud-based segmentation methods into a coarse-to-fine framework. Pan et al. (2021) proposed 3D Dense-U-Net which was further optimized with a focal loss to tackle the imbalance problem.

Recently, some methods have been proposed to incorporate image data into segmentation by converting them into special data structures, such as tree structures and graph structures. Note that the morphological structure of the coronary arteries is tree-like, and generally divides

into a left coronary artery and a right coronary tree with blood flowing from the aorta to the coronary arteries and then to the individual branches. Wolterink et al. (2019) used graph convolutional networks (GCNs) to predict the spatial location of vertices in a tubular surface mesh that segments coronary artery. Kong et al. (2020) proposed a novel tree-structured convolutional gated recurrent unit (ConvGRU) model to learn the anatomical structure of coronary artery.

2.3. Benchmark and datasets

In Table 1 and Table 2 we can easily notice the broad attention of this topic around the world. We can also find that most works did not make a fair and comprehensive comparison with earlier works. For the traditional ML-based approach, many works focused on algorithmic development based on a public dataset (Kirişli et al., 2013) which has only 42 CTA images. The comparisons were also made based on this dataset, and each work (Mohr et al., 2012; Wang et al., 2012) just obtained the performance from other papers without reimplementation. In the meantime, other works (Han et al., 2016) were using another public dataset (Schaap et al., 2009a) for their evaluation. On the other hand, however, even more works prepared in-house datasets for evaluation which failed to compare to other works or only compared with general methods (Lugauer et al., 2014b; Chi et al., 2015; Han et al., 2016; Freiman et al., 2017; Gao et al., 2019; Du et al., 2021). Note that we define general methods here as those usually target at general algorithms for a variety of applications, while oppositely specific methods as thoses designed and optimized for specific applications, e.g., coronary vessel segmentation. The use of proprietary dataset and hence missing or partial comparison problem became even more common when DL-Based approach began to prevail. For example, Kong et al. (2020) did not compare their model with any previous related works but just some popular DNNs such as 3D U-net (Huang et al., 2019). Shen et al. (2019) listed the Dice score achieved by other works on different dataset for the evaluation of theirs. In addition to dataset inconsistency, problems in other benchmarking components are also widespread. Some works (Chi et al., 2015; Han et al., 2016) used different sets of evaluation metrics based on their specific clinical
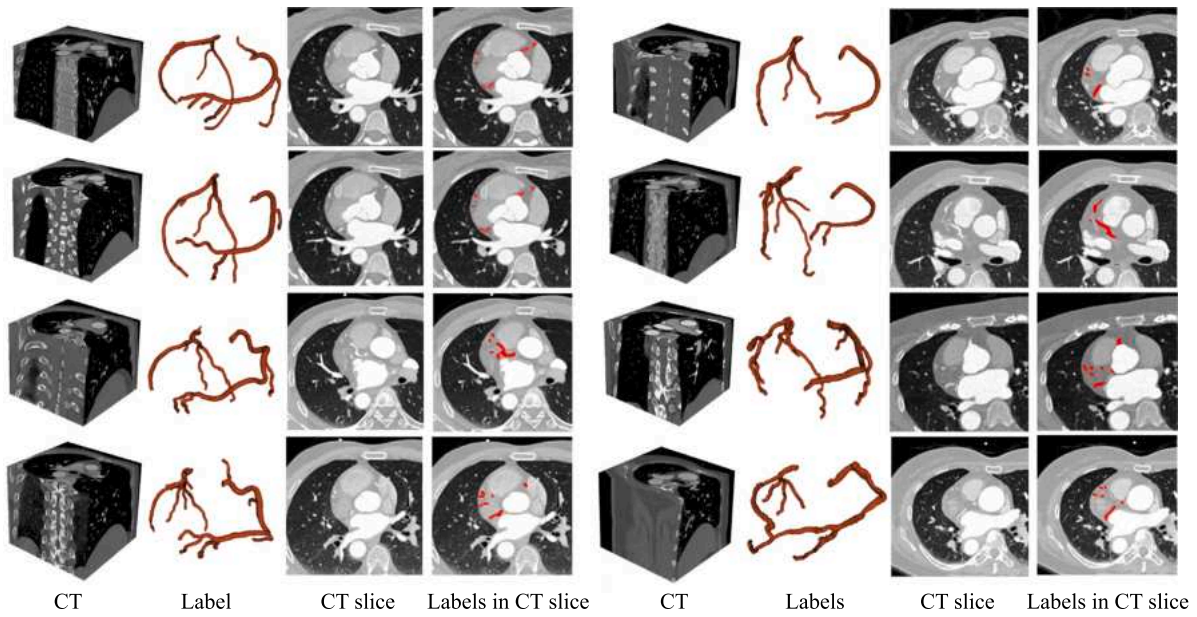
Fig. 1. Examples including CT images, their labels, CT slices and their labels in CT slices in the proposed ImageCAS dataset. Note that subclasses of coronary arteries are not further individually labeled.

needs, while others (Shen et al., 2019) used a different annotation strategy in which the initial aorta is included, leading to a much higher Dice score.

With the above analysis, we can notice that a benchmark and a dataset for further development of this topic are needed. On one hand, most of the methods only used in-house datasets for evaluation, but their image acquisition parameters, reconstruction techniques, and label methods vary significantly. On the other hand, most of the methods just performed comparison with general methods but without existing related works. The situations became even worse for deep learning based methods. Note that the success of deep learning heavily relies on a large dataset with high quality annotations, but the two commonly used datasets (Schaap et al., 2009a; Kirişli et al., 2013) are quite small (only contain just 8 and 18 images respectively for training). In addition, almost all works have not provided their source code, which also brings a lot of difficulties for their fair comparison. Although some works (Moeskops et al., 2016; Kjerland, 2017; Yu et al., 2017) are relatively easy to implement for comparison purpose, most others are intensively complex with too many details in their specifications such as network structure, preprocessing, hyper-parameters, and postprocessing. The above problem has also been identified in some other works. For example, the work (Tian et al., 2021) stated that "since most of the coronary artery segmentation methods use private dataset, there is no suitable public coronary data set for us to test". Both (Huang et al., 2018) and Yang et al. (2019) performed comparison with related works on the dataset (Schaap et al., 2009b), but conducted experiment on a private dataset without comparison which is due the fact that it is not easy to implement the existing methods. The work (Huang et al., 2018) stated that "the training data seemed relatively small.."... Some works (Yang et al., 2019) also tried to solve the small dataset problem. The work (Moeskops et al., 2016) stated that "in future work, we will further investigate the capacity of the current architecture with more data and segmentation tasks".

In order to uniformly investigate the effectiveness of each method, we collected a large dataset and made it public available. This dataset contains 1000 3D CTA images, which is considerably larger than the existing public datasets. Furthermore, we proposed a benchmark in which we tried our best to implement typical existing methods. Although the code of the existing methods is not publicly available and some experimental details are missing, we have conducted similar implementations

based on their core ideas. We further evaluated the performance of each method on the proposed dataset with multiple sets of configurations. In addition, we have also proposed a baseline segmentation framework which achieved better performance than existing works. Details of the proposed dataset and benckmark will be discussed in Sections 3 and 4.

3. ImageCAS dataset

The proposed dataset consists of 3D CTA images captured by Siemens 128-slice dual-source scanner from 1000 patients. For those patients who have previously been diagnosed with coronary artery disease, early revascularization (within 90 days after) are included. The high-dose CTA is performed and during the reconstruction, the 30%–40% phase or the 60%–70% phase is selected to obtain the best coronary artery images. The images have sizes of $512 \times 512 \times (206 - 275)$ voxels, a planar resolution of 0.29–0.43 mm^2, and spacing of 0.25–0.45 mm. The data was collected from realistic clinical cases at the Guangdong Provincial People's Hospital during April 2012 to December 2018. Only the patients older than 18 years and with a documented medical history of ischemic stroke, transient ischemic attack and/or peripheral artery disease are eligible to be included. The index cardiac CTA or low imaging quality of CCTA (assessed by level III radiologist) were also excluded due to the possible influence on the function of the coronary artery. Finally, there are totally 414 females and 586 males included, the average ages being 59.98 and 57.68, respectively. The left and right coronary arteries in each image are independently labeled by two radiologists, and their results are cross-validated. In case of discrepancy, a third radiologist will perform the annotation and the final result is determined by consensus. The labeled coronary artery includes the left main coronary artery, left anterior descending coronary artery, left circumflex coronary artery, right coronary artery, diagonal 1, diagonal 2, diagonal 3, obtuse marginal branch 1, obtuse marginal branch 2, obtuse marginal branch 3, ramus intermedius, posterior descending arteries, acute marginal 1 and other blood vessels according to the AHA naming convention (17 paragraphs). Two examples of the dataset are shown in Fig. 1.

4. Benchmark

As the deep learning based methods in almost all recent works have shown promising performance in cornary artery segmentation,
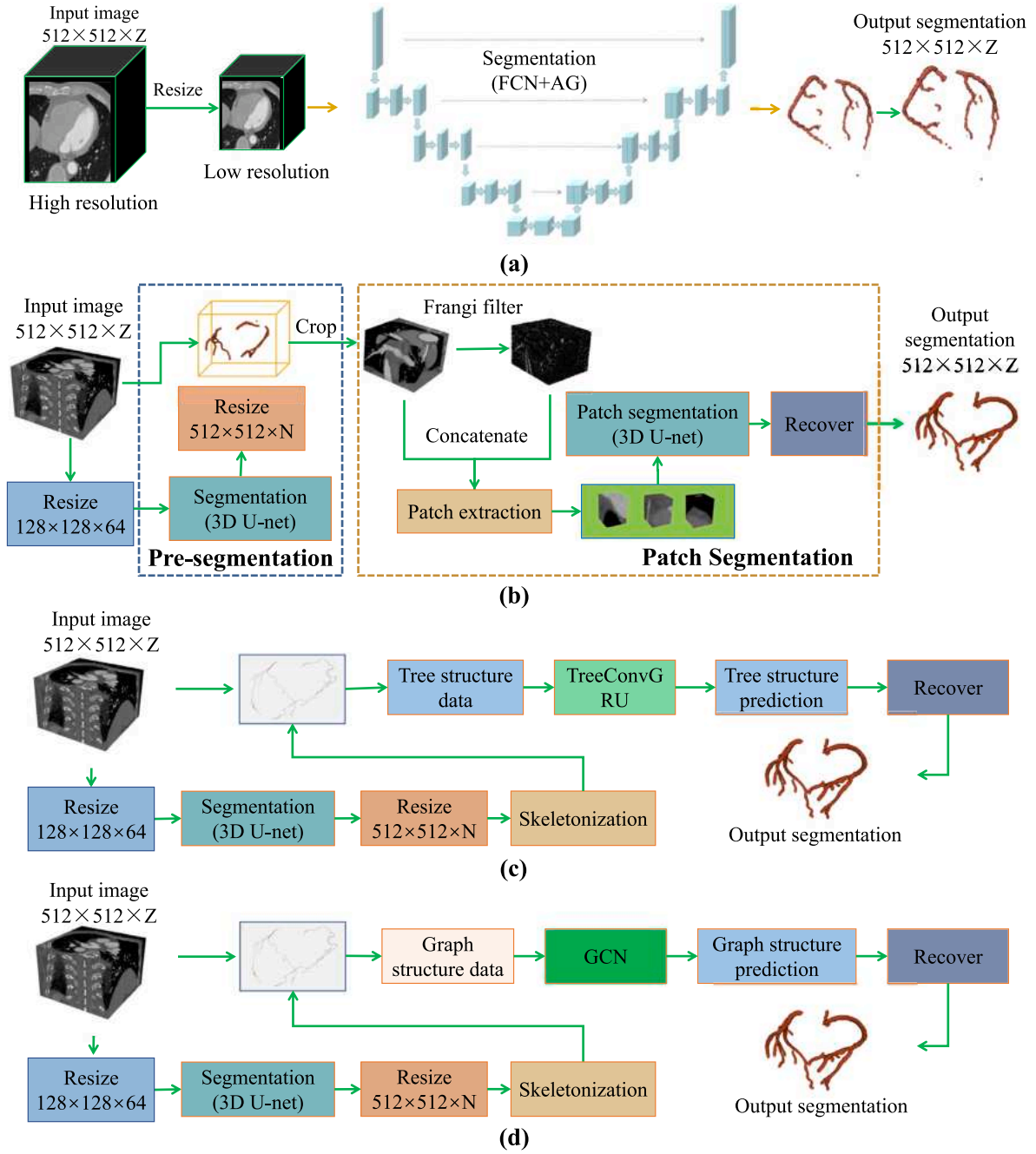
Fig. 2. Overview of the methods in the proposed benchmark including (a) direct segmentation (Shen et al., 2019), (b) patch based segmentation (Huang et al., 2018; Chen et al., 2019), (c) tree data based segmentation (Kong et al., 2020), and (d) graph data based segmentation (Wolterink et al., 2019).

we have implemented several typical deep learning based methods including direct segmentation (Shen et al., 2019), patch based segmentation (Huang et al., 2018; Chen et al., 2019), tree data based segmentation (Kong et al., 2020), graph data based segmentation (Wolterink et al., 2019). In addition, we also propose a baseline method which achieves better performance than existing works on the proposed large-scale dataset. We describe the details of each method in the following sections and more about the their implementations can be found at https://github.com/XiaoweiXu/ImageCAS-A-Large-Scale-Dataset-and-Benchmark-for-Coronary-Artery-Segmentation-based-on-CT.

4.1. Direct segmentation

Direct segmentation only adopts a single neural network, which is a simple but efficient for coronary artery segmentation. Particularly,

the image is fed to the network which outputs the probability map without complex processing. As the images are usually large and thus hard to be accommodated by GPUs, they need to be downsized to a smaller resolution in practice. We select FCN-AG (Shen et al., 2019) as a representative. The details of the method are shown in Fig. 2(A), and the overall structure is an FCN backbone with attention gate modules. The process can be divided into the following three steps:

(1) resize high resolution CTA images into low resolution ones using interpolation;
(2) feed the low resolution image to the FCN-AG network, and then obtain the prediction of coronary artery.
(3) resize low resolution prediction labels into the high resolution ones using interpolation.

For more details of the method, please refer to the original work (Shen et al., 2019).

4.2. Patch based segmentation

To tackle the computational resource constraints and missing details due to downsampling, patch based segmentation (Huang et al., 2018; Chen et al., 2019) has been proposed. We consider the dual-CNN-based framework (Huang et al., 2018; Chen et al., 2019) as an example to evaluate this class of methods (Fig. 2(b)). First, a 3D U-net is used to extract the region of interest (RoI) and remove irrelevant areas. Then, the cutout image is resized from $512 \times 512 \times Z$ to $W \times H \times Z_c$ ($W \leq 512, H \leq 512, Z \leq Z_c$). According to the (Chen et al., 2019), we use Frangi filtering to enhance tubular structures, and the obtained vascular enhancement map is combined with the input image to form a multi-channel input. Next, the multi-channel input is decomposed into small patches which are fed into another CNN for processing. Finally, the segmentation of the patches are combined to obtain the final segmentation result.

4.3. Tree data based segmentation

Tree data based segmentation method is promising as it considers the morphological structure of the coronary arteries. We select the work presented in Kong et al. (2020) to implement a tree convolutional recurrent neural network and details are shown in Fig. 2(c). First, the approximate centerline is obtained by skeletonizing the pre-segmented labels. To simplify the steps of constructing the tree, the point with the largest coordinate in the Z-axis is taken as the root node, and the rest are regarded as the leaf nodes. Then, each centerline point is taken as the center to get the patch which is then used to extract node features. Next, the connections in the tree are constructed based on the neighborhood relationship between centerline points. Finally, the tree structure data is used as the input to TreeConvGRU (Kong et al., 2020) and the predicted labels are obtained as output.

4.4. Graph based segmentation

Graph based segmentation works in a similar way as tree data based segmentation. Referring to the idea in Wolterink et al. (2019), we have designed a similar scheme as shown in Fig. 2(d). First, the centerline is obtained by the same processing pipeline as that in tree data based segmentation. Then, graph structured data is produced. Particularly, a number of rays perpendicular to the tangent of each centerline point are issued while the angle formed by the two tangents between adjacent rays remains the same. A d-dimensional feature is formed from the starting point of the ray in the direction of the tangent line at some step size, depending on the block of voxels through which the ray passes. Each ray intersects the edge of the vessel, and the Euclidean distance between the point of intersection and the centerline point is obtained as its radius. After predicting the radii of all the centerline point, the coronary arteries are reconstructed and we can get the final segmentation.

4.5. Baseline method

The proposed baseline method is a combination of patch segmentation and coarse segmentation as shown in Fig. 3. Such combination is a trade-off between performance and implementation feasibility. On one hand, the direct segmentation of the whole 3D image in the original resolution is not feasible due to the large memory consumption, while the segmentation of the resized image (defined as coarse segmentation) is feasible but leading to a limited performance. On the other hand, patch based segmentation can provide more details but sometimes also has obvious errors due to the loss of global contextual information.

Our baseline method has two main modules: patch segmentation and coarse segmentation. Before being processed by the two modules, each input image is resized from $512 \times 512 \times (206 - 275)$ to $128 \times 128 \times 64$ using linear interpolation. In coarse segmentation, a 3D U-net (Çiçek et al., 2016) is used to obtain a rough segmentation of coronary artery which can usually capture the global structure, e.g., all the related vessels are included with inaccurate boundaries. For the ease of discussion, we denote $\theta(X)$ as the segmentation network, where X is the input image, and $\hat{Y}_c$ is the output. Then, the network is trained with the following similarity coefficient loss:

$$L_c(\theta(X), Y) = 1 - \frac{2 \mid \hat{Y}_c \cap Y \mid + s}{\mid \hat{Y}_c \mid + \mid Y \mid + s}, \tag{1}$$

where Y is the ground truth and s is the smoothing factor.

In patch segmentation, the main problem is how to precisely obtain the relevant patches which contains coronary arteries. If we simply use sliding window to generate patches, all the patches including the ones that contain other small vessels or vessel-like anatomies will be processed. In this way, the segmentation network needs to consider both the target regions and the non-target regions, which introduces much difficulty into the segmentation task. To solve this problem, we adopt another network to roughly obtain the overall target region. Note that the result in the coarse segmentation step can be adopted for such purpose. However, in the prior experiments we find that this can lead to a number of problems, such as interrupted vessels. Thus, we introduce another sub-step, *dilated vessel segmentation* to extract a rough mask of the coronary artery using 3D U-net. Then, the output is dilated, and the vessels become much thicker and less likely to be interrupted. To make the vessels even thicker, the following weighted similarity coefficient loss function (Sudre et al., 2017) is used:

$$L_m(\phi(X), Y) = 1 - \frac{\mid \hat{Y}_d \cap Y \mid + s}{\alpha \mid \hat{Y}_d \mid + (1 - \alpha) \mid Y \mid + s}, \tag{2}$$

where $\phi(X)$ is the segmentation network, $\hat{Y}_d$ is the output, s is the smoothing factor, and α is the class weight ($\alpha \in (0, 1)$). In order to make the segmentation results to cover most of target regions, α is set to 0.01. Thus, the network is biased to obtain oversized vessels in the output. In the network training, $\hat{Y}_d$ is obtained by dilating the coronary artery in the ground truth Y. To obtain the precise position, $\hat{Y}_d$ is further resized to the original size of the input image using interpolation. Note that during the test stage, the results from the dilated vessel segmentation network are further dilated to insure better connectivity. The skeleton of the vessels is extracted using a surface thinning algorithm (Lee et al., 1994).

With the resized skeleton and the input image, we perform patch segmentation as follows: (1) Extracting the two largest connected components through connected component analysis and discarded others. This is due to the domain knowledge that there are usually two vessels (the left and right coronary arteries); (2) Extracting nc sets of cubic patches with a skeleton point as the center and an edge length of r ($nc = 3$, and $r = 16, 32, 64$ in Fig. 3); (3) 3D U-net++ (Zhou et al., 2018) is adopted to process nc sets of patches, as it handles details more accurately than 3D U-net; (4) The segmented patches are then fused to obtain a segmentation image of the same size as the original input image; (5) Finally, the nc segmentation images are fed to the ensemble step to get the final output. Note that majority voting is adopted here for the ensemble.

5. Experiments and discussion

In this section, we first discuss the overall setup for all experiments. Then, each method with specific configurations and their performances are discussed. Finally, all the methods with the optimal configurations are compared and analyzed.
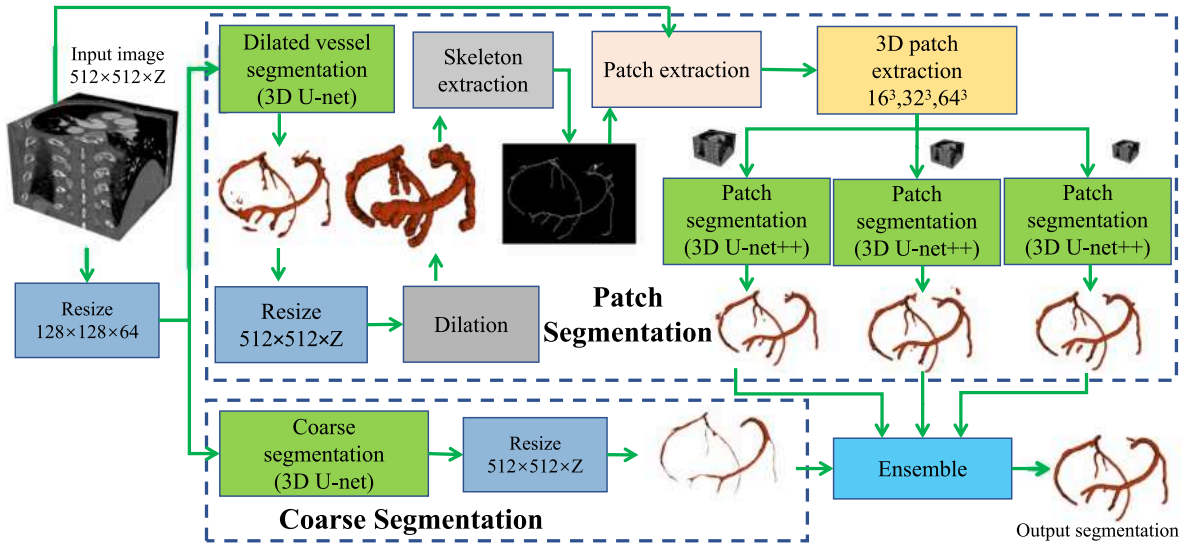
Fig. 3. Framework of the proposed baseline method.

5.1. Experiment setup

All experiments were implemented using PyTorch (Paszke et al., 2019) and DGL (Wang et al., 2019), and performed on a Nvidia RTX 3090 GPU with a 24G memory. In the direct segmentation, patch segmentation and tree-structure segmentation, we used the Dice loss during training. In the graph based segmentation, distance values are cubed in the loss function as in Wolterink et al. (2019), and we adopted the loss functions as shown in Eq. (1) and Eq. (2) for training. In patch based segmentation, the patches with 16,32,64 are discussed, and the Dice loss is used for training. In the baseline method, a spherical structure with a radius of $r = 5$ is used to dilate the vessels. All the networks in our implementation are trained for 30 epochs (about 21,000 iterations), and the Adam optimizer is adopted with a learning rate of 0.002. Due to the limited GPU memory, the batch size various for different input sizes. The batch size is 8, 2, and 1 for an input size of $128 \times 128 \times 128$, $256 \times 256 \times 128$, and $512 \times 512 \times 256$, respectively in the pre-segmentation or coarse segmentation steps. While that in the patch segmentation step is 512, 64, 8 for an input size of 16^3, 32^3, and 64^3, respectively. Experiments were evaluated using a 4-fold cross-validation approach, with a training set of 750 cases (50 cases are used for validation) and a test set of 250 cases. The Dice score is used for evaluation, which is widely used in the community as indicated in Table 1 and Table 2.

5.2. Configuration discussion

5.2.1. Direct segmentation

We discussed some factors that may affect the segmentation performance including the input size, the use of attention mechanism and the number of channels. Input size including $128 \times 128 \times 128$, $256 \times 256 \times 128$, and $512 \times 512 \times 256$ are discussed which are obtained by nearest neighbor interpolation of the original image. The number of channels including 4 and 12 are discussed.

The results are shown in Fig. 4. We can notice that the input size of $512 \times 512 \times 256$ significantly improves the Dice score by 7.38% ($p < 0.0001$) and 12.32% ($p < 0.0001$) compared to that of $256 \times 256 \times 128$ and $128 \times 128 \times 128$, respectively. The addition of the attention gate module significantly improves the performance by 1.34% ($p < 0.0001$), which is also observed by the work (Shen et al., 2019). The number of channel of 12 (parameter number of 5.24M) obtains a significantly better Dice score than that of 4 (parameter number of 0.59M) by 2.13% ($p < 0.0001$) in Dice score.

Visual discussion is shown in Fig. 5. Due to the low contrast around the coronary region in the CTA images, direct segmentation always focus on the whole coronary artery while ignoring the local details. We can notice that there is little difference in voxel intensity between the vascular portion and the rest of the adjacent tissue, and the direct segmentation method cannot correctly identify this portion of coronary artery resulting in segmentation errors. In addition, directly segmentation with a high resolution input needs more computational resources and limits the network size and the model capability.

5.2.2. Patch based segmentation

In the implementation, the cropping method in Chen et al. (2019) was adopted, where real labels are used in the training set for skeletonization. The corresponding patches are cropped at the center of the skeleton points, and the cropped regions are randomly selected so that the ratio of regions with labels to those without is 1:1. We explored three factors including patch size, frangi channel and data augmentation that may affect the performance. For data augmentation, the probability of rotation (randomly 0, 90, 180, 270 degrees) and horizontal flipping are discussed. For ease of discussion, the two probabilities are set to the same value, and three values including 0, 0.2, and 0.5 are discussed.

Quantitative performance is shown in Fig. 6. There is no significant difference between the network with the Frangi channel and that without, with a difference of only 0.01% (p>0.05) in the Dice score. For patch size, we can notice that a larger patch size obtains a significantly higher Dice score than a smaller one, which is expected as a larger patch size has a larger receptive field and thus can capture better context information. For data augmentation, we find that flipping and rotation probabilities of 0/0 (no flipping and rotation) obtains an improvement of 2.63% ($p < 0.0001$) and 2.73% ($p < 0.0001$) in Dice scores than that of 0.2/0.2 and 0.5/0.5. This interesting phenomenon may due to the fact that coronary arteries have its own directions with corresponding surround anatomies, and rotation and flipping operation may produce unrealistic augmented samples which may harm the training process.

Visual discussion of failed cropped regions of interest (RoI) in patch based segmentation is shown in Fig. 7. We can see that several similar vessels are recognized as coronary arteries which may caused by the fact that their appearance are quite similar to the coronary arteries. To tackle this problem, we can adopt a connectivity domain analysis in the post-processing step, and we find that for most of the predicted images, this refinement is effective in removing similar regions. However, this
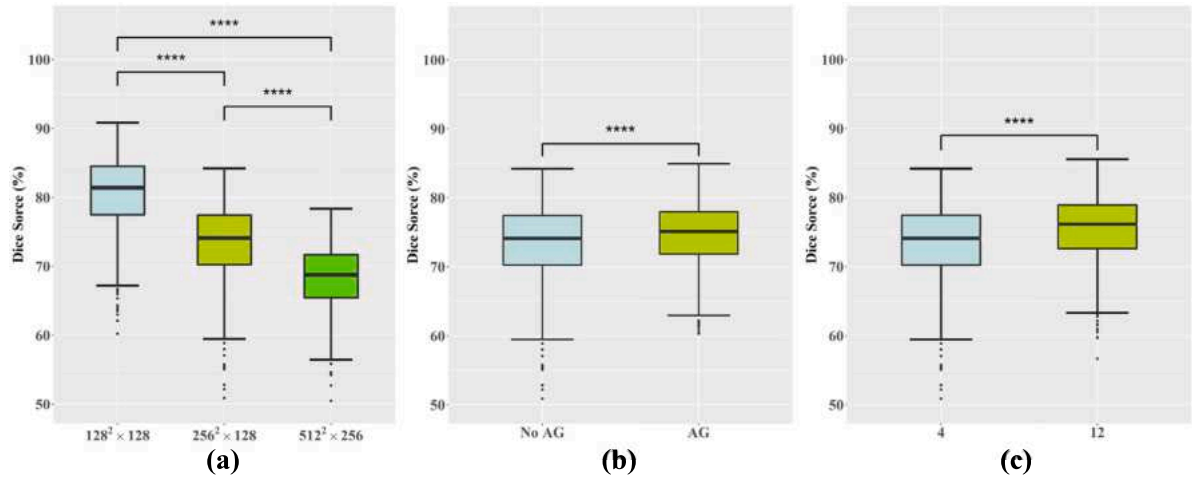
Fig. 4. Performance discussion of direct segmentation with various configurations including (a) input size, (b) attention gate module, and (c) the number of channels. ns stands for not significant ($p>0.05$), and ** stands for p smaller than 0.01, *** stands for p smaller than 0.001, and **** stands for p smaller than 0.0001.
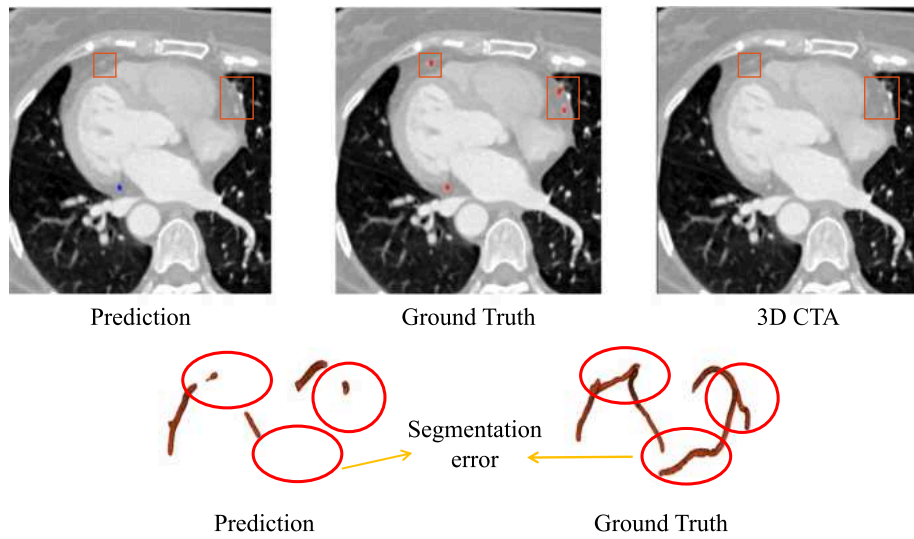

Prediction   Ground Truth   3D CTA

Segmentation error

Prediction   Ground Truth

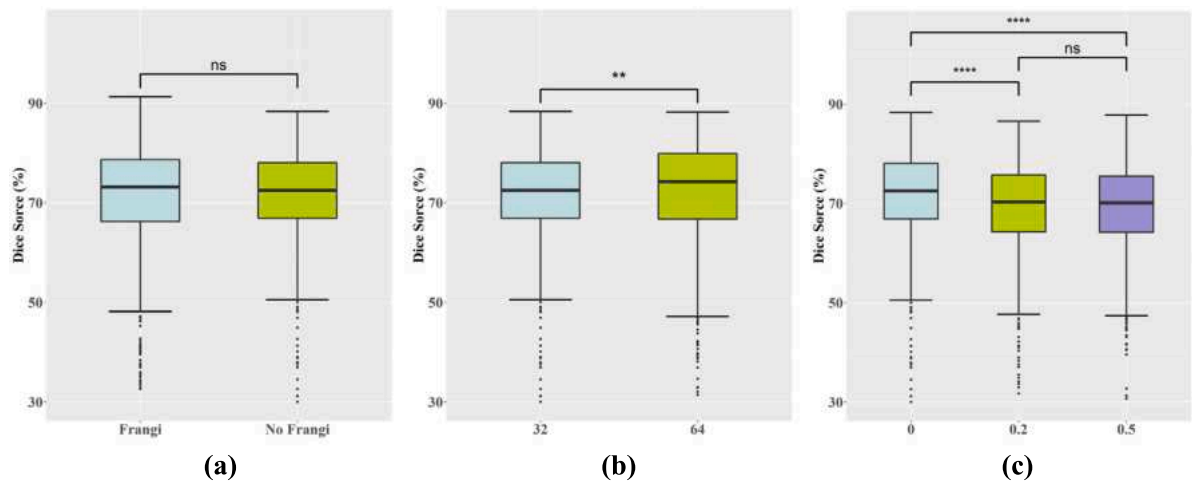**Fig. 5.** Visual discussion of direct segmentation results with failure cases.

Fig. 6. Performance discussion of patch based segmentation with various configurations including (a) Frangi channel, (b) patch size, and (c) data augmentation. ns stands for not significant ($p>0.05$), and ** stands for p smaller than 0.01, *** stands for p smaller than 0.001, and **** stands for p smaller than 0.0001.

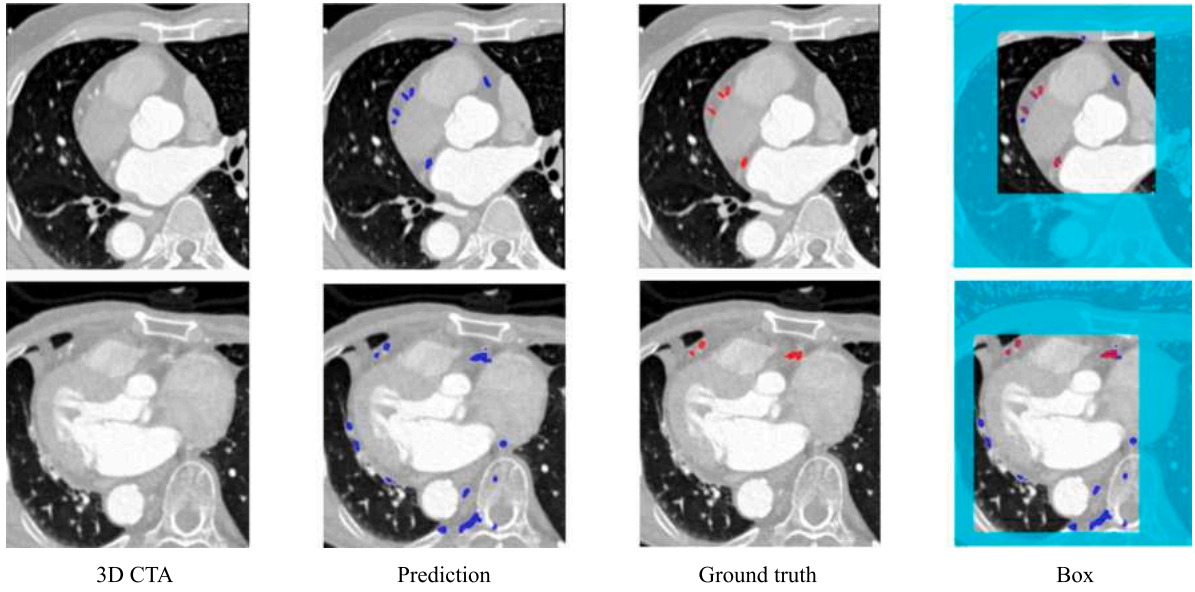|			
3D CTA	Prediction	Ground truth	Box

Fig. 7. Visual discussion of failed cropped RoI in patch based segmentation including cross-sectional view on the input images, predicted labels, ground truths and pre-segmented formed bounding boxes.
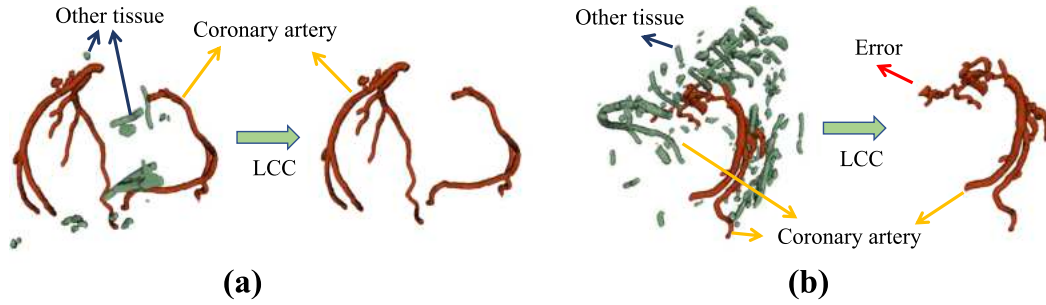

Fig. 8. Visual discussion of (a) successful and (b) failed cases of largest connected component (LCC) post-processing in patch based segmentation.
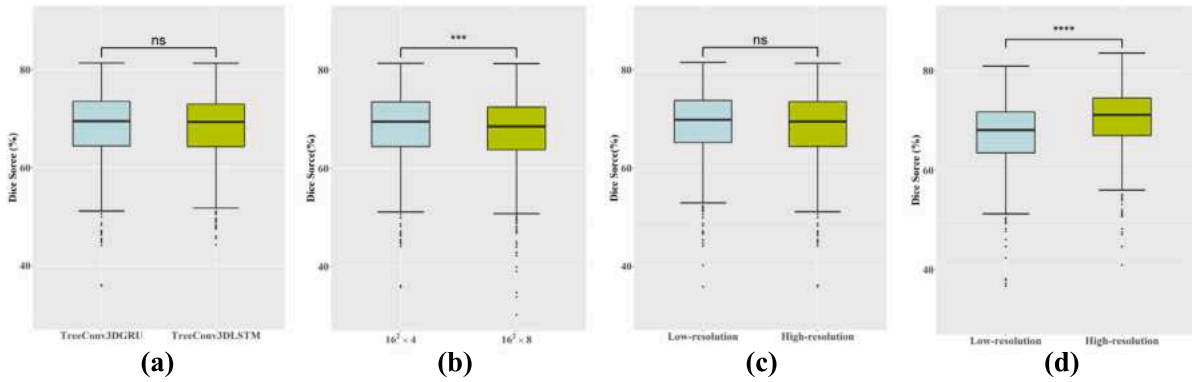

Fig. 9. Performance discussion of tree data based segmentation and graph based segmentation with various configurations including (a) tree models, (b) patch size in tree node construction, input size in (c) tree data based segmentation and (d) graph based segmentation. ns stands for not significant ($p>0.05$), and ** stands for p smaller than 0.01, *** stands for p smaller than 0.001, and **** stands for p smaller than 0.0001.

does not work well for some images, generating large number of small vessels and resulting in the partial removal of the coronary arteries as shown in Fig. 8. In Fig. 8(a), the border is well removed from the tissue around the heart, while in Fig. 8(b) the pre-segmentation holds other structures such as bone tissue in the output.

5.2.3. Tree data based segmentation and graph based segmentation

As tree data based segmentation and graph based segmentation are similar and share the same pre-segmentation module, we put the two together for ease of discussion. In both the two methods, the input size including $128 \times 128 \times 128$ and $512 \times 512 \times 256$ are discussed. In the tree structure segmentation, we adopted two models including TreeConv3FGRU and TreeConv3DLSTM and the patch size including $16 \times 16 \times 4$ and $16 \times 16 \times 8$ during tree node construction for discussion.

Segmentation performance is shown in Fig. 9. We can discover that the two tree models TreeConvGRU and TreeConvLSTM has only
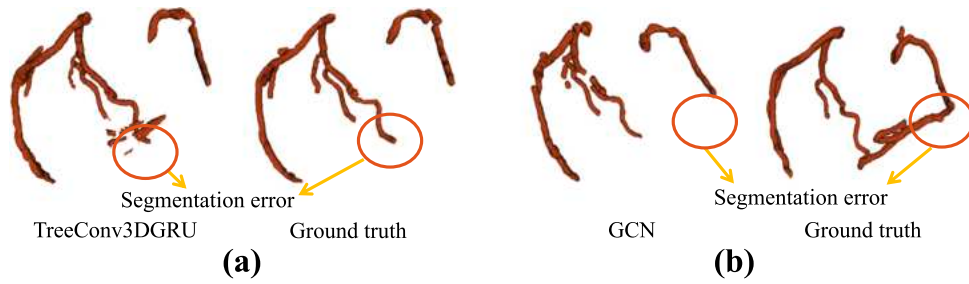
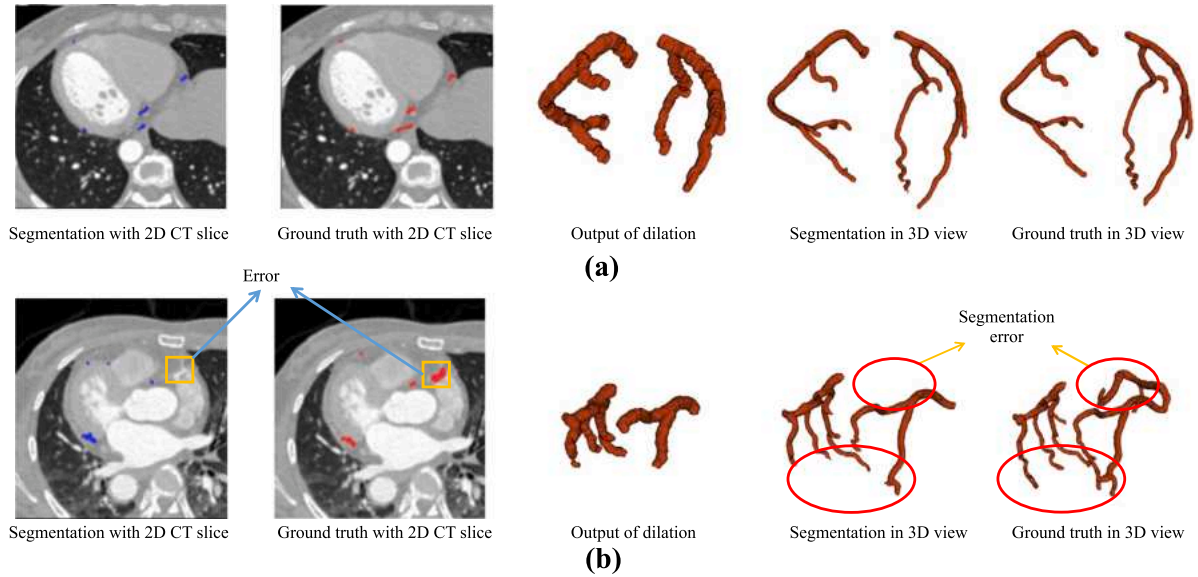**Fig. 10.** Visual discussion of failed cases in tree and graph based segmentation.

Fig. 11. Visual discussion of (a) good and (b) poor segmentation cases in the baseline method.

Table 3
Dice score (in %) of the baseline method. Intermediate results including those from the coarse segmentation step and the branches of the patch segmentation are also included.

Module	Dilation	Patch size	Dice score
Patch segmentation	w/o	16^3	79.56
Patch segmentation	w/o	32^3	81.22
Patch segmentation	w/o	64^3	82.34
Ensemble (Patch)	w/o	$16^3,32^3,64^3$	80.97
Ensemble (Coarse+Patch)	w/o	$16^3,32^3,64^3$	82.21
Patch segmentation	w/	16^3	77.80
Patch segmentation	w/	32^3	82.27
Patch segmentation	w/	64^3	82.70
Ensemble (Patch)	w/	$16^3,32^3,64^3$	81.11
Ensemble (Coarse+patch)	w/	$16^3,32^3,64^3$	82.96

a difference of 0.06% (p>0.05) in Dice score thus with no significant difference. In tree data based segmentation, the patch size of $16 \times 16 \times 4$ obtains a higher Dice score of 1.38% ($p < 0.005$) than that of $16 \times 16 \times 8$, which indicates larger patch size may not always benefit the final performance. Compared Fig. 9(c) and (d), we can find an interesting phenomenon that hat the input size has a very different effect on the performance. For tree data based segmentation, the difference between the performance of the implementation with the high-resolution and that with the low-resolution input is only 0.12% (p>0.05). While for graph based segmentation, the implementation with the high-resolution input outperforms that with the low-resolution inputs by 2.95% ($p < 0.0001$) in Dice score.

We also discuss the failed cases in pre-segmentation which is a critical step for both tree data based segmentation and graph based segmentation as shown in Fig. 10. Note that the two methods are both

centerline-based solutions, relying on the centerline to construct data and segment the vessels. In practical terms, we lack the labels of the true centerline to support the training of a centerline model. Here, we adopt the segmentation of the network which are then labeled and skeletonized to make an approximate centerline to implement both methods. Therefore, the quality of the pre-segmentation becomes a key factor in the accuracy of the subsequent segmentation. As shown in Fig. 10, some coronary arteries are missing in pre-segmentation, and as a result the missing vessels will not be constructed as a tree structure or graph structure. Finally, these coronary arteries are missing in the rest of the processing, resulting in errors in the final segmentation.

5.2.4. Baseline method

We discussed several factors including modules (patch segmentation and coarse segmentation), ensemble, dilation (w/ and w/o), and patch size (16^3, 32^3, and 64^3). The performance is shown in Table 3. For the patch segmentation module, Dice scores of 79.56%, 81.22% and 82.34% are obtained without dilation for patch sizes of 16^3, 32^3 and 64^3, respectively, and their paired comparisons (16^3 vs 32^3 ($p < 0.0001$), 16^3 vs 64^3 ($p < 0.0001$), and 32^3 vs 64^3 ($p < 0.001$)) show statistical significance. We can notice that larger patch sizes indicate larger receptive field, which benefits the segmentation. While with dilation, larger patch sizes still obtains higher Dice score. However, there is no statistical significance (p>0.05) between a patch size of 32^3 and that of 64^3. The main reason is that dilation is also effective to extract the context information. The combination of the patch size of 32^3 and dilation is powerful enough to extract the context information, and larger receptive field with a larger patch size and dilation cannot further extract more context information. We can also note that dilation
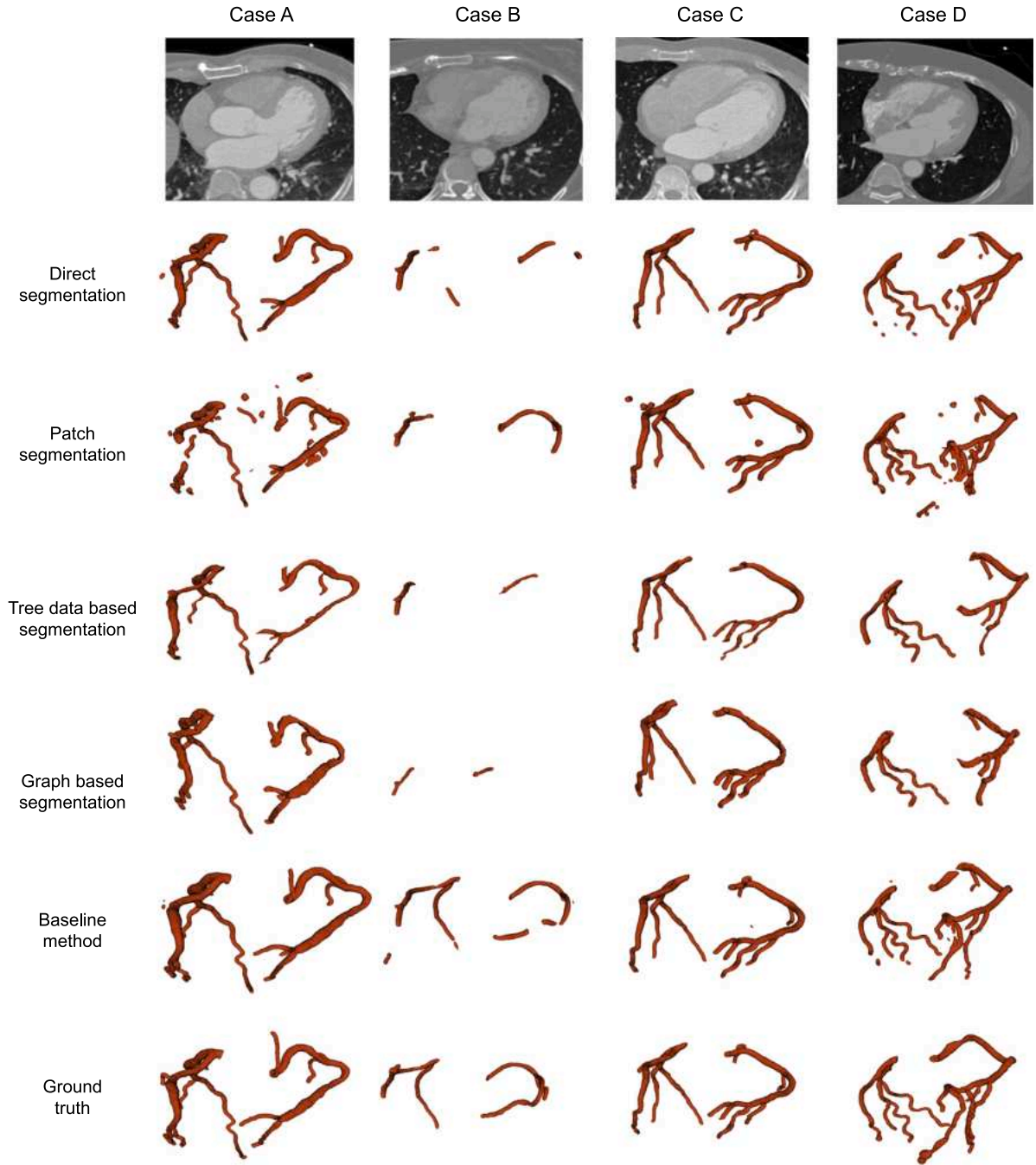
Fig. 12. Visual comparison of the results of the five methods in the benchmark with four cases.

can improve the performance with the patch size of 64^3 and 32^3 but not for 16^3, which may due to the fact that dilation will dismiss some pixels which is with a small quantity in the input with a small patch size. In terms of ensemble, the ensemble using the coarse segmentation module with the patch segmentation module (82.96%) outperforms that of each base classifier (77.80%, 82.27%, 82.70%) in terms of Dice score.

Visual illustration of good and poor segmentation cases is shown in Fig. 11. As shown in Fig. 11(a), the segmentation result is good which can well match the ground truth in both the 2D CT slices and the 3D view. We can also notice that the output of dilation covers all the areas of the ground truth. Fig. 11(b) shows an example of poor segmentation. We can notice that a thick vessel and a long thin vessel are missing which is due to the fact that they are not recognized by the dilated vessel segmentation module. Particularly, the thick vessel as shown in

the 2D CT slice is close to the right atrium and has a similar gray scale value to the right atrium. As a result, the vessel is relatively hard to be recognized correctly.

5.3. Benchmark comparison

Performance comparison of all the methods in the benchmark is shown in Table 4. We can notice that the proposed baseline method achieves the optimal performance on all metrics. We can also discover an interesting phenomenon that patch segmentation, tree data based segmentation and graph based segmentation has a much lower performance than direct segmentation. We can make a rough discussion here. Such phenomenon partially shows the same trend in existing works in which direct segmentation (89 in Cheung et al. (2021)) has

Table 4
Performance comparison of the methods in the benchmark in Dice score (%). Each method is with the optimal result which is selected from implementations with various configurations.

Method	Input type	Input size	Dice score (%)	HD (mm)	AHD (mm)
Direct segmentation (3D FCN) (Shen et al., 2019)	Full image	$512 \times 512 \times 256$	80.58	28.6656	0.8503
Patch segmentation (3D U-net) (Huang et al., 2018; Chen et al., 2019)	Patch	$64 \times 64 \times 64$	72.01	40.9693	3.0686
Tree data based segmentation (3D TreeConvGRU) (Kong et al., 2020)	Tree	$N \times 16 \times 16 \times 4$	68.78	30.3350	1.4336
Graph based segmentation (GCN) (Wolterink et al., 2019)	Graph	$N \times 32$	70.61	27.8718	1.2433
Baseline method (3D U-net and U-net++)	Patch	$128 \times 128 \times 64, 16^3, 32^3, 64^3$	82.96	27.2169	0.8180

higher Dice score than patch segmentation (94 in Pan et al. (2021)), tree data based segmentation (85 in Kong et al. (2020)), and Graph based segmentation (73–75 in Wolterink et al. (2019)). But we can still notice some inconsistencies that patch segmentation has a higher performance than direct segmentation in existing works. There are mainly two reasons. First, the dataset for evaluation is not the same, and the quality of the datasets also varies. Note that our dataset is the largest one currently, which is orders of magnitude larger than most of the datasets used by existing methods. Second, there are much details including hyper parameters, pre-processing and post-processing in the existing works which is relatively hard for implementation. Though we have tried our best to implement existing works, there is some details (sometimes maybe critical details) that are inevitably missed due to the lack of understanding and the limited length of the related papers. Thus, we welcome related researchers in the community to join us to improve the implementations.

Visual comparison of the methods in the proposed benchmark is shown in Fig. 12. For direct segmentation, the performance of Case A and Case C is good. However, when the quality of the images is low and the structure of the coronary arteries varies considerably as shown in Case B and D, areas with low contrast cannot be correctly detected. For patch segmentation, the results are quite similar to that of direct segmentation. The only difference is that patch segmentation can recognize some thin vessels in Case D but cannot in Case A, while direct segmentation has a opposite performance on the vessels in the two cases. Also patch segmentation has a slightly better performance than that of direct segmentation in Case B. It seems that patch segmentation can process local features better than direct segmentation patch which is expected as patch segmentation pay more attention on local feature processing. Tree data based segmentation and graph based segmentation share a rather similar, however, they cannot discover the vessels with low contrast well in Case B and Case D. The main reason is that tree data based and graph based segmentation methods rely heavily on the centerline extracted using the pre-segmentation, which determines the number of nodes in the tree and graph structure. For the proposed baseline method, the results show much better performance especially on Case B where the majority of the vessels are correctly recognized. The main reason is that the baseline method combines both the features from coarse segmentation and patch segmentation with various patch sizes, which can better extract the context information.

5.4. Discussion

We have implemented several typical deep learning based methods including direct segmentation (Shen et al., 2019), patch based segmentation (Huang et al., 2018; Chen et al., 2019), tree data based segmentation (Kong et al., 2020), graph data based segmentation (Wolterink et al., 2019), and our baseline method. The proposed baseline method achieves the optimal performance on Dice score, HD and AHD. On the other hand, we can also notice that the tree data based segmentation method achieves a better performance than direct segmentation, which is contrary to the results in Kong et al. (2020). This may be partially due to the fact that there are many technique details in the training and optimization in Kong et al. (2020). This is also the reason that we published our dataset and code for fair comparison.

Though our dataset is quite large compared with exiting works and well labeled by two to three experienced radiologists, it has limitations. First, our dataset was collected in only one center thus with biased distributions. Second, only one kind of CT machines namely Siemens 128-slice dual-source scanner was used to acquire the CT images, which make the biased problem even worse. Third, no detailed labels is provided. For example, subclasses of coronary artery including left main coronary artery, left anterior descending coronary artery, etc. are not separated. We hope that others can also publish their datasets to mitigate the above limitations and at the same time facilitate related research.

Future directions of our benchmark can be various, and we only name a few here. First, more advanced segmentation networks like nnU-net (Isensee et al., 2021), CoTr (Xie et al., 2021), and UN-ETR (Hatamizadeh et al., 2022) can be employed in existing frameworks to enhance performance. Second, more advanced networks and evaluation metrics can be explored to preserve the topology or connectivity (Shit et al., 2021; Hu et al., 2019, 2021; Saeki et al., 2021) of the coronary vessels which plays a critical role for further analysis and diagnosis. Third, based on our dataset and existing ones (Schaap et al., 2009a; Kirişli et al., 2013), multi-center related topics including federal learning (Rajasree et al., 2022) and domain adaptation (Guan and Liu, 2021) can be further investigated.

6. Conclusion

Segmentation of coronary arteries is a critical task for the diagnosis and quantification of coronary artery disease. In this paper, we propose a benchmark dataset for coronary artery segmentation on CTA images. In addition, we have implemented a benchmark in which we have not only tried our best to implement several typical existing methods but also proposed a strong baseline method. We have performed a comprehensive evaluation of the methods in the benchmark, and the results show that the proposed baseline method achieves the optimal performance with a Dice score of 82.96%. However, the performance still has room of improvement for accurate diagnosis and stenosis quantification for clinical practice. The benchmark and the dataset is published at https://github.com/XiaoweiXu/ImageCAS-A-Large-Scale-Dataset-and-Benchmark-for-Coronary-Artery-Segmentation-based-on-CT. We hope that the proposed dataset and benchmark can stimulate further research in the community.

CRediT authorship contribution statement

An Zeng: Conceptualization, Methodology, Software. **Chunbiao Wu:** Data collection, Writing – original draft. **Guisen Lin:** Conceptualization, Writing – reviewing & editing. **Wen Xie:** Writing – reviewing & editing. **Jin Hong:** Writing – reviewing & editing. **Meiping Huang:** Data collection, Investigation. **Jian Zhuang:** Data collection. **Shanshan Bi:** Data collection, Validation. **Dan Pan:** Data collection, Writing – reviewing & editing. **Najeeb Ullah:** Writing – reviewing & editing. **Kaleem Nawaz Khan:** Writing – reviewing & editing. **Tianchen Wang:** Data collection, Writing – reviewing & editing. **Yiyu Shi:** Writing – reviewing & editing. **Xiaomeng Li:** Conceptualization, Writing – reviewing & editing. **Xiaowei Xu:** Conceptualization, Writing – reviewing and editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We have published the dataset with a link.

Ethical and information governance approvals

This work and the collection of data of retrospective data on implied consent received Research Ethics Committee (REC) approval from Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences under Protocol No. 2019324H. It complies with all relevant ethical regulations. Deindentification was performed in which all CT files are transformed into NIfTI format, and sensitive information of the patients including name, birther day, admission year, admission number, and CT number is removed.

References

Altunbay, D., Cigir, C., Sokmensuer, C., Gunduz-Demir, C., 2010. Color graphs for automated cancer diagnosis and grading. IEEE Trans. Biomed. Eng. 57 (3), 665–674.

Aylward, S.R., Bullitt, E., 2002. Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. IEEE Trans. Med. Imaging 21 (2), 61–75.

Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39 (12), 2481–2495.

Broersen, A., Kitslaar, P., Frenay, M., Dijkstra, J., 2012. FrenchCoast: fast, robust extraction for the nice challenge on coronary artery segmentation of the tree. In: Proc. of MICCAI Workshop" 3D Cardiovascular Imaging: A MICCAI Segmentation Challenge.

Cetin, S., Unal, G., 2015. A higher-order tensor vessel tractography for segmentation of vascular structures. IEEE Trans. Med. Imaging 34 (10), 2172–2185.

Chen, H., Dou, Q., Yu, L., Qin, J., Heng, P.-A., 2018a. VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images. NeuroImage 170, 446–455.

Chen, F., Li, Y., Tian, T., Cao, F., Liang, J., 2018b. Automatic coronary artery lumen segmentation in computed tomography angiography using paired multi-scale 3D CNN. In: Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging, Vol. 10578. International Society for Optics and Photonics, p. 105782R.

Chen, Y.-C., Lin, Y.-C., Wang, C.-P., Lee, C.-Y., Lee, W.-J., Wang, T.-D., Chen, C.-M., 2019. Coronary artery segmentation in cardiac CT angiography using 3D multi-channel U-net. In: Medical Imaging with Deep Learning 2019.

Cheung, W.K., Bell, R., Nair, A., Menezes, L.J., Patel, R., Wan, S., Chou, K., Chen, J., Torii, R., Davies, R.H., et al., 2021. A computationally efficient approach to segmentation of the aorta and coronary arteries using deep learning. IEEE Access 9, 108873–108888.

Chi, Y., Huang, W., Zhou, J., Zhong, L., Tan, S.Y., Felix, K.Y.J., Sheon, L.C.S., San Tan, R., 2015. A composite of features for learning-based coronary artery segmentation on cardiac CT angiography. In: International Workshop on Machine Learning in Medical Imaging. Springer, pp. 271–279.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 424–432.

Collet, C., Onuma, Y., Andreini, D., Sonck, J., Pompilio, G., Mushtaq, S., La Meir, M., Miyazaki, Y., de Mey, J., Gaemperli, O., et al., 2018. Coronary computed tomography angiography for heart team decision-making in multivessel coronary artery disease. Eur. Heart J. 39 (41), 3689–3698.

Cooper, R., Cutler, J., Desvigne-Nickens, P., Fortmann, S.P., Friedman, L., Havlik, R., Hogelin, G., Marler, J., McGovern, P., Morosco, G., et al., 2000. Trends and disparities in coronary heart disease, stroke, and other cardiovascular diseases in the United States: findings of the national conference on cardiovascular disease prevention. Circulation 102 (25), 3137–3147.

Diakogiannis, F.I., Waldner, F., Caccetta, P., Wu, C., 2020. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS J. Photogramm. Remote Sens. 162, 94–114.

Douc, R., Moulines, E., Olsson, J., 2009. Optimality of the auxiliary particle filter. Probab. Math. Statist. 29 (1), 1–28.

Doyle, S., Madabhushi, A., Feldman, M., Tomaszeweski, J., 2006. A boosting cascade for automated detection of prostate cancer from digitized histology. MICCAI 504–511.

Du, H., Shao, K., Bao, F., Zhang, Y., Gao, C., Wu, W., Zhang, C., 2021. Automated coronary artery tree segmentation in coronary CTA using a multiobjective clustering and toroidal model-guided tracking method. Comput. Methods Programs Biomed. 199, 105908.

Duan, J., Bello, G., Schlemper, J., Bai, W., Dawes, T.J., Biffi, C., de Marvao, A., Doumoud, G., O'Regan, D.P., Rueckert, D., 2019. Automatic 3D bi-ventricular segmentation of cardiac images by a shape-refined multi-task deep learning approach. IEEE Trans. Med. Imaging 38 (9), 2151–2164.

Duan, Y., Feng, J., Lu, J., Zhou, J., 2018. Context aware 3D fully convolutional networks for coronary artery segmentation. In: International Workshop on Statistical Atlases and Computational Models of the Heart. Springer, pp. 85–93.

Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A., 1998. Multiscale vessel enhancement filtering. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 130–137.

Freiman, M., Nickisch, H., Prevrhal, S., Schmitt, H., Vembar, M., Maurovich-Horvat, P., Donnelly, P., Goshen, L., 2017. Improving CCTA-based lesions' hemodynamic significance assessment by accounting for partial volume modeling in automatic coronary lumen segmentation. Med. Phys. 44 (3), 1040–1049.

Friman, O., Hindennach, M., Kühnel, C., Peitgen, H.-O., 2010. Multiple hypothesis template tracking of small 3D vessel structures. Med. Image Anal. 14 (2), 160–171.

Fu, Y., Guo, B., Lei, Y., Wang, T., Liu, T., Curran, W., Zhang, L., Yang, X., 2020. Mask R-CNN based coronary artery segmentation in coronary computed tomography angiography. In: Medical Imaging 2020: Computer-Aided Diagnosis, Vol. 11314. International Society for Optics and Photonics, p. 113144F.

Fu, H., Qiu, G., Shu, J., Ilyas, M., 2014. A novel polar space random field model for the detection of glandular structures. IEEE Trans. Med. Imaging 33 (3), 764–776.

Gao, Z., Liu, X., Qi, S., Wu, W., Hau, W.K., Zhang, H., 2019. Automatic segmentation of coronary tree in CT angiography images. Internat. J. Adapt. Control Signal Process. 33 (8), 1239–1247.

Glover, G., Pelc, N., 1980. Nonlinear partial volume artifacts in x-ray computed tomography. Med. Phys. 7 (3), 238–248.

Gu, L., Cai, X.-C., 2021. Fusing 2D and 3D convolutional neural networks for the segmentation of aorta and coronary arteries from CT images. Artif. Intell. Med. 121, 102189.

Gu, J., Fang, Z., Gao, Y., Tian, F., 2020. Segmentation of coronary arteries images using global feature embedded network with active contour loss. Comput. Med. Imaging Graph. 86, 101799.

Guan, H., Liu, M., 2021. Domain adaptation for medical image analysis: a survey. IEEE Trans. Biomed. Eng. 69 (3), 1173–1185.

Gunduz-Demir, C., Kandemir, M., Tosun, A.B., Sokmensuer, C., 2010. Automatic segmentation of colon glands using object-graphs. Med. Image Anal. 14 (1), 1–12.

Han, D., Doan, N.-T., Shim, H., Jeon, B., Lee, H., Hong, Y., Chang, H.-J., 2014. A fast seed detection using local geometrical feature for automatic tracking of coronary arteries in CTA. Comput. Methods Programs Biomed. 117 (2), 179–188.

Han, D., Shim, H., Jeon, B., Jang, Y., Hong, Y., Jung, S., Ha, S., Chang, H.-J., 2016. Automatic coronary artery segmentation using active search for branches and seemingly disconnected vessel segments from coronary CT angiography. PLoS One 11 (8), e0156837.

Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H.R., Xu, D., 2022. Unetr: Transformers for 3d medical image segmentation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 574–584.

Hu, X., Li, F., Samaras, D., Chen, C., 2019. Topology-preserving deep image segmentation. Adv. Neural Inf. Process. Syst. 32.

Hu, X., Wang, Y., Fuxin, L., Samaras, D., Chen, C., 2021. Topology-aware segmentation using discrete morse theory. arXiv preprint arXiv:2103.09992.

Huang, C., Han, H., Yao, Q., Zhu, S., Zhou, S.K., 2019. 3D U²-Net: A 3D universal U-Net for multi-domain medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 291–299.

Huang, W., Huang, L., Lin, Z., Huang, S., Chi, Y., Zhou, J., Zhang, J., Tan, R.-S., Zhong, L., 2018. Coronary artery segmentation by deep learning neural networks on computed tomographic coronary angiographic images. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. EMBC, IEEE, pp. 608–611.

Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H., 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods 18 (2), 203–211.

Kirişli, H., Schaap, M., Metz, C., Dharampal, A., Meijboom, W.B., Papadopoulou, S.-L., Dedic, A., Nieman, K., de Graaf, M.A., Meijs, M., et al., 2013. Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography. Med. Image Anal. 17 (8), 859–876.

Kjerland, Ø., 2017. Segmentation of Coronary Arteries from CT-Scans of the Heart using Deep Learning (Master's thesis). NTNU.

Kong, B., Wang, X., Bai, J., Lu, Y., Gao, F., Cao, K., Xia, J., Song, Q., Yin, Y., 2020. Learning tree-structured representation for 3D coronary artery segmentation. Comput. Med. Imaging Graph. 80, 101688.

Kroon, D.-J., 2009. Hessian Based Frangi Vesselness Filter. MATLAB Central File Exchange, https://www.mathworks.com/matlabcentral/fileexchange/24409-hessian-based-frangi-vesselness-filter.

Lee, T.-C., Kashyap, R.L., Chu, C.-N., 1994. Building skeleton models via 3-D medial surface axis thinning algorithms. CVGIP: Graph. Models Image Process. 56 (6), 462–478.

Lee, M.C.H., Petersen, K., Pawlowski, N., Glocker, B., Schaap, M., 2019. Template transformer networks for image segmentation. In: International Conference on Medical Imaging with Deep Learning–Extended Abstract Track.

Lei, Y., Guo, B., Fu, Y., Wang, T., Liu, T., Curran, W., Zhang, L., Yang, X., 2020. Automated coronary artery segmentation in coronary computed tomography angiography (CCTA) using deep learning neural networks. In: Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications, Vol. 11318. International Society for Optics and Photonics, 1131812.

Lesage, D., Angelini, E.D., Funka-Lea, G., Bloch, I., 2016. Adaptive particle filtering for coronary artery segmentation from 3D CT angiograms. Comput. Vis. Image Underst. 151, 29–46.

Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.-W., Heng, P.-A., 2018a. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Trans. Med. Imaging 37 (12), 2663–2674.

Li, X., Dou, Q., Chen, H., Fu, C.-W., Qi, X., Belavỳ, D.L., Armbrecht, G., Felsenberg, D., Zheng, G., Heng, P.-A., 2018b. 3D multi-scale FCN with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality MR images. Med. Image Anal. 45, 41–54.

Li, Z., Zhang, X., Müller, H., Zhang, S., 2018c. Large-scale retrieval for medical image analytics: A comprehensive review. Med. Image Anal. 43, 66–84.

Liang, R., Ma, J., Ma, G., Wang, K., et al., 2021. 3D U-Net with attention and focal loss for coronary tree segmentation.

Lin, Y., Yao, H., Li, Z., Zheng, G., Li, X., 2022. Calibrating label distribution for class-imbalanced barely-supervised knee segmentation. arXiv preprint arXiv:2205.03644.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3431–3440.

Lugauer, F., Zhang, J., Zheng, Y., Hornegger, J., Kelm, B.M., 2014a. Improving accuracy in coronary lumen segmentation via explicit calcium exclusion, learning-based ray detection and surface optimization. In: Medical Imaging 2014: Image Processing, Vol. 9034. International Society for Optics and Photonics, p. 90343U.

Lugauer, F., Zheng, Y., Hornegger, J., Kelm, B.M., 2014b. Precise lumen segmentation in coronary computed tomography angiography. In: International MICCAI Workshop on Medical Computer Vision. Springer, pp. 137–147.

Milletari, F., Navab, N., Ahmadi, S.-A., 2016. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV). IEEE, pp. 565–571.

Mirunalini, P., Aravindan, C., Nambi, A.T., Poorvaja, S., Priya, V.P., 2019. Segmentation of coronary arteries from CTA axial slices using deep learning techniques. In: TENCON 2019-2019 IEEE Region 10 Conference. TENCON, IEEE, pp. 2074–2080.

Moeskops, P., Wolterink, J.M., van der Velden, B.H., Gilhuijs, K.G., Leiner, T., Viergever, M.A., Išgum, I., 2016. Deep learning for multi-task medical image segmentation in multiple modalities. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 478–486.

Mohr, B., Masood, S., Plakas, C., 2012. Accurate lumen segmentation and stenosis detection and quantification in coronary CTA. In: Proceedings of 3D Cardiovascular Imaging: A MICCAI Segmentation Challenge Workshop.

Mortazi, A., Karim, R., Rhode, K., Burt, J., Bagci, U., 2017. CardiacNET: Segmentation of left atrium and proximal pulmonary veins from MRI using multi-view CNN. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 377–385.

Nguyen, K., Sarkar, A., Jain, A.K., 2012. Structure and context in prostatic gland segmentation and classification. In: MICCAI. Springer, pp. 115–123.

Nickisch, H., Lamash, Y., Prevrhal, S., Freiman, M., Vembar, M., Goshen, L., Schmitt, H., 2015. Learning patient-specific lumped models for interactive coronary blood flow simulations. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 433–441.

Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., et al., 2018. Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.

Organization, W.H., et al., 2009. Cardiovascular diseases (cvds). http://www.who.int/mediacentre/factsheets/fs317/en/index.html.

Pan, L.-S., Li, C.-W., Su, S.-F., Tay, S.-Y., Tran, Q.-V., Chan, W.P., 2021. Coronary artery segmentation under class imbalance using a U-Net based architecture on computed tomography angiography images. Sci. Rep. 11 (1), 1–7.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al., 2019. Pytorch: An imperative style, high-performance deep learning library. Adv. Neural Inf. Process. Syst. 32.

Patravali, J., Jain, S., Chilamkurthy, S., 2017. 2D-3D fully convolutional neural networks for cardiac MR segmentation. In: International Workshop on Statistical Atlases and Computational Models of the Heart. Springer, pp. 130–139.

Rajasree, R., Gopika, G., Romero, C.A.T., et al., 2022. The role and impact of federal learning in digital healthcare: A useful survey. In: Handbook of Research on Technical, Privacy, and Security Challenges in a Modern World. IGI Global, pp. 127–147.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.

Saeki, Y., Saito, A., Cousty, J., Kenmochi, Y., Shimizu, A., 2021. Statistical modeling of pulmonary vasculatures with topological priors in CT volumes. In: Interpretability of Machine Intelligence in Medical Image Computing, and Topological Data Analysis and its Applications for Medical Data: 4th International Workshop, iMIMIC 2021, and 1st International Workshop, TDA4MedicalData 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 4. Springer, pp. 108–118.

Schaap, M., Metz, C.T., van Walsum, T., van der Giessen, A.G., Weustink, A.C., Mollet, N.R., Bauer, C., Bogunović, H., Castro, C., Deng, X., et al., 2009a. Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms. Med. Image Anal. 13 (5), 701–714.

Schaap, M., Metz, C., van Walsum, T., Niessen, W., 2009b. Rotterdam coronary artery algorithm evaluation framework.

Shahzad, R., Kirişli, H., Metz, C., Tang, H., Schaap, M., van Vliet, L., Niessen, W., van Walsum, T., 2013. Automatic segmentation, detection and quantification of coronary artery stenoses on CTA. Int. J. Cardiovasc. Imaging 29 (8), 1847–1859.

Shen, Y., Fang, Z., Gao, Y., Xiong, N., Zhong, C., Tang, X., 2019. Coronary arteries segmentation based on 3D FCN with attention gate and level set function. IEEE Access 7, 42826–42835.

Shit, S., Paetzold, J.C., Sekuboyina, A., Ezhov, I., Unger, A., Zhylka, A., Pluim, J.P., Bauer, U., Menze, B.H., 2021. clDice-a novel topology-preserving loss function for tubular structure segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 16560–16569.

Sirinukunwattana, K., Snead, D.R., Rajpoot, N.M., 2015a. A novel texture descriptor for detection of glandular structures in colon histology images. In: SPIE Medical Imaging. International Society for Optics and Photonics, p. 94200S.

Sirinukunwattana, K., Snead, D.R., Rajpoot, N.M., 2015b. A stochastic polygons model for glandular structures in colon histology images. IEEE Trans. Med. Imaging 34 (11), 2366–2378.

Skare, Ø., Bølviken, E., Holden, L., 2003. Improved sampling-importance resampling and reduced bias importance sampling. Scand. J. Stat. 30 (4), 719–737.

Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Cardoso, M.J., 2017. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 240–248.

Tabesh, A., Teverovskiy, M., Pang, H.-Y., Kumar, V.P., Verbel, D., Kotsianti, A., Saidi, O., 2007. Multifeature prostate cancer diagnosis and Gleason grading of histological images. IEEE Trans. Med. Imaging 26 (10), 1366–1378.

Tian, F., Gao, Y., Fang, Z., Gu, J., 2021. Automatic coronary artery segmentation algorithm based on deep learning and digital image processing. Appl. Intell. 51 (12), 8881–8895.

Wang, C., Moreno, R., Smedby, Ö., 2012. Vessel segmentation using implicit model-guided level sets. In: "MICCAI Workshop" 3D Cardiovascular Imaging: A MICCAI Segmentation Challenge, Nice France, 1st of October 2012.

Wang, M., Yu, L., Zheng, D., Gan, Q., Gai, Y., Ye, Z., Li, M., Zhou, J., Huang, Q., Ma, C., et al., 2019. Deep graph library: Towards efficient and scalable deep learning on graphs.

Wang, Q., Zhao, W., Yan, X., Che, H., Ye, K., Lu, Y., Li, Z., Cui, S., 2021. Geometric morphology based irrelevant vessels removal for accurate coronary artery segmentation. In: 2021 IEEE 18th International Symposium on Biomedical Imaging. ISBI, IEEE, pp. 757–760.

Wolterink, J.M., Leiner, T., Išgum, I., 2019. Graph convolutional networks for coronary artery segmentation in cardiac CT angiography. In: International Workshop on Graph Learning in Medical Imaging. Springer, pp. 62–69.

Wu, D., Wang, X., Bai, J., Xu, X., Ouyang, B., Li, Y., Zhang, H., Song, Q., Cao, K., Yin, Y., 2019. Automated anatomical labeling of coronary arteries via bidirectional tree LSTMs. Int. J. Comput. Assist. Radiol. Surg. 14 (2), 271–280.

Xie, Y., Zhang, J., Shen, C., Xia, Y., 2021. Cotr: Efficiently bridging cnn and transformer for 3d medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III 24. Springer, pp. 171–180.

Yang, H., Chen, J., Chi, Y., Xie, X., Hua, X., 2019. Discriminative coronary artery tracking via 3d cnn in cardiac ct angiography. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 468–476.

Yu, L., Cheng, J.-Z., Dou, Q., Yang, X., Chen, H., Qin, J., Heng, P.-A., 2017. Automatic 3D cardiovascular MR segmentation with densely-connected volumetric convnets. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 287–295.

Yu, L., Yang, X., Qin, J., Heng, P.-A., 2016. 3D FractalNet: dense volumetric segmentation for cardiovascular MRI volumes. In: Reconstruction, Segmentation, and Analysis of Medical Images. Springer, pp. 103–110.

Zhang, D.P., 2010. Coronary artery segmentation and motion modelling.

Zhang, S., Fu, H., Yan, Y., Zhang, Y., Wu, Q., Yang, M., Tan, M., Xu, Y., 2019. Attention guided network for retinal image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 797–805.

Zheng, Y., Loziczonek, M., Georgescu, B., Zhou, S.K., Vega-Higuera, F., Comaniciu, D., 2011. Machine learning based vesselness measurement for coronary artery segmentation in cardiac CT volumes. In: Medical Imaging 2011: Image Processing, Vol. 7962. International Society for Optics and Photonics, p. 79621K.

Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J., 2018. Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 3–11.

Zhu, H., Song, S., Xu, L., Song, A., Yang, B., 2021. Segmentation of coronary arteries images using spatio-temporal feature fusion network with combo loss. Cardiovasc. Eng. Technol. 1–12.