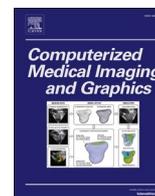


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Automated three-dimensional vessel reconstruction based on deep segmentation and bi-plane angiographic projections

D.M. Bappy^a, Ayoung Hong^{a,c}, Eunpyo Choi^b, Jong-Oh Park^{a,b,**}, Chang-Sei Kim^{a,b,*}^a Korea Institute of Medical Microrobotics, Gwangju, 61011, South Korea^b School of Mechanical Engineering, Chonnam National University, Gwangju, 61186, South Korea^c Robotics Engineering Convergence, Chonnam National University, Gwangju, 61186, South Korea

ARTICLE INFO

Keywords:

Automated medical image reconstruction
 Vessel segmentation
 3-D reconstruction
 Bi-plane angiography
 Lightweight multi-resolution U-shaped architecture

ABSTRACT

Automated three-dimensional (3D) blood vessel reconstruction to improve vascular diagnosis and therapeutics is a challenging task in which the real-time implementation of automatic segmentation and specific vessel tracking for matching artery sequences is essential. Recently, a deep learning-based segmentation technique has been proposed; however, existing state-of-the-art deep architectures exhibit reduced performance when they are employed using real in-vivo imaging because of serious issues such as low contrast and noise contamination of the X-ray images. To overcome these limitations, we propose a novel methodology composed of the de-haze image enhancement technique as pre-processing and multi-level thresholding as post-processing to be applied to the lightweight multi-resolution U-shaped architecture. Specifically, (1) bi-plane two-dimensional (2D) vessel images were extracted simultaneously using the deep architecture, (2) skeletons of the vessels were computed via a morphology operation, (3) the corresponding skeleton structure between image sequences was matched using the shape-context technique, and (4) the 3D centerline was reconstructed using stereo geometry. The method was validated using both in-vivo and in-vitro models. The results show that the proposed technique could improve the segmentation quality, reduce computation time, and reconstruct the 3D skeleton automatically. The algorithm accurately reconstructed the phantom model and the real mouse vessel in 3D in 2 s. Our proposed technique has the potential to allow therapeutic micro-agent navigation in clinical practice, thereby providing the 3D position and orientation of the vessel.

1. Introduction

Navigation of a microrobot or therapeutic micro-sized agents to perform targeted therapy inside the cardiovascular system is a prolific research area for minimally invasive surgeries (MIS) and efficient treatments through early diagnosis of disease (Abbott et al., 2007; Cavalcanti et al., 2009; Ishiyama et al., 2002). Owing to the micro- or nano- size robot body, the microrobot can move and target to numerous distal locations through the blood circulatory system (Belharet et al., 2010). Where the 3D reconstructed blood vessel image is a key requirement to implement the microrobot navigation for path generation, simulation, control, visualization, and monitoring.

Clinically, advanced non-invasive imaging systems, such as cardiac and coronary computed tomography (CCTA) and magnetic resonance angiography (MRA), are available for the three-dimensional (3D)

visualization of vessels. However, 2D invasive X-ray coronary angiography is still the best choice for clinicians or experts to diagnose diseases and provide therapy guidance (Mark et al., 2010). It is a familiar device and can produce the highest quality projections in terms of spatial and temporal resolution (Chen and Schäfer, 2009), but, 2D projections of 3D anatomies are not optimal for effective analysis.

There are many ongoing studies of the 3D representation of blood vessel trees using X-ray angiography. Different reconstruction techniques have been developed for various applications such as angiography systems, cardiac handling strategies, and respiratory motion. These methods can be mainly categorized into two groups: tomographic reconstruction and model-based reconstruction. Tomographic reconstruction creates the vessel volume using X-ray angiography. This technique uses X-ray absorption coefficients to compute the vessel volume and can deal with unusual anatomies with little or no prior

* Corresponding author at: 77 Yongbong-ro, Buk-gu, Gwangju, 61197, South Korea.

** Corresponding author at: 43-26, Cheomdangwagi-ro 208beon-gil, Buk-gu, Gwangju, 61011, South Korea.

E-mail addresses: jop@kimiro.re.kr (J.-O. Park), ckim@jnu.ac.kr (C.-S. Kim).

<https://doi.org/10.1016/j.compmedimag.2021.101956>

Received 10 December 2019; Received in revised form 18 August 2020; Accepted 16 July 2021

Available online 21 July 2021

0895-6111/© 2021 Elsevier Ltd. All rights reserved.

information (Hansis et al., 2008). The vessel surface can be obtained accurately via tomographic reconstruction and used independently without any manual interaction (Schoonenberg et al., 2009). However, the primary limitation of tomographic reconstruction arises from its need for a significant amount of data and large computation time and cannot be computed in real time. On the other hand, the model-based reconstruction is a flexible tool that utilizes X-ray angiography from different imaging systems to compute a 3D centerline of a vessel and construct its surfaces. The model-based techniques can be categorized into two types of algorithms: forward-projection and back-projection. The forward-projection algorithm does not necessitate any correspondence matching between 2D centerlines; further, even 2D segmentations are not required in some methods. However, there are major drawbacks associated with the forward-projection algorithm, such as the manual selection of corresponding points between vessel branches from projection sequences (Zheng et al., 2010; Yang et al., 2014; Cong et al., 2016). The back-projection-based method can reconstruct the vessel using the 2D information gathered from X-ray projections. This technique is appropriate for different types of medical imaging systems with calibrated and non-calibrated data (Cañero et al., 2002; Hoffmann et al., 2000).

In addition, other relevant methods are also widely researched as; a graph-cut-based reconstruction technique (Liu et al., 2015; Liao et al., 2010) has also been developed to smooth the depths in 3D. These methods divide the 3D space between the X-ray source and detector into parallel planes of equal depth. The reconstructed image is then obtained by assigning 2D centerline points to a plane. A 3D centerline reconstruction method was developed from the intersection of surfaces through corresponding branches (Galassi et al., 2018). This technique uses multiple projections to extract the vessels for the purpose of generating 3D models of the stenotic vessel to assess the stenosis severity. Zhang et al., formulated structure from motion as a learning problem and developed an end-to-end learning framework to compute the depth information from a sequence of images using a single camera (Zhang et al., 2019). Joe et al., proposed 3D model reconstruction methodology by using two sonar images in different viewpoints (Joe et al., 2020), and Lee et al., have conducted interesting research on 3D reconstruction and pose estimation for robust navigation (Lee et al., 2020). They introduced the recurrent convolutional neural network to learn spatio-temporal information using convolutional gated recurrent unit cells that store weights from previous layers.

As important sub algorithms in the reconstruction processes, segmentation and vesselness response are critically handled. Segmentation of the vessel and an adequate number of 2D point correspondences are commonly required to compute the geometry containing the position and orientation between projection sequences (Yang et al., 2009; James Chen and Carroll, 2000). Vesselness response highlights the vessel in 2D projection images. Subsequently, these responses are back-projected with the help of projection geometry to estimate the vessel in 3D. Segmentation in 2D (Andriotis et al., 2008), tubular shape response (Law et al., 2004), and distance map to centerline (Jandt et al., 2009) are some methods to obtain vessel response. However, these techniques cannot segment the target vessel automatically. Jerman et al., proposed a novel enhancement filter based on the ratio of multiscale Hessian eigenvalues that could enhance the vascular structure together with unwanted tubular structures, such as soft tissues, during segmentation (Jerman et al., 2015). Recently, Cui et al., developed a hybrid technique for a fast and accurate vascular enhancement filter by using vesselness diffusion and improved the vesselness filter based on the ratio of eigenvalues (Cui et al., 2019).

By virtue of recent advances in artificial intelligence, deep convolutional networks outperformed most of the state-of-the-art segmentation techniques. However, deep convolutional networks have not been popular, because they require a large number of training data, which are difficult to be obtained clinically. This is recently addressed by using transfer learning (Ibragimov et al., 2018) and pre-trained models

(Romero et al., 2019). Ronneberger et al. proposed an elegant architecture that works with very few training images to achieve precise segmentation (Ronneberger et al., 2015). There exist other improved U-net models for medical image segmentation: Hybrid Densely Connected UNet (Li et al., 2018) combined 2D and 3D DenseUNet for liver and tumor segmentation. The 3-D feature-enhanced network successfully segmented 3-D femur by introducing two feature enhancement modules that include the edge detection task and the fusion of multiscale features (Chen et al., 2017). Ibtehaz et al. determined that the U-Net architecture faces problems while segmenting different scale images and proposed a novel architecture called MultiResUNet to handle biomedical images at different scales (Ibtehaz and Rahman, 2019). In which, correspondence formation is a challenging and vital step in the back-projection technique as the 3D vessel centerline is reconstructed from the triangulation of correspondences. For this, an epipolar constraint technique can be used to determine the corresponding points (Li and Cohen, 2011). However, isocenter offsets and mechanical distortions of the X-ray system lead to the matching points being incorrectly computed. A soft epipolar constraint (Blondel et al., 2006) may solve this problem to some extent. Alternatively, branch-to-branch matching (Cardenas et al., 2012) dealt with the problem caused by the point-by-point matching methods in previous researches. Nevertheless, this type of matching faces major problems for vessel skeletons, where similar vessels in different images may produce varied branch points. During the automatic process, these points will lead to a complete mismatch of branches. Shape context (Belongie et al., 2002) is an effective way to solve the correspondence problem. This technique uses a shape descriptor to explain the coarse distribution of the rest of the shape compared to a specific point on the shape.

On this background, the automated reconstruction of vessels is still a challenging task requiring complex processing steps to generate the final 3D reconstruction model. Herein, we present a fully automated 3D reconstruction technique using projections through a biplane X-ray system. We address each step carefully and introduce methodologies to perform all computations automatically, and then validate the algorithm using experimental datasets. First, we propose a lightweight MultiResUNet architecture to reduce computation time together with pre- and post-processing steps. The primary difference between the proposed lightweight architecture and the original MultiResUNet is the amount of layer reduction. It decreases the computation time while increasing the segmentation quality. By incorporating de-hazing to input datasets and thresholding to output data, the segmentation quality can be improved while providing fast computation. Because the accuracy of the reconstruction is highly dependent on the segmentation of the target vessel, this technique successfully extracts the target vessel with a higher accuracy and reduced computation time. Second, we propose a top-to-bottom strategy for accessing the endpoints as a feature to select features and branch points. Without any manual handling, the proposed method can overcome the difficulty of feature point selection and skeleton matching caused by displacement and orientation discrepancies. Finally, we present details of in-vivo and in-vitro experiments conducted under the bi-plane system to validate the proposed reconstruction methodology. We also analyze the results by incorporating state-of-the-art techniques to validate the segmentation performance. Thus, we show that the reconstruction process is fully automated and that the proposed segmentation method outperforms existing ones.

The fundamental contributions of this paper can be summarized as follows: First, we identify two key issues that prevent the 3D reconstruction process from being automated: target vessel segmentation and appropriate feature selection. Second, we propose a 3D reconstruction algorithm to include deep learning-based segmentation to extract the accurate target vessel and skeleton endpoints as the only features for the reconstruction process. Third, a lightweight MultiResUNet architecture appended with de-haze pre-processing and a thresholding as post-processing to segment the target vessel are newly developed; the highest accuracy and lowest computation time achieved are 98.68 % and

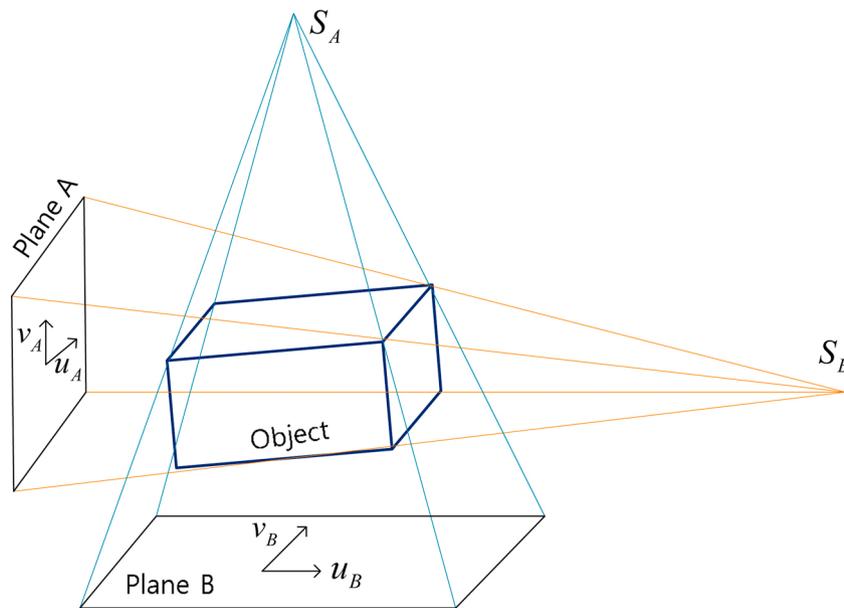


Fig. 1. Geometric relationship between two angiographic views.

1.46 s, respectively, as compared to other existing methods. Finally, we implement a fully automated 3D reconstruction process with a computation time of only 2.05 s. This contribution may advance automated 3D vessel reconstruction technology for cardiovascular disease diagnosis and therapeutics development in clinical settings.

2. Method

The proposed method is composed of three processes. The first process is projection image acquisition, in which biplane system calibrations and projection transformations are performed. The second process deals with the preprocessing of the acquired images: contrast similarity between image sequences is determined and the contrast is enhanced. The final process is segmentation and automated reconstruction. Detail explanations for each step are presented below.

2.1. Projection acquisition

The imaging system in Fig. 1 shows two planes that are perpendicular to each other. During acquisition, the model had a rotation angle to obtain a better view from both sides. The acquisition time was set to 1 fps to synchronize both images simultaneously. The X-ray image sequences $I_{A,j}$ at image plane A and $I_{B,j}$ at image plane B could then be obtained simultaneously from the biplane X-ray system. Image sequences are denoted as $j = 1, 2, 3, \dots$. As shown in Fig. 1, a spatial point on the vessel centerline with coordinates $W_i = (x_i, y_i, z_i)^T$ is projected to image plane A as a 2D point $u_{i,A} = (u_{i,A}, v_{i,A})^T$, where the number of 2D points is defined as $i = 1, 2, 3, \dots$. Finally, the projection on image plane A can be computed as follows:

$$(x_i, y_i, z_i)^T \rightarrow \left(SID_A * \frac{x_i}{z_i} + u_{c,A}, SID_A * \frac{y_i}{z_i} + v_{c,A} \right)^T = (u_{i,A}, v_{i,A})^T \quad (1)$$

where $(u_{c,A}, v_{c,A})^T$ represents the coordinates of the principal point and SID_A is the distance between the source and image intensifier A. Similarly, the projection on image plane B can also be calculated by Eq. (1).

The projections of a spatial point W_i on image planes A and B through projection matrices are denoted as $P_{i,A}$ and $P_{i,B}$, respectively. These can be formulated to homogenous forms as follows (Hartley and Zisserman, 2003):

$$\begin{aligned} u_{i,A} &= P_{i,A} W_i = K_A [I|0] W_i \\ u_{i,B} &= P_{i,B} W_i = K_B [R|t] W_i \end{aligned} \quad (2)$$

where K_A and K_B are the intrinsic matrices of image planes A and B, respectively. These intrinsic matrices are known that were obtained by calibrating the bi-plane system in hardware level. Matrices $[I|0]$ and $[R|t]$ represent the extrinsic parameters.

2.2. Preprocessing

In the preprocessing, a midway equalization between two images in different view planes and a selection of a specific vessel to reconstruct are implemented.

Projections from two different imaging planes were obtained under different lighting conditions, and several image factors related to luminance were varied with time. Therefore, all images should maintain similar luminance qualities for the fully automated process. We utilized histogram information from image sequences, thereby allowing similar luminance to be maintained over time. This technique (Sánchez, 2016) initially computes the cumulative normalized histogram and then calculates the change in contrast functions f_j depending on the temporal Gaussian convolutions. The cumulative normalized histogram of an image $I_{A,j}$ from image plane A and $I_{B,j}$ from image plane B can be computed as

$$\begin{aligned} H_{I_{A,j}}(\gamma) &= \frac{1}{|\mu_j|} |\{(ii, jj) \in \mu_j; I_{A,j}(ii, jj) \leq \gamma\}| \\ H_{I_{B,j}}(\gamma) &= \frac{1}{|\mu_j|} |\{(ii, jj) \in \mu_j; I_{B,j}(ii, jj) \leq \gamma\}| \end{aligned} \quad (3)$$

where μ_j denotes the discrete domain $\{1, \dots, n\} \times \{1, \dots, m\}$ and $|\mu_j| = n \times m$ represents total number of pixels of image $I_{A,j}$ or $I_{B,j}$. Then, f_j can be computed as

$$f_j = \left(0.5 (H_{I_{A,j}}^{-1} + H_{I_{A,j+1}}^{-1})^{-1} \right) = \left(0.5 (H_{I_{B,j}}^{-1} + H_{I_{B,j+1}}^{-1})^{-1} \right) \quad (4)$$

Subsequently, the algorithm transforms the sequence of images from planes A and B by applying Eq. (3):

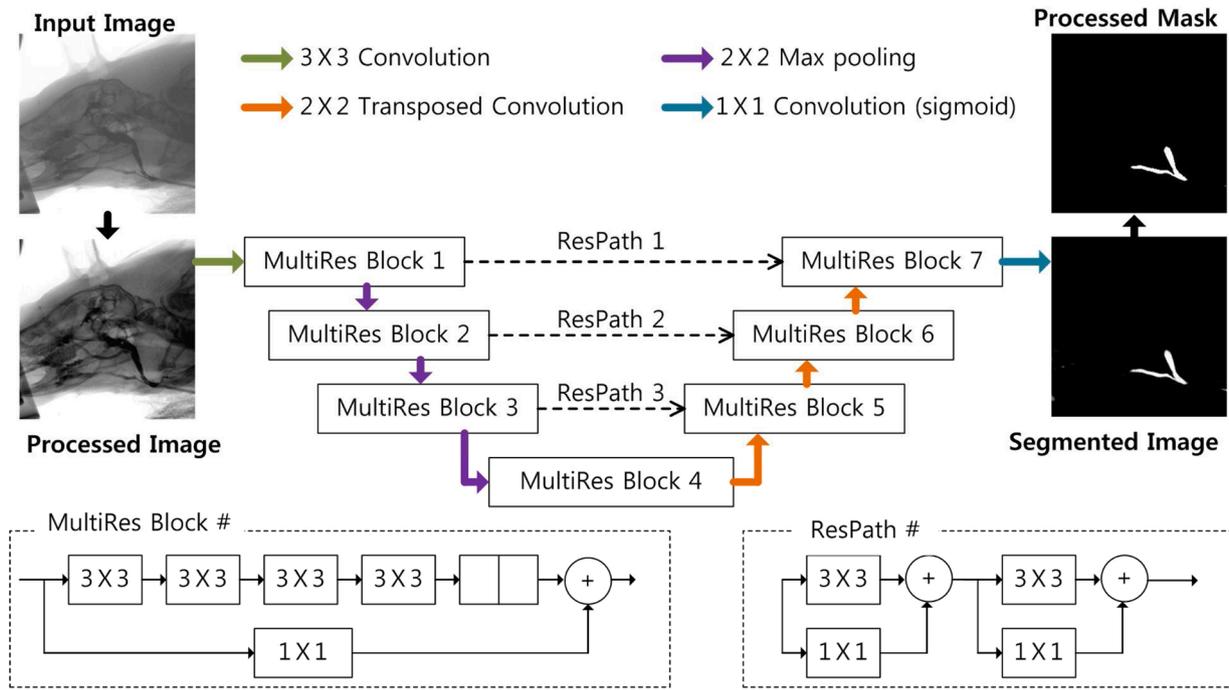


Fig. 2. Proposed segmentation technique.

$$\begin{aligned} I_{A,j} &= f_j(I_{A,j}) \\ I_{B,j} &= f_j(I_{B,j}) \end{aligned} \quad (5)$$

In addition, the problem of low contrast is resolved in the pre-processing. Low-contrast in-vivo images are frequently obtained because of image modality, lighting condition, and complex organ structures. These are primary issues that result in noisy and unqualified imaging. As the final preprocessing step to resolve the low contrast problem, the X-ray images are inverted, and then, the de-haze technique is applied to enhance the contrast of the image. The enhanced images are obtained by

$$I_{enA,j} = (I_{A,j,inv} - A) / T + A \quad (6)$$

where $I_{A,j,inv}$ is the inverted image, A is the global atmospheric light, and T is the medium transmission explaining the portion of the light that is not scattered and reaches the camera.

2.3. Lightweight MultiResUNet segmentation

Vessel segmentation is essential but difficult to automate because target vessels have different shapes and scales depending on their positions in vascular networks. To segment these vessels, we propose a lightweight MultiResUNet architecture where two MultiRes blocks and one corresponding ResPath between the blocks are utilized to reduce computation time for segmentation, as shown in Fig. 2. The architecture is symmetric; it contains an Encoder that extracts spatial features from the image and a Decoder that builds the segmentation map using the encoded features. The Encoder consists of three MultiRes blocks for down-sampling and the Decoder has three MultiRes blocks to perform up-sampling. The Encoder and Decoder are connected by a single MultiRes block. These blocks include convolutional layers, where the number of filters is increased gradually in successive layers, and a residual block is connected. Each MultiRes block from the Encoder connects with each corresponding MultiRes block of the Decoder through ResPath.

The input images and their corresponding segmentation masks are used to train the network by implementing the stochastic gradient descent method. The energy function is computed by a pixel-wise soft-

max on the final feature map gathered with the cross-entropy loss function. The cross entropy can be defined as

$$S = \sum w(x) - \log(p(x)) \quad (7)$$

where x is the pixel position; $w(x)$ is the weight map, which is pre-computed for ground truth segmentation; and the soft-max function is denoted as $p(x)$.

The resolution of the original X-ray images was 2352×2944 , but they were resized to 256×256 for training. In addition, our approach did not use any data augmentation in this work. For training and testing, we split the images into 20 for training and 20 for testing. There was no overlap between the training and test images. The model is optimized by training for over 50 epochs using the stochastic gradient optimizer with a learning rate of 0.01. The patches are extracted from full-size images with dimensions of 48×48 . Thus, the total number of patches became 190,000 for 20 training images. We set the batch size of the network as 32.

2.4. 3D reconstruction

As two plane images are obtained in this method, one view at a time from a primary plane is used to reconstruct the vessel centerline. The segmented vessel contains several random noises. Using maximum thinning, these noises are suppressed during vessel skeleton extraction. Mathematical morphology based on the thinning algorithm (Kerschitzki et al., 2013; Lee et al., 1994) is applied to vessels to obtain the vascular skeleton. However, this skeleton contains some small extended branches and disconnected parts due to noises in the segmentation mask. Thus, the extended branches are removed by pruning the skeleton. Finally, by thresholding out smaller regions, the disconnected parts can be removed from the final skeleton. A threshold value of 10 is set to remove any skeleton branch of fewer than 10 pixels in length. In this study, the threshold value is obtained by trial and error, but this number can be adjusted depending on user requirements.

The accurate topology and feature points of the vessel are crucial characteristics of the image required for the automated 3D reconstruction of the selected vessel. Feature points are defined as points placed on the vessel skeleton, such as at the endpoint, bifurcation point, and cross

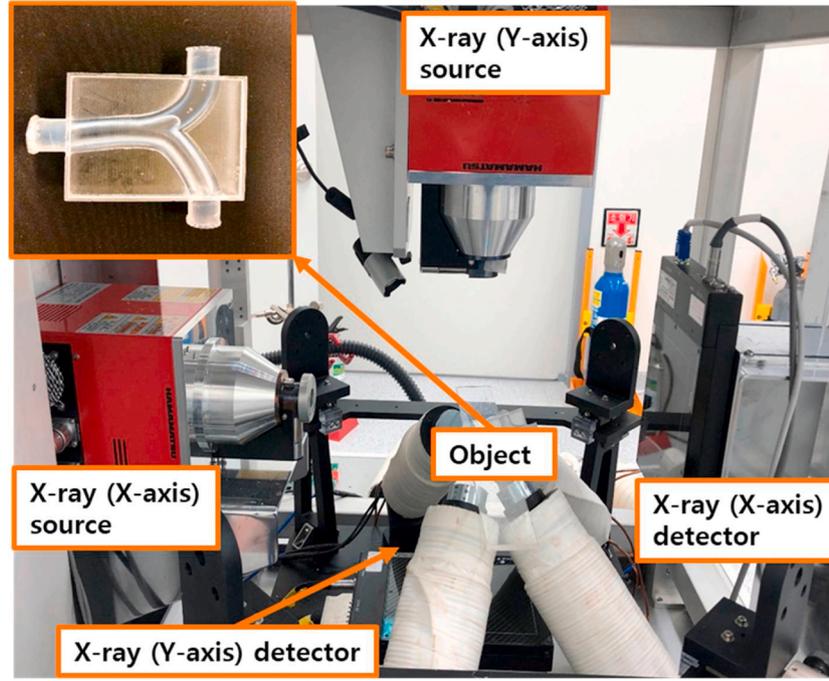


Fig. 3. Experimental Setup; bi-plane X-ray system with phantom model and support.

points. Bifurcation refers to the source point where a single vessel separates into two or more branches along its orientation. Cross points do not exist in reality: rather, they appear in angiograms because of the overlapping shadows of two vessels. These two feature points are not stable for automated reconstruction, because a tiny change in the position of these points may introduce large reconstruction errors. In contrast, endpoints simply represent the terminals of vessels. Therefore, our proposed technique utilizes the endpoints to match the shapes between corresponding skeletons.

We treat each skeleton as a point set and assume the shape of a skeleton is captured by a finite subset of its points. Specifically, a skeleton shape is denoted by a discrete set of points from the internal or external contours of the skeleton. These can be obtained from the locations of each pixel of the skeleton. The skeleton is obtained by a thinning operation that is related to the hit-and-miss transform. The thinning of the image from plane A and plane B is computed via structuring element J.

$$\begin{aligned} u_{skl,A} &= I_{en,A,j} - hit_and_miss(I_{en,A,j}, J) \\ u_{skl,B} &= I_{en,B,j} - hit_and_miss(I_{en,B,j}, J) \end{aligned} \quad (8)$$

Consider a point $u_{skl,i,A}$ on the skeleton from image plane A and a point $v_{skl,j,B}$ on image plane B.

Let, $C_{i,j} = C(u_{skl,i,A}, v_{skl,j,B})$ define the cost of matching these two points. Then, the shape contexts become distributions that can be represented as histograms:

$$C_{i,j} = C(u_{skl,i,A}, v_{skl,j,B}) = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \quad (9)$$

where $h_i(k)$ and $h_j(k)$ define the K-bin normalized histogram at $u_{skl,i,stack_A}$ and $v_{skl,j,stack_B}$, respectively.

According to bipartite graph matching for a given set of costs $C_{i,j}$ between all pairs of points $u_{skl,i,stack_A}$ on image plane A and $v_{skl,j,stack_B}$ on image plane B, the total cost of matching is minimized by

$$H(\pi) = \sum_i C(u_{skl,i,A}, v_{\pi(sk_{l,j},B)}) \quad (10)$$

Eq. (10) is subject to the constraint that the correspondence follows

one-to-one matching. Here, π is defined as the permutation. Eq. (11) is minimized to compute permutation $\pi(sk_{l,j}, B)$. Consequently, the above correspondences $u_{skl,i,A}$ and $v_{skl,j,B}$ transform into global coordinates $u_{skl,i,A}^w$ and $v_{skl,j,B}^w$ using intrinsic camera parameters, K_{img}

$$u^w = \begin{bmatrix} u^w \\ v^w \\ 1 \end{bmatrix} = K_{img}^{-1} \begin{bmatrix} u_{skl} \\ v_{skl} \\ 1 \end{bmatrix} \quad (11)$$

As the top view image has a different size than the side view, the sizes of the images must be adjusted. This system maintains the acquisition time at 1 fps such that it is synchronized with our proposed technique. We assume SID_{img} is the distance between the X-ray source and the detector plane, and SOD_{img} is the distance between the detector and the object. Therefore, the depth z can be computed as $SID_{img} - SOD_{img}$.

According to the theorem of similar triangles, we can obtain the following relationship.

$$\frac{s}{d} = \frac{SID_{img} - SOD_{img}}{SOD_{img}} = \frac{z}{SID_{img} - z} \quad (12)$$

where d is the distance between the two corresponding points and s defines the distance between the two X-ray sources. Finally, the 3D point can be estimated as

$$\begin{bmatrix} u \\ v \\ z \end{bmatrix} = \frac{1}{2} \times (u_{skl,i,A}^w + v_{skl,j,B}^w) + \begin{bmatrix} 0 \\ 0 \\ s \times \frac{SID}{d+s} - 1 \end{bmatrix} \quad (13)$$

The following algorithm represents the pseudocode of the reconstruction technique.

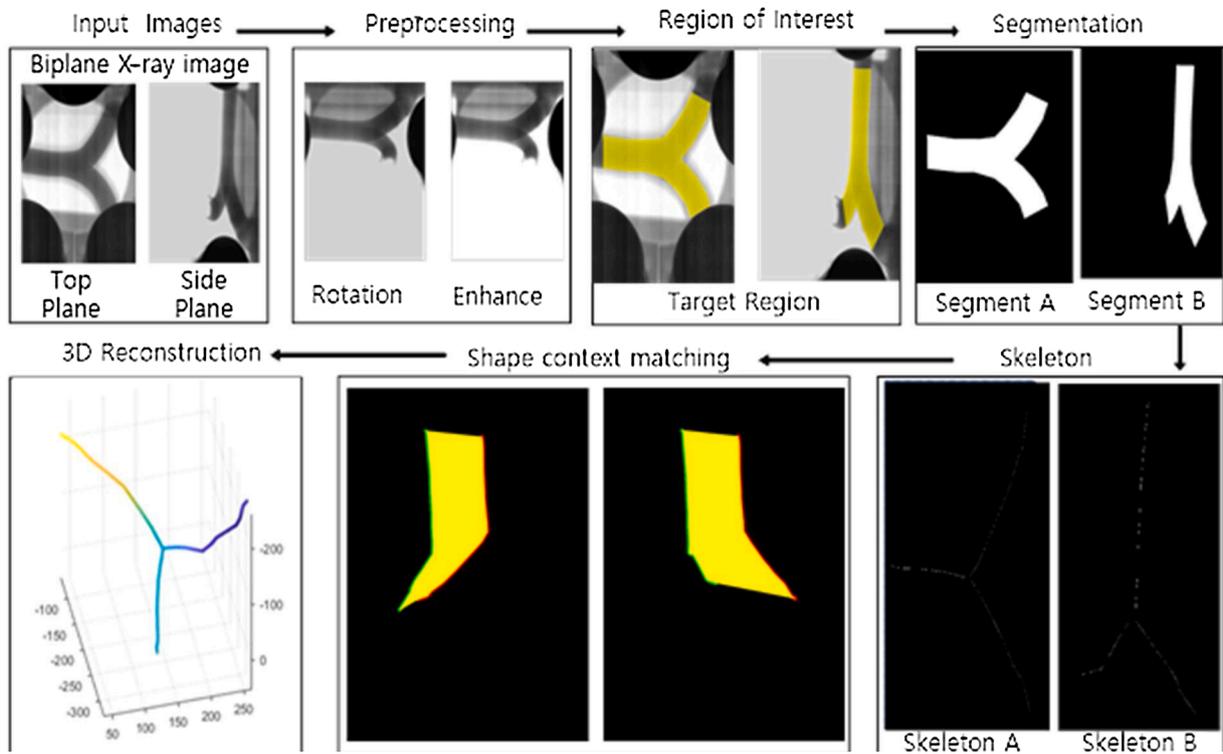


Fig. 4. System workflow with examples for each step.

Algorithm: Deep segmentation-based 3D reconstruction		
Acquisition	Receive X-ray image sequence from top plane A	Receive X-ray image sequence from side plane B
Initialization	Initialize a 2D array I_A with zero values	Initialize a 2D array I_B with zero values
Preprocessing	for each input image I_A do estimate midway equalized image I_A using eq. 5. estimate enhanced image $I_{en,A}$ using eq. 6. update deep segmentation dataset A	for each input image I_B do estimate midway equalized image using eq. 5. estimate enhanced image $I_{en,B}$ using eq. 6. update deep segmentation dataset B
Reconstruction	<p> predict segmented mask $I_{A,j}$ predict segmented mask $I_{B,j}$ estimate skeleton using morphology $u_{skl,j,A}$ estimate skeleton using morphology $v_{skl,j,B}$ estimate feature points from $u_{skl,j,A}$ estimate feature points from $v_{skl,j,B}$ estimate shortest path between feature points on $u_{skl,j,A}$ estimate shortest path between feature points on $v_{skl,j,B}$ estimate matching points between skeletons shortest paths using eq. 9 and eq. 10 estimate 3D points of skeleton using eq. 13 clear dataset end for </p>	

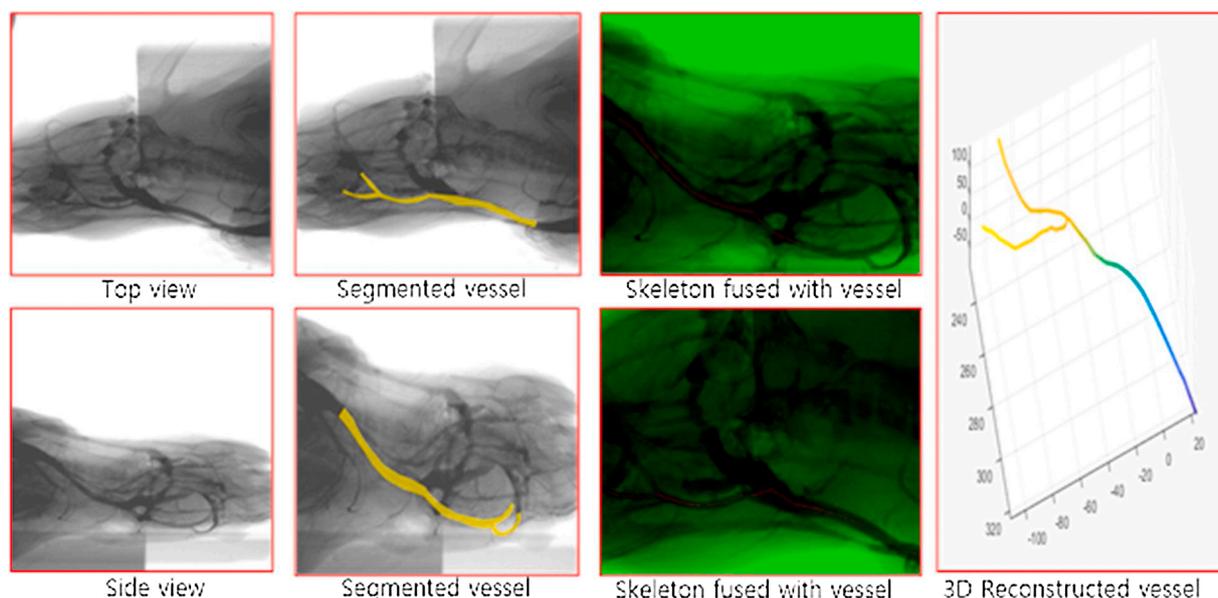


Fig. 5. 3D reconstruction using mouse dataset 1.

3. Experimental results

3.1. Experimental setup

Biplane angiograms are obtained using biplane X-ray equipment similar to that used in our previous work (Nguyen et al., 2018). Fig. 3 shows the biplane X-ray system. This system consists of two X-ray sources and two detectors. A set of X-ray devices was configured perpendicular to the other set of X-ray devices. For the experiments, the object position was fixed at a stationary position under the flat detectors. The performance of the proposed algorithm was assessed using phantom and clinical data. According to the X-ray angiography imaging technique, 2D vessels were imaged by passing X-rays through the vessel-like structure at different angles. A series of vessel-like phantom and mouse angiographic images were utilized to verify the feasibility and reliability of the proposed automatic reconstruction technique. The algorithm was implemented using C++ and MATLAB. The experiments were

conducted on a mid-range PC with an Intel Core i7 4.2 GHz processor, NVIDIA TITAN XP GPU, and 16 GB memory.

Both in-vitro and in-vivo experiments were conducted. For the in-vitro experiments, the phantom model was prepared with Vaeroclear RGD810 material. This model consisted of a main vessel with two branches. The size of the main vessel was 9 mm with 7 mm branches. A contrast agent was inserted inside the phantom to visualize the vessel-like structure through X-ray images. A mouse was used for the in-vivo experiment. During imaging, the focal length of the bi-plane system was 400 mm, the pixel resolution was 2352×2944 pixels, and the spatial resolution was 0.019 mm/pixel. To avoid complexities related to the large angular difference between X-ray sources, and for comparison with the phantom model, we selected a vessel with two branches near the neck. The original image size was scaled down to 256×256 for computational efficiency. The mouse was approximately 10 weeks old. Contrast agent insertion and flow were controlled by experts during real-time data acquisition.

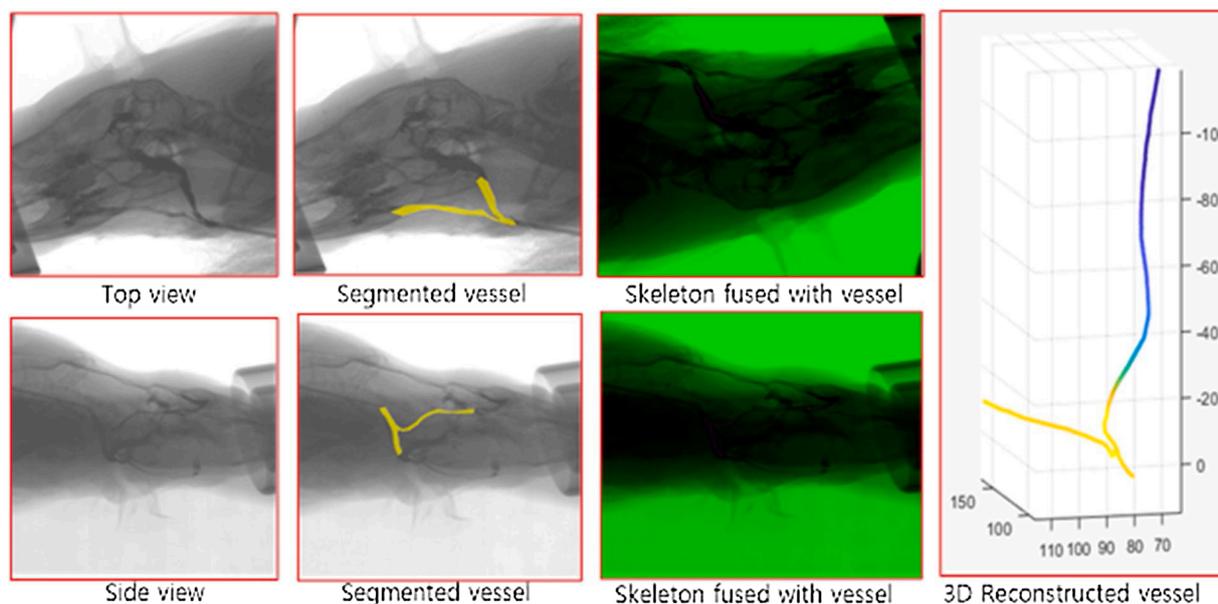


Fig. 6. 3D reconstruction using mouse dataset 2.

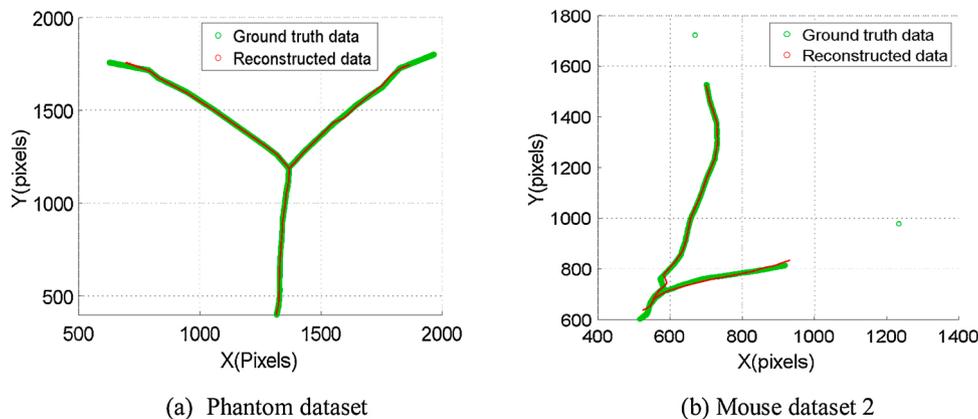


Fig. 7. Re-projection of reconstructed data and ground truth for phantom dataset (left) and mouse dataset2 (right).

3.2. Segmentation performance evaluation metrics

We validated our architecture using different metrics: accuracy (ACC) and area under the curve (AUC) of receiver operating characteristic (ROC). ACC is a metric that measures the ratio between the correctly classified pixels and the total pixels in the dataset. Precision (PPV) specifies the proportion of true positive samples among the total predicted positive samples. Sensitivity (TPR) estimates the proportion of positives that are correctly identified. Specificity (TNR) computes the proportion of negatives that are correctly classified. To relate the similarity and diversity of test datasets, the F-measure ($F1$) (Sasaki, 2015) and Jaccard similarity (JS) (Jaccard, 1912) are estimated.

3.3. Evaluation with in-vitro experiments

Fig. 4 illustrates the workflow used in this study. We use the following steps to automatically reconstruct the 3D vessels from bi-plane X-ray image sequences. First, we receive two X-ray images simultaneously from the bi-plane system, and we enhance the contrast of these images. Second, we train our proposed lightweight architecture with the target vessel. Third, we use the trained architecture to segment the specific vessels. Fourth, we compute the centerline of the segmented mask from the previous step and obtain the corresponding points from shape-context-based matching. Finally, we compute the 3D point cloud from the triangulation of two sets of matching points. The reconstructed 3D centerline of the phantom model shown in the top view of Fig. 4 represents a shape similar to that of the 2D skeleton extracted using top-view X-rays. The similarity between reconstructed and original skeleton shapes in Fig. 7 (left) proves that our proposed technique performs well compared to the gold-standard phantom model structure. We have repeated this experiment three times to confirm the final reconstruction.

3.4. Evaluation with in-vivo experiments

Fig. 5 illustrates the reconstruction technique with mouse dataset 1. Because of the large angular difference between the two views, it is difficult to obtain a high quality visualization of the vessels. We chose the neck vessel to evaluate our reconstruction technique; this prevented overlap with other vessels and organs as vessels from other body parts are avoided. Our proposed algorithm simultaneously receives two images from the bi-plane system and passes these through the deep architecture to obtain the segmented vessel. Then, a satisfactory 3D reconstruction can be achieved by computing the centerline and corresponding points. The vessel centerline and angiography image are fused to validate the accuracy of centerline extraction. The experiment with dataset 1 was repeated at least five times to reconstruct the target vessel clearly. Fig. 6 presents an evaluation of our reconstruction technique with mouse dataset 2. The mouse position was maintained on a slope to

Table 1

euclidean distance between different reconstruction methods.

Methods	Branch 1 (pixels)	Branch 2 (pixels)
Proposed	8.44	7.87
MultiResUNet	13.27	14.24
U-Net	16.22	15.98

obtain a better view of the same vessel. Fig. 6 indicates that the 3D reconstructed skeleton from the side view is similar to the target vessel shapes obtained from the lateral X-ray image. The comparison between the reconstructed and 2D skeleton shapes, as shown in Fig. 7 (right), indicates slight dissimilarity in some regions. Although the large angular difference between two views introduced many variations in vessel shapes, our proposed method could reconstruct a skeleton shape that was similar to the original one. The experiment with dataset 2 was repeated approximately six times for accurate vessel reconstruction.

3.5. Quantitative evaluation

To evaluate the accuracy of our proposed algorithm, we compare the 3D reconstructed centerline with the ground truth centerline. We consider the re-projection of 3D reconstructed centerline points to their corresponding 2D locations in the image. Furthermore, we calculate projection and transformation matrixes to project the 3D centerline points to the image coordinates. Fig. 7 presents the accuracy analysis of two types of datasets. In both cases, the green markers represent the ground truth, and the red markers indicate the projected centerline. We intentionally used a bigger size for the green marker compared to that of the red marker to ensure that the projected and ground truth centerlines could be visualized. It is evident that the phantom dataset on the left exhibits better re-projection as compared to the mouse dataset 2 on the right. A better segmentation ensures higher accuracy in 3D reconstructions. It is known that the position and orientation of the mouse in a bi-plane system are difficult to handle during angiography. In such situations, the segmentation process is significantly complex. On the contrary, segmenting a fixed vessel, such as from a phantom dataset, is an easier task.

In addition, we have performed reconstruction on dataset 2 with ground truth, MultiResUNet, and U-Net. Subsequently, we compare the ground truth reconstruction with that performed using the proposed and other techniques discussed herein. The results listed in Table 1 confirm that our proposed reconstruction technique better approximates ground truth than other methods.

4. Discussions and conclusion

We evaluate the proposed technique by comparing it with state-of-

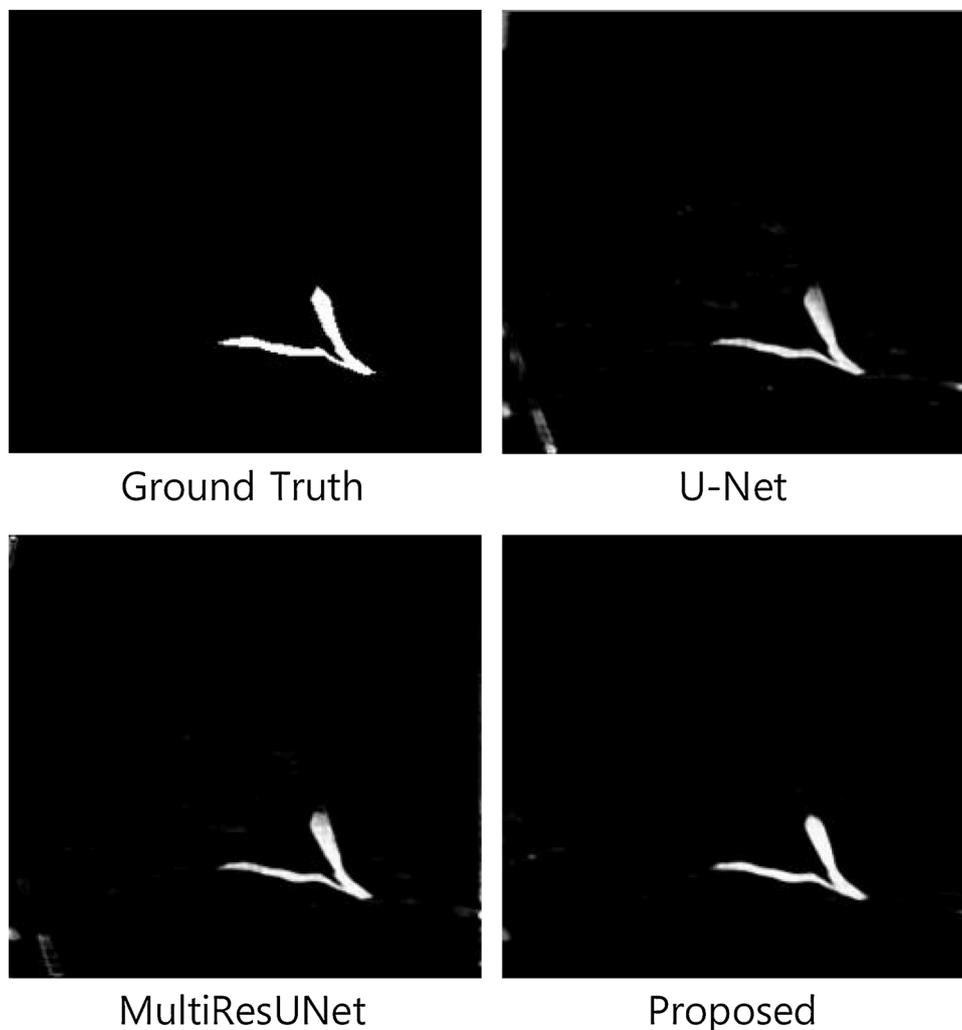


Fig. 8. Segmenting target vessel using different techniques.

Table 2

Performance comparison between different techniques.

Models	U-Net	MultiResUNet	Proposed
ACC	0.9717	0.9775	0.9868
F1	0.8519	0.8878	0.9272
JS	0.9367	0.9533	0.9784
AUC	0.9514	0.9791	0.9888
Time [sec]	1.65	1.72	1.46

the-art vessel segmentation techniques, i.e., U-Net and MultiResUNet, on mouse dataset 2. There are two versions of dataset 2. The first version is the original X-ray image from a bi-plane system, and the second version is an enhanced X-ray image obtained after applying an image enhancement algorithm. The first version of the dataset is used for other segmentation techniques, whereas the second version is prepared for the proposed technique. Both datasets are divided into a training set and a test set. We newly trained all three architectures again using the corresponding training datasets. We set the batch size to 32 and the total number of training epochs to 100. We utilized the Adam optimizer and cross-entropy as the loss function.

Fig. 8 presents the segmentation masks of different techniques, along with the ground truth mask. Segmentation masks from U-Net and MultiResUNet contain noisy structures, in contrast with that from our proposed method. These noises will increase substantially for complex vessels. Moreover, segmentation performance between different

techniques was evaluated via different quantitative analyses, the results of which are presented in Table 1.

A comparison of the segmentation performance and computation time of different vessel segmentation techniques is presented in Table 2. The results in the table confirm that the highest accuracy of 98.68 % was achieved on enhanced mouse dataset 2. Furthermore, we compared the architectures with performance metrics such as F1, JS, and AUC. The table shows that the proposed technique achieves the best performance for these metrics. Additionally, the proposed architecture is capable of segmenting a single image in the lowest time of 1.46 s compared to the computation times of the other existing architectures.

Table 3 presents a comparison of the proposed method with recent 3D reconstruction algorithms, in terms of reconstruction methods and computation times. The primary goal of our research is to reconstruct 3D vessels in real-time. To realize such a real-time system, it is necessary to create a fully automated reconstruction process. Except for the technique proposed by (Liu et al., 2015), all other techniques require manual operation. Hence, these methods cannot be evaluated in real time. However, the proposed methodology is fully automated and offers the best computation time, as compared to other recent methods.

Automated 3D reconstruction is essential for diagnosing cardiovascular network diseases, therapeutic device development, and micro-robot navigation in blood vessels. Moreover, tracking the corresponding vessel, which is a crucial task in reconstructions, has been a challenging problem due to complicated in-vivo networks and DICOM® medical imaging quality. Hence, in a majority of state-of-the-art techniques,

Table 3
Comparison between different reconstruction techniques.

Algorithm	Methods	Reconstruction Time [s]	Reconstruction Process
Galassi [11]	C-arm X-ray Manual selection of control points across projections Multiple 2-dimensional angiogram	60–120	Manual
Yang [32]	C-arm X-ray Force back-projection technique	N/A	Manual
Liu [25]	Manual selection of corresponding points C-arm X-ray Energy optimization More than two projections	3	Automatic
Liao [24]	C-arm X-ray Graph-cut based technique Manual selection extraction	10–20	Manual
Yang [33]	C-arm X-ray Perspective projection method Manual selection of bifurcation points	4.5	Manual
Proposed	Bi-plane X-ray Deep-learning-based segmentation Automatic selection of feature points	2.05	Automatic

either corresponding vessels or feature points are manually selected during the reconstruction process. However, manual processes cannot be employed to implement real-time image reconstruction applications. In this study, we propose a lightweight deep architecture to segment corresponding vessels. Our technique is capable of segmenting target vessels in complex imaging environments with better performance than that of existing state-of-the-art methods. The proposed reconstruction technique is highly adaptable for real-time applications. The proposed algorithm can achieve automated 3D reconstruction of a vessel by resolving the real-time reconstruction problem and may advance the medical image processing application to diagnosis and therapeutics for blood vessels.

Considering that segmenting vessels on a bi-plane X-ray system is a challenging task, especially due to the large angular difference between the two views, the difficulties of tracking the corresponding vessel still remain. To address this issue, future works will aim to overcome the angular differences between two views by remodeling the bi-plane system and implementing a more advanced vessel recognition algorithm. In addition, we aim to employ these 3D reconstructed vessels to implement microrobot navigation under in-vivo environments.

In future work, we will employ a variety of complex target areas such as heart and liver vessels. For these kinds of complex situations, we will require large datasets and data augmentation. Further, cross-validation will be suitable for such large datasets.

CRedit authorship contribution statement

D.M. Bappy: Methodology, Software, Writing - original draft.
Ayoung Hong: Investigation, Validation.
Eunpyo Choi: Resources, Formal analysis.
Jong-Oh Park: Supervision, Funding acquisition.
Chang-Sei Kim: Conceptualization, Supervision, Writing - review & editing.

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.

Acknowledgements

This research was supported by a grant of the Korea Health Technology Development R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI19C0642).

References

- Abbott, J.J., Nagy, Z., Beyeler, F., Nelson, B.J., 2007. Robotics in the small. *IEEE Robot. Autom. Mag.* 14 (2), 92–103. <https://doi.org/10.1109/MRA.2007.380641>.
- Andriotis, A., Zifan, A., Gavaises, M., Liatsis, P., Pantos, I., Theodorakakos, A., Efstathiopoulos, E.P., Katrasis, D., 2008. A new method of three-dimensional coronary artery reconstruction from X-ray angiography: validation against a virtual phantom and multislice computed tomography. *Catheter. Cardiovasc. Interv.* 71 (1), 28–43. <https://doi.org/10.1002/ccd.21414>.
- Belharet, K., Folio, D., Ferreira, A., 2010. MRI-based microrobotic system for the propulsion and navigation of ferromagnetic microcapsules. *Minim. Invasive Ther. Allied Technol.* 19 (3), 157–169. <https://doi.org/10.3109/13645706.2010.481402>.
- Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (4), 509–522. <https://doi.org/10.1109/34.993558>.
- Blondel, C., Malandain, G., Vaillant, R., Ayache, N., 2006. Reconstruction of coronary arteries from a single rotational X-ray projection sequence. *IEEE Trans. Med. Imaging* 25 (5), 653–663. <https://doi.org/10.1109/TMI.2006.873224>.
- Cañero, C., Vilarino, F., Mauri, J., Radeva, P., 2002. Predictive (Und)distortion model and 3-D reconstruction by biplane snakes. *IEEE Trans. Med. Imaging* 21 (9), 1188–1201. <https://doi.org/10.1109/TMI.2002.804421>.
- Cardenes, R., Novikov, A., Gunn, J., Hose, R., Frangi, A.F., 2012. 3D reconstruction of coronary arteries from rotational X-ray angiography. *Proceedings - International Symposium on Biomedical Imaging* 618–621. <https://doi.org/10.1109/ISBI.2012.6235624>.
- Cavalcanti, A., Shirinzadeh, B., Fukuda, T., Ikeda, S., 2009. Nanorobot for brain aneurysm. *Int. J. Rob. Res.* 28 (4), 558–570. <https://doi.org/10.1177/0278364908097586>.
- Chen, S.J., Schäfer, D., 2009. Three-dimensional coronary visualization, part 1: modeling. *Cardiol. Clin.* 27 (3), 433–452. <https://doi.org/10.1016/j.ccl.2009.03.004>.
- Chen, F., Liu, J., Zhao, Z., Zhu, M., Liao, H., 2017. Three-dimensional feature-enhanced network for automatic femur segmentation. *IEEE J. Biomed. Heal. Informatics* 23 (1), 243–252. <https://doi.org/10.1109/JBHI.2017.2785389>.
- Cong, W., Yang, J., Ai, D., Chen, Y., Liu, Y., Wang, Y., 2016. Quantitative analysis of deformable model-based 3-D reconstruction of coronary artery from multiple angiograms. *IEEE Trans. Biomed. Eng.* 63 (7), 1389–1400. <https://doi.org/10.1109/TBME.2014.2347058>.
- Cui, H., Xia, Y., Zhang, Y., 2019. 2D and 3D vascular structures enhancement via improved vesselness filter and vessel enhancing diffusion. *IEEE Access* 7, 123969–123980. <https://doi.org/10.1109/ACCESS.2019.2938392>.
- Galassi, F., Alkhalil, M., Lee, R., Martindale, P., Kharbanda, R.K., Channon, K.M., Grau, V., Choudhury, R.P., 2018. 3D reconstruction of coronary arteries from 2D angiographic projections using nonuniform rational basis splines (NURBS) for accurate modelling of coronary stenoses. *PLoS One* 13 (1), e0190650. <https://doi.org/10.1371/journal.pone.0190650>.
- Hansis, E., Schäfer, D., Dössel, O., Grass, M., 2008. Evaluation of iterative sparse object reconstruction from few projections for 3-D rotational coronary angiography. *IEEE Trans. Med. Imaging* 27 (11), 1548–1555. <https://doi.org/10.1109/TMI.2008.2006514>.
- Hartley, R., Zisserman, A., 2003. *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge UK Cambridge Univ Press.
- Hoffmann, K.R., Sen, A., Lan, L., Chua, K.G., Esthappan, J., Mazzucco, M., 2000. A system for determination of 3D vessel tree centerlines from biplane images. *Int. J. Card. Imaging* 16 (5), 315–330. <https://doi.org/10.1023/A:1026528209003>.
- Ibragimov, B., Toesca, D., Chang, D., Yuan, Y., Koong, A., Xing, L., 2018. Development of deep neural network for individualized hepatobiliary toxicity prediction after liver SBRT. *Med. Phys.* 45 (10), 4763–4774.
- Ibtehaz, N., Rahman, M.S., 2019. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Netw.* 1–25. <https://doi.org/10.1016/j.neunet.2019.08.025>.
- Ishiyama, K., Sendoh, M., Arai, K.L., 2002. Magnetic micromachines for medical applications. *J. Magn. Mater.* 242, 41–46. [https://doi.org/10.1016/S0304-8853\(01\)01181-7](https://doi.org/10.1016/S0304-8853(01)01181-7).
- Jaccard, P., 1912. The distribution of the Flora in the Alpine Zone. *New Phytol.* 11 (2), 37–50. <https://doi.org/10.1111/j.1469-8137.1912.tb05611.x>.
- James Chen, S., Carroll, J.D., 2000. 3-D reconstruction of coronary arterial tree to optimize angiographic visualization. *IEEE Trans. Med. Imaging* 19 (4), 318–336. <https://doi.org/10.1109/42.848183>.
- Jandt, U., Schäfer, D., Grass, M., Rasche, V., 2009. Automatic generation of 3D coronary artery centerlines using rotational X-ray angiography. *Med. Image Anal.* 13 (6), 846–858. <https://doi.org/10.1016/j.media.2009.07.010>.

- Jerman, T., Pernuš, F., Likar, B., Špiclin, Ž., 2015. Beyond Frangi: an improved multiscale vesselness filter. in: *Medical Imaging: Image Processing* 9413, 94132A. <https://doi.org/10.1117/12.2081147>.
- Joe, H., Kim, J., Yu, S.-C., 2020. 3D reconstruction using two sonar devices in a monte-carlo approach for AUV application. *Int. J. Control Autom. Syst.* 18 (3), 587–596.
- Kerschitzki, M., Kollmannsberger, P., Burghammer, M., Duda, G.N., Weinkamer, R., Wagermaier, W., Fratzl, P., 2013. Architecture of the osteocyte network correlates with bone material quality. *J. Bone Miner. Res.* 28 (8), 1837–1845. <https://doi.org/10.1002/jbmr.1927>.
- Law, A.K.W., Zhu, Hui, Chan, F.H.Y., 2004. 3D Reconstruction of Coronary Artery Using Biplane Angiography, 1, pp. 533–536. <https://doi.org/10.1109/iembs.2003.1279790>.
- Lee, T.C., Kashyap, R.L., Chu, C.N., 1994. Building skeleton models via 3-D medial surface Axis thinning algorithms. *CVGIP Graph. Model. Image Process* 56 (6), 462–478. <https://doi.org/10.1006/cgip.1994.1042>.
- Lee, S.J., Choi, H., Hwang, S.S., 2020. Real-time depth estimation using recurrent CNN with sparse depth cues for SLAM system. *Int. J. Control Autom. Syst.* 18 (1), 206–216.
- Li, J., Cohen, L.D., 2011. Reconstruction of 3D tubular structures from cone-beam projections. *Proceedings - International Symposium on Biomedical Imaging* 1162–1166. <https://doi.org/10.1109/ISBI.2011.5872608>.
- Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.W., Heng, P.A., 2018. H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes. *IEEE Trans. Med. Imaging* 37 (12), 2663–2674. <https://doi.org/10.1109/TMI.2018.2845918>.
- Liao, R., Luc, D., Sun, Y., Kirchberg, K., 2010. 3-D reconstruction of the coronary artery tree from multiple views of a rotational X-ray angiography. *Int. J. Cardiovasc. Imaging* 26 (7), 733–749. <https://doi.org/10.1007/s10554-009-9528-0>.
- Liu, X., Hou, F., Hao, A., Qin, H., 2015. A parallelized 4D reconstruction algorithm for vascular structures and motions based on energy optimization. *Vis. Comput.* 31 (11), 1431–1446. <https://doi.org/10.1007/s00371-014-1024-4>.
- Mark, D.B., Berman, D.S., Budoff, M.J., Carr, J.J., Gerber, T.C., Hecht, H.S., Hlatky, M.A., Hodgson, J.M., Lauer, M.S., Miller, J.M., Morin, R.L., Mukherjee, D., Poon, M., Rubin, G.D., Schwartz, R.S., ACCF/ACR/AHA/NASCI/SAIP/SCAI/SCCT, 2010. Expert consensus document on coronary computed tomographic angiography. *Circulation*. 121 (2010), 2509–2543. <https://doi.org/10.1161/cir.0b013e3181d4b618>.
- Nguyen, P.B., Kang, B., Bappy, D.M., Choi, E., Park, S., Ko, S.Y., Park, J.O., Kim, C.S., 2018. Real-time microrobot posture recognition via biplane X-ray imaging system for external electromagnetic actuation. *Int. J. Comput. Assist. Radiol. Surg.* 13 (11), 1843–1852. <https://doi.org/10.1007/s11548-018-1846-z>.
- Romero, M., Interian, Y., Solberg, T., Valdes, G., 2019. Training Deep Learning Models With Small Datasets. *arXiv Prepr. arXiv1912.06761*.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
- Sánchez, J., 2016. Midway video equalization. *Image Process. Line.* 6, 114–129.
- Sasaki, Y., 2015. The truth of the F-measure. *Teach Tutor Mater.*
- Schoonenberg, G.A.F., Garcia, J.A., Carroll, J.D., 2009. Left coronary artery thrombus characterized by a fully automatic three-dimensional gated reconstruction. *Catheter. Cardiovasc. Interv.* 74 (1), 97–100. <https://doi.org/10.1002/ccd.21972>.
- Yang, J., Wang, Y., Liu, Y., Tang, S., Chen, W., 2009. Novel approach for 3-D reconstruction of coronary arteries from two uncalibrated angiographic images. *IEEE Trans. Image Process.* 18 (7), 1563–1572. <https://doi.org/10.1109/TIP.2009.2017363>.
- Yang, J., Cong, W., Chen, Y., Fan, J., Liu, Y., Wang, Y., 2014. External force back-projective composition and globally deformable optimization for 3-D coronary artery reconstruction. *Phys. Med. Biol.* 59 (4), 975. <https://doi.org/10.1088/0031-9155/59/4/975>.
- Zhang, J.-N., Su, Q.-X., Liu, P.-Y., Ge, H.-Y., Zhang, Z.-F., 2019. MuDeepNet: Unsupervised learning of dense depth, optical flow and camera pose using multi-view consistency loss. *Int. J. Control Autom. Syst.* 17 (10), 2586–2596.
- Zheng, S., Meiying, T., Jian, S., 2010. Sequential reconstruction of vessel skeletons from X-ray coronary angiographic sequences. *Comput. Med. Imaging Graph.* 34 (5), 333–345. <https://doi.org/10.1016/j.compmedimag.2009.12.004>.