Engineering 34 (2024) 83–91

Contents lists available at ScienceDirect

Engineering

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

Research
Artificial Intelligence—Article

A Dual-Functional System for the Classification and Diameter Measurement of Aortic Dissections Using CTA Volumes via Deep Learning

Zhihui Huang [a,#], Rui Wang [b,#], Hui Yu [c,#], Yifan Xu [a], Cheng Cheng [a], Guangwei Wang [a], Haosen Cao [a], Xiang Wei [b,*], Hai-Tao Zhang [a,*]

[a] School of Artificial Intelligence and Automation & the MOE Engineering Research Center of Autonomous Intelligent Unmanned Systems & the State Key Laboratory of Intelligent Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China
[b] Division of Cardiovascular Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China
[c] Wuhan Children's Hospital: Wuhan Women and Children Medical Care Center, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430010, China

ARTICLE INFO

Article history:
Received 6 June 2023
Revised 29 September 2023
Accepted 23 November 2023
Available online 4 January 2024

Keywords:
Aortic dissections
Computed tomography angiography
Classification
Deep learning

ABSTRACT

Acute aortic dissection is one of the most life-threatening cardiovascular diseases, with a high mortality rate. Its prevalence ranges from 0.2% to 0.8% in humans, resulting in a significant number of deaths due to being missed in manual examinations. More importantly, the aortic diameter—a critical indicator for surgical selection—significantly influences the outcomes of surgeries post-diagnosis. Therefore, it is an urgent yet challenging mission to develop an automatic aortic dissection diagnostic system that can recognize and classify the aortic dissection type and measure the aortic diameter. This paper offers a dual-functional deep learning system called aortic dissections diagnosis-aiding system (DDAsys) that enables both accurate classification of aortic dissection and precise diameter measurement of the aorta. To this end, we created a dataset containing 61 190 computed tomography angiography (CTA) images from 279 patients from the Division of Cardiovascular Surgery at Tongji Hospital, Wuhan, China. The dataset provides a slice-level summary of difficult-to-identify features, which helps to improve the accuracy of both recognition and classification. Our system achieves a recognition F_1 score of 0.984, an average classification F_1 score of 0.935, and the respective measurement precisions for ascending and descending aortic diameters are 0.994 mm and 0.767 mm root mean square error (RMSE). The high consistency (88.6%) between the recommended surgical treatments and the actual corresponding surgeries verifies the capability of our system to aid clinicians in developing a more prompt, precise, and consistent treatment strategy.

© 2024 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Aortic dissection (AD) is a life-threatening disease [1,2] of the aorta with three layers [3]. Rupture of the intima layer leads to infiltration of blood into the middle layer, and AD occurs [4]. Immediately after symptoms begin, the mortality rate of untreated patients with acute aortic dissection (AAD) increases by 1%–2% per hour [5]. Hence, rapid and precise diagnosis is essential for reducing preoperative time. Successful aorta diameter measurement can be used to prescribe a surgical strategy [6], assisting physicians in making timely and accurate treatment decisions. Several attempts have been made to use machine intelligence to diagnose ADs, focusing on identifying ADs or measuring diameters. Unfortunately, due to the low incidence rate of ADs, at approximately 3–35 cases per 100 000 inhabitants per year [7], there is still a lack of publicly available annotated data. Therefore, it has become an urgently needed mission to design a niche artificial intelligence (AI)-based combination method that can provide complete and accurate surgical recommendations.

Computed tomography angiography (CTA) is an effective method for the identification of ADs [8]. CTA-based machine learning ADs diagnosis has thus been the subject of extensive research over the past few years, with the aim of providing more accurate

* Corresponding authors.
E-mail addresses: xiangwei@tjh.tjmu.edu.cn (X. Wei), zht@mail.hust.edu.cn (H.-T. Zhang).
# These authors contributed equally to this work.

https://doi.org/10.1016/j.eng.2023.11.014
2095-8099/© 2024 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

results [9]. Most of these studies have investigated the segmentation of true and false lumen of specific types of AAD [10], while other researchers have focused on the locating of the aorta [11] and the extraction of the central line [12]. These works have demonstrated the effectiveness of machine learning in terms of extracting aortic contours, paving the way from AI theory and methods to their application in the detection and classification of ADs. Pepe et al. [13] focused on achieving the robust extraction of aortic cross-sections near landmarks using a convolutional neural network (CNN) with DropoutBlock. They also noted that the inter-operator consistency bias in aortic diameter measurements made by different human operators can be up to 5.33 mm, which provides a reference for future researchers in the development of machine measurement algorithms. Unfortunately, the accuracy of the associated diameter measurements has not yet been specified.

In recent years, tremendous development has been achieved in research on the automatic identification of ADs. In a pioneering work, Zhao et al. [2] combined morphological constraints with a deep encoding network to achieve efficient and accurate segmentation of aortas. Hahn et al. [14] developed a five-step segmentation pipeline based on TernausNet [15] to achieve the segmentation of aortas as well as that of the true and false lumen of a specific type of dissection. Chen et al. [16] designed a two-stage deep network consisting of three dense and deconvolution blocks to segment true and false lumen. Furthermore, they proposed an aortic centerline straightening method to improve the accuracy of the segmentation of aorta, thus facilitating subsequent aorta interior feature-extracting algorithms. To achieve automatic recognition of ADs, other scholars [17–20] have performed abnormal detection at the slice level, thereby synthesizing all slices' information to yield a patient-level diagnosis. However, these studies were mostly conducted on small datasets (only 10–20 cases), and more general versions are still ongoing.

Promisingly, Yellapragada et al. [21] designed a CNN to directly diagnose acute aortic syndrome, which includes three kinds of lesions—penetrating aortic ulcer, intramural hematoma, and ADs—at the patient level. Xiong et al. [22] proposed a cascaded multi-task generative framework to detect ADs using non-contrast-enhanced computed tomography with an F_1 score of 0.847, which is still far from ready for practical application. To date, it remains challenging for most existing relevant models to determine the types of detected ADs.

To address this issue, the present study develops a two-stage segmentation network with dilation blocks to construct a diagnosis-aiding system for multicategory ADs. The system is trained on our established dataset, composed of 61 190 images gathered from 279 patients, as shown in Fig. 1. We have named the system the "aortic dissections diagnosis-aiding system (DDAsys)". Considering that some patient cases do not present with an obvious true or false lumen, our proposed DDAsys directly segments the intimal flap, which is the most distinctive feature for diagnosing ADs. By training the model on low-quality images, we have achieved the high-precision recognition and classification of ADs according to potential interferences-based DeBakey classification, which categorizes ADs into three types according to their specific ranges (Fig. 1). In addition, the proposed DDAsys can measure the aortic diameters and thereby assist the choice of both artificial blood vessels and coated stent grafts [23].

Finally, a surgical method is suggested; the possible recommendations are ascending aortic/arch surgery, thoracic endovascular aortic repair (TEVAR) [24], or their combination. Moreover, the measurement of the patient's aortic diameter can provide a basis for the selection of artificial blood vessels and coated stent grafts. This system is expected to pave the way from the deep learning-based CTA image feature recognition method to real applications of AI-aided AD diagnosis.

In brief, the main contributions of this study are three-fold, including: ① providing a dataset consisting of 61 190 images gathered from 279 patients and thereby generating six types of interference features, in order to substantially reduce the number of false positives; ② developing a high-accuracy dual-functional system for ADs diagnosis at the patient and slice level; and ③ achieving a high rate (88.6%) of consistency when comparing the recommended surgery methods of the DDAsys system with the actual methods used.

2. The method of ADs diagnosis

2.1. Data collection and annotation

The self-established ADs dataset contains the retrospectively anonymized data of adult patients undergoing medical examination at the Division of Cardiovascular Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology (HUST), Wuhan, Hubei, China, from 16 January 2014 to 24 December 2020.

The CTA volumes were labeled and categorized into four classes: normal, type I, type II, and type III. Their iconographies are shown in Appendix A Fig. S1, and age and gender information is recorded in Appendix A Table S1. The annotation procedure consisted of two steps:

Step 1: The aortas in the training, validation, and testing scans were pixel-marked by an experienced physician from the Division of Cardiovascular Surgery of Tongji Hospital using the medical image processing software 3D Slicer 4.11 (National Institutes of Health, USA) [25].

Step 2: All the CTA volumes were interpreted by the physician according to DeBakey's classification. For the case used to verify the accuracy of the diameter measurement, it was necessary for the physician to manually measure the diameters of the aortas at intervals of every five slices for the selected sections.

2.2. Overview of the AI system

Two sequential CNNs were used to construct a two-stage network framework for the segmentation of the aorta and intimal flap in CTA scans. The measurement of the aortic diameter was implemented in the segmentation of the aorta, and the speckle-reducing anisotropic diffusion (SRAD) [26,27] algorithm was adopted to denoise the images; the resolution of the images was uniformly resized to 256×256 to match the input layer of the first CNN. Preprocessed images are fed into the first CNN to segment the aorta, whose outputs are converted to 256×256 images as well and then fed into the second CNN to produce the segmentation of the intimal flaps. The outputs of the first network can be divided into three cases, corresponding to different segments of the aorta, as shown in Fig. 2. According to the number and shape of the detected areas, DDAsys can concisely and efficiently determine which segment the current output image belongs to.

To reduce the errors caused by segmentation inaccuracy (e.g., getting two aortic regions in a segment where only a descending aorta exists), we use a sliding window with a size of five slices to screen out the junctions of the aorta's different segments, thereby dividing the whole CTA volume into three segments. Each slice is classified into one of these three segments according to the number and shape of the aortic regions detected within the slice. Once a slice's category does not fit the segment it belongs to, the slice will be discarded. For each CTA volume, we calculate the frequency of dissections in both ascending and descending aortas, respectively, and thereby conduct the subsequent diagnosis. An overview of the DDAsys method is shown in Fig. 2.

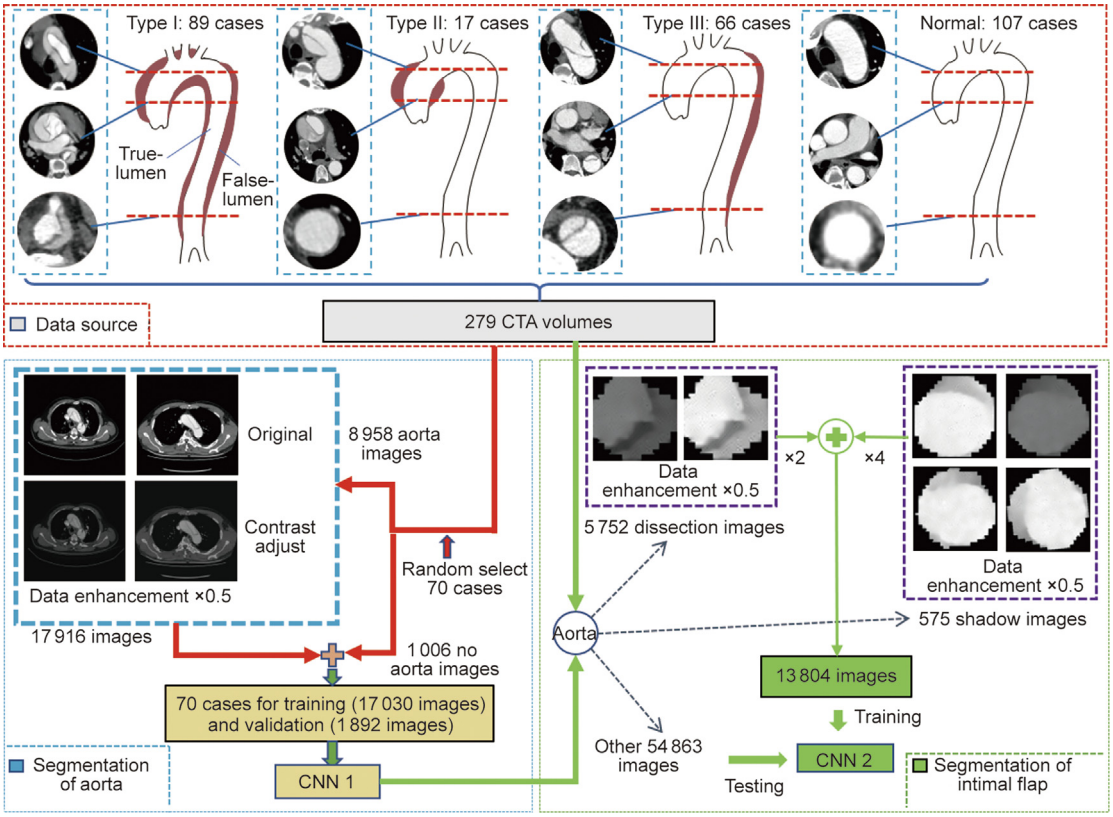

Fig. 1. Sketch map of the datasets. The distribution of each AD category: Two datasets are used for the segmentation of aorta and intimal flap, respectively. Data enhancement is conducted in both datasets (contrast adjustment is used in aorta segmentation, whereas both flip and rotation are conducted in intimal flap segmentation).

Filtered images are input to the second-stage network for the segmentation of intimal flaps. Slices in which intimal flaps are detected are recorded; finally, a classification score is calculated via a sigmoid function.

2.2.1. Two stage-CNN architecture

As shown in Fig. 3, the model uses CTA slice input $\in \mathbb{R}^{256\times256}$ as the model input, which generates the aorta segmentation result $y_1 \in \mathbb{R}^{256\times256}$ and the intimal flaps segmentation result $y_2 \in \mathbb{R}^{256\times256}$.

In order to enhance the accuracy and avoid overfitting, we set 21 hidden layers in the first CNN for aorta segmenting. To prevent gradient vanishing, residual blocks are employed at each block. The ascending and descending aorta are distant from each other, and they are simple in shape (forming a circle). Expanding the receptive field to better extract information on the location of the aorta can prevent other somewhat circular structures from being identified as false positives. Li et al. [28] demonstrated that dilated convolution can effectively expand the receptive field of the network. Therefore, as shown in Fig. 3, we added dilated convolution at the bottleneck of the first CNN [29], which improves the CNN's ability to capture global information in order to locate the aortas more accurately.

The second CNN was established according to the U-Net architecture [30], which adopts standard convolutional layers. Rectified linear unit (ReLU) with dropout is used in both CNNs to increase the convergence rate.

The networks used for aorta and intimal flap segmentation were trained with the Dice loss function. Output layers generate segmentation results with a sigmoid function $1/(1+e^{-x_f})$ that returns the distribution of the targets and backgrounds. Here, x_f denotes the output of the last convolutional layer.

2.2.2. Loss function

In the first stage of extracting the aortic region from the entire scan and the second stage of extracting the intimal flaps from the aortic region, the target occupies only a small part of the entire image. Dice loss [31], which compares the predicted results with the ground truth, can cope well with this situation. Thus, the two sequential CNNs were both trained with Dice loss, as in Refs. [32,33]:

$$\text{Dice loss} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (1)$$

where X and Y represent the predicted results and the ground truth, respectively. No background pixels will be included in the calculation; even if the target is only a small part of the image, Dice loss can return an appropriate loss function to drive the training of the network.

3. Experiment and results

3.1. Training and testing data

The data we used were collected at Tongji Hospital; they consisted of 279 CTA volumes—including 89 type I, 17 type II, 66 type III, and 107 normal cases—of which 137 were used for testing. We constructed two image datasets according to these CTA volumes for aorta segmentation and intimal flap segmentation. The aorta segmentation dataset for training consists of 1006 + 8958 × 2 = 18 922 images from 70 CTA volumes, among which 8958 images were sampled from aortic segments. Considering the different dosages of contrast-enhancement agents used in patient examinations, we adjusted the image brightness to half for data enhancement. The other 1006 images were sampled outside the aortic area to prevent the misidentification of other structures as the aorta. A sketch map is provided in Fig. 1.

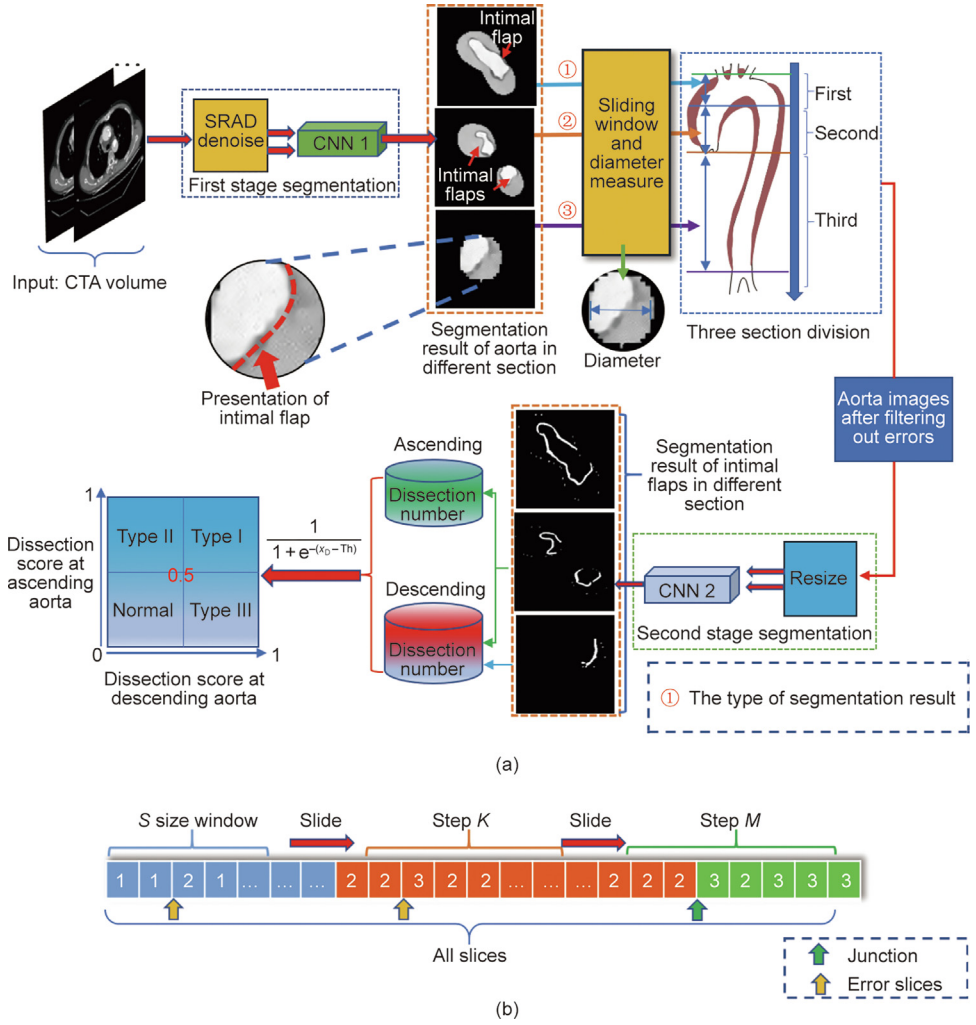

Fig. 2. The identification procedure. (a) Overview of the approach: After the first segmentation stage, each slice is categorized into three types according to features of the segmented areas, corresponding to three segments of the aorta: ① an aortic arch, ② coexistence of an ascending aorta and descending aorta, and ③ a descending aorta only. The value of threshold (Th) is 10, which is obtained by evaluating the point of inflexion on the receiver operating characteristic (ROC) curves. A detailed description of Th can be found in the second paragraph of Section 3.3. The red arrows point to the location of intimal flaps. (b) Illustration of the sliding window. Slices resembling an "8" shape are categorized as class 1. Slices with double- and single-circular patterns are categorized as classes 2 and 3, respectively. All other cases are considered to be errors. These three different classifications correspond to the cross-sectional characteristics of the aortic arch, ascending aorta, and descending aorta, respectively. SRAD: speckle-reducing anisotropic diffusion; x_D: the number of CTA slice with AD; S: the size of sliding window; K and M: the number of steps for sliding window, respectively.

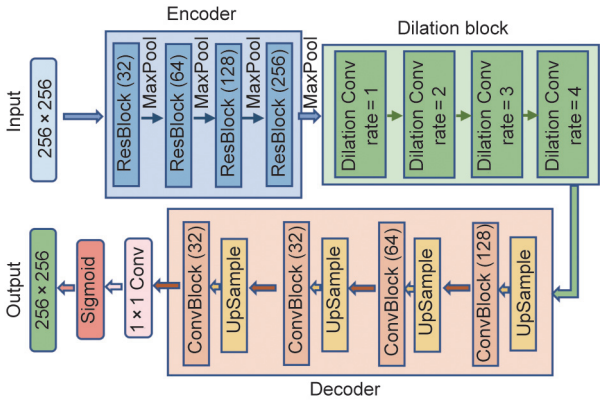

Fig. 3. Network architecture of the aorta segmentation network. The numbers inside the brackets of the figure denote the number of channels for each convolutional module. Different dilated rate dilation convolutions are used to enhance the localization capability of the network to avoid the identification of other somewhat circular structures as aorta. All convolutions other than the 1 × 1 convolution are set to stride 1 and kernel size 3. We cascade two convolutions to form the ConvBlock. Conv: convolution.

The second stage of the proposed approach is to segment the intimal flap inside the aorta. We fed all the images to CNN 1 to extract the aortas and uniformly resized the images to a resolution of 256 × 256 as the data sources. After filtering and removing incorrect segmentation results, we obtained 61 190 images. These images were classified into two categories: AD (positive) and normal (negative). Motivated by Han et al.'s work [34], in which a semi-supervised approach is used to train the segmentation network, we used the Hessian matrix (HM)-based method [18] to perform a preliminary detection of the intimal flaps on each image; 5752 images, matching the ground truth, were selected as part of the training set at this stage, and the associated results were used as their labels. In addition, 575 negative images with shadows near the edges were added to the training set. The remaining 54 863 images were used for testing at the slice level. The 5752 and 575 images used for training were partially obtained from the 70 samples previously used for CNN 1 and partially obtained from an additional 72 samples. Therefore, a total of 142 cases were used for training, and the remaining 137 cases were used for testing. Data enhancement (Fig. 1) was conducted on the training images; in addition to halving the brightness adjustment, we rotated the

image counterclockwise by 90°. In this way, a training set comprising 13 804 images was constructed.

The remaining 137 patients were used for testing. We sampled 54 863 images from these cases to assess the network's ability to recognize difficult-to-identify images at the slice level. Six imageology features that induce identification errors were generalized. A more detailed distribution of the 54 863 images at the slice level was obtained according to their features in iconography, as shown in Fig. 4.

It should be noted that 19 type I and 26 type III cases were used to evaluate the diameter measurement accuracy via the root mean square error (RMSE). At the end, the system recommends a surgical method, which is ascending aortic/arch surgery, TEVAR [35], or their combination. The feasibility of the recommended surgical method according to the proposed DDAsys was verified on 65 patients undergoing 65 surgeries.

We used the deep learning CNNs (as shown in Fig. 2) with the optimal performance in the validation set as the final model. The aorta segmentation and the intimal flap in the final model achieved the Dice coefficients [36] of 0.958 and 0.913, respectively.

3.2. Statistical analysis

The precision, sensitivity and recall, specificity, F_1 score, and area under the curve (AUC) of the receiver operating characteristic (ROC) curve, with two-sided 95% confidence intervals (CIs), were used to quantify the system's performance in AD recognition [37]. The precision, sensitivity and recall, specificity, and F_1 score of each class are related to the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) rates [38], which have the following formulas:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Sensitivity/recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{4}$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Sensitivity/recall}}{\text{Precision} + \text{Sensitivity/recall}} \tag{5}$$

In addition, we employed an average F_1 score to evaluate the overall classification performance across all classes of data, as follows:

$$\bar{F_1} = \frac{1}{N_c}\sum_{n=1}^{N_c} F_{1n} \tag{6}$$

where N_c is the number of classes.

The RMSE and mean absolute error (MAE) were used to evaluate the accuracy of the diameter measurements:

$$\text{RMSE} = \sqrt{\frac{1}{N_d}\sum_{i=1}^{N_d}\left(D_i - \widehat{D}_i\right)^2} \tag{7}$$

$$\text{MAE} = \frac{1}{N_d}\sum_{i=1}^{N_d}\left|D_i - \widehat{D}_i\right| \tag{8}$$

where D_i is the actual average diameter of the ith test cases, $\widehat{D}_i$ is the measured average diameter of the ith test cases, and N_d is the number of test cases.

3.3. Performance at the patient level

We evaluated DDAsys in terms of both recognition and dissections classification; the results are provided in Tables 1 and 2, respectively. It can be seen that DDAsys achieves an F_1 score of 0.984, a sensitivity of 0.980, and a specificity of 0.976 for recognition. To further evaluate the effectiveness of DDAsys, we compared it with a commonly used deep classification network, ResNet18 [39]. The results show that DDAsys outperforms ResNet18 in terms of both higher recognition and higher classification accuracies of ADs at the patient level due to its substantially increased true positive rate at the slice level (Table 3).

A sigmoid function $1/(1+e^{-(x_D-Th)})$(where x_D is the number of CTA slice with AD, and Th represents the threshold for patient-level diagnosis) was used to transform the number of abnormal

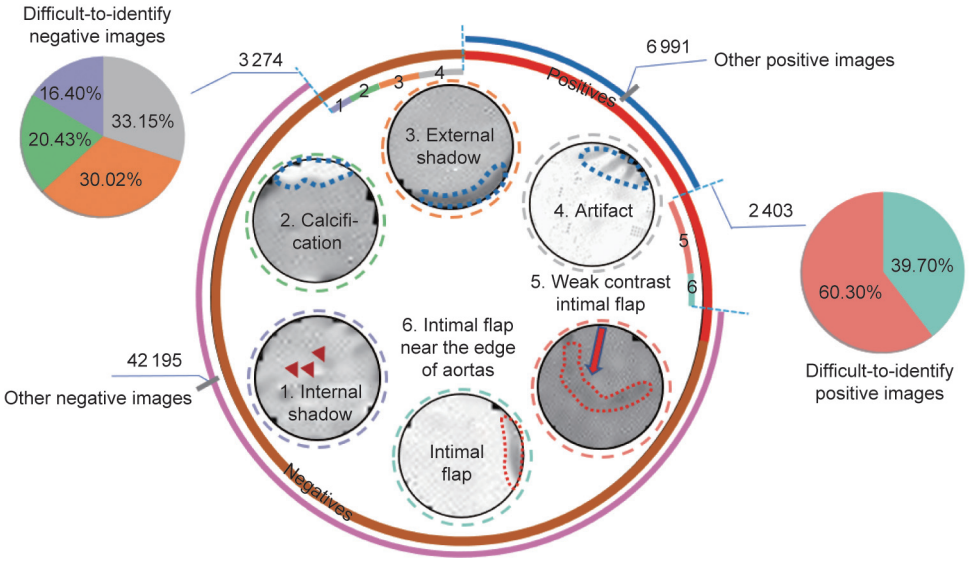

Fig. 4. Distribution of positive and negative images. The middle ring represents the proportions of positive to negative images. Six kinds of images (1. internal shadow, 2. calcification, 3. external shadow, 4. artifact, 5. weak contrast intimal flap, and 6. intimal flap near the edge of aortas), which are difficult to identify, are shown in the center of the ring. The dotted red lines represent the locations of intimal flaps that are difficult to detect due to weak contrast or imperceptible position; the dotted blue lines represent the difficult-to-identify features produced by imaging manifestations similar to intimal flap; the red arrows point to the location of the internal shadow.

Table 1
Performance comparison of the proposed DDAsys method and ResNet18 regarding recognition of ADs and normal cases.

Class of cases	DDAsys sensitivity	DDAsys specificity	DDAsys precision	DDAsys F_1 score	ResNet18 F_1 score
Normal cases	0.979 (0.975, 0.984)	0.977 (0.975, 0.980)	0.956 (0.951, 0.962)	0.967 (0.964, 0.971)	0.841
Dissection cases	0.980 (0.977, 0.983)	0.976 (0.972, 0.980)	0.988 (0.986, 0.990)	0.984 (0.982, 0.986)	0.899

Table 2
Classification rate comparison of DDAsys and ResNet18 for three types of AD.

Type of cases	DDAsys sensitivity	DDAsys specificity	DDAsys AUC	DDAsys F_1 score	ResNet18 F_1 score
Type I	0.971 (0.965–0.977)	0.980 (0.978–0.983)	0.959 (0.953–0.964)	0.956 (0.951–0.961)	0.793
Type II	0.857 (0.845–0.895)	0.992 (0.991–0.994)	0.855 (0.833–0.877)	0.857 (0.845–0.886)	0.333
Type III	0.941 (0.935–0.947)	0.988 (0.986–0.990)	0.942 (0.936–0.948)	0.960 (0.956–0.963)	0.813
Normal	0.978 (0.973–0.981)	0.978 (0.975–0.981)	0.966 (0.962–0.971)	0.967 (0.963–0.970)	0.841

Table 3
Comparison of the sensitivity of DDAsys, HM, and ResNet18 in terms of recognition of ADs.

Method	Weak contrast intimal flap	Intimal flap near the edge of aortas	Other intimal flaps
DDAsys	29.9%	39.6%	100.0%
HM	30.4%	50.5%	99.6%
ResNet18	24.4%	6.9%	67.4%

slices in the ascending and descending aortas into dissection scores, thereby producing a diagnosis result at the patient level (Fig. 2). By adjusting the thresholds for the ascending and descending aortas, respectively, we obtained the ROCs of different classes in order to further evaluate the classification performance of DDAsys, as shown in Fig. 5(a). The classification sensitivities reached 97.1%, 85.7%, 94.1%, and 97.8%, respectively (Fig. 5(b)). The AUCs for types I–III and normal were 0.959 (95% CIs 0.953–0.964), 0.855 (0.833–0.877), 0.942 (0.936–0.948), and 0.966 (0.962–0.971), respectively (Fig. 5(a)). We evaluated the point of inflexion on the ROC curves so as to yield the threshold.

The rapid rise of the curves in Fig. 5(a) indicates that DDAsys is able to maintain high sensitivity and specificity at the same time. The performance soon reaches a bottleneck, and changes in thresholds no longer affect it; this indicates that there are fewer false positives or false negatives.

3.4. Recognition performance at the slice level

Six imageology features at the slice level that induce identification errors were generalized (Fig. 4); moreover, in the training set, the proposed algorithm considers the external shadow feature (which usually appears along the aortic edge but originates from the connective tissue). This feature helps prevent the identification of any anomalies as intimal flaps, which reduces the potential errors caused by these six features and avoids overfitting while ensuring true dissection detection. Subsequent experiments demonstrated the precision of DDAsys at the slice level, further revealing its performance on these obscure features.

Four imageology features are often encountered in false positive samples: internal shadows, calcification, artifacts, and external shadows. Images with artifacts have similar imageology characteristics with intimal flap. External shadows in images are often misidentified as false lumen and are hence prone to incur false positives. Moreover, internal shadows and calcification may cause noticeable gray value differences in the image, which usually appear between normal aortas and dissection-bearing aortas and hence intensify the challenge of recognition.

Table 4 shows the performance of the proposed DDAsys method in identifying the four features. DDAsys achieves an average sensitivity of 93.8% (with the highest and lowest sensitivity rates at 98.9% and 90.3%, respectively), which ensures the accuracy of subsequent dissection classifications.

The other two features leading to false negatives are considered here—that is, the weak contrast intimal flaps and the intimal flaps near the edge of aortas—whose imageologies are shown in Fig. 4. The corresponding results are shown in Table 3, which indicates the arduousness of achieving a high recognition accuracy with both factors.

The performance of DDAsys at the slice level was further compared with those of HM and ResNet18, respectively. Compared with HM, DDAsys has a higher average sensitivity on negatives, at 95.0% in comparison with 75.3% (Table 4).

Although ResNet18 slightly outperforms DDAsys for negatives, ResNet18's cost of accuracy loss in detecting positives is even larger than that of DDAsys (Table 3); hence, DDAsys still performs better, with a higher average accuracy in intimal flaps detection of 56.5% in comparison with 32.9% for ResNet18. The high average precision on both positive and negative slices explains DDAsys's high average accuracy at the patient level because it is not rational to rashly discard the positive slices, whereas a mass of false positives will lead to both identification and classification errors, which will lead to a decrease in the recognition accuracy at the patient level.

The results in Tables 3 and 4 show the advantages of DDAsys over HM and ResNet18. These advantages are because DDAsys is based on deep convolutional networks, and because part of its training data is extracted using the HM algorithm, which enables a balance of sensitivity and specificity.

The distribution of the features extracted by DDAsys is shown in Fig. 5(c). Apart from a few misclassification results, there are separating boundaries among different classes of case clusters. Moreover, most of the misclassification results orient to adjacent vertical areas. This implies that misjudgments usually occur in the detection of ascending ADs, which may be a key to further improving accuracy in subsequent studies.

3.5. Diameter measurement and surgical methods recommendation

Furthermore, according to the extracted section of the aorta, the aorta's diameter can be measured by DDAsys as well. For the ascending aorta, we calculated the mean diameter from the root of the aorta to the aortic arch. For the descending aorta, we detected the section between the aortic arch and the abdominal aorta. The results are provided in Table 5.

In order to verify the value of the system in clinical application, DDAsys provides a recommendation to the cardiologist to adopt ascending aortic/arch surgery and antegrade TEVAR for patients with type I and ascending aortic/arch surgery for type II dissection. In contrast, for patients with type III dissection, TEVAR will be recommended. We tested the DDAsys's recommended surgical

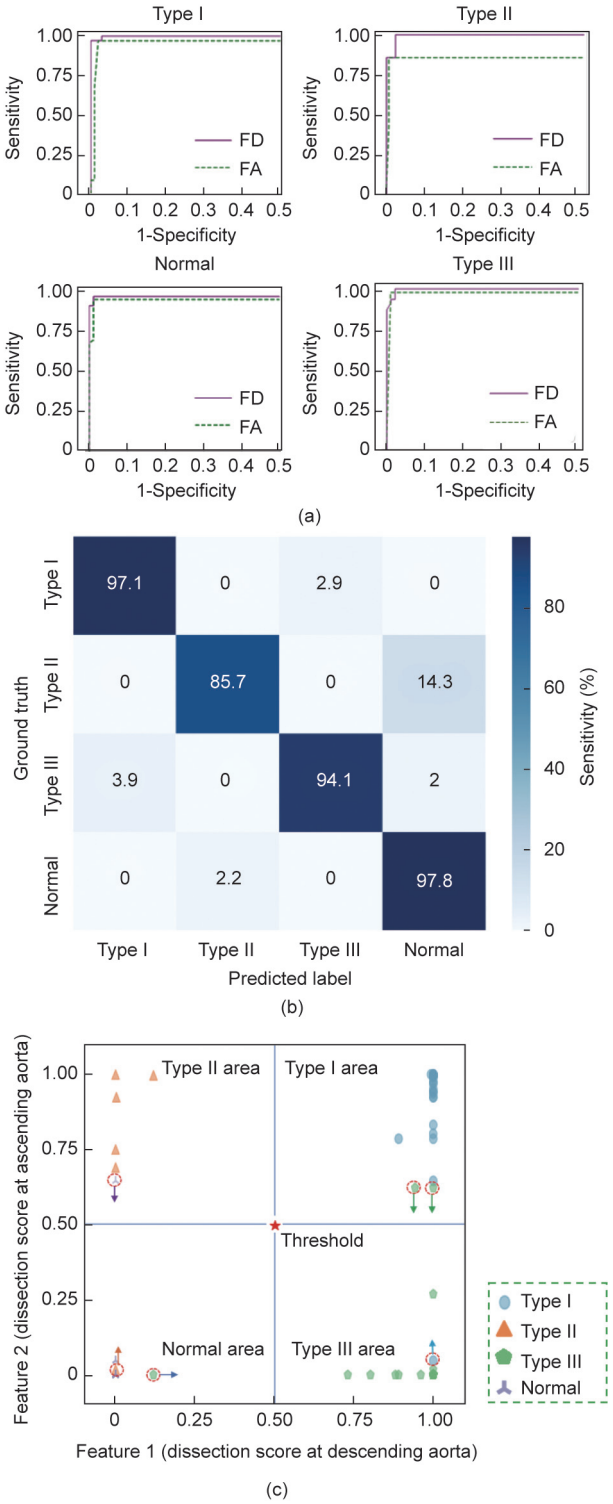

Fig. 5. The performance of DDAsys. (a) The ROC of different classes. Each class has two curves, which are obtained by adjusting the thresholds of the ascending and descending aortas, respectively. The average AUC of the two curves is used for the evaluation. (b) Confusion matrix on the classification. Values greater than 0 for non-shaded fields show misclassification. (c) Feature distribution of the test cases. Red dotted circles represent the misclassification results, and arrows indicate the category aggregation region to which they belong. Different classes of cases have clustered in different areas. FA: fix the threshold of the ascending aorta; FD: fix the threshold of the descending aorta.

treatments on 65 patients, comprising seven cases in which the cardiologist adopted ascending aortic/arch surgery and antegrade TEVAR, four cases in which ascending aortic/arch surgery was performed, and 54 cases in which TEVAR was performed. The recommendations given by DDAsys reached an average consistency of 88.6% with the actual surgical treatments performed by experienced clinicians (Table S2 in Appendix A).

4. Discussion

In this study, we propose a dual-functional system that aims to assist non-cardiologists in diagnosing ADs. Our system demonstrates high accuracy in both classification and diameter measurement. In this way, patients can be classified and treated in a timely manner. The results in Tables 3 and 4 indicate that traditional image-processing algorithms and deep classification networks cannot maintain a high level of accuracy on both positive and negative challenging samples, and the accuracy of our system on positive challenging images is still insufficient, as shown in Fig. 6. We performed a detailed analysis of CTAs at the slice level, and challenging samples have been included in our dataset, which helps in facilitating the analysis of indistinguishable dissection situations and in breaking through performance bottlenecks.

Artifacts [40,41] often appear around the aorta because the present dataset was collected without electrocardiogram (ECG) gating. Experienced physicians can distinguish artifacts, but for AI-aided diagnostic systems, such interference inevitably increases the probability of misjudgment. In addition, due to the variances in the dosages of angiographic contrast enhancer used by different patients, the contrast of CTA volumes varies considerably. Our deep learning model can distill the experiences of physicians in dealing with artifacts and extract high-order features of CTA images in such a way that the model's effectivity for diverse images can be expected to be considerably enhanced.

It should be noted that DDAsys is capable of identifying abnormalities at the slice level, thereby enabling comprehensive analysis of the entire CTA volume. DDAsys maintains a low false positive rate at the slice level, which improves the sensitivity of diagnosing aortic abnormalities. We compared the number of false positive slices detected by DDAsys and HM in some cases, as presented in Fig. 7. The results show the average values $\overline{\mathrm{FP}_T} = \left(\sum_j^{N_T} \mathrm{FP}_j\right)/N_T$ of false positives in different aortic segments that should not contain positive slices for different types of samples, where T denotes the type of cases, N_T denotes the number of cases used for comparison for the Tth type, and FP_j represents the number of false positive slices detected by the algorithm in the ascending aorta or descending aorta of the jth sample. The number of false positives detected by DDAsys at each aortic segment is significantly lower than that detected by HM. Since type II has the lowest number, its mean $\overline{\mathrm{FP}_T}$ has the most significant impact on a large number of false positives. Correspondingly, in Fig. 7, both HM and DDAsys have the highest average false positives for type II.

Importantly, DDAsys's refined DeBakey recognition F_1 score of 0.984 and average classification F_1 score of 0.935 can aid clinicians in choosing appropriate surgical methods. To validate our system's diagnostic advice capability, we examined the actual surgical methods used on 65 test patients, seven of whom underwent ascending aorta replacement, 54 underwent TEVAR and four cases underwent their combination. The average consistency between the recommended surgical methods and those actually applied by experienced clinicians was 88.6%. Thus, DDAsys can not only effectively reduce the workload of non-cardiologist physicians (e.g., those in the intensive care unit (ICU)) but also provide valuable treatment recommendations to surgeons. Moreover, DDAsys provides highly precise measurements of the average aortic diameter, enabling a thorough analysis of a greater number of slices within the same timeframe. Therefore, DDAsys holds the potential to aid surgeons in accurately selecting the appropriate artificial

Table 4
Recognition comparison among DDAsys, HM, and ResNet18 for negative images.

Method	Identification sensitivity of the images (%)				
	Internal shadows	Calcification	Artifact	External shadow	Other negative images
DDAsys	92.2	98.9	93.6	90.3	99.9
HM	64.1	71.0	67.9	74.3	99.0
ResNet18	98.5	100.0	98.3	99.1	99.9

Table 5
Diameter measuring accuracy.

Aortic trunks	RMSE (mm)	MAE (mm)
Ascending	0.994	0.797
Descending	0.767	0.541

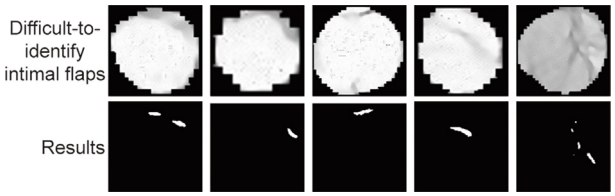

Fig. 6. Examples of the segmentation of challenging slices. The intimal flaps in the first three examples are all located at the edge of the aortas, while the last two have weak contrast.

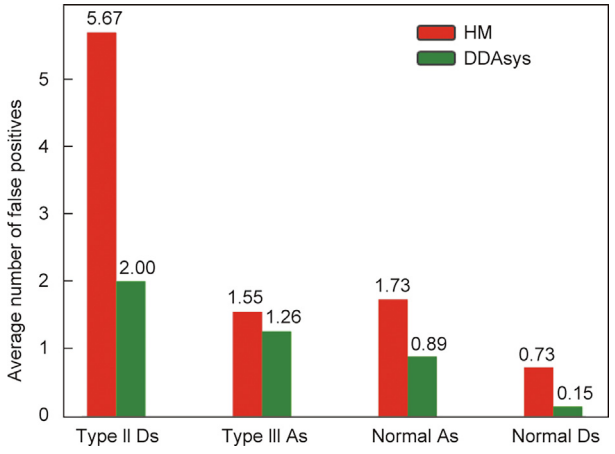

Fig. 7. Comparison of the HM-based method and DDAsys in terms of the numbers of false positives. The average number of false positives (the smaller the better) on the sections which should not contain positive slices: ① descending aorta (Ds) in cases of type II, ② ascending aorta (As) in cases of type III, ③ ascending aorta in normal cases, and ④ descending aorta in normal cases are calculated.

blood vessel and coated stent graft during surgery, thereby reducing the duration of intraoperative circulation stoppage and enhancing surgical outcomes.

The DDAsys presented here can be expected to help meet the efficiency and accuracy requirements of clinical applications. However, this study has several limitations. For images with weak contrast or those in which the intimal flap is located at the aortic edge, DDAsys may not be precise enough to fully avoid misidentification. Furthermore, the system can only offer recommendations based on the DeBakey classification, and thus lacks analysis of the breach location, aortic root condition, and the blood supply to vital branches. Future research should focus on enhancing the comprehensiveness of the data distribution and incorporating inter-center analysis.

5. Conclusions and future work

This study presents a dual-functional deep learning system, DDAsys, that aids in the recognition, classification of ADs, and the measurement of aortas, achieving high accuracy at both the slice and patient levels. The system identifies six types of interfering features, significantly reducing false positives. Moreover, it offers practical recommendations for clinicians regarding surgical methods, as well as the selection of artificial blood vessels and coated stent grafts. This could be a valuable reference for clinicians with patients undergoing cardiac surgery, potentially reducing operation times. Future work includes further increasing the diagnostic accuracy with weak contrast and intimal flaps close to the edge of the aortas. We also hope to implement the detection of aortic ulcers and intramural hematomas to achieve early warning.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (U2141235, 62225306, 82070488, and 82000440), the Natural Science Foundation of Hubei Province (2021CFB084), and the Guangdong Basic and Applied Research Foundation (2022B1515120069).

Compliance with ethics guidelines

Zhihui Huang, Rui Wang, Hui Yu, Yifan Xu, Cheng Cheng, Guangwei Wang, Haosen Cao, Xiang Wei, and Hai-Tao Zhang declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

The test dataset used to support the findings of this study is publicly available at Mendeley Data, V1(doi: 10.17632/5b3wx7mkk6.1, 10.17632/4yfcyv5xsk.1, and 10.17632/2yxbmchmgg.1). Any data use will be restricted to noncommercial research purposes.

The total images in slice level could be visited through https://pan.baidu.com/s/1gcCAw5ZPbNP5tR4dE0n45w with code w7fi.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2023.11.014.

References

[1] Kamel H, Roman MJ, Pitcher A, Devereux RB. Pregnancy and the risk of aortic dissection or rupture: a cohort-crossover analysis. Circulation 2016;134(7):527–33.
[2] Zhao J, Zhao J, Pang S, Feng Q. Segmentation of the true lumen of aorta dissection via morphology-constrained stepwise deep mesh regression. IEEE Trans Med Imaging 2022;41(7):1826–36.
[3] He D, Mao A, Zheng CB, Kan H, Zhang K, Zhang Z, et al. Aortic heterogeneity across segments and under high fat/salt/glucose conditions at the single-cell level. Natl Sci Rev 2020;7(5):881–96.
[4] Nguyen CT, Hall CS, Wickline SA. Characterization of aortic microstructure with ultrasound: implications for mechanisms of aortic function and dissection. IEEE Trans Ultrason Ferroelectr Freq Control 2002;49(11):1561–71.

[5] Isselbacher EM, Preventza O, Black III JH, Augoustides JG, Beck AW, Bolen MA, et al. 2022 ACC/AHA guideline for the diagnosis and management of aortic disease: a report of the American heart association/American college of cardiology joint committee on clinical practice guidelines. J Am Coll Cardiol 2022;80(24):e223–393.
[6] Cheung WK, Bell R, Nair A, Menezes LJ, Patel R, Wan S, et al. A computationally efficient approach to segmentation of the aorta and coronary arteries using deep learning. IEEE Access 2021;9:108873–88.
[7] Brunet J, Pierrat B, Badel P. A parametric study on factors influencing the onset and propagation of aortic dissection using the extended finite element method. IEEE Trans Biomed Eng 2021;68(10):2918–29.
[8] Nazerian P, Mueller C, Soeiro AM, Leidel BA, Salvadeo SAT, Giachino F, et al. Diagnostic accuracy of the aortic dissection detection risk score plus D-dimer for acute aortic syndromes: the ADvISED prospective multicenter study. Circulation 2018;137(3):250–8.
[9] Rong G, Mendez A, Assi EB, Zhao B, Sawan M. Artificial intelligence in healthcare: review and prediction case studies. Engineering 2020;6(3):291–301.
[10] Pepe A, Li J, Rolf-Pissarczyk M, Gsaxner C, Chen X, Holzapfel GA, et al. Detection, segmentation, simulation and visualization of aortic dissections: a review. Med Image Anal 2020;65:101773.
[11] Avila-Montes OC, Kurkure U, Nakazato R, Berman DS, Dey D, Kakadiaris IA. Segmentation of the thoracic aorta in noncontrast cardiac CT images. IEEE J Biomed Health Inform 2013;17(5):936–49.
[12] Zhao J, Feng Q. Automatic aortic dissection centerline extraction via morphology-guided CRN tracker. IEEE J Biomed Health Inform 2021;25(9):3473–85.
[13] Pepe A, Egger J, Codari M, Willemink MJ, Gsaxner C, Li J, et al. Automated cross-sectional view selection in CT angiography of aortic dissections with uncertainty awareness and retrospective clinical annotations. 2021. arXiv:2111.11269.
[14] Hahn LD, Mistelbauer G, Higashigaito K, Koci M, Willemink MJ, Sailer AM, et al. CT-based true- and false-lumen segmentation in type B aortic dissection using machine learning. Radiol Cardiothorac Imaging 2020;2(3):e190179.
[15] Vladimir I, Alexey S. TernausNet: U-Net with VGG11 encoder pre-trained on imageNet for image segmentation. 2018. arXiv:1801.05746.
[16] Chen D, Zhang X, Mei Y, Liao F, Xu H, Li Z, et al. Multi-stage learning for segmentation of aortic dissections using a prior aortic anatomy simplification. Med Image Anal 2021;69:101931.
[17] Gayhart M, Arisawa H. Automated detection of healthy and diseased aortae from images obtained by contrast-enhanced CT scan. Comput Math Methods Med 2013;2013:107871.
[18] Dehghan E, Wang H, Syeda-Mahmood T. Automatic detection of aortic dissection in contrast-enhanced CT. In: Proceedings of the 14th International Symposium on Biomedical Imaging (ISBI); 2017 Apr 18–21; Melbourne, VIC, Australia. New York City: IEEE; 2017. p. 557–60.
[19] Xu X, He Z, Niu K, Zhang Y, Tang H. An automatic detection scheme of acute Stanford type A aortic dissection based on DCNNs in CTA images. In: Proceedings of the 4th International Conference on Multimedia Systems and Signal Processing (ICMSSP); 2019 May 10–12; Guangzhou, China; 2019. p. 16–20.
[20] Cheng J, Tian S, Yu L, Ma X, Xing Y. A deep learning algorithm using contrast-enhanced computed tomography (CT) images for segmentation and rapid automatic detection of aortic dissection. Biomed Signal Process Control 2020;62:102145.
[21] Yellapragada MS, Xie Y, Graf B, Richmond D, Krishnan A, Sitek A. Deep learning based detection of acute aortic syndrome in contrast CT images. In: Proceedings of the 17th International Symposium on Biomedical Imaging (ISBI); 2020 Apr 3–7; Iowa City, IA, USA. New York City: IEEE; 2020. p. 1474–7.
[22] Xiong X, Ding Y, Sun C, Zhang Z, Guan X, Zhang T, et al. A cascaded Multi-task generative framework for detecting aortic dissection on 3D non-contrast-enhanced computed tomography. IEEE J Biomed Health Inform 2022;26(10):5177–88.
[23] Wörz S, von Tengg-Kobligk H, Henninger V, Rengier F, Schumacher H, Böckler D, et al. 3-D quantification of the aortic arch morphology in 3-D CTA data for endovascular aortic repair. IEEE Trans Biomed Eng 2010;57(10):2359–68.
[24] Wang J, Zhao J, Ma Y, Huang B, Yuan D, Yang Y. Midterm prognosis of type B aortic dissection with and without dissecting aneurysm of descending thoracic aorta after endovascular repair. Sci Rep 2019;9(1):8870.
[25] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 2012;30(9):1323–41.
[26] Yu Y, Acton ST. Speckle reducing anisotropic diffusion. IEEE Trans Image Process 2002;11(11):1260–70.
[27] Krissian K, Westin CF, Kikinis R, Vosburgh KG. Oriented speckle reducing anisotropic diffusion. IEEE Trans Image Process 2007;16(5):1412–24.
[28] Li J, Yu ZL, Gu Z, Liu H, Li Y. Dilated-Inception Net: multi-scale feature aggregation for cardiac right ventricle segmentation. IEEE Trans Biomed Eng 2019;66(12):3499–508.
[29] Vesal S, Ravikumar N, Maier A. A 2D dilated residual U-Net for multi-organ segmentation in thoracic CT. 2018. arXiv:1905.07710.
[30] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2015 Oct 5–9; Munich, Germany; 2015. p. 234–41.
[31] Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26(3):297–302.
[32] Golla AK, Bauer DF, Schmidt R, Russ T, Norenberg D, Chung K, et al. Convolutional neural network ensemble segmentation with ratio-based sampling for the arteries and veins in abdominal CT scans. IEEE Trans Biomed Eng 2021;68(5):1518–26.
[33] Shi F, Yang Q, Guo X, Qureshi TA, Tian Z, Miao H, et al. Intracranial vessel wall segmentation using convolutional neural networks. IEEE Trans Biomed Eng 2019;66(10):2840–7.
[34] Han K, Liu L, Song Y, Liu Y, Qiu C, Tang Y, et al. An effective semi-supervised approach for liver CT image segmentation. IEEE J Biomed Health Inform 2022;26(8):3999–4007.
[35] Bodell BD, Taylor AC, Patel PJ. Thoracic endovascular aortic repair: review of current devices and treatments options. Tech Vasc Interv Radiol 2018;21(3):137–45.
[36] Preetha CJ, Meredig H, Brugnara G, Mahmutoglu MA, Foltyn M, Isensee F, et al. Deep-learning-based synthesis of post-contrast T1-weighted MRI for tumour response assessment in neuro-oncology: a multicentre, retrospective cohort study. Lancet Digit Health 2021;3(12):e784–94.
[37] Xu X, Jiang X, Ma C, Du P, Li X, Lv S, et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering 2020;6(10):1122–9.
[38] Zhu H, Cheng C, Yin H, Li X, Zuo P, Ding J, et al. Automatic multilabel electrocardiogram diagnosis of heart rhythm or conduction abnormalities with deep learning: a cohort study. Lancet Digit Health 2020;2(7):e348–57.
[39] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR); 2016 Jun 26–Jul 1; Las Vegas, NV, USA. IEEE; 2016. p. 770–8.
[40] Chin AS, Fleischmann D. State-of-the-art computed tomography angiography of acute aortic syndrome. Semin Ultrasound CT MR 2012;33(3):222–34.
[41] Liu S, Wang Y, Yang X, Lei B, Liu L, Li SX, et al. Deep learning in medical ultrasound analysis: a review. Engineering 2019;5(2):261–75.