



## Simultaneous vessel segmentation and unenhanced prediction using self-supervised dual-task learning in 3D CTA (SVSUP)



Wenjian Huang<sup>a,1</sup>, Weizheng Gao<sup>a</sup>, Chao Hou<sup>b</sup>, Xiaodong Zhang<sup>b</sup>, Xiaoying Wang<sup>a,b,\*</sup>, Jue Zhang<sup>a,c,\*</sup>

<sup>a</sup> Academy for Advanced Interdisciplinary Studies, Peking University, No.5 Yiheyuan Rd., Beijing, 100871, China

<sup>b</sup> Department of Radiology, Peking University First Hospital, No.8 Xishiku Street, Beijing, 100034, China

<sup>c</sup> College of Engineering, Peking University, No.5 Yiheyuan Rd., Beijing, 100871, China

### ARTICLE INFO

#### Article history:

Received 23 August 2021

Revised 5 June 2022

Accepted 1 July 2022

#### Keywords:

Head-neck CTA  
Dual-task learning  
All-vessel segmentation  
Unenhanced prediction  
Self-supervised learning  
Virtual unenhanced image

### ABSTRACT

**Background and objective:** The vessel segmentation in CT angiography (CTA) provides an important basis for automatic diagnosis and hemodynamics analysis. Virtual unenhanced (VU) CT images obtained by dual-energy CT can assist clinical diagnosis and reduce radiation dose by obviating true unenhanced imaging (UECT). However, accurate segmentation of all vessels in the head-neck CTA (HNCTA) remains a challenge, and VU images are currently not available from conventional single-energy CT imaging. **Methods:** In this paper, we proposed a self-supervised dual-task deep learning strategy to fully automatically segment all vessels and predict unenhanced CT images from single-energy HNCTA based on a developed iterative residual-sharing scheme. The underlying idea was to use the correlation between the two tasks to improve task performance while avoiding manual annotation for model training. **Results:** The feasibility of the strategy was verified using the data of 24 patients. For vessel segmentation task, the proposed model achieves a significantly higher average Dice coefficient (84.83%,  $P$ -values  $10^{-3}$  in paired  $t$ -test) than the state-of-the-art segmentation model, vanilla VNet (78.94%), and several popular 3D vessel segmentation models, including Hessian-matrix based filter (62.59%), optically-oriented flux (66.33%), spherical flux model (66.91%), and deep vessel net (66.47%). For the unenhanced prediction task, the average ROI-based error compared to the UECT in the artery tissue is  $6.1 \pm 4.5$  HU, similar to previously reported  $6.4 \pm 5.1$  HU for VU reconstruction. **Conclusions:** Results show that the proposed dual-task framework can effectively improve the accuracy of vessel segmentation in HNCTA, and it is feasible to predict the unenhanced image from single-energy CTA, providing a potential new approach for radiation dose saving. Moreover, to our best knowledge, this is the first reported annotation-free deep learning-based full-image vessel segmentation for HNCTA.

© 2022 Elsevier B.V. All rights reserved.

### 1. Introduction

Cerebrovascular diseases (CBVDs) are among the leading causes of death worldwide [1]. Computed tomography angiography (CTA) is a routine imaging method for the detection of cerebral atherosclerosis, arterial stenosis, subarachnoid hemorrhage (SAH), and other vascular anomalies in the head and neck arteries [2–4]. The CTA imaging of the intracranial and extracranial vessels even have largely replaced gold standard digital subtraction angiography (DSA) for diagnostics of many vascular diseases [3–7] since DSA is time-consuming, expensive and requires invasive arterial access

with the risk of neurologic complications [3]. In CTA images, the vascular regions are of high intensity due to the presence of iodine contrast agent, whereas in plain CT images (unenhanced computed tomography, UECT), the vascular regions are not enhanced.

Blood vessel segmentation from the CTA image is essential for the diagnosis and analysis of vascular diseases. For example, the automatic quantification of stenosis and hemodynamic research all rely on the accurate segmentation of blood vessels [8–12]. However, manual segmentation of the blood vessel requires a massive amount of time. Especially for 3D images, manual segmentation is impractical as the vessels present very complex 3D structures [13]. The automatic segmentation of blood vessels from CTA images is of critical importance but is also a challenging task. For head-neck CTA (HNCTA) images, the main reason for the difficulty is that the cortical bone near the blood vessels has similar CT values [10,14,15]. This constitutes the main difference between automatic

\* Corresponding authors.

E-mail addresses: [wjhuang@pku.edu.cn](mailto:wjhuang@pku.edu.cn) (W. Huang), [cjr.wangxiaoying@vip.163.com](mailto:cjr.wangxiaoying@vip.163.com) (X. Wang), [zhangjue@pku.edu.cn](mailto:zhangjue@pku.edu.cn) (J. Zhang).

<sup>1</sup> [\[orcid:0000-0003-2408-8302\]](https://orcid.org/0000-0003-2408-8302)

vessel segmentation in head-neck CTA and other images, such as MRA, pulmonary or coronary CTA.

### 1.1. Geometric modeling based vessel segmentation in HNCTA

Previously, many researchers have developed various modeling techniques for the automatic or semi-automatic vessel segmentation of head-neck CTA images. Yi et al. presented a region growing algorithm based on an iterative procedure, which tracked and segmented of small local cube at each iteration, to segment vascular structures. The researchers validated their method on the extracranial internal carotid artery (ICA) in CTA images [16]. Scherl et al. presented a level-set based segmentation algorithm for extracranial ICA near carotid bifurcation using CTA images [8]. Cheng et al. proposed a constrained B-Snake, which provides size and shape constraints on the vascular cross-sectional planes to perform vessel segmentation [17]. The model was used to segment the carotid bifurcation from CTA images. In addition, the review papers [18,19] have summarized several other geometric models for carotid bifurcation segmentation from CTA images. However, both carotid bifurcation and extracranial ICA are just a segment of the carotid artery without bone tissues nearby sharing similar pixel intensity in CTA. Deshpande et al. combined Hessian-based vessel-enhancing filtering with an active-contour-based technique to segment cerebral vasculature. The method was tested on brain CTA; nevertheless, vessels adjacent to bony structures outside the intracranial vasculature were excluded in the region of interest [20].

In the whole image area, the proximity to bones of some other vessels makes the all-vessel segmentation more complex and challenging than the segmentation of specific vascular segments. These segmentation algorithms designed for the carotid bifurcation or intracranial vessels have not been validated for other arteries in head and neck regions, for example, carotid siphon, which is also a section of carotid artery prone to atherosclerosis as carotid bifurcation but closer to the bone tissues. For the challenge of all-vessel segmentation in HNCTA, i.e., the vessel boundaries are very close to non-vessel tissues with similar intensity signals, Yang et al. proposed a progressive contrast enhancement algorithm and a boundary refinement algorithm in their study [21] following the traditional filter-based enhancement approach for vessel segmentation. However, the complex algorithm pipelines had many hyperparameters to tune, and the filter-based vessel enhancement may suffer from the sensitivity of local deformations [22], such as aneurysms, stenosis, and bifurcations. These filters may also yield substantial computational costs for large selected regions [22]. We refer the readers to the review articles [19,22–24] for more modeling techniques for the vessel segmentation in CTA images.

It is worth mentioning that most of the above-mentioned modeling methods rely on many strong geometry assumptions on the vessels, for example, local tubularity of the vessel surface, limited curvature of centerline and radial intensity profile [19,22,25]. Nevertheless, in the presence of vascular diseases, these geometry assumptions made on normal blood vessels are not anymore valid [23]. This inherent disadvantage may lead to poor model performance in the presence of pathology [21,23].

### 1.2. Bone removal strategy and virtual unenhanced reconstruction

Since the bone signal is the main factor affecting the vessel segmentation in HNCTA, from this perspective, some researchers attempted to explore another direction by removing the bones as an alternative to directly extracting the vessels in HNCTA [10,15]. CTA bone removal can improve the delineation of vessels closely adjacent to bony structures [26]. Bozkurt et al. called this bone-removing strategy an inverse approach [10]. In their work, they firstly eliminated the bone region in the 2D image using region

growing or random walker algorithm with experimentally detected seed point through the analysis of local histograms. Afterward, the same process was conducted for an initial segmentation of vessel followed by a very complex post-processing step, including the elimination of noises and non-vessel fragments in the 2D image, vessel tracking, and 3D reconstruction using 2D image set. Most operations in their method were essentially based on 2D image, ignoring the 3D image information, and was difficult to be generalized for 3D cases mainly due to the computational expense of the algorithms, such as region growing and random walk algorithms. Moreover, some key operations in their solution, for example, morphological opening and roundness detection, may not be appropriate for small blood vessels. Other researchers tried to eliminate the bones by jointly analyzing the HNCTA and UECT images [27–31], i.e., using the UECT images for bone suppression [32]. In the pre/post-contrast subtraction images, the background tissues are effectively removed and blood vessels are clearly visualized. Nevertheless, not every patient has both CTA and UECT scans, and this background subtraction strategy requiring pre/post-contrast images is not applicable if there is a non-negligible mismatch between the CTA and NECTA images.

In recent years, dual-energy CT (DECT) offers the potential to perform almost simultaneous acquisitions with different tube voltages, allowing automated bone removal (distinguishing iodine from bone) based on DECT attenuation information [26,33–36]. The main idea of bone removal in DECT is that materials can be differentiated by their attenuation differences depending on tube voltage [34]. Another similar application of DECT is iodine removal aiming at generating virtual unenhanced (VU) images similar to UECT, which may potentially obviate the need for the pre-contrast UECT scan for radiation-dose reduction [5,37–39]. However, commercially available DECT systems have at least one of the following intrinsic disadvantages depending on different techniques: motion misregistration, limited adjustability with relatively high overlap in energy spectra, limited FOV, susceptible to motion artifact, and scattered radiation [40,41]. Furthermore, it is more expensive than traditional CT while requiring similar radiation doses [42]. Due to these problems, at present, DECT scanners are not widely available [42–44].

### 1.3. CNN-based vessel segmentation in HNCTA

Most recently, the convolutional neural network (CNN) based learning method has attracted increasing attention in the biomedical image processing community and has shown great success in tasks such as image segmentation [45–48] and image prediction/conversion [49–53]. It is mainly because deep neural networks can effectively extract robust task-specific features automatically [48], compared with traditional modeling techniques, which highly rely on manually defined rules and features [54]. The strong geometric assumptions and the complex geometric modeling process can also be avoided in the deep learning (DL)-based approach.

For the task of intracranial vessel segmentation from CTA, Ni et al. proposed a global channel attention network that introduced the channel attention mechanism aggregating high-level and shallow features [55]. For automatic head and neck vessel segmentation and labeling, Yao et al. combined the convolutional V-Net model with graph convolutional network (GCN)-based point cloud model to improve vessel segmentation accuracy [56]. Tetteh et al. proposed DeepVesselNet tailored for vessel segmentation in 3D angiography, which use 2D orthogonal cross-hair filters to reduce the computational burden and introduce a class balancing cross-entropy loss function to handle the high-class imbalance problem [57]. To segment extracranial carotid arteries from CTA, Zhou et al. proposed a two-stage cascade network including an ROI detection network and a segmentation network based on a modified U-

Net architecture with residual connections, dilated convolution and classing-balancing hybrid objective function [58]. Fu et al. applied three cascaded ResU-Net models to sequentially remove the skull and segment the head and neck vessel from CTA [14]. Nevertheless, these supervised DL models have a prominent limitation, i.e., requiring a large amount of training data, which involve intensive manual labeling efforts, especially for 3D images.

#### 1.4. The motivation of this study

In this study, regarding extracting vessels from 3D HNCTA, we relate this task to another task which was not investigated before, prediction of unenhanced CT image by removing the hyperintensity in the vascular area in HNCTA. It is easy to see these two tasks are highly relevant. For both tasks, the key challenge is to reliably distinguish the vessel from other tissues with similar signal intensity, for example, the cortical bone. These two learning tasks are functionally resembling the bone and iodine removal process in DECT though they are theoretically different. In the deep learning field, it has been considered as an effective strategy to improve model performance by learning several related tasks together [59–61], which is generally called multi-task learning (MTL). By sharing representations between related tasks subject to constraints determined by domain expertise, MTL can enable the model to generalize better on the original task [62,63].

Here, we attempt to propose a dual-task CNN framework for fully-automatic simultaneous vessel segmentation and unenhanced prediction from 3D HNCTA, with the hope of improving the single-task performance. Additionally, to resolve the DL limitation of requiring large, labeled training sets, we adopt the patch-wise learning strategy and also try to automatically generate the training labels, i.e., the segmentation labels, by analyzing the UECT and HNCTA training data, avoiding intensive manual labeling efforts. This learning paradigm, automatically exploiting labels that are freely available within the data, can be considered as self-supervised learning [64].

To verify the feasibility and effectiveness of the proposed self-supervised dual-task learning framework, the study was performed on a real image dataset from 24 patients with both HNCTA and UECT scans. The performance of automatic segmentation was mainly assessed in terms of Dice's coefficient with the reference segmentation, and the unenhanced prediction was evaluated using the mean absolute error (MAE) with the real UECT images. Furthermore, we also compared our vessel segmentation performance with the state-of-the-art medical image segmentation model [46], and several available open source vascular segmentation methods [57,65–67].

## 2. Materials and methods

### 2.1. Data description

The local institutional review board approved this retrospective study and waived the informed consent requirement. The study included the image data of 24 patients with both HNCTA and UECT examinations performed with the CT system (Revolution CT, GE Medical Systems, Milwaukee, WI). Sixteen patients had atherosclerosis with calcified plaques or mild or moderate stenosis in various arteries. Four patients had thinner vessels in the vertebral artery or the anterior arteries. Additionally, one patient had intracranial hemorrhage, and one patient had a craniopharyngioma. No obvious lesion was found in three patients. The smallest field-of-view (FOV) was from the cranial end of the skull to the inferior margin of the mandible. The largest FOV was from the cranial end of the skull to the aortic arch. The image resolution was  $512 \times 512$  for each axial slice, with pixel spacing of  $0.488\text{ mm}$ . The number of

axial images ranged from 241 to 661 for different patients, with slice thickness of  $0.625\text{ mm}$ .

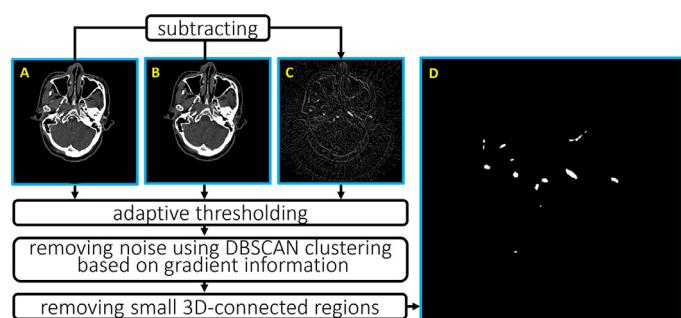
### 2.2. The idea of self-supervision and generation of training labels

Almost all CTA-based vessel segmentation studies use manual segmentation as the reference segmentation [10,17–19,21–23,55,68–71]. However, it is unfeasible to manually delineate all the blood vessels in each complete 3D scan as they present very complex 3D structures [13,15,55,72–74], and for small vessels, manual annotation is prone to high inter- and intra-user variability [23,70]. Previously, manual annotations were usually performed in several 2D slices [73], or for a specific vascular segment [19] or small parts of the vessel tree [74,75]. The difficulty of obtaining the manual annotation hinders the applicability of the state-of-the-art deep learning-based or some other supervised learning approaches in this area due to the lack of training labels. Therefore, automatically/semi-automatically generating the reference segmentation is very meaningful. Previous investigators sought to create the reference segmentation of intracranial vessels from TOF-MRA images and used it as training labels for a CTA segmentation model [76]. However, this strategy is prone to registration errors for the extracranial area due to the challenge of accurate registration between MR and CT [77].

In recent years, the self-supervised learning paradigm has become a promising approach for annotation-efficient learning [64], which is a form of unsupervised learning in which the data itself provides supervision. In many previous studies, it refers to automatically obtaining the training set/training labels without manual labeling [64,76,78–85]. Currently, most DL-based self-supervised learning methods create training labels of various pretext tasks (different from the target task), such as rotation prediction [86] and context restoration [87], to learn transferable feature representation for the target task. However, such a strategy still requires manual annotation of the target task to fine-tune the pretext task model. To avoid the need for manual annotation, in this study, we treat the pretext and target task as the same task and use contrasting modalities (pre/post-contrast imaging) to construct the pseudo training label for the task.

Here, motivated by the idea of background subtraction, we proposed a novel approach to generate the segmentation label by analyzing the HNCTA and UECT images consecutively acquired with the patients in a fixed position. The generation pipeline was shown in Fig. 1.

The UECT images were registered to the HNCTA images by 3D rigid registration using affine transformation. The subtraction images between the HNCTA images and the UECT images, and the gradient images, defined as  $G(I^*) = \left| \frac{\partial I^*}{\partial x} \right| + \left| \frac{\partial I^*}{\partial y} \right| + \left| \frac{\partial I^*}{\partial z} \right|$  were calculated ( $I^*$  represented the UECT image  $I^{UECT}$  or CTA image  $I^{CTA}$ ).



**Fig. 1.** Proposed pipeline for the generation of training label: (A) HNCTA images; (B) registered UECT images; (C) subtraction images between (A) and (B); (D) subtraction-based segmentation label.

In the subtraction image  $I^{CTA} - I^{UECT}$ , the vascular regions were enhanced and the signals of the background tissues, such as bone tissue, were suppressed. It is worth mentioning that there was a small amount of residual signal remaining on the surfaces of some tissues (e.g., skin tissue) in the subtraction image due to the influence of partial volume effects. The vascular area and the small non-vascular areas (e.g., the tissue surfaces) with residual signals in the subtraction images can be distinguished by analyzing the CTA signal, UECT signal, and the corresponding gradient signals, as described below.

In Fig. 1, the subtraction images, HNCTA images, and the UECT images were first thresholded by adaptive upper and lower thresholds to create three binary maps containing blood vessels. Because  $S^{UECT} \approx S^{CTA}$  holds for nonvascular regions ( $S$  denotes pixel-wise CT value), the CT-value of vessels distributes in the asymmetric part of the distribution  $P(S^{UECT}, S^{CTA})$ . Therefore, we determined the adaptive upper and lower thresholds by detecting the asymmetric part of  $P(S^{UECT}, S^{CTA})$ . The thresholding step is detailed in Supplementary Material 1 Section 2.1. Afterward, the three binary maps were multiplied to obtain the intersection map, which extracted the arterial and venous vessels. To exclude noise voxels, the density-based DBSCAN algorithm [88] was subsequently used to cluster the voxels in the intersection map based on gradient information ( $G(I^{CTA})$ ,  $G(I^{UECT})$ ) (see Supplementary Material 1 Section 2.1 for experimental details). Except for the major cluster containing the most voxels, other minor clusters and outliers were removed. Finally, the small 3-D connected regions were considered as noises, and were thus eliminated to automatically generate the subtraction-based segmentation used for model training.

### 2.3. Reference segmentation

For the HNCTA images in our dataset, the contrast-enhanced venous vasculatures were different for different patients. To ensure the vessels of different patients were consistent, the transverse sinus and the veins below it, if extracted by subtraction-based segmentation, were manually removed. In addition, two experienced radiologists with over 3 yr of experience further verified and finely adjusted the subtraction-based segmentation to provide a reference segmentation through the use of a locally developed MATLAB-based software. The reference segmentation was used for model evaluation. It is worth mentioning that although the reference label used for segmentation evaluation was manually labeled, the training label was automatically generated. Thus, our method required no manual annotation.

### 2.4. The dual-task framework for simultaneous vessel segmentation and unenhanced prediction

We leveraged the hard parameter sharing strategy to design our dual-task framework, which is one of the most commonly used approaches to MTL in neural networks [62]. This strategy is generally applied by sharing the hidden layers between multiple tasks and using task-specific branching layers for different tasks [62,89,90], and the shared layers for global feature extraction are usually referred to as the backbone network [90].

Since our goal was to achieve unenhanced prediction and all-vessel segmentation in the full imaging volume containing vessels of different sizes, we preferably use a multi-scale pyramidal convolutional network as the backbone network. There are numerous such networks proposed in different fields, such as the feature pyramid network in object detection research and UNet/VNet for biomedical image segmentation, which all extract the multi-scale features using several down-sampling layers in combination with convolution layers. Since UNet based architecture is one of the most popular networks for semantic segmentation problems in

biomedical imaging [91,92] and commonly regarded as the baseline segmentation model [93–96], we referred to a 3D-UNet architecture (VNet model) [46], and build our 3D backbone network with two down-sampling and up-sample operations, as shown in Fig. 2 (B).

As to the branching network for vessel segmentation (vSG) and unenhanced prediction (uEP), we used the stacked convolutional blocks (ConvBlock) for the two tasks. As mentioned in the introduction, predicted unenhanced CT can be generated by removing the hyperintense vascular signal in CTA images, and the vessels are highlighted by subtracting the unenhanced CT images from CTA images. Therefore, we designed the following iterative scheme, as in Eq. (1), to share the extracted features for the two tasks (iterative residual-sharing scheme).

$$\begin{cases} G_{ijkl}^n = E_{ijk} - V_{ijkl}^n; \\ H_{ijkl}^n = E_{ijk} - U_{ijkl}^n; \\ U^{n+1} = C_u^{n+1}([G^n, U^n]); \\ V^{n+1} = C_v^{n+1}([H^n, V^n]); \end{cases} \quad (1)$$

where  $V^1 = C_v^1(F)$  and  $U^1 = C_u^1(F)$ .  $F$  represents the features learned by the backbone network.  $U^n$  and  $V^n$  represent the learned features for the uEP task and the vSG task, respectively, after  $n$  iterations.  $E$  denotes the CTA images. The subscripts  $i, j, k$  denotes the three-dimensional coordinates, and  $l$  denotes the channel index. The operator  $C_*(*)$  stands for a ConvBlock, which includes one or more convolution operations. The operator  $[*, *]$  stands for the concatenation of two features.

As the blood vessels have multi-scale self-similarity [97], i.e., the vessels of different local regions, though of different sizes, exhibit similar morphological structure, it is reasonable to partition the original image into many overlapping patches and train the model with these small patches. This patch-wise training strategy we adopted has two main advantages: greatly reducing the required GPU memory for training and effectively augmenting the training dataset. The flowchart of the overlapping patch-wise learning is shown in Fig. 2 (A).

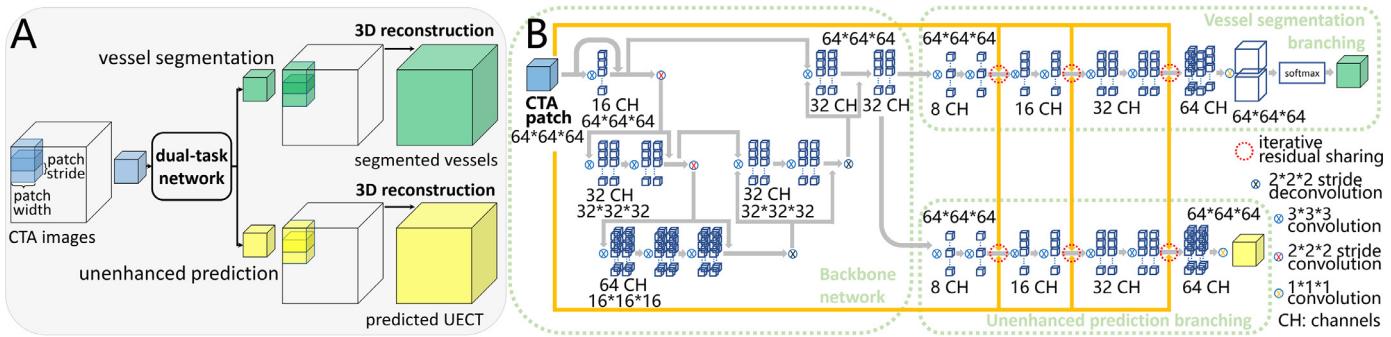
The whole structure of the proposed dual-task network was shown in Fig. 2 (B). The size of the image patch in Fig. 2 was empirically set to  $64 \times 64 \times 64$ . Since the predicted unenhanced images contains negative CT values, PReLU (Parametric Rectified Liner Units) was used as the activation function for the dual-task network. To investigate the benefit of dual-task learning, we also analyzed the performance of the single-task network for vessel segmentation and unenhanced prediction, respectively, in which only one branching was reserved with the other branching discarded, as was shown in Supplementary Material 1 Section 1 Fig. S1. In the single-task network, there was no feature sharing between the two tasks and iterative scheme for the branching network became Eq. (2).

$$\begin{cases} G_{ijkl}^n = E_{ijk}; \\ H_{ijkl}^n = E_{ijk}; \\ U^{n+1} = C_u^{n+1}([G^n, U^n]); \\ V^{n+1} = C_v^{n+1}([H^n, V^n]); \end{cases} \quad (2)$$

Moreover, we also used the vanilla VNet to conduct vessel segmentation, which was regarded as the baseline model in this study. For fair comparison, PReLU was chosen as the activation function for all compared networks.

### 2.5. Implementation and evaluation

The implementation details of the dual-task learning model, including the generation of training labels in Fig. 1, the training and testing processes, and the 3D reconstruction from patch-wise prediction in Fig. 2 (A) are all provided in the Supplementary Material



**Fig. 2.** Overview of the proposed dual-task model: (A) the flowchart of overlapping patch-wise learning; (B) the framework of the proposed dual-task network, the yellow lines and the red dotted circles indicate the iterative residual-sharing process for the extracted features. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

1 Section 2. The performance of vSG and uEP tasks were mainly assessed by the Dice coefficient and the mean absolute error (MAE), respectively, under both patient-wise analysis and patch-wise analysis of large-, medium-, and small-vessel patches. For vSG, we also calculated other commonly used evaluation metrics, including precision, recall, intersection-over-union (IoU), the area under the precision-recall curve (PR AUC), and the area under the receiver operating characteristic curve (ROC AUC). The detailed description of the evaluation indexes is in Supplementary Material 1 Section 3.

## 2.6. Comparison with other vessel segmentation methods

In addition to the comparison with the popular supervised segmentation model, VNet model, and the single-task segmentation model, we further compared the performance of the proposed dual-task model with several available open-source methods for 3D vessel segmentation. The compared methods included: vascular enhancement based multiscale Hessian-matrix based filter (HMBF) [67], geometric modeling based optimally oriented flux (OOF) model [66], spherical flux model (SFM) [65], and deep vessel net (DVN) model [57]. The first three methods are widely recognized for their good performance in unsupervised 3D vessel segmentation [21], and DVN is a supervised DL architecture tailored to segment vessels from 3D angiographic volumes. It is worth noting that although OOF and SFM were published ten years ago, they are still widely used unsupervised methods for 3D vessel segmentation in recent years.

For the first three methods, we found from our experiments that blood vessels cannot be reliably separated from background areas (especially bone tissue) if the original CTA image was directly used as the model input. To improve their segmentation results, a dual scan masking method [35] was first applied to remove bone tissue from the CTA image. Subsequently, the grid search method was used to find the best values for the hyper-parameters, under which the three models achieved their best performance (highest Dice coefficient) in our dataset. For the DVN model, the best performed DeepVesselNet-FCN architecture was selected [57] and was trained in the same way as the original study [57] using our training data.

### 3. Results

### 3.1. Ablation study: The unenhanced prediction and vessel segmentation of dual-task framework

**Fig. 3** presents typical vessel segmentation and unenhanced prediction results for five axial slices (the results for all the slices of the patient are given in Supplementary Material 2). The predicted unenhanced CT images show good agreement with the real

UECT images, for both vessel and non-vessel regions. The automatic vessel segmentation also match very well with the reference segmentation, even in some regions where vessels are closely attached to the bones, for example, carotid artery lies close to the temporal and sphenoid bone, and the vertebral artery is adjacent to the vertebra.

We further demonstrate the performance of the model in the lesion area. Fig. 4 shows the results of vessel segmentation and unenhanced reconstruction for the regions close to calcified plaques. The proposed method can effectively differentiate the calcified plaques from the vessel lumen, i.e., the calcified plaques are excluded from the segmented regions while their original forms are preserved in the predicted unenhanced images. Fig. 5 shows the results of vessel segmentation and unenhanced prediction in patients with craniopharyngioma and cerebral hemorrhage at the corresponding lesion sites. These lesions do not affect the vascular segmentation, and the appearance of these lesions in the predicted unenhanced image are also in good agreement with that in the real UECT image. Fig. 6 shows the 3D visualization of typical segmentation results (rendering movies in Supplementary Material 3).

We further evaluate the performance of the dual-task model by quantitative metrics. The results comparing the dual- and single-task networks are shown in Tables 1 and 2. Table 1 presents the statistical results of the Dice coefficient for vessel segmentation by the patient- and patch-wise analysis. The Dice coefficients of the dual-task network are significantly higher than that of the single-task network, for both patient- and patch-wise analysis (paired *t*-test,  $P < 10^{-3}$ ).

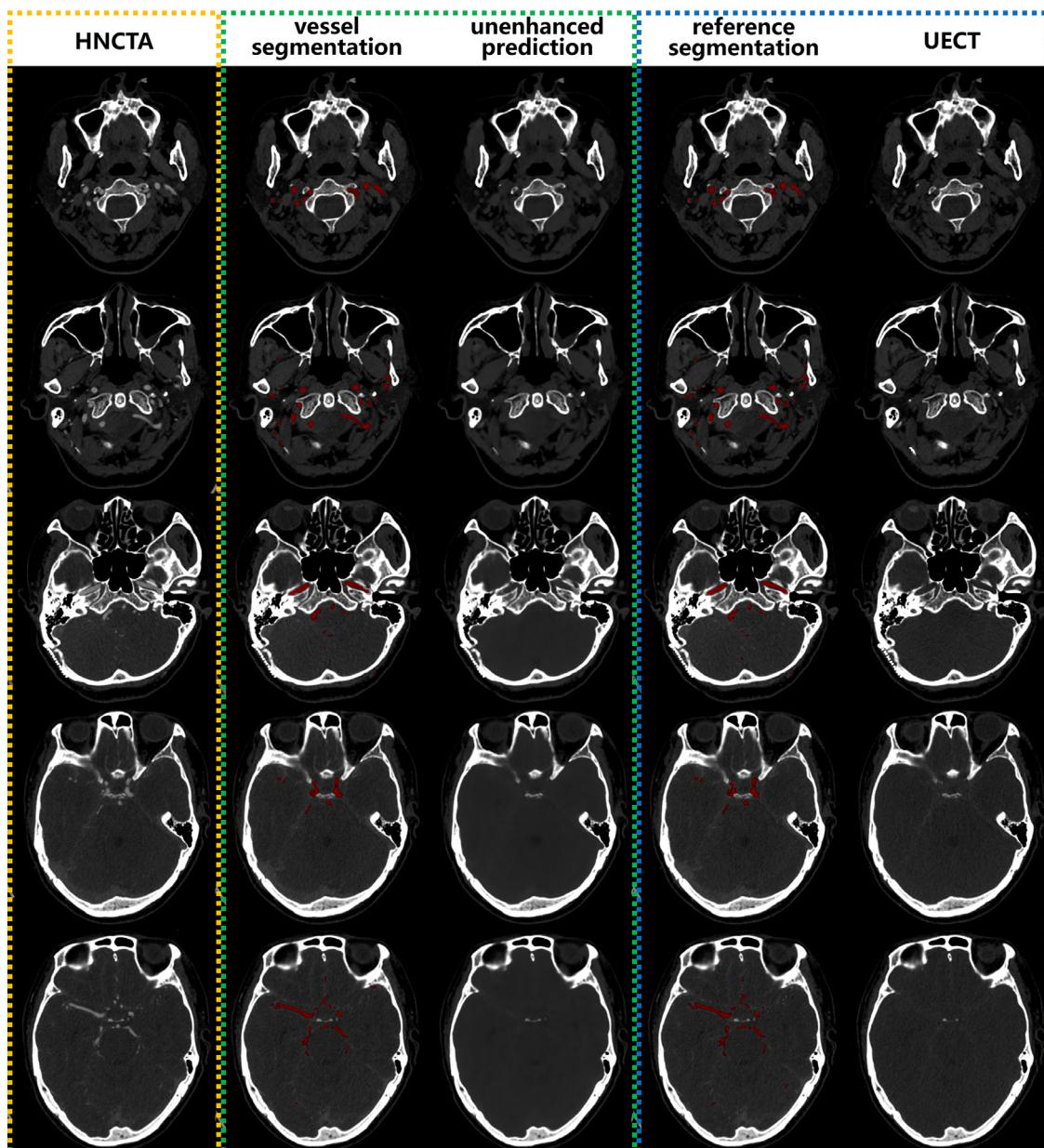
To evaluate the performance of the model in segmenting blood vessels adjacent to bones, radiologists manually annotated the bounding boxes of the carotid C2~3 segment, which are attached to the temporal bone, and carotid C4~7 segment, which lie close to the sphenoid bone. The calculated region-wise Dice coefficients are shown in Fig. 7. The averaged Dice coefficients are also included in Table 1. The DTM achieves significant higher segmentation accuracy than the STM (paired *t*-test,  $P < 10^{-3}$ ).

**Table 2** summarizes the statistical results of the MAE of unenhanced prediction in the vessel regions from reference segmentation, here denoted as V-MAE, as well as the MAE of unenhanced prediction over the entire image area, here abbreviated as E-MAE.

$$\text{Error}_{\text{V-MAE}} = \frac{1}{\sum_{ijk} M_{ijk}^V} \sum_{ijk} M_{ijk}^V |I_{ijk}^{\text{UECT}} - \hat{I}_{ijk}^{\text{UECT}}| \quad (3)$$

$$\text{Error}_{\text{E-MAE}} = \frac{1}{\sum_{ijk} 1} \sum_{ijk} |l_{ijk}^{\text{UECT}} - \hat{l}_{ijk}^{\text{UECT}}| \quad (4)$$

$I^{\text{UECT}}$  and  $\hat{I}^{\text{UECT}}$  are the ground-truth and predicted unenhanced CT images, respectively.  $M^V$  is the patient- or patch-wise binary



**Fig. 3.** Typical vessel segmentation and unenhanced prediction in one case: from left to right, columns display real HNCTA images, vessel segmentation results overlaid on the HNCTA, unenhanced prediction results, reference segmentation, real UECT images.

**Table 1**

The patient-wise, patch-wise and region-wise analysis of the dice coefficient (unit: %) of the single-task model (STM) and dual-task model (DTM) for vessel segmentation.

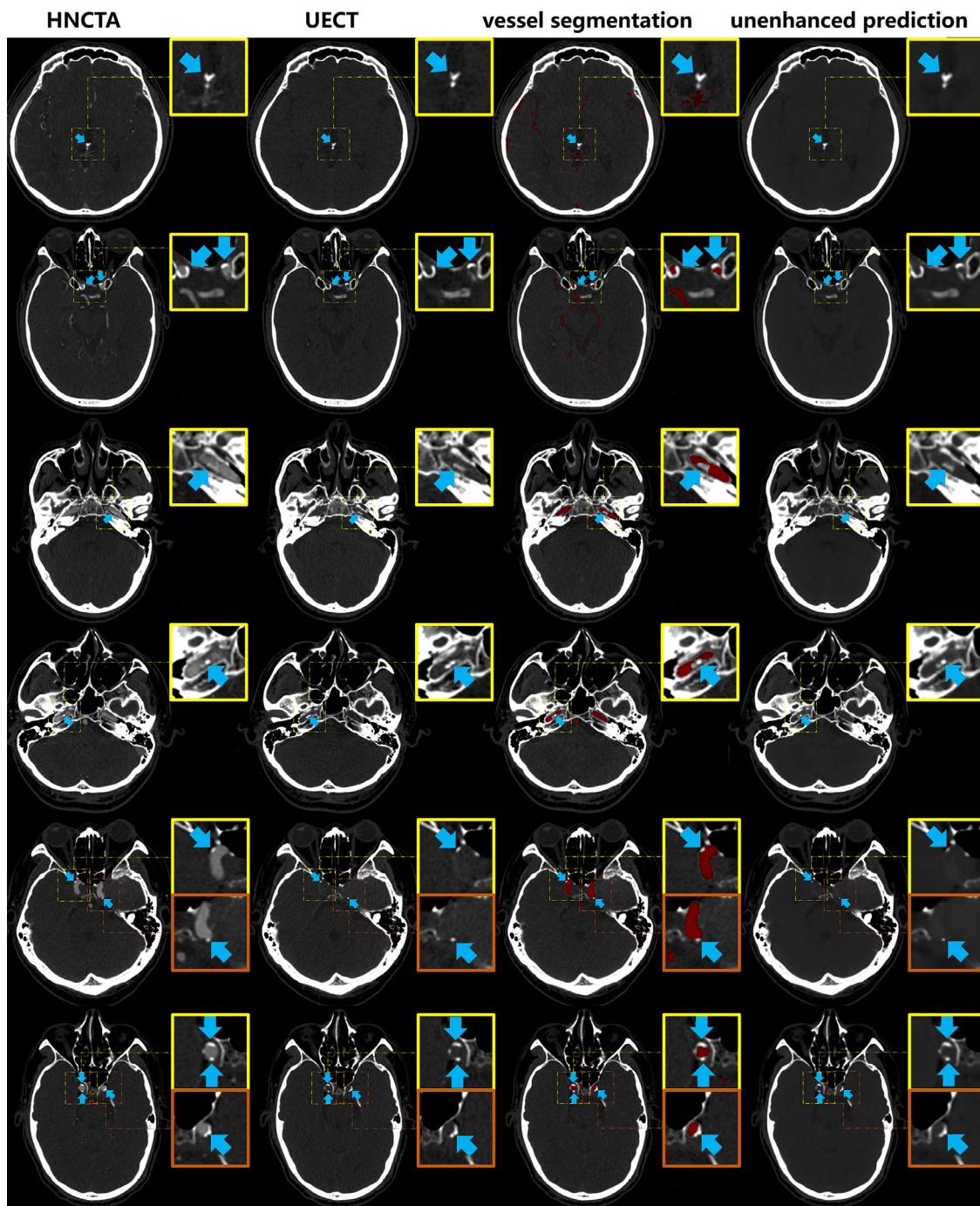
		N	ADC of STM	ADC of DTM	P-value
Patient-wise analysis		24	79.77	84.83	$10^{-3}$
Patch-wise analysis	Large-vessel patch	645	85.53	89.42	$10^{-3}$
	Medium-vessel patch	3080	81.98	84.56	$10^{-3}$
	Small-vessel patch	3610	68.19	73.47	$10^{-3}$
Region-wise analysis	Carotid C2~C3	24	84.32	90.93	$10^{-3}$
	Carotid C4~C7	24	81.84	89.27	$10^{-3}$

ADC, averaged Dice coefficient; N, sample size.

mask indicating the vascular regions. As shown in Table 2, the V-MAE and E-MAE of the dual-task network are both significantly smaller than that of the single task network (paired *t*-test,  $P < 0.01$ ). Table 2 reveals that the improvement of the unenhanced prediction task brought by the supervision of the vessel segmentation task is particularly prominent in the vessel regions.

### 3.2. Comparison with vanilla VNet and other vessel segmentation models

For the vSG task, the performance of the proposed dual-task model (DTM) was also compared with that of the state-of-the-art medical image segmentation model, vanilla VNet, and several pop-



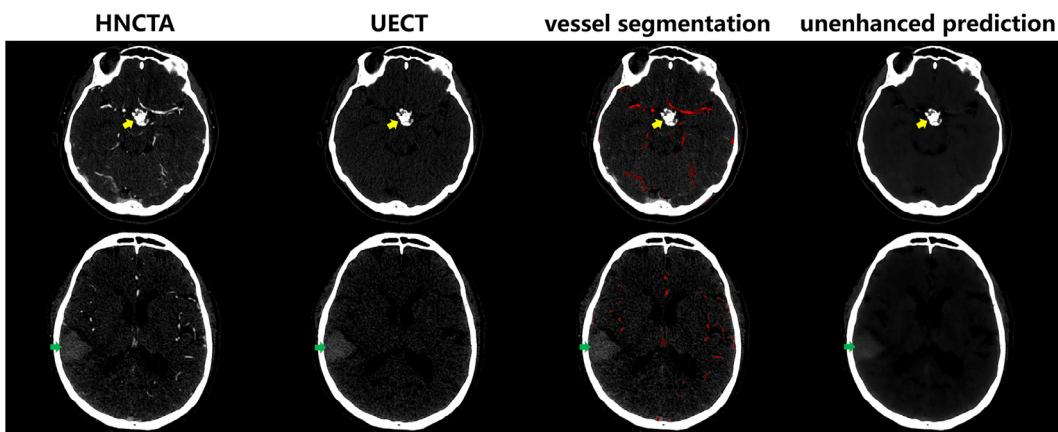
**Fig. 4.** Vessel segmentation and unenhanced reconstruction in the regions close to calcified plaques: from left to right, columns display real HNCTA images, real UECT images, vessel segmentation results overlaid on the HNCTA, unenhanced prediction results. Yellow and orange boxes indicate different magnified local regions. The red areas represent segmented vessels. The proposed method can effectively differentiate the vessels and the calcified plaques (blue arrows). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 2**

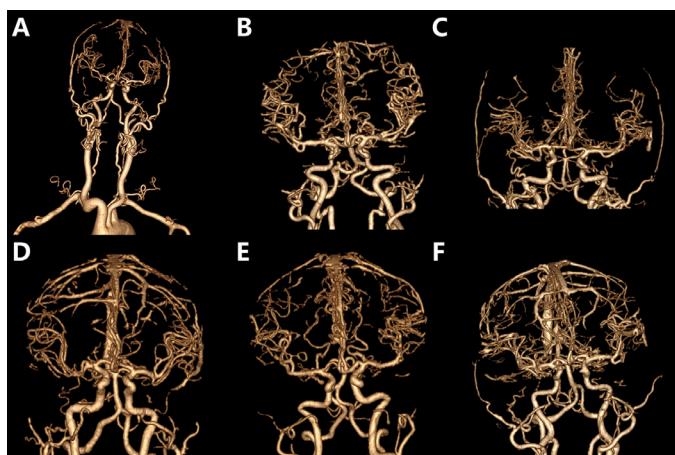
The patient-wise and patch-wise analysis of the mean absolute error (unit: Hounsfield unit, HU) for unenhanced prediction.

		N	V-MAE of STM	V-MAE of DTM	P-value	E-MAE of STM	E-MAE of DTM	P-value
Patient-wise analysis		24	19.54	15.07	$10^{-3}$	7.23	7.01	$10^{-3}$
Patch-wise analysis	Large-vessel patch	645	21.31	15.26	$10^{-3}$	21.57	20.40	$10^{-3}$
	Medium-vessel patch	3080	18.07	14.30	$10^{-3}$	18.10	17.54	$10^{-3}$
	Small-vessel patch	3610	17.36	14.87	$10^{-3}$	16.67	16.15	$10^{-3}$

V-MAE, mean absolute error in vessel region; E-MAE, mean absolute error over entire image area; N, sample size.



**Fig. 5.** The vessel segmentation and unenhanced reconstruction for the patients with craniopharyngioma (yellow arrow) and hemorrhage (green arrow): from left to right, columns display real HNCTA images, real UECT images, vessel segmentation results overlaid on the HNCTA, unenhanced prediction results. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 6.** 3D visualization of typical segmentation results from six patients (A~F).

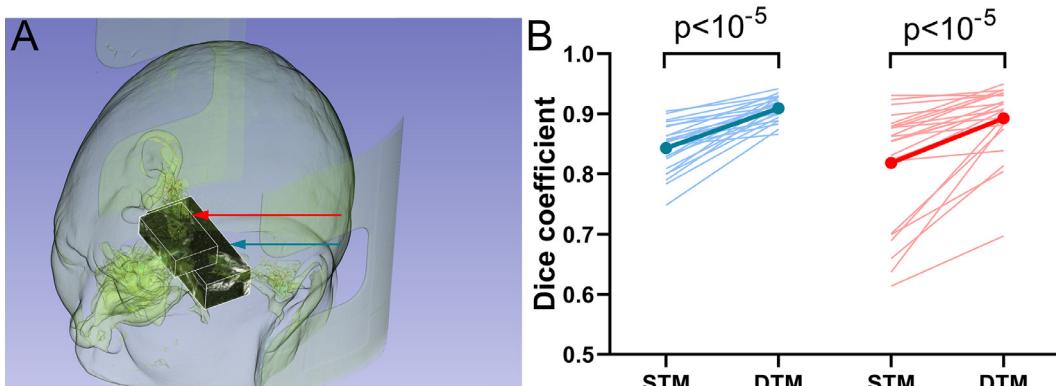
ular vascular segmentation models, including HMBF, OOF, SFM, and DVN models. The mean and 95% confidence interval (CI) of the patient-wise, region-wise, and patch-wise Dice coefficients of different models are given in Fig. 8. The patient-wise and region-wise (regions containing carotid C2~3 segment and C4~7 segment) Dice coefficient of the proposed DTM are significantly higher than those of vanilla VNet, HMBF, OOF, SFM, and DVN methods (paired *t*-test,

$P < 10^{-3}$ ). The proposed DTM also achieves a significantly higher patch-wise Dice coefficient than other models for both the category of large-, medium- and small-vessel patches (paired *t*-test,  $P < 10^{-3}$ ). It is worth noting that, for most methods, the segmentation accuracy decreases as the vessel size becomes smaller. Additionally, Table 3 compares different segmentation methods by other performance measures, including ROC AUC, PR AUC, IoU, precision and recall. The proposed DTM model attains the highest performance in all metrics for both patient- and region-wise analysis. Fig. 9 shows typical vessel segmentation results near bony regions by different models. The results demonstrate that the proposed method can more effectively distinguish vessels from bones.

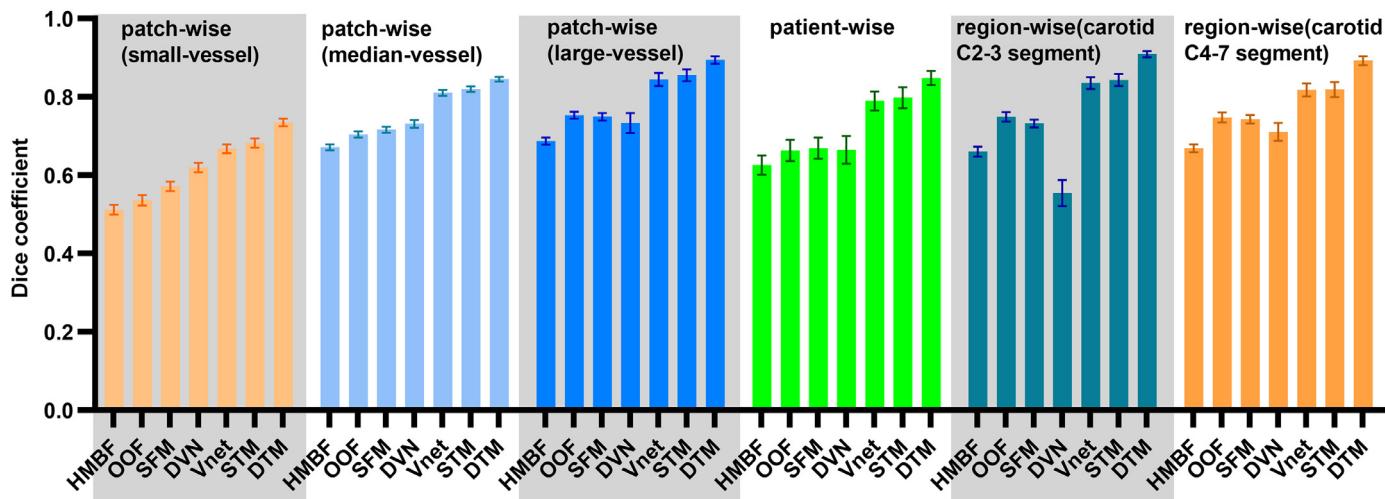
#### 4. Discussion

Automated segmentation of the 3D arteries is an important prerequisite for observing vascular structures, evaluating vascular lesions such as stenosis and occlusion, and studying vascular hemodynamics. In this study, we proposed a self-supervised dual-task deep learning framework to automatically segment head and neck blood vessels from CT angiography images. We introduced a concept of unenhanced prediction (uEP) task and used this task as a second task to improve the performance of the vessel segmentation (vSG) task based on a developed residual-sharing scheme.

Our results showed that the proposed dual-task model can achieve significantly higher vessel segmentation and unenhanced



**Fig. 7.** (A) Visualization of manually annotated bounding boxes for carotid C2~3 segment (blue arrow) and C4~7 segment (red arrow); (B) vessel segmentation accuracy in the corresponding regions: the blue and red lines show the difference between the single-task model (STM) and dual-task model (DTM), the points indicate the average values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 8.** The patch-wise, patient-wise and region-wise Dice coefficients of different segmentation models.

**Table 3**

Comparison of segmentation accuracy of different methods via six evaluation metrics (mean $\pm$ 95%CI).

		HMBF	OOF	SFM	DVN	VNet	STM	DTM
ROC AUC	All arteries	0.828 $\pm$ 0.013	0.848 $\pm$ 0.015	0.813 $\pm$ 0.018	0.984 $\pm$ 0.007	0.972 $\pm$ 0.011	0.976 $\pm$ 0.008	0.989 $\pm$ 0.006
	Carotid C2~C3	0.765 $\pm$ 0.007	0.827 $\pm$ 0.008	0.695 $\pm$ 0.015	0.952 $\pm$ 0.008	0.971 $\pm$ 0.008	0.975 $\pm$ 0.007	0.995 $\pm$ 0.003
	Carotid C4~C7	0.776 $\pm$ 0.004	0.838 $\pm$ 0.005	0.746 $\pm$ 0.007	0.968 $\pm$ 0.004	0.965 $\pm$ 0.009	0.972 $\pm$ 0.005	0.993 $\pm$ 0.002
PR AUC	All arteries	0.626 $\pm$ 0.043	0.687 $\pm$ 0.035	0.659 $\pm$ 0.035	0.701 $\pm$ 0.042	0.850 $\pm$ 0.031	0.856 $\pm$ 0.036	0.917 $\pm$ 0.017
	Carotid C2~C3	0.654 $\pm$ 0.014	0.739 $\pm$ 0.010	0.633 $\pm$ 0.012	0.543 $\pm$ 0.046	0.874 $\pm$ 0.009	0.884 $\pm$ 0.009	0.929 $\pm$ 0.006
	Carotid C4~C7	0.632 $\pm$ 0.013	0.713 $\pm$ 0.010	0.628 $\pm$ 0.010	0.653 $\pm$ 0.031	0.809 $\pm$ 0.014	0.809 $\pm$ 0.015	0.874 $\pm$ 0.008
Precision	All arteries	0.702 $\pm$ 0.048	0.729 $\pm$ 0.050	0.717 $\pm$ 0.050	0.654 $\pm$ 0.043	0.805 $\pm$ 0.038	0.799 $\pm$ 0.037	0.863 $\pm$ 0.020
	Carotid C2~C3	0.884 $\pm$ 0.027	0.887 $\pm$ 0.025	0.876 $\pm$ 0.028	0.526 $\pm$ 0.048	0.879 $\pm$ 0.026	0.884 $\pm$ 0.025	0.905 $\pm$ 0.013
	Carotid C4~C7	0.856 $\pm$ 0.026	0.854 $\pm$ 0.028	0.850 $\pm$ 0.026	0.643 $\pm$ 0.032	0.823 $\pm$ 0.022	0.813 $\pm$ 0.027	0.879 $\pm$ 0.017
Recall	All arteries	0.569 $\pm$ 0.019	0.614 $\pm$ 0.022	0.634 $\pm$ 0.025	0.678 $\pm$ 0.032	0.779 $\pm$ 0.031	0.800 $\pm$ 0.027	0.837 $\pm$ 0.027
	Carotid C2~C3	0.530 $\pm$ 0.015	0.651 $\pm$ 0.016	0.633 $\pm$ 0.014	0.605 $\pm$ 0.024	0.806 $\pm$ 0.033	0.815 $\pm$ 0.033	0.917 $\pm$ 0.016
	Carotid C4~C7	0.556 $\pm$ 0.008	0.676 $\pm$ 0.010	0.670 $\pm$ 0.007	0.821 $\pm$ 0.009	0.828 $\pm$ 0.021	0.842 $\pm$ 0.020	0.913 $\pm$ 0.009
IoU	All arteries	0.457 $\pm$ 0.026	0.498 $\pm$ 0.031	0.505 $\pm$ 0.032	0.501 $\pm$ 0.038	0.654 $\pm$ 0.033	0.666 $\pm$ 0.036	0.734 $\pm$ 0.030
	Carotid C2~C3	0.494 $\pm$ 0.014	0.600 $\pm$ 0.015	0.578 $\pm$ 0.012	0.388 $\pm$ 0.032	0.719 $\pm$ 0.021	0.731 $\pm$ 0.023	0.835 $\pm$ 0.013
	Carotid C4~C7	0.504 $\pm$ 0.011	0.601 $\pm$ 0.016	0.594 $\pm$ 0.013	0.562 $\pm$ 0.026	0.700 $\pm$ 0.022	0.703 $\pm$ 0.026	0.811 $\pm$ 0.017
Dice	All arteries	0.626 $\pm$ 0.024	0.663 $\pm$ 0.027	0.669 $\pm$ 0.027	0.665 $\pm$ 0.035	0.789 $\pm$ 0.024	0.798 $\pm$ 0.027	0.848 $\pm$ 0.018
	Carotid C2~C3	0.661 $\pm$ 0.012	0.749 $\pm$ 0.012	0.732 $\pm$ 0.010	0.555 $\pm$ 0.033	0.835 $\pm$ 0.015	0.843 $\pm$ 0.016	0.909 $\pm$ 0.008
	Carotid C4~C7	0.669 $\pm$ 0.010	0.747 $\pm$ 0.013	0.743 $\pm$ 0.011	0.711 $\pm$ 0.023	0.818 $\pm$ 0.017	0.819 $\pm$ 0.019	0.893 $\pm$ 0.011

prediction accuracies than the single-task models. This may be mainly attributed to the high correlation between the vessel segmentation task and the unenhanced prediction task, i.e., the signal difference between the CTA image and the unenhanced CT image is mainly from the vessel area. This perspective was also verified by the result that the improvement of unenhanced prediction obtained by the dual-task model was especially prominent in the vessel regions.

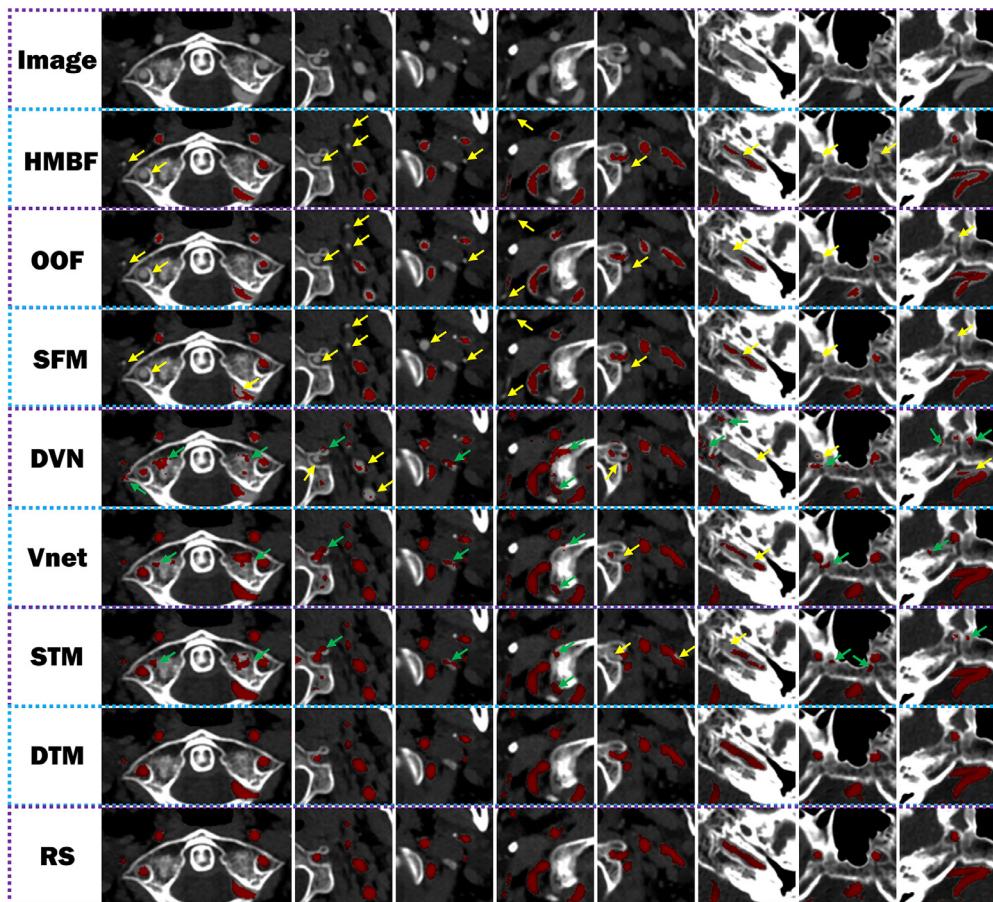
For the vSG task, the performance of the dual-task model was also compared with that of the state-of-the-art biomedical image segmentation model, VNet, and several popular vascular segmentation models. The proposed method achieved the highest segmentation accuracy.

For the uEP task, the results demonstrated a high consistency between the predicted image and the real unenhanced CT images. In this study, the MAE between the predicted pixel values and real pixel values in the vascular region was calculated and further averaged over the patients, termed as patient-wise V-MAE. The patient-wise V-MAE was only 15.07 Hounsfield unit (HU) for our dual-task network. It was even smaller than the standard deviation of the pixel values of real UECT in the vascular region, which varied between 17.31 HU and 33.81 HU for different patients, and the average value was 22.95 HU. Additionally, for some of the lesions seen in our patients, including calcified plaque, intracranial

hemorrhagic lesion, and craniopharyngioma, the form of these lesions in the predicted unenhanced image was in good agreement with that in real UECT image.

Here, to make a reasonable comparison between the unenhanced prediction and the virtual unenhanced reconstruction (VU) using dual-energy CT angiography, we referred to the evaluation method given in [98], in which the researcher calculated the mean attenuation of a specific tissue by averaging the pixels values in a selected ROI and evaluate the difference of the mean attenuations in VU and real UECT images. For virtual unenhanced reconstruction, the absolute difference averaged over the patient for the artery tissue was 6.4 $\pm$ 5.1 HU (the mean ROI volume for UECT was 0.63 $\pm$ 0.27 cm<sup>3</sup> and for VU was 0.63 $\pm$ 1.05 cm<sup>3</sup>) [98]. In this study, we extracted all image patches with vessel volume ranging between 0.63 $\pm$ 0.27 cm<sup>3</sup>, and used the local vascular region as the ROI to calculate the above absolute difference. The calculated result for our unenhanced prediction was 6.1 $\pm$ 4.5 HU. This indicated that the performance of our unenhanced prediction was comparable to that of the virtual unenhanced reconstruction in terms of the error level in vascular region.

For our dual-task learning framework, we now provide an intuitive explanation for the proposed residual-sharing scheme: for well-predicted uEP feature map  $U$ ,  $E - U$  should have values significantly distant from zero in the vascular area and values close



**Fig. 9.** Typical segmentation results by different models and the reference segmentation (RS) for some vessels close to bones: the green arrows indicate the bones are misclassified as vessels, and the yellow arrows point to the vessels without being segmented. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to zero in the background ( $E$  denotes the HNCTA image), so  $E - U$  can be regarded as a useful feature for vSG task. Additionally, for well-predicted vSG feature map  $V$ , there will be a difference in signal intensity between the vascular area and the non-vascular area (the intensity range of the vascular area or the non-vascular area is not restricted due to the presence of the final softmax layer in the vSG branching network). Therefore,  $E - V$  can provide different levels of signal suppression (or enhancement) for vascular area and the non-vascular area. In principle, the unenhanced image can be readily estimated, i.e., the hyperintense signals in the vascular area can be removed from  $E$ , by performing a single convolution to the multi-channel  $E - V$  if appropriate convolution kernel is learned. Thus,  $E - V$  is also considered as a good feature for uEP task. The key challenge for both tasks lies in effectively distinguishing the vascular and non-vascular areas. By dynamically sharing the useful features between the two related tasks, in the form of the residual-based scheme, the vascular and non-vascular areas are expected to be better identified, and thereby the performance of the vSG and uEP tasks will be improved. The dual-task learning based on residual-sharing scheme proposed in this paper has the potential to be applied to the analysis of other contrast-enhanced images.

This study also has the following contributions. Although the deep learning models have exhibited extraordinary performance in medical image processing, the main drawback is the requirement of large training datasets and the corresponding class labels [99]. This shortcoming is especially serious for the 3D vessel segmentation task. To effectively solve this problem, we proposed a self-supervised label-generation strategy and utilized the patch-wise

learning methodology. In our vascular segmentation task, we initially tried to manually label the vessels directly from CTA images, but we found that it caused significant labeling errors for small vessels, and the labeling time for all vessels in a single patient was much longer than 24 h. The difficulty in obtaining the segmentation label presents challenges not only to the training but also to the evaluation of the segmentation algorithms [80,84]. The proposed self-supervised learning strategy, which automatically generates pseudo labels by using both the enhanced and unenhanced images, is of great importance for reducing the labeling effort to obtain reference segmentation, and promoting the application of deep learning in the field of vessel segmentation task. We expect to extend this strategy for vessel segmentation from CT angiography or MR angiography of other tissues, such as the heart, lung, or liver.

Our experimental results also suggest that the patch-wise learning paradigm is a very effective approach for 3D vessel segmentation and unenhanced prediction. This may be mainly due to the multi-scale self-similarity of blood vessels. For 3D data analysis, patch-wise learning can greatly reduce the GPU memory requirement and effectively augment the dataset. This is the main reason why our deep learning model can be successfully trained even under a relatively small sample size. It is worth mentioning that we observed that the patch-wise learning scheme generally does not perform very well in the boundary of the patch. It can help to reduce the reconstruction error by integrating the prediction results of multiple partially-overlapped patches (see Supplementary Material 1 Section 2.4 Fig. S4). Therefore, we recommend

using overlapping partitioning to generate the dataset for model training and prediction in the patch-wise learning scheme. In the real-data analysis, the researcher needs to consider the tradeoff between the training time and GPU memory requirement when determining the appropriate patch size, and the path stride (overlapping ratio).

In future studies, modification to the loss functions may help to improve the performance of the dual-task model. For example, for segmentation tasks, Dice loss has been shown to be more effective than binary cross-entropy loss in case of class imbalance [46,100]. For the unenhanced prediction task, the MAE loss can be replaced by multi-class cross-entropy and generalized cross-entropy, which can yield good performance in noisy label scenarios [101]. In this study, the magnitudes of the MAE loss for unenhanced prediction and the CE loss for vessel segmentation were in the same order. Therefore, we set the loss weight to be 1:1. However, for other loss function combinations, the magnitudes of the two losses may differ greatly. Then we may need to use other adaptive balancing techniques to determine the loss weight [102,103].

Another potential future work is the analysis of the hard instances in the patch-wise learning scheme. Some vessels tend to be more difficult to segment, such as the vessels adjacent to bones. The image patches containing these vessels can be regarded as the hard instances. For these hard instances, increasing their sampling rate or giving them larger weight in the loss function may help to improve the performance of the model further.

Analysis of disease-specific data is also a valuable topic for future research, such as the dataset of intracranial bleeding. Previous studies have shown that the virtual unenhanced (VU) images are sufficient for the detection of intracranial bleeding when compared with true UECT images [5,39]. If the proposed unenhanced prediction can replace the true UECT in certain situations, just as the VU replaced the UECT in some conditions, the radiation dose can potentially be reduced.

In conclusion, we proposed a dual-task deep learning framework for simultaneous vessel segmentation and unenhanced prediction using a self-supervised strategy. The dual-task framework can significantly improve the performances of two single-tasks. The proposed method achieved significantly better vessel segmentation performance in HNCTA compared with the state-of-the-art biomedical image segmentation model (VNet), and several popular vascular segmentation models. The unenhanced prediction, which is introduced in this study, also had the potential to achieve comparable performance to the virtual unenhanced reconstruction in terms of the error level in the vascular region. Moreover, to the best of our knowledge, we present the first DL based model for the segmentation of all vessels in HNCTA that requires no manual annotation for model training.

## Declaration of Competing Interest

The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.cmpb.2022.107001](https://doi.org/10.1016/j.cmpb.2022.107001).

## References

- [1] C.O. Johnson, M. Nguyen, G.A. Roth, E. Nichols, T. Alam, D. Abate, F. Abd-Allah, A. Abdelalim, H.N. Abraha, N.M. Abu-Rmeileh, et al., Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the global burden of disease study 2016, *Lancet Neurol.* 18 (5) (2019) 439–458.
- [2] J. Yang, M. Xie, C. Hu, O. Alwalid, Y. Xu, J. Liu, T. Jin, C. Li, D. Tu, X. Liu, et al., Deep learning for detecting cerebral aneurysms with ct angiography, *Radiology* 298 (1) (2021) 155–163.
- [3] R. Fronczek, M.P. dos Santos, M.A. van Walderveen, F.J. Meijer, P.W. Willems, et al., High sensitivity and specificity of 4D-CTA in the detection of cranial arteriovenous shunts, *Eur. Radiol.* 29 (11) (2019) 5961–5970.
- [4] A.A. Postma, M. Das, A.A. Stadler, J.E. Wildberger, Dual-energy CT: what the neuroradiologist should know, *Curr. Radiol. Rep.* 3 (5) (2015) 1–16.
- [5] M. Bonatti, F. Lombardo, G.A. Zamboni, P. Pernter, R. Pozzi Mucelli, G. Bonatti, Dual-energy CT of the brain: comparison between DECT angiography-derived virtual unenhanced images and true unenhanced images in the detection of intracranial haemorrhage, *Eur. Radiol.* 27 (7) (2017) 2690–2697.
- [6] M.C. Patel, D.C. Levin, L. Parker, V.M. Rao, Have CT and MR angiography replaced catheter angiography in diagnosing peripheral arterial disease? *J. Am. Coll. Radiol.* 12 (9) (2015) 909–914.
- [7] C. Denby, K. Chatterjee, R. Pullicino, S. Lane, M. Radon, K. Das, Is four-dimensional ct angiography as effective as digital subtraction angiography in the detection of the underlying causes of intracerebral haemorrhage: a systematic review, *Neuroradiology* 62 (3) (2020) 273–281.
- [8] H. Scherl, J. Hornegger, M. Prümmer, M. Lell, Semi-automatic level-set based segmentation and stenosis quantification of the internal carotid artery in 3D CTA data sets, *Med. Image Anal.* 11 (1) (2007) 21–34.
- [9] R. Shahzad, H. Kirişli, C. Metz, H. Tang, M. Schaap, L. van Vliet, W. Niessen, T. van Walsum, Automatic segmentation, detection and quantification of coronary artery stenoses on CTA, *Int. J. Cardiovasc. Imaging* 29 (8) (2013) 1847–1859.
- [10] F. Bozkurt, C. Köse, A. Sarı, An inverse approach for automatic segmentation of carotid and vertebral arteries in CTA, *Expert Syst. Appl.* 93 (2018) 358–375.
- [11] J. Gounley, M. Vardhan, A. Randles, A framework for comparing vascular hemodynamics at different points in time, *Comput. Phys. Commun.* 235 (2019) 1–8.
- [12] Y. Sen, Y. Qian, A. Avolio, M. Morgan, Image segmentation methods for intracranial aneurysm haemodynamic research, *J. Biomech.* 47 (5) (2014) 1014–1019.
- [13] R.P. Kumar, F. Albregtsen, M. Reimers, B. Edwin, T. Langø, O.J. Elle, Three-dimensional blood vessel segmentation and centerline extraction based on two-dimensional cross-section analysis, *Ann. Biomed. Eng.* 43 (5) (2015) 1223–1234.
- [14] F. Fu, J. Wei, M. Zhang, F. Yu, Y. Xiao, D. Rong, Y. Shan, Y. Li, C. Zhao, F. Liao, et al., Rapid vessel segmentation and reconstruction of head and neck angiograms using 3D convolutional neural network, *Nat. Commun.* 11 (1) (2020) 1–12.
- [15] S. Behrens, Automatic level set based cerebral vessel segmentation and bone removal in CT angiography data sets, in: *GCPR*, 2013, pp. 237–242.
- [16] J. Yi, J.B. Ra, A locally adaptive region growing algorithm for vascular segmentation, *Int. J. Imaging Syst. Technol.* 13 (4) (2003) 208–214.
- [17] Y. Cheng, X. Hu, J. Wang, Y. Wang, S. Tamura, Accurate vessel segmentation with constrained b-snake, *IEEE Trans. Image Process.* 24 (8) (2015) 2440–2455.
- [18] R. Manniesing, M. Schaap, S. Rozie, R. Hameeteman, D. Vukadinovic, A. van der Lugt, W. Niessen, Robust CTA lumen segmentation of the atherosclerotic carotid artery bifurcation in a large patient population, *Med. Image Anal.* 14 (6) (2010) 759–769.
- [19] K. Hameeteman, M.A. Zuluaga, M. Freiman, L. Joskowicz, O. Cuisenaire, L.F. Valencia, M.A. Gülsün, K. Krissian, J. Mille, W.C. Wong, et al., Evaluation framework for carotid bifurcation lumen segmentation and stenosis grading, *Med. Image Anal.* 15 (4) (2011) 477–488.
- [20] A. Deshpande, N. Jamilpour, B. Jiang, P. Michel, A. Eskandari, C. Kidwell, M. Wintermark, K. Laksari, Automatic segmentation, feature extraction and comparison of healthy and stroke cerebral vasculature, *NeuroImage Clin.* 30 (2021) 102573.
- [21] X. Yang, C. Liu, H. Le Minh, Z. Wang, A. Chien, K.-T.T. Cheng, An automated method for accurate vessel segmentation, *Phys. Med. Biol.* 62 (9) (2017) 3757.
- [22] D. Lesage, E.D. Angelini, I. Bloch, G. Funka-Lea, A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes, *Med. Image Anal.* 13 (6) (2009) 819–845.
- [23] S. Moccia, E. De Momi, S. El Hadji, L.S. Mattos, Blood vessel segmentation algorithms-review of methods, datasets and evaluation metrics, *Comput. Methods Programs Biomed.* 158 (2018) 71–91.
- [24] A. Ajam, A.A. Aziz, V.S. Asirvadam, A.S. Muda, I. Faye, S.J.S. Gardezi, A review on segmentation and modeling of cerebral vasculature for surgical planning, *IEEE Access* 5 (2017) 15222–15240.
- [25] Y. Zheng, M. Loziczonek, B. Georgescu, S.K. Zhou, F. Vega-Higuera, D. Comaniciu, Machine learning based vesselness measurement for coronary artery segmentation in cardiac CT volumes, in: *Medical Imaging 2011: Image Processing*, vol. 7962, 2011, pp. 489–500.
- [26] G. Ma, Y. Yu, H. Duan, Y. Dou, Y. Jia, X. Zhang, C. Yang, X. Chen, D. Han, C. Guo, et al., Subtraction CT angiography in head and neck with low radiation and contrast dose dual-energy spectral CT using rapid kv-switching technique, *Br. J. Radiol.* 91 (1086) (2018) 20170631.

- [27] D. Morhard, C. Fink, C. Becker, M.F. Reiser, K. Nikolaou, Value of automatic bone subtraction in cranial CT angiography: comparison of bone-subtracted vs. standard CT angiography in 100 patients, *Eur. Radiol.* 18 (5) (2008) 974–982.
- [28] W. Chen, W. Xing, Z. He, Y. Peng, C. Wang, Q. Wang, Accuracy of 320-detector row nonsubtracted and subtracted volume CT angiography in evaluating small cerebral aneurysms, *J. Neurosurg.* 127 (4) (2016) 725–731.
- [29] Z. Shi, C. Miao, U.J. Schoepf, R.H. Savage, D.M. Dargis, C. Pan, X. Chai, X.L. Li, S. Xia, X. Zhang, et al., A clinically applicable deep-learning model for detecting intracranial aneurysm in computed tomography angiography images, *Nat. Commun.* 11 (1) (2020) 1–11.
- [30] M. Lell, K. Anders, E. Klotz, H. Ditt, W. Bautz, B. Tomandl, Clinical evaluation of bone-subtraction ct angiography (BSCTA) in head and neck imaging, *Eur. Radiol.* 16 (4) (2006) 889–897.
- [31] P. Aulbach, D. Mucha, K. Engellhardt, K. Hädrich, M. Kuhn, R. von Kummer, Diagnostic impact of bone-subtraction CT angiography for patients with acute subarachnoid hemorrhage, *AJNR Am. J. Neuroradiol.* 37 (2) (2016) 236–243.
- [32] R. Manniesing, M.A. Viergever, A. van der Lugt, W.J. Niessen, Cerebral arteries: fully automated segmentation from CT angiography'a feasibility study, *Radiology* 247 (3) (2008) 841–846.
- [33] M.J. Siegel, J.C. Ramirez-Giraldo, Dual-energy CT in children: imaging algorithms and clinical applications, *Radiology* 291 (2) (2019) 286–297.
- [34] M. Patino, A. Prochowski, M.D. Agrawal, F.J. Simeone, R. Gupta, P.F. Hahn, D.V. Sahani, Material separation using dual-energy CT: current and emerging applications, *Radiographics* 36 (4) (2016) 1087–1105.
- [35] M. van Straten, M. Schaap, M.L. Dijkshoorn, M.J. Greuter, A. van der Lugt, G.P. Kreftin, W.J. Niessen, Automated bone removal in CT angiography: comparison of methods based on single energy and dual energy scans, *Med. Phys.* 38 (11) (2011) 6128–6137.
- [36] R. Gupta, M. Poirot, R. Bergmans, Clinical applications of dual energy CT in neuroradiology, in: *Spectral, Photon Counting Computed Tomography: Technology and Applications*, 2020, pp. 17–34.
- [37] K. Kalisz, S. Halliburton, S. Abbala, J.A. Leipsic, M.H. Albrecht, U.J. Schoepf, P. Rajiah, Update on cardiovascular applications of multienergy CT, *Radiographics* 37 (7) (2017) 1955–1974.
- [38] M. Meyer, R.C. Nelson, F. Vernuccio, F. González, A.E. Farjat, B.N. Patel, E. Samei, T. Henzler, S.O. Schoenberg, D. Marin, Virtual unenhanced images at dual-energy CT: influence on renal lesion characterization, *Radiology* 291 (2) (2019) 381–390.
- [39] X. Jiang, S. Zhang, Q. Xie, Z. Yin, Q. Liu, M. Zhao, X. Li, X. Mao, Evaluation of virtual noncontrast images obtained from dual-energy CTA for diagnosing subarachnoid hemorrhage, *AJR Am. J. Roentgenol.* 195 (5) (2015) 855–860.
- [40] T.R.C. Johnson, Dual-energy CT: general principles, *AJR Am. J. Roentgenol.* 199 (5\_supplement) (2012) S3–S8.
- [41] R. Symons, T.E. Cork, M.N. Lakshmanan, R. Evers, C. Davies-Venn, K.A. Rice, M.L. Thomas, C.-Y. Liu, S. Kappler, S. Ulzheimer, et al., Dual-contrast agent photon-counting computed tomography of the heart: initial experience, *Int. J. Card. Imaging* 33 (8) (2017) 1253–1261.
- [42] J. Oldan, M. He, T. Wu, A.C. Silva, J. Li, J.R. Mitchell, W.M. Pavlicek, M.C. Roarke, A.K. Hara, Pilot study: evaluation of dual-energy computed tomography measurement strategies for positron emission tomography correlation in pancreatic adenocarcinoma, *J. Digit. Imaging* 27 (6) (2014) 824–832.
- [43] N. Dalbeth, H.K. Choi, Dual-energy computed tomography for gout diagnosis and management, *Curr. Radiol. Rep.* 15 (1) (2013) 1–7.
- [44] Y. Liu, J. Cheng, L. Zhang, Y. Xing, Z. Chen, P. Zheng, A low-cost dual energy CT system with sparse data, *Tsinghua Sci. Technol.* 19 (2) (2014) 184–194.
- [45] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, in: *MICCAI*, vol. 9351, 2015, pp. 234–241.
- [46] V-Net: fully convolutional neural networks for volumetric medical image segmentation, *Int. Conf. 3DV*, 2016, pp. 565571.
- [47] T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, et al., U-Net: deep learning for cell counting, detection, and morphometry, *Nat. Methods* 16 (1) (2019) 67–70.
- [48] H.R. Roth, C. Shen, H. Oda, M. Oda, Y. Hayashi, K. Misawa, K. Mori, Deep learning and its application to medical image segmentation, *Med. Imaging Technol.* 36 (2) (2018) 63–71.
- [49] E.M. Christiansen, S.J. Yang, D.M. Ando, A. Javaherian, G. Skibinski, S. Lipnick, E. Mount, A. O'neil, K. Shah, A.K. Lee, et al., In silico labeling: predicting fluorescent labels in unlabeled images, *Cell* 173 (3) (2018) 792–803.
- [50] X. Han, Mr-based synthetic ct generation using a deep convolutional neural network method, *Med. Phys.* 44 (4) (2017) 1408–1419.
- [51] Y. Rivenson, H. Ceylan Koydemir, H. Wang, Z. Wei, Z. Ren, H. Günaydin, Y. Zhang, Z. Gorocs, K. Liang, D. Tseng, et al., Deep learning enhanced mobile-phone microscopy, *ACS Photonics* 5 (6) (2018) 2354–2364.
- [52] Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, A. Ozcan, Deep learning microscopy, *Optica* 4 (11) (2017) 1437–1443.
- [53] C. Ounkomol, S. Seshamani, M.M. Maleckar, F. Collman, G.R. Johnson, Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy, *Nat. Methods* 15 (11) (2018) 917–920.
- [54] D. Shen, G. Wu, H.-I. Suk, Deep learning in medical image analysis, *Annu. Rev. Biomed. Eng.* 19 (2017) 221–248.
- [55] J. Ni, J. Wu, H. Wang, J. Tong, Z. Chen, K.K. Wong, D. Abbott, Global channel attention networks for intracranial vessel segmentation, *Comput. Biol. Med.* 118 (2020) 103639.
- [56] L. Yao, P. Jiang, Z. Xue, Y. Zhan, D. Wu, L. Zhang, Q. Wang, F. Shi, D. Shen, Graph convolutional network based point cloud for head and neck vessel labeling, in: *International Workshop on Machine Learning in Medical Imaging*, vol. 12436, 2020, pp. 474–483.
- [57] G. Tetteh, V. Efremov, N.D. Forkert, M. Schneider, J. Kirschke, B. Weber, C. Zimmer, M. Piraud, B.H. Menze, DeepVesselNet: vessel segmentation, centerline prediction, and bifurcation detection in 3-D angiographic volumes, *Front. Neurosci.* (2020) 1285.
- [58] T. Zhou, T. Tan, X. Pan, H. Tang, J. Li, Fully automatic deep learning trained on limited data for carotid artery segmentation from large image volumes, *Quant. Imaging. Med. Surg.* 11 (1) (2021) 67.
- [59] Z. Kang, K. Grauman, F. Sha, Learning with whom to share in multi-task feature learning, in: *ICML*, 2011, pp. 521–528.
- [60] Y. Zhang, Q. Yang, An overview of multi-task learning, *Natl. Sci. Rev.* 5 (1) (2018) 30–43.
- [61] M. Marsden, K. McGuinness, S. Little, N.E. O'Connor, ResnetCrowd: a residual deep learning architecture for crowd counting, violent behaviour detection and crowd density level classification, in: *Int. Conf. AVSS*, 2017, pp. 1–7.
- [62] S. Ruder, An overview of multi-task learning in deep neural networks, *arXiv preprint arXiv:1706.05098* (2017).
- [63] I. Misra, A. Shrivastava, A. Gupta, M. Hebert, Cross-stitch networks for multi-task learning, in: *CVPR*, 2016, pp. 3994–4003.
- [64] W. Huang, H. Li, R. Wang, X. Zhang, X. Wang, J. Zhang, A self-supervised strategy for fully automatic segmentation of renal dynamic contrast-enhanced magnetic resonance images, *Med. Phys.* 46 (10) (2019) 4417–4430.
- [65] M.W. Law, A.C. Chung, Efficient implementation for spherical flux computation and its application to vascular segmentation, *IEEE Trans. Med. Imaging* 18 (3) (2009) 596–612.
- [66] M.W.K. Law, A.C.S. Chung, Three dimensional curvilinear structure detection using optimally oriented flux, in: *ECCV*, vol. 5305, 2008, pp. 368–382.
- [67] T. Jerman, F. Pernuš, B. Likar, Ž. Špiclin, Enhancement of vascular structures in 3D and 2D angiographic images, *IEEE Trans. Med. Imaging* 35 (9) (2016) 2107–2118.
- [68] M. Schneider, S. Hirsch, B. Weber, G. Székely, B.H. Menze, Joint 3-D vessel segmentation and centerline extraction using oblique hough forests with steerable filters, *Med. Image Anal.* 19 (1) (2015) 220–249.
- [69] D. Han, H. Shim, B. Jeon, Y. Jang, Y. Hong, S. Jung, S. Ha, H.-J. Chang, Automatic coronary artery segmentation using active search for branches and seemingly disconnected vessel segments from coronary CT angiography, *PLoS ONE* 11 (8) (2016) e0156837.
- [70] F. Lugauer, Y. Zheng, J. Hornegger, B.M. Kelm, Precise lumen segmentation in coronary computed tomography angiography, in: *Int. MICCAI Workshop on Medical Computer Vision*, vol. 8848, 2014, pp. 137–147.
- [71] A. Kerkeni, A. Benabdallah, A. Manzanera, M.H. Bedoui, A coronary artery segmentation method based on multiscale analysis and region growing, *Comput. Med. Imaging Graph* 48 (2016) 49–61.
- [72] M. Chung, J. Lee, J.W. Chung, Y.-G. Shin, Accurate liver vessel segmentation via active contour model with dense vessel candidates, *Comput. Methods Programs Biomed.* 166 (2018) 61–75.
- [73] M. Orkisz, M.H. Hoyos, V.P. Romanello, C.P. Romanello, J. Prieto, C. Revol-Muller, Segmentation of the pulmonary vascular trees in 3D CT images using variational region-growing, *IRBM* 35 (1) (2014) 11–19.
- [74] H.M. Luu, C. Klinik, A. Moelker, W. Niessen, T. Van Walsum, Quantitative evaluation of noise reduction and vesseness filters for liver vessel segmentation on abdominal CTA images, *Phys. Med. Biol.* 60 (10) (2015) 3905.
- [75] H. Kırılış, M. Schaap, C. Metz, A. Dharampal, W.B. Meijboom, S.-L. Papadopoulou, A. Dedic, K. Nieman, M.A. de Graaf, M. Meijls, et al., Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography, *Med. Image Anal.* 17 (8) (2013) 859–876.
- [76] X. Chen, Y. Lu, J. Bai, Y. Yin, K. Cao, Y. Li, H. Chen, Q. Song, J. Wu, Train a 3D U-Net to segment cranial vasculature in CTA volume without manual annotation, in: *ISBI*, 2018, pp. 559–563.
- [77] Q. Chu, Y. Zhan, F. Guo, M. Song, R. Yang, Automatic 3D registration of CT-MR head and neck images with surface matching, *IEEE Access* 7 (2019) 78274–78280.
- [78] Á.S. Hervella, J. Rouco, J. Novo, M. Ortega, Self-supervised deep learning for retinal vessel segmentation using automatically generated labels from multimodal data, in: *IJCNN*, 2019, pp. 1–8.
- [79] Y. Wang, J. Zhang, M. Kan, S. Shan, X. Chen, Self-supervised equivariant attention mechanism for weakly supervised semantic segmentation, in: *CVPR*, 2020, pp. 12272–12281.
- [80] K. Gong, X. Liang, D. Zhang, X. Shen, L. Lin, Look into person: self-supervised structure-sensitive learning and a new benchmark for human parsing, in: *CVPR*, 2017, pp. 6757–6765.
- [81] S. Zhou, J. Xi, M.W. McDaniel, T. Nishihata, P. Salesses, K. Iagnemma, Self-supervised learning to visually detect terrain surfaces for autonomous robots operating in forested terrain, *J. Field Robot.* 29 (2) (2012) 277–297.
- [82] F. Taherkhani, A. Dabouei, S. Soleymani, J.M. Dawson, N.M. Nasrabadi, Self-supervised Wasserstein pseudo-labeling for semi-supervised image classification, in: *CVPR*, 2021, pp. 12267–12277.
- [83] C.A. Brooks, K. Iagnemma, Self-supervised terrain classification for planetary surface exploration rovers, *J. Field Robot.* 29 (3) (2012) 445–468.
- [84] L. Gomez-Bigorda, Y. Patel, M. Rusiñol, D. Karatzas, C.V. Jawahar, Self-supervised learning of visual features through embedding images into text topic spaces, in: *CVPR*, 2017, pp. 2017–2026.
- [85] S. Achar, B. Sankaran, S. Nuske, S.A. Scherer, S. Singh, Self-supervised segmentation of river scenes, in: *ICRA*, 2011, pp. 6227–6232.

- [86] S. Gidaris, P. Singh, N. Komodakis, Unsupervised representation learning by predicting image rotations, ICLR, 2018.
- [87] L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara, D. Rueckert, Self-supervised learning for medical image analysis using image context restoration, *Med. Image Anal.* 58 (2019) 101539.
- [88] E. Schubert, J. Sander, M. Ester, H.P. Kriegel, X. Xu, DBSCAN revisited, revisited: why and how you should (still) use DBSCAN, *ACM Trans. Database Syst.* 42 (3) (2017) 1–21.
- [89] S. Ren, K. He, R.B. Girshick, J. Sun, Faster R-CNN: towards real-time object detection with region proposal networks, in: NeurIPS, 2015, pp. 91–99.
- [90] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, in: CVPR, 2017, pp. 2961–2969.
- [91] J. Li, K.V. Sarma, K.C. Ho, A. Gertych, B.S. Knudsen, C.W. Arnold, A multi-scale U-Net for semantic segmentation of histological images from radical prostatectomies, in: AMIA Annual Symposium Proc., 2017, p. 1140.
- [92] L. Song, J. Lin, Z.J. Wang, H. Wang, Dense-residual attention network for skin lesion segmentation, in: Int. Workshop on Machine Learning in Medical Imaging, 2019, pp. 319–327.
- [93] S. Kumawat, S. Raman, Local phase U-Net for fundus image segmentation, in: ICASSP, 2019, pp. 1209–1213.
- [94] K. van Garderen, M. Smits, S. Klein, Multi-modal segmentation with missing MR sequences using pre-trained fusion networks, in: Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data, 2019, pp. 165–172.
- [95] M. Wu, C. Zhang, J. Liu, L. Zhou, X. Li, Towards accurate high resolution satellite image semantic segmentation, *IEEE Access* 7 (2019) 55609–55619.
- [96] D. Bermejo-Peláez, Y. Okajima, G.R. Washko, M.-J. Ledesma-Carbayo, R.S.J. Estépar, A SR-NET 3D-to-2D architecture for paraseptal emphysema segmentation, in: ISBI, 2019, pp. 303–306.
- [97] P. Nardelli, D. Jimenez-Carretero, D. Bermejo-Pelaez, G.R. Washko, F.N. Rahaghi, M.J. Ledesma-Carbayo, R.S.J. Estépar, Pulmonary artery–vein classification in CT images using deep learning, *IEEE Trans. Med. Imaging* 37 (11) (2018) 2428–2440.
- [98] M. Toepker, T. Moritz, B. Krauss, M. Weber, G. Euller, T. Mang, F. Wolf, C.J. Herold, H. Ringl, Virtual non-contrast in second-generation, dual-energy computed tomography: reliability of attenuation values, *Eur. J. Radiol.* 81 (3) (2012) e398–e405.
- [99] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88.
- [100] L. Fidon, W. Li, L.C. García-Peraza-Herrera, J. Ekanayake, N. Kitchen, S. Ourselin, T. Vercauteren, Generalised Wasserstein dice score for imbalanced multi-class segmentation using holistic convolutional networks, in: Int. MICCAI Brainlesion Workshop, 2017, pp. 64–76.
- [101] Z. Zhang, M.R. Sabuncu, Generalized cross entropy loss for training deep neural networks with noisy labels, in: NeurIPS, 2018, pp. 8792–8802.
- [102] Z. Chen, V. Badrinarayanan, C.-Y. Lee, A. Rabinovich, GradNorm: gradient normalization for adaptive loss balancing in deep multitask networks, in: ICML, 2018, pp. 794–803.
- [103] A. Kendall, Y. Gal, R. Cipolla, Multi-task learning using uncertainty to weigh losses for scene geometry and semantics, in: CVPR, 2018, pp. 7482–7491.