US Patent & Trademark Office
Patent Public Search | Text View

United States Patent	12387337
Kind Code	B2
Date of Patent	August 12, 2025
Inventor(s)	Siemionow; Kris et al.

Autonomous segmentation of contrast filled coronary artery vessels on computed tomography images

Abstract

A computer-implemented method for autonomous segmentation of contrast-filled coronary artery vessels includes receiving a CT scan volume representing a 3D volume of a region of anatomy that includes a pericardium; preprocessing the CT scan volume to output a preprocessed scan volume; dividing the CT scan volume into a first set of subvolumes; extracting a region of interest by autonomous segmentation of the heart region as outlined by the pericardium, by means of a neural network trained on 3D subvolumes and combining the results of the individual subvolume predictions for the first set to output a mask denoting a heart region as delineated by the pericardium; combining the preprocessed scan volume with the mask to obtain a masked volume; converting the masked volume to a second set of 3D subvolumes; and performing autonomous coronary vessel segmentation to output a mask denoting the coronary vessels.

Inventors: **Siemionow; Kris (Chicago, IL), Kraft; Marek (Poznan, PL), Pieczynski; Dominik (Tulce, PL), Lewicki; Paul (Tulsa, OK), Malota; Zbigniew (Zabrze, PL), Sadowski; Wojciech (Zabrze, PL), Kania; Jacek (Rogozno, PL)**

Applicant: **Kardiolytics Inc.** (Tulsa, OK)

Family ID: **1000008750780**

Assignee: **Kardiolytics Inc. (Tulsa, OK)**

Appl. No.: **17/714170**

Filed: **April 06, 2022**

Prior Publication Data

Document Identifier	Publication Date
US 20220230320 A1	Jul. 21, 2022

Related U.S. Application Data

continuation-in-part parent-doc US 16895024 20200608 US 11315293 child-doc US 17714170
us-provisional-application US 62830441 20190406

Publication Classification

Int. Cl.: **G06T7/10** (20170101); **G06T5/75** (20240101); **G06T7/00** (20170101); **G06T7/62** (20170101); **G06T15/08** (20110101)

U.S. Cl.:

CPC **G06T7/10** (20170101); **G06T5/75** (20240101); **G06T7/0012** (20130101); **G06T7/62** (20170101); G06T15/08 (20130101); G06T2200/04 (20130101); G06T2207/20021 (20130101); G06T2207/20024 (20130101); G06T2207/20084 (20130101)

Field of Classification Search

CPC: G06T (5/75); G06T (7/0012); G06T (7/10); G06T (7/11); G06T (7/62); G06T (15/08); G06T (2200/04); G06T (2207/10081); G06T (2207/20024); G06T (2207/20081); G06T (2207/20084); G06T (2207/30048); G06T (2207/30101)

References Cited

U.S. PATENT DOCUMENTS

Patent No.	Issued Date	Patentee Name	U.S. Cl.	CPC
9968257	12/2017	Burt	N/A	A61B 5/0035
2018/0061059	12/2017	Xu	N/A	G06T 7/174
2020/0219609	12/2019	Harte	N/A	G16H 50/20
2020/0410687	12/2019	Siemionow	N/A	G06T 7/62
2022/0237801	12/2021	Kaufman	N/A	G06N 3/045

OTHER PUBLICATIONS

Sharobeem et al. “Validation of a whole heart segmentation from computed tomography imaging using a deep-learning approach.” Journal of Cardiovascular Translational Research 15.2 (2021): 427-437. (Year: 2021). cited by examiner
Payer et al. “Multi-label whole heart segmentation using CNNs and anatomical label configurations.” International Workshop on Statistical Atlases and Computational Models of the Heart. Cham: Springer International Publishing, 2017. (Year: 2017). cited by examiner
Harms et al. “Automatic delineation of cardiac substructures using a region-based fully convolutional network.” Medical Physics 48.6 (2021): 2867-2876. (Year: 2021). cited by examiner
Jerman et al. “Enhancement of vascular structures in 3D and 2D angiographic images.” IEEE transactions on medical imaging 35.9 (2016): 2107-2118. (Year: 2016). cited by examiner
Liu et al. “Automatic whole heart segmentation using a two-stage u-net framework and an adaptive threshold window.” IEEE Access 7 (2019): 83628-83636. (Year: 2019). cited by examiner
Wang et al. “Automatic whole heart segmentation using deep learning and shape context.” Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Sep. 10-14, 2017, Revised Selected Papers 8. Springer International Publishing, 2018 (Year: 2018). cited by examiner
Tong et al. “3D deeply-supervised U-net based whole heart segmentation.” Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop,

STACOM 2017, Sep. 10-14, 2017, Revised Selected Papers 8. Springer International Publishing, 2018 (Year: 2018). cited by examiner
Galea et al. “Region-of-interest-based cardiac image segmentation with deep learning.” Applied Sciences 11.4 (2021): 1965. (Year: 2021). cited by examiner
Morris et al. “Cardiac substructure segmentation with deep learning for improved cardiac sparing.” Medical physics 47.2 (2019): 576-586. (Year: 2019). cited by examiner

---

Primary Examiner: Fujita; Katrina R

Attorney, Agent or Firm: Rivka Friedman

Background/Summary

TECHNICAL FIELD
(1) The invention generally relates to autonomous segmentation of contrast filled coronary artery vessels on computed tomography images, useful in particular for the field of computer assisted diagnosis, treatment, and monitoring of coronary artery diseases.
BACKGROUND
(2) Specialized computer systems can be used to process the CT images to develop three-dimensional models of the anatomy fragments. For this purpose, various machine learning technologies are developed, such as a convolutional neural network (CNN) that is a class of deep, feed-forward artificial neural networks. CNNs use a variation of feature detectors and/or multilayer perceptrons designed to require minimal preprocessing of input data.
SUMMARY OF THE INVENTION
(3) So far, the image processing systems were not capable of efficiently providing autonomous segmentation of contrast filled coronary artery vessels on CT images and, therefore, Applicant has recognized a need to provide improvements in this area.
(4) Certain embodiments disclosed herein relate to machine learning based detection of vascular structures in medical images, and more particularly, to machine learning based detection of coronary vessels in computed tomography (CT) images. Automatic detection and segmentation of contrast filled coronary arteries CT scans facilitates the diagnosis, treatment, and monitoring of coronary artery diseases.
(5) In one aspect, there is disclosed a computer-implemented method for autonomous segmentation of contrast-filled coronary artery vessels, the method comprising: a) receiving a CT scan volume representing a 3D volume of a region of anatomy that includes a pericardium; b) preprocessing the CT scan volume to output a preprocessed scan volume; c) converting the CT scan volume to three sets of two-dimensional slices, wherein the first set is arranged along the axial plane, the second set is arranged along the sagittal plane and the third set is arranged along the coronal plane; d) extracting a region of interest by autonomous segmentation of the heart region as outlined by the pericardium, by means of three individually trained ROI extraction convolutional neural networks, each trained to process a particular one of the three sets of two-dimensional slices to output a mask denoting a heart region as delineated by the pericardium; e) combining the preprocessed scan volume with the mask to obtain a masked volume; f) converting the masked volume to three groups of sets of two-dimensional masked slices, wherein the first group is arranged along the axial plane, the second group is arranged along the sagittal plane and the third group is arranged along the coronal plane and each group includes at least three sets, wherein the first set corresponds to the principal plane of the set and at least two other sets are tilted with respect to the principal plane; g) performing autonomous coronary vessel segmentation by autonomous segmentation of the sets of

the two-dimensional masked slices by means of three individually trained segmentation convolutional neural networks, each trained to process a particular one of the sets of the two-dimensional masked slices to output a mask denoting the coronary vessels.
(6) In another aspect, there is disclosed a computer-implemented method for autonomous segmentation of contrast-filled coronary artery vessels, the method comprising: a) receiving a CT scan volume representing a 3D volume of a region of anatomy that includes a pericardium; b) preprocessing the CT scan volume to output a preprocessed scan volume; c) dividing the CT scan volume into a first set of subvolumes; d) extracting a region of interest by autonomous segmentation of the heart region as outlined by the pericardium, by means of a neural network trained on 3D subvolumes and combining the results of the individual subvolume predictions for the first set to output a mask denoting a heart region as delineated by the pericardium; e) combining the preprocessed scan volume with the mask to obtain a masked volume; f) converting the masked volume to a second set of 3D subvolumes; g) performing autonomous coronary vessel segmentation by autonomous segmentation of the subvolumes of the second set by means of a trained segmentation convolutional neural network, trained to process a subvolume to output a mask denoting the coronary vessels.
(7) The step of preprocessing the CT scan may include performing at least one of: windowing, filtering and normalization.
(8) The step of preprocessing the CT scan may include computing a 3D Jerman filter response.
(9) The method may further comprise combining the 3D Jerman filter response with the mask to obtain a masked Jerman-filtered volume, converting the masked Jerman-filtered volume to three groups of sets of two-dimensional masked Jerman-filtered slices and providing the two-dimensional masked Jerman-filtered slices as an input to a second channel of the segmentation convolutional neural networks.
(10) The method may further comprise combining the masks denoting the coronary vessels to a segmented 3D data set representing the shape, location and size of the coronary vessels.
(11) In another aspect, there is disclosed a computer-implemented system, comprising: at least one nontransitory processor-readable storage medium that stores at least one of processor-executable instructions or data; and at least one processor communicably coupled to the at least one nontransitory processor-readable storage medium, wherein the at least one processor is configured to perform the steps of the method in accordance with any of the embodiments described above.
(12) These and other features, aspects and advantages of the invention will become better understood with reference to the following drawings, descriptions and claims.

Description

BRIEF DESCRIPTION OF DRAWINGS
(1) Various embodiments are herein described, by way of example only, with reference to the accompanying drawings, wherein:
(2) FIG. **1** shows an overview of a procedure for segmentation of the artery vessels;
(3) FIG. **2**A shows an example of a 3D image and individual planes location in accordance with a first embodiment of the segmentation procedure;
(4) FIG. **2**B shows an example of an actual image of a heart and coronary arteries in accordance with the first embodiment of the segmentation procedure;
(5) FIG. **2**C shows an example of an image processed using Jerman filter in accordance with the first embodiment of the segmentation procedure;
(6) FIG. **2**D shows examples of slices of the 3D image in 3 individual planes in accordance with the first embodiment of the segmentation procedure;
(7) FIG. **2**E shows an example of a masked raw volume in accordance with the first embodiment of

the segmentation procedure;
(8) FIG. **2**F shows an example of a masked volumetric Jerman filter response in accordance with the first embodiment of the segmentation procedure;
(9) FIG. **3** shows examples of final results of the first and second embodiment of the segmentation procedure;
(10) FIG. **4** shows a schematic representation of a ROI extraction CNN for use in the first or second embodiment of the segmentation procedure;
(11) FIG. **5** shows an outline of a training procedure for the ROI extraction CNN in accordance with one embodiment of the training procedure;
(12) FIG. **6**A shows a set of planes in accordance with the first embodiment of the segmentation procedure;
(13) FIG. **6**B shows projection of planes on a plane analogous to the regular plane in accordance with the first embodiment of the segmentation procedure;
(14) FIG. **6**C shows sample annotation and desired result in accordance with the first embodiment of the segmentation procedure;
(15) FIG. **7** shows the structure of a computer system for implementing the method of FIG. **1** in accordance with one embodiment of a computer system.

DETAILED DESCRIPTION OF THE EMBODIMENTS

(16) The description is not to be taken in a limiting sense, but is made merely for the purpose of illustrating the general principles of the invention.
(17) The overview of a segmentation method, including a first embodiment (with steps **103**A, **104**A, **106**A, **107**A) and a second embodiment (with steps **103**B, **104**B, **106**B, **107**B) is presented in detail in FIG. **1**. In step **101**, a computer tomography (CT) volumetric scan (also called a three-dimensional (3D) scan or a volume) is received. The CT volume comprises a set of medical scan images of a region of the anatomy, such as a set of DICOM (Digital Imaging and Communications in Medicine) images. The set **201** represents consecutive slices of the region of the anatomy, such as illustrated in FIG. **2**A.
(18) The region of the anatomy should be selected such that it contains the heart and the coronary arteries **202**, such as shown in FIG. **2**B.
(19) In step **102**, the 3D volume is autonomously preprocessed to prepare the images for region of interest (ROI) extraction. This preprocessing step may comprise raw 3D CT data windowing, filtering and normalization, as well as computing the 3D Jerman filter response for the whole volume. Computing the Jerman filter can be performed in accordance with the article "Enhancement of Vascular Structures in 3D and 2D Angiographic Images" (by T. Jerman, et al., IEEE Transactions on Medical Imaging, 35(9), p. 2107-2118 (2016)). The Jerman filter emphasizes elongated structures in images and volumes. An example of applying the filter on infrared hand vessel pattern image (left) **203** is shown in FIG. **2**C, wherein the right image **204** shows the output, processed image.
(20) Next, in accordance with a first embodiment of the segmentation procedure, in step **103**A the 3D volume is converted to 3 sets of two-dimensional (2D) slices, wherein the first set is arranged along the axial plane, the second set is arranged along the sagittal plane and the third set is arranged along the coronal plane (as marked in FIG. **2**A). Next, in step **104**A a region of interest (ROI) is extracted by autonomous segmentation of the heart region as outlined by the pericardium. The procedure is performed by three individually trained convolutional neural networks (CNNs), each for processing a particular one of the three sets of 2D slices, namely an axial plane ROI extraction CNN, a sagittal plane ROI extraction CNN and a coronal plane ROI extraction CNN. These three CNNs are trained by training data that consists of pairs of CT volume slices in its corresponding plane and its corresponding binary, expert-annotated mask, denoting the heart region as delineated by the pericardium. Direct correspondence of binary masks and CT scan data enables their direct use for segmentation training. Sample annotations **205**, **206**, **207** and desired results **208**, **209**, **210**

for the three imaging planes for two different slices in each plane are shown in FIG. **2**D. The training procedure for all the three networks is identical, though each one uses a different set of data. A part of the training set is held out as a validation set.

(21) A schematic representation of the ROI extraction CNN in accordance with one embodiment is shown in FIG. **4**. It will be described herein for use with the first embodiment of the segmentation method, while modifications for use in the second embodiment will be described later on. The input data represents a CT volume slice in a particular plane. The left side of the network is the encoder **401**, which is a convolutional neural network, and the right side is the decoder **402**. The encoder **401** may include a number of convolutional layers and a number of pooling layers, each pooling layer preceded by at least one convolutional layer. The encoder might be either pretrained, or trained from random initialisation. The decoder **402** path may include a number of convolutional layers and a number of upsampling layers, each upsampling layer preceded by at least one convolutional layer, and may include a transpose convolution operation which performs upsampling and interpolation with a learned kernel. The network may include a number of residual connections bypassing groups of layers in both the encoder and the decoder.

(22) The residual connections may be either unit residual connections, or residual connections with trainable parameters. The residual connections can bypass one or more layers. Furthermore, there can be more than one residual connection in a section of the network. The network may include a number of skip connections connecting the encoder and the decoder section. The skip connections may be either unit connections or connections with trainable parameters. Skip connections improve the performance through information merging enabling the use of information from the encoder stages to train the deconvolution filters to upsample. The number of layers and number of filters within a layer is also subject to change, depending on the requirements of the application. The final layer for segmentation outputs a mask denoting the heart region as delineated by the pericardium (such as shown in FIG. **2**D)—for example, it can be a binary mask.

(23) The convolution layers can be of a standard kind, the dilated kind, or a combination thereof, with ReLU, leaky ReLU, Swish or Mish activation attached.

(24) The upsampling or deconvolution layers can be of a standard kind, the dilated kind, or a combination thereof, with ReLU, leaky ReLU, Swish or Mish activation attached.

(25) During training, the network may repeatedly perform the following steps: 1. the step of prediction output binary mask based on the input CT imaging data, 2. the computation of the difference between the ground truth mask (as given in the training data) and the predicted mask with the difference computed as dice loss, cross-entropy loss, Tversky loss, . . . ; 3. The update of weights according to the gradient backpropagation method based on the steepest descent gradient algorithm or one of its variants (Adam, Nadam, adagrad, . . .)

(26) Doing so, the network adjusts its parameters and improves its predictions over time. During training, the following means of improving the training accuracy can be used: learning rate scheduling (fixed, cyclic learning rate changes, cosine annealing, . . .) early stopping regularization by dropout L2 regularization, batch normalization, group normalization data augmentation (by random rotations, intensity changes, noise introduction, affine and elastic transformations etc.)

(27) The training process may include periodic check of the prediction accuracy using a held out input data set (the validation set) not included in the training data. If the check reveals that the accuracy on the validation set is better than the one achieved during the previous check, the complete neural network weights are stored for further use. The early stopping function may terminate the training if there is no improvement observed during the last CH checks. Otherwise, the training is terminated after a predefined number of steps S.

(28) The training procedure may be performed according to the outline shown in FIG. **5** in accordance with one embodiment of the training procedure. The training starts at **501**. At **502**, batches of training images are read from the training set, one batch at a time.

(29) At **503** the images can be augmented. Data augmentation is performed on these images to

make the training set more diverse. The input/output data pair is subjected to the same combination of transformations from the following set: rotation, scaling, movement, horizontal flip, additive noise of Gaussian and/or Poisson distribution and Gaussian blur, elastic transform, brightness shift, contrast/gamma changes, grid/optical distortion, batch-level samples averaging, random dropout, etc.

(30) At **504**, the images and generated augmented images are then passed through the layers of the CNN in a standard forward pass. The forward pass returns the results, which are then used to calculate at **505** the value of the loss function—the difference between the desired output and the actual, computed output. The difference can be expressed using a similarity metric, e.g.: mean squared error, mean average error, categorical cross-entropy or another metric.

(31) At **506**, weights are updated as per the specified optimizer and optimizer learning rate. The loss may be calculated using a per-pixel cross-entropy loss function and the Adam update rule.

(32) The loss is also back-propagated through the network, and the gradients are computed. Based on the gradient values, the network's weights are updated. The process (beginning with the image batch read) is repeated continuously until an end of the training session is reached at **507**.

(33) Then, at **508**, the performance metrics are calculated using a validation dataset—which is not explicitly used in training set. This is done in order to check at **509** whether not the model has improved. If it isn't the case, the early stop counter is incremented at **514** and it is checked at **515** if its value has reached a predefined number of epochs. If so, then the training process is complete at **516**, since the model hasn't improved for many sessions now, so it can be concluded that the network started overfitting to the training data.

(34) If the model has improved, the model is saved at **510** for further use and the early stop counter is reset at **511**. As the final step in a session, learning rate scheduling can be applied. The session at which the rate is to be changed are predefined. Once one of the session numbers is reached at **512**, the learning rate is set to one associated with this specific session number at **513**.

(35) Once the training is complete, the network can be used for inference, i.e. utilizing a trained model for prediction on new input data.

(36) Upon the completion of the training, the weights of the neural network are stored and can be used for prediction. The input data for the prediction process are CT scan data of the heart volume with contrast filled coronary arteries. For prediction of the location of the heart in individual slices in the form of the binary mask, the data is propagated through all the layers of the networks, successively, until it reaches the final layer. The output of the final layer is a binary image containing the location of the heart as delineated by the pericardium.

(37) The individual prediction of each neural network is an image. As the networks make predictions in a slice by slice manner, the volumetric information can be reconstructed simply by combining the predictions by stacking the slices.

(38) The volumetric predictions in the 3 axes are then combined by averaging the individual results (e.g. calculating a sum of the components divided by the number of components) and applying a threshold and postprocessing by nonlinear filtering (morphological, median). The final result in 3D looks as shown below (a few different samples), as shown in FIG. **2**D.

(39) Next, in step **105**, the preprocessed scan volume (as in FIG. **2**B) and its corresponding 3d Jerman filter response (as in FIG. **2**C) are additionally masked with the volumetric pericardium segmentation mask (as in FIG. **2**D) with the addition of a safety margin (equal to e.g. a few voxels, preferably 3 voxels), so that all the information outside the pericardium region is cancelled out, wherein the result is a masked volume. For example, the pericardium can be postprocessed with dilation using ball-shaped kernel (radius=3 voxels). This is essential, since coronary arteries occupy only a part of the overall volume. Constraining the overall volume registered by the CT scanner to the subvolume that actually contains the heart improves the accuracy of the next step (the coronary vessel segmentation).

(40) Both the masked raw volume (as shown in FIG. **2**E—the top image **211** represents the 3D

mask shape and the bottom image **212** represents the masked raw volume) and (optionally) the corresponding masked volumetric Jerman filter response **213** (as shown in FIG. **2**F) are used as the input data for coronary vessel segmentation by the segmentation CNN. If the Jerman filter output is used, it is included as an additional channel of the input images.

(41) In step **106**A, the masked volume is converted to three groups of two-dimensional slices, wherein each groups corresponds to a particular principal plane (the axial plane, the sagittal plane and the coronal plane) and the sets within the group correspond to planes tilted at an angle with respect to the principal plane.

(42) FIG. **6**A shows three such planes **601**, **602**, **603**, one typical, aligned with the wall of the volume (i.e. along a principal plane), and two or more additional planes, each tilted at a small angle. For example, if 2N+1 planes are used, then the angles may be +/−(N*10) —which results in for example: for 3 planes, they are arranged at angles −10 deg., 0 deg, +10 deg. for 5 planes, they are arranged at angles −20 deg., −10 deg., 0 deg., +10 deg., +20 deg.

(43) Such a set of planes moves along an axis of each principal plane instead of just one, as shown in FIG. **6**A. The slanted planes are projected on a plane **604** analogous to the regular plane as shown in FIG. **6**B. Therefore, three groups of sets of two-dimensional slices are obtained.

(44) Next, in step **107**A, the coronary vessel segmentation is performed, preferably individually for each plane, by segmentation CNNs in a similar way as for pericardium, for the two-dimensional slices obtained in the previous step. Therefore, preferably (2N+1)*3 networks are used. All the networks share the same input size and the masks do not degrade due to interpolation, since Bresenham discretization pattern is used for sampling.

(45) Prior to use, each of the neural networks used for prediction needs to be trained. The training data consists of pairs of CT volume slices in its corresponding plane and their corresponding binary, expert-annotated mask, denoting the coronary vessels. Direct correspondence of binary masks and CT scan data enables their direct use for segmentation training.

(46) Sample annotation and desired result **605**, **606**, **607** for three imaging planes in each plane are shown in FIG. **6**C—the left column **605** corresponds to the axial plane, the middle column **606** corresponds to the coronal plane and the right column **607** corresponds to the sagittal plane. The top row represents slices from the volume and the bottom row represents binary coronary vessel masks corresponding to those slices.

(47) The training procedure for networks corresponding to the planes is identical, though each one uses a different set of data. The training is performed using a pair of corresponding CT-segmentation mask images for individual slices or alternatively subvolumes. A set of those forms a single batch. A part of the training set is held out as a validation set.

(48) The segmentation CNN in accordance with one embodiment has a structure as discussed with reference to FIG. **4**, with the following differences. The input data is a masked raw volume and (optionally) the corresponding masked volumetric Jerman filter response. If both types of inputs are used, then the segmentation CNNs should include an additional channel for processing that data. The output is a binary coronary vessels mask image containing the location of the coronary vessels.

(49) Next, the coronary vessels masks output for the masked slices for the different planes can be combined to a segmented 3D data set representing the shape, location and size of the coronary vessels.

(50) The training procedure in accordance with one embodiment is equivalent to that discussed in FIG. **5**. Upon the completion of the training, the weights of the neural network are stored and can be used for prediction. The input data for the prediction process are CT scan slices of the heart volume with contrast filled coronary arteries masked with the heart region as predicted in the previous steps. For prediction of the location of the coronary vessels slices in the form of the binary mask, the data is propagated through all the layers of the networks, successively, until it reaches the final layer. The output of the final layer is a binary image containing the location of the coronary vessels. The individual prediction of each neural network is an image. As the network makes

predictions in a slice by slice, we can reconstruct the volumetric information simply by accumulating the response from all the networks slice by slice in all directions. The volumetric predictions are then combined by averaging the individual results and applying a threshold or other means (majority voting, reconstruction neural network, e.g. 3D Unet).

(51) The predicted binary volume representing the coronary vessels can be subjected to additional post-processing: Deletion of small blobs to minimize the number of false positive responses Frangi/Jerman/median/morphological filtering to improve the smoothness and continuity of segmented vessels Transformation of volumetric data into surface mesh for visualization of volumetric mesh for further processing using e.g. finite element method for mechanical modeling

(52) Alternatively, in accordance with a second embodiment of the segmentation procedure, in step **103**B the 3D volume is converted to a first set of subvolumes. Next, in step **104**B, a region of interest is extracted by autonomous segmentation of the heart region as outlined by the pericardium, by means of a neural network trained on 3D subvolumes and combining the results of the individual subvolume predictions for the first set to output a mask denoting a heart region as delineated by the pericardium. In step **106**B, the masked volume is divided into a second set of subvolumes. In step **107**B, the coronary vessel segmentation is performed by a segmentation CNN in a similar way as for pericardium, for the 3D subvolumes of the second obtained in the previous step **106**B. A single segmentation neural network can be used. The output of the final layer of the segmentation network is a binary volume containing the location of the coronary vessels. As the segmentation network makes predictions in a subvolume by subvolume manner, we can reconstruct the volumetric information simply by combining the subvolume predictions. Therefore, the coronary vessels mask output for the different subvolumes can be combined to a segmented 3D data set representing the shape, location and size of the coronary vessels. The other steps of the procedure are equivalent to that described above with respect to the first embodiment of the segmentation procedure.

(53) The CNNs as shown in FIG. **4**, for use in the first embodiment of the segmentation procedure, involving 2D image slices, can be configured as follows: the convolutional layers and pooling layers are two-dimensional; the convolutions are 2D convolutions operating on slices; the CNN adjusts its internal parameters, which include the weights in the internal convolutional layers of the dimensions W×H, W and H being positive integers; input/output data pairs are 2D slices; three CNNs are used.

(54) The CNN as shown in FIG. **4**, for use in the second embodiment of the segmentation procedure, involving 3D subvolumes, can be configured as follows: the convolutional layers and pooling layers are three-dimensional; the convolutions are 3D convolutions operating on volumes; the CNN adjusts its internal parameters, which include the weights in the internal convolutional layers of the dimensions W×H×D, W, H and D being positive integers; input/output data pairs are volumes; as the CNN makes predictions in a volume by volume manner, the volumetric information can be reconstructed simply by combining the predictions by combining the subvolumes; a single CNN is used.

(55) Examples of desired final results **301**, **302**, **303** for both embodiments of the segmentation procedure are given in FIG. **3**.

(56) The functionality described herein in accordance with one embodiment can be implemented in a computer system **700**, such as shown in FIG. **7**. The system **700** may include at least one nontransitory processor-readable storage medium **710** that stores at least one of processor-executable instructions **715** or data; and at least one processor **720** communicably coupled to the at least one nontransitory processor-readable storage medium **710**. The at least one processor **720** may be configured to (by executing the instructions **715**) perform the procedure of FIG. **1**.

(57) While the invention has been described with respect to a limited number of embodiments, it will be appreciated that many variations, modifications and other applications of the invention may

be made. Therefore, the claimed invention as recited in the claims that follow is not limited to the embodiments described herein.

Claims

1. A computer-implemented method for autonomous segmentation of contrast-filled coronary artery vessels, the method comprising: a) receiving a computed tomography (CT) scan volume representing a three-dimensional (3D) volume of a region of anatomy that includes a pericardium; b) preprocessing the CT scan volume to output a preprocessed scan volume; c) dividing the preprocessed scan volume into a first set of subvolumes, wherein the division is performed at a resolution that corresponds to the resolution of the CT scan volume, wherein the first set of subvolumes comprises a plurality of first set subvolumes of the resolution that corresponds to the resolution of the CT scan volume; d) extracting a region of interest (ROI) corresponding to a heart region outlined by the pericardium, by autonomous segmentation of the heart region by means of a first neural network trained on 3D subvolumes to delineate the pericardium, the extracting step comprising: feeding each first set subvolume of the first set of subvolumes to the first neural network, receiving an individual subvolume prediction for each first set subvolume of the first set of subvolumes and combining the individual subvolume predictions for the first set of subvolumes to generate an ROI mask that delineates the heart region based on the pericardium; e) combining the preprocessed scan volume with the ROI mask to obtain a masked volume; f) dividing the masked volume to a second set of subvolumes, wherein the division is performed at a resolution that corresponds to the resolution of the CT scan volume, wherein the second set of subvolumes comprises a plurality of second set subvolumes of the resolution that corresponds to the resolution of the CT scan volume; and g) performing autonomous coronary vessel segmentation by autonomous segmentation of each second set subvolume of the second set of subvolumes individually by means of a trained segmentation convolutional neural network, trained to process a subvolume to output a mask denoting the coronary vessels, the segmentation being performed at a resolution that corresponds to the resolution of the masked volume, to output a mask denoting the coronary vessels having the resolution of the masked volume.
2. The method according to claim 1, wherein the step of preprocessing the CT scan includes performing at least one of: windowing, filtering and normalization.
3. The method according to claim 1, wherein the step of preprocessing the CT scan includes computing a 3D Jerman filter response.
4. The method according to claim 1, further comprising combining the masks denoting the coronary vessel to a segmented 3D data set representing the shape, location and size of the coronary vessels.
5. A computer-implemented system, comprising: at least one nontransitory processor-readable storage medium that stores at least one of processor-executable instructions or data; and at least one processor communicably coupled to the at least one nontransitory processor-readable storage medium, wherein the at least one processor is configured to perform the steps of the method of claim 1.