HHS Public Access

Author manuscript

Phys Med. Author manuscript; available in PMC 2022 May 13.

Published in final edited form as:

Phys Med. 2021 May ; 85: 107–122. doi:10.1016/j.ejmp.2021.05.003.

A Review of Deep Learning based Methods for Medical Image Multi-Organ Segmentation

Yabo Fu, **Yang Lei**, **Tonghe Wang**, **Walter J. Curran**, **Tian Liu**, **Xiaofeng Yang**

Department of Radiation Oncology and Winship Cancer Institute, Emory University, Atlanta, GA

Abstract

Deep learning has revolutionized image processing and achieved the-state-of-art performance in many medical image segmentation tasks. Many deep learning-based methods have been published to segment different parts of the body for different medical applications. It is necessary to summarize the current state of development for deep learning in the field of medical image segmentation. In this paper, we aim to provide a comprehensive review with a focus on multi-organ image segmentation, which is crucial for radiotherapy where the tumor and organs-at-risk need to be contoured for treatment planning. We grouped the surveyed methods into two broad categories which are 'pixel-wise classification' and 'end-to-end segmentation'. Each category was divided into subgroups according to their network design. For each type, we listed the surveyed works, highlighted important contributions and identified specific challenges. Following the detailed review, we discussed the achievements, shortcomings and future potentials of each category. To enable direct comparison, we listed the performance of the surveyed works that used thoracic and head-and-neck benchmark datasets.

1. Introduction

Medical image segmentation is one of the most important medical image analysis tasks. It has a wide range of applications in imaging systems such as microscopy, X-ray, ultrasound, computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET). Medical image segmentation plays an essential role in radiotherapy, which is the standard care for certain cancers [1]. The success of radiotherapy depends highly on accurate irradiation to the target and sparing of organs-at-risk (OARs) [2, 3]. Therefore, accurate structure delineation is crucial for radiotherapy, especially for highly conformal radiotherapy such as intensity modulated radiotherapy (IMRT), proton therapy and stereotactic body radiotherapy (SBRT). These highly conformal treatments are designed to shape radiation to target volume while sparing healthy OARs, and are usually planned

Corresponding author: Xiaofeng Yang, PhD, Department of Radiation Oncology, Emory University School of Medicine, 1365 Clifton Road NE, Atlanta, GA 30322, xiaofeng.yang@emory.edu.

Disclosures

The authors declare no conflicts of interest.

with sharp dose drop-off. Misdelineation of anatomical structures could result in severe misadministration of radiation doses to the target and OAR. In current clinical practice, structure contours are manually delineated by physicisans. The manual contouring process is tedious, time consuming and laborious. Manual delineation of soft tissues on CT images is challenging due to low soft tissue contrast, which makes the contours prone to errors and inter/intra-observer variabilities [4–9]. In the past decades, researchers have spent enormous effort to develop automatic contouring methods for accurate and consistent organ delineation.

Traditional medical image segmentation [10–14] usually involves handcrafted image feature detection such as line/edge detection and mathematical models to trace image gradient along object boundaries such as graph cuts, active contours, level-set and so on. Atlas-based method is another commonly used approach for automatic segmentation [15–18]. The atlas-based methods propagate the predefined structure contours to the images to be segmented using image registration. The segmentation accuracy of this technique highly relies on the accuracy of the image registration. Model-based methods which make use of statistical shape models for automated segmentation have also been proposed [19–21]. The accuracy of those methods depends on the reliability and generalizability of the models. Models that are built based on normal anatomical structures have shown limited success on irregular structure segmentation.

Recently, machine learning (ML) has gained a lot of interest in medicine [22–24]. Artificial Neural Network (ANN), a subfield of ML, utilizes multiple layers of connected neurons with learnable weights and biases to simulate human brains to accomplish high-level tasks [25–30]. Deep Learning (DL) is a new term for ANN arising from advances in the ANN architectures and algorithms since 2006, referring to ANN with many hidden layers. Since there is no consensus on the number of layers required to be count as deep, the distinction between ANN and DL is not clearly defined [31]. DL has demonstrated enormous potential in computer vision [32]. DL uses a data-driven approach to explore vast image features to facilitate various vision tasks, such as image classification [33], object detection [34] and segmentation [35]. Inspired by the success of DL in computer vision, researchers have proposed various methods to extend the use of DL techniques to medical imaging. To date, DL has been extensively studied in medical image segmentation [36–72], image synthesis [73–96], image enhancement and correction [97–107], and registration [108–122]. DL-based multi-organ segmentation technique represents a significant potential in daily practices of radiation therapy since it can expedite the contouring process, improve contour accuracy and consistency and promote compliance to delineation guidelines [39, 45, 123–127]. Furthermore, rapid DL-based multi-organ segmentation could facilitate online adaptive radiotherapy to improve clinical outcomes. After studying 80 online MRI-guided adaptive radiotherapy cases, Lamb et al. [128] reported the median time of adaptive process prior to beam delivery was 54 minutes, out of which the re-contouring process took up to 22 minutes. To expedite the contouring process for adaptive radiotherapy, DL-based abdominal multi-organ segmentation has been proposed and tested on Viewray MR images [129]. Though the DL-based contour process took only several minutes, post manual correction was often needed to meet the physicians' satisfaction. The time analysis reported in this study shows that the average contouring time, which is the automatic contouring time plus

the post manual correction time, was only a quarter of the total time needed to manually contour from scratch. CT-based multi-organ segmentation which includes eight organs has been proposed for pancreatic radiotherapy [130]. The eight organs include large bowel, small bowel, duodenum, left kidney, right kidney, liver spinal cord and stomach. This CT-based method could be used with on-rail CT to facilitate fast contouring for adaptive radiotherapy. Besides the CT-based segmentation, CBCT-based multiorgan segmentation has also been proposed for prostate adaptive radiotherapy [131].

DL-based methods [130, 132–136] have achieved the-state-of-art performances in medical image segmentation, especially in multi-organ segmentation. In contrast to traditional methods that ultilize handcrafted features, DL-based methods adaptively explore representative features from medical images [137]. In this paper, we reviewed deep learning-based methods for medical image segmentation with a focus on multi-organ segmentation. We classified the methods into two broad categories which are pixelwise classification and end-to-end segmentation. Each category was reviewed in details to study its latest developments, contributions and chellenges. We provided benchmark evaluations of recently published multi-organ segmentation methods for CT thoracic and Head and Neck (HN) segmentations.

2. Deep Learning in Multi-Organ Segmentation

DL-based multi-organ segmentation methods can be categorized by network architecture, training process (supervised, semi-supervised, unsupervised, transfer learning), input image types (patch-based, whole volume-based, 2D, 3D) and so on. In this paper, we first classify them into two broad categories which are pixelwise classification and end-to-end segmentation since the two category represent the major steps of development in DL-based image segmentation. Based on the network structure design, we further divided the pixelwise classification methods in to 1) Auto-Encoder (AE) and 2) Convolutional Neural Network (CNN). Similarly, we divided the end-to-end segmentaiton methods into 1) Fully Convolutional Network (FCN), 2) Region-based FCN (R-FCN), 3) Generative Adversarial Network (GAN) and 4) Synthetic Image-aided Segmentation. For each sub-category, we provided a comprehansive list of the surveyed works followed by a short discussion.

Works cited in this review were collected from various databases, including Google Scholar, PubMed, Web of Science, Semantic Scholar and so on. Keywords used to search literature include but not limited to deep learning, multi-organ, medical image segmentation, convolutional neural network and so on. Over 60 papers that are closely related to multi-organ segmentation were collected. Most of these works were published between the year of 2017 and 2020. In this paper, we also included some single organ segmentation papers for the ease of description since many multi-organ segmentation methods were developed by replacing the last layer for multi-class classification. The number of multi-organs publications is plotted against year in Fig. 1.

Dice similarity coefficient (DSC), 95% Hausdorff distance (HD95) and mean surface distance (MSD) were often used to evaluate the performance of the segmentation methods.

The DSC is a measure of the volumetric overlap between the predicted and ground truth segmentation.

$$DSC = \frac{2 \times |X \cap Y|}{|X| + |Y|}$$

where X and Y are the predicted and ground truth segmentation, respectively.

The HD95 and MSD measures the surface distances between the predicted and ground truth segmentations. The HD95 quantifies the maximum 95 percentile distance while the MSD quantifies the mean surface distance.

$$HD95 = max_{k95\%}[d(X,Y), d(Y,X)]$$

$$MSD = \frac{1}{|X| + |Y|}\left(\sum_{x \in X} d(x,Y) + \sum_{y \in Y} d(y,X) \right)$$

where $d(x,Y) = min_{y \in Y}\|x - y\|_2$. $d(X,Y)$ is the total surface distance between X and Y.

2.1 Pixel-wise Classification

Early DL-based methods performed image segmentation by repeatedly classifying the center pixels of sliding image patches that cover the whole image. Two major types of network for pixelwise classification are the AE and CNN. Fig. 2 shows the common network components for AE and CNN based methods.

2.1.1 Auto-Encoder—AE consists of a neural network encoder that encodes the input into a latent representation by minimizing the reconstruction errors between the input and the output. The output represents the restored input from the low-dimensional latent representation. By constraining the dimension of the latent space, AE can effectively compress the input into patterned latent space representation. To prevent the AE from learning an identity function, stacked AE (SAE) was proposed. SAE is constructed by stacking AEs on top of each other, where the output of each layer is wired to the inputs of its successive layers [138]. The benefit of SAE is that it can benefit from deeper network, which has higher level of feature representation [138]. Denoising autoencoder (DAEs) is another variant of the AE which prevent the model from learning a trivial solution by training the network to reconstruct a clean input from the corrupted input [139]. Stacked denoising autoencoder (SDAE) is another type of AE that utilizes the power of DAE [140].

Overview of works: Ahmad *et al.* proposed a deep SAE (DSAE) for CT liver segmentation [141]. First, deep features were extracted from unlabeled data using the AE. Second, these features are fine-tuned to classify the liver among other abdominal organs. An average DSC of 0.9 was achieved on 659 2D liver images. Vaidhya *et al.* used SDAE to overcome the challenge of varying shape and texture of glioma tissue in MRI for this segmentation task [140]. 3D image patches were extracted from multiple sequences MRI and then were fed into the SDAE model to obtain the glioma segmentation. Two SDAE models were trained,

one for high grade glioma (HGG) data, the other one for a combination of HGG and low-grade glioma (LGG) data. During testing, the segmentation was obtained by a combination of predictions from the two networks via maximum a posteriori (MAP) estimation. The network has achieved mean DSC of 0.82 ± 0.14 and 0.72 ± 0.21 for whole tumor segmentation on the HGG data and LGG data, respectively. Alex *et al.* applied SDAE for brain lesion detection, segmentation, and false-positive reduction [139]. SDAE was pretrained using many unlabeled patient volumes and fine-tuned with 2D patches drawn from a limited number of patients. LGG segmentation was achieved using a transfer learning approach in which the pretrained SDAE network was fine-tuned using the LGG data. The method was able to achieve a mean DSC of 0.86 ± 0.12 for brain whole tumor segmentation on the BraTS challenge datasets.

Accurate vertebrae segmentation in the spine is essential for spine assessment, surgical planning and diagnosis. Qadri *et al.* proposed a stacked SAE (SSAE) model for the segmentation of vertebrae from CT images [142]. High-level features were extracted from the 2D image patches using the SSAE model, which was trained in an unsupervised way. To improve the network performance, the authors fine-tuned the network using supervised training. The SSAE model was validated on the 2014 MICCAI CSI challenge datasets with an average DSC of 0.86.

Discussion: SDAE has been shown to be working for the segmentation of brain MRI tumor on public BraTS 2013 and BraTS 2015 data [139, 145]. DSAE has been shown to have high classification accuracy and speed for liver segmentation on CT images [141]. AE can learn medical image deep contextual features from large-range input samples to improve their contextual discrimination ability [144]. Validated on the 98 spine CT scans from the public MICCAI CSI 2014 dataset, the SSAE method could effectively and automatically locate and identify spinal targets in CT scans, and achieve high localization accuracy without making any assumptions about visual field in CT scans [142].

Although AE has many advantages, it faces some challenges and limitations in medical multi-organ segmentations. One of the limitations is related to data regularity. AE-based segmentation methods work quite well for anatomical structures that have small shape variability such as lung, heart and liver. However, it remains challenging for the unsupervised AE methods to segment irregular lesions and tumors that have large shape variability. The number of layers used in AE could be limited due to large computation complexity. Unlike CNN which uses convolution kernels with shared learnable parameters, AE methods cannot be easily extended to large number of layers which limits its learning ability.

2.1.2 Convolutional Neural Networks—

A typical CNN consists of convolutional layers, activation functions, max pooling layers, batch normalization layers, dropout layers and fully connected layers. The last layer of a CNN is typically a sigmoid or softmax layer for classification and tanh layer for regression. The convolution layers can learn to extract various feature maps depending on the task. Pooling layers are used to reduce the spatial size of the feature maps using maximum/average down-sampling operations. Activation functions such as Rectified linear unit (ReLU) and Leaky ReLU are used to simulate neuron

activation by clipping any negative input values to zero and passing positive input values to the connected neurons [146]. Fully connected layer connects every neuron in previous layer to every neuron in next layer. They are placed before the final classification layer to flatten the feature maps. The final classification layers are used to predict the possibility of the center image pixel of belonging to one of the classes.

During the training, gradient based optimization methods such as stochastic gradient descent (SGD) and Adam gradient descent are commonly used to update the learnable parameters of the CNN architecture through back-propagation. Cross-entropy is one of the most widely used loss functions. LeNet was first proposed by Lecun *et al.* to classify hand-written digits [147]. LeNet is composed of convolution layers, pooling layers and fully connected layers. As computers get more powerful and more data are available for network training, Krizhevsky *et al.* proposed AlexNet in 2012 and won the ILSVRC-2012 image classification competition [148] by a large margin [33]. Since the introduction of AlexNet, CNNs started to gain widespread attention, which has led to the development of various types of CNNs that achieved the-state-of-art performances in many image processing tasks. The improvements of AlexNet over LeNet include 1) ReLU layer for nonlinearity and sparsity, 2) data augmentation to enlarge the dataset variety, 3) dropout layer to reduce learnable parameters and prevent overfitting, 4) GPU for parallel computing, 5) local response normalization and 6) overlapping pooling. In 2014, Zeiler and Fergus proposed ZFNet to improve the performance of AlexNet [149] and showed that shallow network can learn edge, color and texture features of images and deep network can learn abstract features of images. They demonstrated that better performance can be achieved via deeper network. The main improvement of ZFNet is deconvolution network used to visualize the feature map. To evaluate the network performance with respect to network depths, VGG was proposed to extend the network depth to 19 layers by Simonyan and Zisserman [150]. GoogLeNet was proposed to introduce the inception module [151]. The inception module allows broader perception field and deeper network which improves the network's ability of feature extraction. As a result, GoogLeNet won the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14). As the network gets deeper, training of the network gets harder due to gradient vanishing/exploding. To alleviate the problem, He et al. proposed a residual network (ResNet) which allows even deeper network to be trained for image recognition [152]. Huang *et al.* later proposed a densely connected convolutional network (DenseNet) by connecting each layer to every other layers [153] in order to combine both low-frequency and high-frequency feature maps.

Overview of works: Roth *et al.* proposed a multi-level deep CNN approach for abdominal CT image pancreas segmentation [154]. A dense local image patches and labels were obtained by extracting 2D image patches in the axial, coronal and sagittal plane using a sliding window. The proposed CNN learns to assign class probabilities to the center voxels of the image patches. The proposed CNN architecture consists of five convolutional layers followed by max-pooling layers, three fully connected layers, two dropout layers and a soft-max operator to perform binary classification. Evaluated on 82 patient's CT images using 4-fold cross-validation, an average DSC of 0.84 ± 0.06 and 0.72 ± 0.11 was obtained for the training and testing, respectively. For volumetric datasets, it is beneficial to explore the 3D

images directly rather than 2D images. Therefore, Hamidian *et al.* proposed to use 3D patch-based CNN to detect lung pulmonary nodules for chest CT images [155]. Volumes of interest image patches were extracted from the 3D lung image database consortium (LIDC) dataset [156]. They demonstrated that 3D CNN is more suitable for volumetric CT data than 2D CNN.

In radiotherapy, it is common to segment multiple organs near the tumor for treatment planning. For nasopharyngeal carcinoma (NPC), it is very challenging to automatically segment the surrounding adhesion tissues of the parotids, thyroids and optic nerves due to low soft tissue contrast of the CT images. To overcome this challenge, Zhong *et al.* proposed a cascaded CNN network to delineate these three organs for NPC radiotherapy using a boosting algorithm which includes three cascaded CNNs [157]. The first network was trained with the traditional approach. The second one was trained on patterns (pixels) filtered by the first network. Finally, the third network was trained on the new patterns (pixels) that were jointly extracted by the first and second networks. The outputs of the three nets were combined to obtain the final output. 2D patch-based ResNet [152] was used to build the cascaded CNNs. CT images of 140 NPC patients treated with radiotherapy were collected. Manual contours of the three organs were used as learning targets. The mean DSC values were above 0.92 for the parotids, above 0.92 for the thyroids, and above 0.89 for the optic nerves. For thoracic radiotherapy treatment, Harten *et al.* proposed a combination of 2D and 3D CNNs for automatic segmentation of organs including esophagus, heart, trachea, and aorta on simulation CT scans of patients diagnosed with lung, breast or esophageal cancer [158]. The 3D patch-based network contains a deep segment of residual blocks [159] with sigmoid layer to perform multi-class binary classification. The 2D patch-based (2D patch extracted from axial, coronal and sagittal planes) network contains dilated convolutions [160] with softmax layer to perform classification. 40 data were used for training and 20 data were used for testing.

Discussion: In the study of [158], researchers evaluated the performance of 2D CNN, 3D CNN and a combination of the two and demonstrated that the combined network produced the best results. The DSC of the esophagus, heart, trachea, and aorta were 0.84±0.05, 0.94±0.02, 0.91±0.02, and 0.93±0.01, respectively. These results demonstrated potential for automating segmentation of OARs in routine radiotherapy treatment planning. A major drawback of the pixelwise classification methods is that classification need to be performed for every pixel repeatedly. This approach is inefficient since it requires repeated forward network prediction on every voxel of the image. To make the segmentation more efficient, Kamnissas *et al.* proposed a dense-inference technique that predicts the segmentation on a smaller patch rather than only the center pixel [161]. However, this method is still relatively inefficient as compared to end-to-end segmentation that utilized the transposed convolution kernel to directly predict a segmentation map that is in the same size as the input image.

2.2 End-to-end Segmentation

2.2.1 FCN methods—

For pixelwise classification-based methods, the center voxel of the input image is classified by fully connected layers based on the flattened feature maps that are passed down through multiple convolutional layers. Shelhamer *et al.* first proposed a

CNN that replaces the fully connected layer by a convolutional layer. Since all layers in the network are convolutional layers, the new network is named as fully convolutional network (FCN). Thanks to the deconvolution kernels that were used to up-sample the feature maps, FCN allows the model to predict a dense segmentation map that has the same size as the input image, which was referred to 'end-to-end segmentation' [35]. By using FCN, the segmentation of whole image can be achieved in just one forward network inference.

U-Net is one of the most well-known FCN structures for medical image segmentation that utilizes the concept of deconvolution and skip connection [177]. As a variant of the FCN, the U-Net is a 19 layer -deep network that includes an encoding path and a decoding path. To preserve the spatial high-resolution information, the U-Net used long skip connections between the layers of equal resolution in the encoding path and decoding path. Milletari *et al.* proposed an variant of U-Net, called V-Net [178]. Unlike U-Net, V-Net involves residual block as short skip connection between early and later convolutional layers. This architecture improves convergence rate as compared to non-residual learning network, such as U-Net. To cope with class imbalance problem, V-Net used Dice loss instead of binary cross entropy loss.

Deep supervision is commonly used to train the FCN. The main idea of deep supervision [43, 46] is to provide supervision over not only the final output layer but also the intermediate hidden layers. The direct supervision was extended to multiple deep layers, which could enhance the network's discriminative ability. Attention gate was used in FCN to improve performance in image classification and segmentation [179] by highlighting salient features and suppressing irrelevant features for a specific task.

Overview of works: Zhou *et al.* proposed a 2.5D FCN segmentation method to automatically segment 19 organs in CT images of whole body [180]. In this work, 2.5D image patches, which consists of several consecutive axial slices, were used as multi-channel input for the FCN. Individual FCNs were also trained for the coronal and sagittal views, resulting in a total of three FCNs. Final segmentation was obtained from the three networks. Transrectal ultrasound (TRUS) is commonly used in image-guided prostate cancer interventions (e.g., biopsy and brachytherapy). Accurate segmentation of the prostate is very important for biopsy needle placement, brachytherapy treatment planning, and motion management. However, the prostate segmentation of TRUS image is challenging due to low image contrast and image noise. Lei *et al.* proposed a deeply supervised V-Net for accurate prostate segmentation [43]. A deep supervision strategy with hybrid loss function (logistic and Dice loss) was used at different stages of the decoding path. To improve the segmentation accuracy at the prostate apex and base, a multi-directional contour refinement model was introduced to fuse transverse, sagittal and coronal plane-based segmentation. Tested on 44 patients' TRUS images, this method has a mean DSC of 0.92±0.03 for the prostate segmentation. Wang *et al.* proposed a 3D FCN with deep supervision and group dilated convolution to segment the prostate on MRI [46]. In this method, deep supervision mechanism was introduced into FCN to effectively alleviate the common gradient exploding and vanishing problems in training deep models. A group dilated convolution which aggregated multi-scale contextual information for dense prediction was proposed to enlarge the effective receptive field. A combined loss which included cosine and cross entropy was

used to improve the segmentation accuracy. Tested on 40 patients' T2 MR images, this method has a mean DSC of 0.86±0.04 for the prostate segmentation.

Glands segmentation is essential in cancer diagnosis. However, accurate automated DL-based segmentation of glands is challenging due to the large variability in glandular morphology across tissues and pathological subtypes. Many accurate gland annotations are required for network training. Binder *et al.* investigated cross-domain (-organ type) approximation to reduce the need of organ-specific annotations [181]. Two proposed Dense-U-Nets are trained on Hematoxylin and Eosin (H&E) strained colon adenocarcinoma samples focusing on the gland and stroma segmentation. Unlike U-Net, Dense-U-Nets use asymmetric encoder and decoder. The encoder is designed to automatically and adaptively learn the spatial hierarchies of features from low to high level patterns coded within the image. The decoder uses transition layer (convolution with stride size 2) and dense convolution blocks consecutively to extract the encoded feature representation. The dense-convolution blocks from DenseNet [153] are used to strengthen feature propagation, encourage feature reuse and substantially reduce the total number of required parameters in the network. The decoder is composed of deconvolution layers and convolution blocks. The skip connection between the encoder and the decoder allows for feature reuse. The architecture has two decoders, one to predict the relevant gland locations, and a second to predicts the gland contours. Thus, the decoders output a gland probability map and a contour probability map. The network is supervised to predict both the gland locations and the gland contours. The model trained via Gland-approach achieves DSC of 0.92 and 0.78 on the colon and breast test datasets, respectively.

Discussion: The FCN usually has a fixed receptive size which struggles to detect the target once its size changes. One solution is multi-scale networks, where input images are resized first before feeding to the network. Multi-scale techniques can alleviate the problem caused by fixed receptive size in the FCN [191]. However, sharing the parameters of the same network on a resized image may not be very effective as the object of different scales requires different parameters to process. Another solution is to perform multiple predictions on a sliding window across the entire image if the reception size is smaller than the image to be segmented [155].

Multiscale FCN-based segmentation [43, 46] can achieve good performance at the cost of higher computation complexity than the U-Net and V-Net methods. The problem of using 2.5D patch images as input is that the segmented contours of the axial, coronal and sagittal plane may not perfectly align. Though the FCN methods segment the object in an end-to-end fashion, each voxel is classified independently. For example, the pixel-wise cascaded CNN [157] outperformed U-Net [177] in segmenting three OARs. This could be due to that the U-Net lacks spatial relationship modeling among the voxels, resulting in unreasonable object shapes. Therefore, post-processing such as conditional random field and graph cut are often adopted to refine the results [129].

2.2.2 Region-based FCN (R-FCN)—

Multi-organ segmentation is more challenging than single object segmentation in that multi-class classification is more difficult than binary classification. To improve the segmentation accuracy, multi-organ segmentation can be

divided into two-steps: 1) to localize the targets of interest and 2) to perform binary classification for every target separately. We call this type of methods region-based FCN.

Cascaded FCN is one type of R-FCN that stacks two FCN where the first FCN is used to locate the targets of interest and the second FCN is used to perform binary classification for each target. The cascaded FCN could help to alleviate the class imbalance problem in 3D FCN by the first FCN that locates the target and balances the foreground and background.

Another type of R-FCN takes a different approach to locate the ROIs. Regional proposal networks are integrated to the FCN [211]. A selective search [212] method was used to extract many candidate regions from image. The locations of the region proposals were represented by multi-scale bounding boxes. After training, the region proposal network can predict the offsets and scales of the bounding boxes to refine the its locations and sizes to better encompass the targets of interest. However, the large number of regional proposals makes the network computationally demanding. To expedite the region detection process, Fast R-CNN [213] was proposed. The Fast R-CNN used a backbone network to identify the regional proposals, which were then processed using ROI pooling layers. Unlike R-CNN, the Fast R-CNN does not need to feed many region proposals to the network for each feeding image. Instead, the convolution operation is performed only once per image in Fast R-CNN. Both R-CNN and Fast R-CNN use selective search to identify the region proposals, which can be time-consuming. To make the algorithm faster, Ren *et al.* proposed Faster R-CNN [214] to replace the selective search with learnable network proposals. A separate network was used to predict the region proposals which were then reshaped using ROI pooling layers for bounding box refinement and classification. The Fast and Faster R-CNN were proposed for only quick object detection and localization. To perform segmentation on the detected bounding box at the same time, He et al. proposed Mask R-CNN which integrated two more convolution layers to perform semantic segmentation within the bounding box [215]. One major contribution of Mask R-CNN is the introduction of ROI align, which makes the target feature maps to have consistent size for better image segmentation.

Overview of works: Christ *et al.* performed liver lesions segmentation using cascaded FCNs, where the first FCN detects the liver location, and the second FCN extracts features from the detected ROI to obtain the liver lesions segmentation [216]. A Dice of 0.823 was achieved for lesion segmentation in CT images and 0.85 in MRI images. Similarly, Wu *et al.* investigated the cascaded FCN to improve the performance for fetal boundary detection in ultrasound images [217]. Their results have shown better performance compared to other boundary refinement techniques. The cascaded FCN usually outperforms the single FCN since separate sets of filters can be learned and applied to each ROI. Trullo *et al.* proposed two collaborate FCNs to jointly segment multi-organ in thoracic CT image, one was used for organ localization and the other one was used to segment the organ within that ROI [218]. The drawbacks of cascaded FCN are that the performance of the second FCN largely depends on the performance of the first FCN to accurately localize the target of interest. Due to the two-step process, the cascaded FCN usually takes longer to train and segment.

Xu *et al.* proposed an efficient detection method for multi-organ localization in CT image using 3D regional proposal network (RPN) [219]. Since the proposed RPN is implemented in a 3D manner, it can take advantage of the spatial context information in CT image. AlexNet was used to build the backbone network architecture that can generate high-resolution feature maps to further improve the localization performance of small organs. The method was evaluated on abdomen and brain site datasets and achieved high detection precision and localization accuracy with fast inference speed. Xu *et al.* proposed a novel heart segmentation framework, called CFUN, which combined Faster R-CNN and U-Net [220]. The CFUN can detect and segment the whole heart with good results at reduced computational cost. CFUN introduces a new loss function based on edge information named 3D Edge-loss to accelerate the network training and improve the segmentation results. The proposed CFUN takes less than 15 seconds to segment the heart with an average DSC of 0.86 on the MM-WHS2017 challenge datasets. Similarly, Bouget *et al.* proposed a combination of Mask R-CNN and U-Net for the segmentation and detection of mediastinal lymph nodes and anatomical structures in CT data for lung cancer staging [221]. Li *et al.* proposed a lung nodule detection method based on Faster R-CNN for thoracic MRI in a transfer learning manner [222]. A false positive (FP) reduction scheme based on anatomical characteristics is designed to reduce the FPs and preserve the true nodule. Similarly, Faster R-CNN was also used for pulmonary nodule detection on CT image [223].

Discussion: The Region-based FCN methods are useful tools for multi-organ segmentation and detection tasks. One drawback of the cascaded FCN is that it is a two-step process which may slow down the segmentation. The segmentation accuracy also largely depends on the accuracy of the region localization. The introduction of region proposal networks such as Faster R-CNN has made progress toward fast bounding box detection and localization. However, small and low contrast targets such as nodules and esophagus may be missed by the region proposal network, resulting in erroneous segmentation. Due to the higher data dimensionality and larger number of weight parameters, training 3D R-FCN based models is more time-consuming than the 2D version. However, significant advantages such as higher localization and segmentation accuracy still encourage us to handle this problem in 3D. To speed up the training procedure of the proposed method, one potential solution is to apply batch normalization after each convolutional layer in the backbone network to improve the model convergence, and conduct most calculations on GPU in parallel [219].

2.2.3 GAN—GAN has gained a lot of attention in image processing due to its data generation capability without explicitly modelling the probability density function. A typical GAN consists of two competing networks, a generator and a discriminator [234]. The generator is trained to generate artificial data that approximate the target data distribution. The discriminator is trained to distinguish the artificial data from the true data. The discriminator encourages the generator to predict realistic data by penalizing unrealistic predictions. The adversarial loss could be considered as a trainable network-based loss term. The two networks compete in a zero-sum game [235]. GAN has been shown to be useful in many applications, such as image reconstruction [236], image enhancement [97, 99], segmentation [45, 237], classification and detection [238], augmentation [239], and cross-modality image synthesis [82].

Overview of works: As discussed above in the FCN-based methods, one challenge in medical image segmentation is that these methods may have boundary leakage in low contrast regions. Using adversarial loss introduced via a discriminator can take into account high order potentials to solve this problem [240]. The adversarial loss can be regarded as a learned similarity measurement between the segmented contours and the annotated ground truth (manual contours) for medical image segmentation tasks. Instead of only measuring the voxel classification loss such as Dice loss and cross entropy loss, the discriminator in GAN can map the segmented and ground truth masks to a latent space and measure the global similarity. Logistic loss between the latent space features of the segmented and ground truth masks can be used to measure the shape similarity. The idea is analogous to the perceptual loss that is widely used in natural images processing.

Dai *et al.* proposed a structure correcting adversarial network (SCAN) to segment the lung and the heart in Chest X-Ray (CXR) images [237]. SCAN used an FCN to generate the binary mask of the segmented organs and a critic network to judge whether the segmented structure is reasonable from a physiological perspective. The critic network was trained to discriminate between the ground truth organ annotations from the segmented masks generated by the network. The critic network helps to regularize the appearance of the segmentation result to achieve realistic segmentation outcomes.

GAN can be used to alleviate the problem of training data shortage. It is common that only limited datasets are available for network training since it is very time-consuming and laborious to manually generate large datasets especially for multi-organ segmentation. To overcome this challenge, Mondal *et al.* proposed a GAN-based method that is capable of learning from a few labeled images. The network was used to perform 3D multimodal brain MRI segmentation from a few-shot learning perspective [241]. The proposed adversarial network encouraged the segmentation to have a similar distribution of outputs for images with and without annotations, thereby improving the generalization ability of the network.

Dong *et al.* proposed a conditional GAN to train deep neural networks for the segmentation of multiple organs on thoracic CT images [45]. The proposed U-Net-generative-adversarial-network (U-Net-GAN) utilized a set of U-Nets as generators and a fully convolutional networks (FCNs) as discriminators. Specifically, the U-Nets were trained to produce image segmentation map of multiple organs. The discriminator, structured as FCN, discriminated between the ground truth and segmented organs produced by the generator. The generator and discriminator competed against each other in an adversarial learning process to improve the segmentation results of multiple organs. For multi-organ segmentation, training a universal network to segment all targets usually result in reduced segmentation accuracy. GAN method in [45] grouped OARs of similar dimensions, and utilized three sub-networks for segmentation, one for lungs and heart, and the other two for esophagus and spinal cord, respectively. This approach improved segmentation accuracy at the cost of computation efficiency. This segmentation technique was applied to delineate the left and right lungs, spinal cord, esophagus, and heart using 35 patients' chest CTs. The averaged DSC for the above five OARs are 0.97, 0.97, 0.90, 0.75, and 0.87, respectively.

Discussion: GAN can improve the segmentation accuracy by training a discriminator to generate an adversarial loss. However, GAN based network can be difficult to train since the generator and discriminator needs to be trained simultaneously to reach Nash equilibrium. Binary classifications of the results as fake or true provide stepped and unsmooth gradient, which make it difficult to train the discriminator. To alleviate this problem, Wasserstein GAN [258] was proposed to use Earth-Mover distance based metrics to replace the binary classification to improve the gradient back propagation during discriminator training. The training stage of GAN can be difficult and time-consuming. However, once trained, the GAN uses only the generator to perform segmentation. Using adversarial loss as a shape regulator can benefit more when the learning target (organ) has a regular distinctive shape, e.g., for lung and heart, but can be less useful for small tubular objects, such as vessels and catheters.

2.2.4 Synthetic image-aided segmentation—

Multi-modal images could be used to improve the segmentation accuracy since each imaging modality has its own advantages and disadvantages. For example, CT images have high bony structure definition but low soft-tissue contrast. MR images have high soft-tissue contrast but lower image spatial resolution. Therefore, it is beneficial to take multi-modal images for segmentation. However, multi-modal images are not always available for the images to be segmented. Even if other modality images exist, they need to be co-registered at first in order to be content-consistent. As an alternative, cross-modality image synthesis has been used to aid the segmentation process [131, 133, 259].

Overview of works: Accurate segmentation of the pelvic OARs on CT image for treatment planning is challenging due to the poor soft-tissue contrast [85, 260]. MRI has been used to aid CT prostate delineation, but it is not as accessible as CT for radiation therapy [261, 262]. Lei *et al.* developed a deep attention-based segmentation strategy to segment CT pelvic organs with the help of synthetic MRI (sMRI), which were generated by cycle generative adversarial network (CycleGAN) [82] [40]. This method includes two steps: first, a CycleGAN was used to estimate sMRI from CT images. Second, a deep attention FCN was trained based on the sMRI and manual contours deformed from MRIs. Attention models were introduced to focus on prostate boundary. Inspired by this method, Dong *et al.* developed a novel sMRI-aided segmentation method for mail pelvic CT multi-organ [39]. The DSC between the segmented bladder, prostate, and rectum manual contours were 0.95 ± 0.03, 0.87 ± 0.04 and 0.89 ± 0.04 respectively. Similarly, Lei *et al.* extended this method to multi-organ segmentation of cone-beam computed tomography (CBCT) pelvic data for potential CBCT-guided adaptive radiation therapy workflow [38]. The DSC between the segmented and physicians' manual contours (bladder, prostate, and rectum) were 0.95 ± 0.02, 0.86 ± 0.06 and 0.91 ± 0.04, respectively.

Head-and-neck (HN) multi-organ segmentation is very challenging for radiotherapy because many vital and small structures exist in the area. CT/CBCT HN images suffer from low soft tissue contrast and image artifacts which make is difficult to accurately segment all the OARs. FCN methods often result in boundary leakage due to the low soft tissue contrast. To alleviate this problem and increase the automatic segmentation accuracy, Liu *et al.* proposed

a synthetic MRI-aided CT image segmentation method using dual pyramid network [133]. A CycleGAN image synthesis network was first trained using manually registered MRI-CT image pairs. Synthesized MRI was generated based on the original CT images using the CycleGAN. The synthetic MRI and CT images were both taken as input images to train a dual pyramid network. The authors have achieved superior results as compared to the 2015 HN challenge results. Similarly, Dai *et al.* proposed a synthetic MRI-aided CBCT HN segmentation method for adaptive radiotherapy [259].

Discussion: Compared to using only the CT/CBCT images, sMRI-aided image segmentation has higher segmentation accuracy. This is because the sMRI provides complementary information for the network to learn. For prostate, this improves the prostate segmentation accuracy and alleviate the issue of prostate volume overestimation when using CT images alone. For HN, sMRI improves the boundary definition of the small organ structures due to the sMRI's superior soft tissue contrast. However, the image quality of the sMRI largely depends on the quality of the training MRI-CT image pairs, which are usually obtained using deformable image registration. Therefore, the performance of this method can be affected by the quality of image registration. It is very difficult to perform robust and accurate image registration of the abdomen especially when the MRI and CT images were acquired at different days.

3. Benchmark

Due to the different datasets used for evaluation, it is difficult for the readers to compare the accuracy of the surveyed methods. To facilitate direct comparison, we have summarized the performance of many surveyed papers that used the same benchmark datasets. Two benchmark datasets surveyed here are from the 2017 AAPM thoracic auto-segmentation challenge [263] and 2015 MICCAI Head and Neck Auto-segmentation Challenge [264].

3.1 2017 AAPM Thoracic Auto-segmentation Challenge

The 2017 AAPM Thoracic Auto-segmentation Challenge provided a benchmark dataset to evaluate the performance of automatic multi-organ segmentation methods for thoracic CT images. The OARs included left and right lungs, heart, esophagus, and spinal cord. Sixty thoracic CT scans provided by three institutions were separated into 36 training, 12 offline testing, and 12 online testing scans. Clinical contours used for treatment planning were quality checked and edited to adhere to the RTOG 1106 contouring guidelines. From the report of this challenge, there are 7 participants who completed the online challenge. Five out of the 7 participants used DL-based methods. In addition to the participating methods, other algorithms that used the datasets were also listed in Table 7.

There are no significant differences in terms of DSC of lung, heart and spinal cord for the listed DL-based methods. In comparison, the DSC of esophagus has larger variations due to its low soft tissue contrast. The DSC metric could be biased since it tends to favor organs with large volume, such as the lung and heart. Similarly, it is important to note whether extensive post-processing was performed when interpreting the MSD and HD95 since the post-processing could significantly affect the values of surface agreement metrics. With no post-processing, the U-Net-GAN method shows better surface agreement with ground truth.

This is due to the use of GAN which enforced additional regularization to the shape of the segmented organs.

3.2 2015 MICCAI Head and Neck Auto-segmentation Challenge

The 2015 MICCAI Head and Neck Auto-segmentation Challenge [264] provided a benchmark dataset to evaluate the performance of automatic multi-organ segmentation methods for head & neck CT images. The OARs were brainstem, mandible, chiasm, bilateral optic nerves, bilateral parotid glands, and bilateral submandibular glands. The datasets included 40 images, out of which, 25 images were used as training data, 10 images were used for off-site testing, and 5 images were used for on-site testing. The datasets were chosen to ensure good image quality, complete OAR coverage and minimal tumor overlap with OARs without any age or gender requirements. The report of this challenge [264] did not include any DL-based method. We studied recent DL-based multi-organ segmentation methods that used this benchmark dataset. The performance of these methods in terms of DSC and HD95 were listed in Table 8.

For HN organ segmentation, the synthetic MRI-aided method has near consistent improvement over other DL-based methods. This demonstrated the efficacy of synthetic MRI in image segmentation. Significant improvement has been achieved for the chiasm segmentation using synthetic MRI. However, as Liu *et al.* pointed out, synthetic MRI may not be as effective for some other organs such as the parotid, for which CT has good contrast. Therefore, the synthetic MRI-aided method has similar performance for the parotid segmentation as other methods. To train a synthetic image generation network, well-aligned CT-MRI image pairs are required for training. The unavailability of well-aligned image pairs poses additional challenge for synthetic image generation.

4. Other considerations

Data collected directly from clinical databases are usually not ready for network training. It is necessary to perform data preprocessing such as image resizing, image cropping, image normalization and data augmentation prior to network training. Other pre-processing techniques include registration [108], bias/scatter/attenuation correction [270, 271], voxel intensity normalization [272] and cropping [52]. Data augmentation is used to increase the amount of training samples and reduce network over-fitting. Typical data augmentation techniques include image rotation, translation, scaling, flipping, distortion, linear warping, elastic deformation, and noise contamination. During training, the ground truth contours are obtained by manual delineation by physicians. Depending on the manual contours generation, it is likely that the DL-based method is biased toward physicians' contouring style as a system error, and contouring uncertainty as a random error. This limitation is expected in all supervised learning-based methods.

Depending on the network design and GPU availability, some methods use the whole image volume as input to train the network [216] whereas some methods uses 2D image slices [177]. The 2D-based approaches can reduce the memory requirement and computational cost; however, it fails to utilize the spatial information in the third dimension. Another way to exploit the 3D feature information while avoid computer memory overflow is to use 2D

kernels on multidirectional 2D images. Segmentation results from different image planes such as axial, coronal and sagittal planes can be combined using a surface-based contour refinement method [43]. 3D image patches are also widely used as network input [139, 140] to keep 3D spatial information and reduce computational cost.

Post-processing is often applied to refine the DL segmentation results to generate realistic structures with smooth boundary. Morphological operations are widely used to remove small erroneous labels. Conditional random field is also widely adopted as a post-processing step to refine segmentation results [216]. Conditional random field models the pixelwise spatial relationships in order to preserve the plausibility of the segmented structures. As GAN-based methods are increasingly used to penalize implausible structures and preserve the spatial integrity of the segmentation results, conditional random field post-processing is expected to be used less in the future.

Another challenge in multi-organ segmentation is the class imbalance problem. For example, the esophagus and spinal cord are often much smaller than the lung for the segmentation of thorax CT organs. Training a network with class imbalanced datasets would bias the result towards the classes of large organs. Therefore, the choice of loss functions is crucial for these tasks [273]. Johnson *et al.* published a survey on deep learning with class imbalance [274].

5. Conclusion

We provide a comprehensive review of recently published deep learning-based multi-organ segmentation methods. There is a clear trend of using fully convolution network to perform end-to-end multi-organ segmentation. This is because the end-to-end segmentation takes the advantage of the rapid inference of DL network which significantly expedites the segmentation process. Region-based FCN methods are also gaining popularity to first identify the regions of targets and then performing region-based FCN segmentation for each target. However, the region-based FCN has not shown wide adoption because the one-step multi-organ segmentation network can also obtain good segmentation results without having to localize the target first. The necessity of the first organ localization step in region-based FCN is yet to be justified on more datasets. Another recent trend is the introduction of synthetic image to provide enhanced tissue contrast and to improve the realism of segmentation. However, the image synthesis step in synthetic image-aided segmentation is not trivial which prevent it from wide adoption. Apart from the unavailability of well-aligned paired images for synthesis network training, how much benefit the synthesis images can bring to the segmentation task remains to be further investigated. One challenge for segmentation is to predict reasonable organ boundaries at regions with poor image contrast and few intensity gradients. Additional organ shape constraints are necessary to regularize the segmentation results. The GAN-based methods have shown promising results in increasing the plausibility and fidelity of the segmented structures by penalizing unreasonable segmentation results. Therefore, we expect to see a steady growth of shape constrained multi-organ segmentation. To enable direct comparison, we have listed the performance of many surveyed studied that used the 2017 AAPM thoracic contouring challenge datasets and the 2015 MICCAI head & neck contouring challenge datasets.

Acknowledgements

This research is supported in part by the National Cancer Institute of the National Institutes of Health under Award Number R01CA215718, and Emory Winship Cancer Institute pilot grant.

References

[1]. Liao ZX, Lee JJ, Komaki R, Gomez DR, O'Reilly MS, Fossella FV, et al. Bayesian Adaptive Randomization Trial of Passive Scattering Proton Therapy and Intensity-Modulated Photon Radiotherapy for Locally Advanced Non-Small-Cell Lung Cancer (vol 36, pg 1813, 2018). Journal of Clinical Oncology. 2018;36:2570-. [PubMed: 31329708]

[2]. Molitoris JK, Diwanji T, Snider JW, Mossahebi S, Samanta S, Badiyan SN, et al. Advances in the use of motion management and image guidance in radiation therapy treatment for lung cancer. Journal of Thoracic Disease. 2018;10:S2437–S50. [PubMed: 30206490]

[3]. Vyfhuis MAL, Onyeuku N, Diwanji T, Mossahebi S, Amin NP, Badiyan SN, et al. Advances in proton therapy in lung cancer. Therapeutic Advances in Respiratory Disease. 2018;12.

[4]. Hurkmans CW, Borger JH, Pieters BR, Russell NS, Jansen EP, Mijnheer BJ. Variability in target volume delineation on CT scans of the breast. Int J Radiat Oncol Biol Phys. 2001;50:1366–72. [PubMed: 11483349]

[5]. Rasch C, Steenbakkers R, van Herk M. Target definition in prostate, head, and neck. Seminars in Radiation Oncology. 2005;15:136–45. [PubMed: 15983939]

[6]. Van de Steene J, Linthout N, de Mey J, Vinh-Hung V, Claassens C, Noppen M, et al. Definition of gross tumor volume in lung cancer: inter-observer variability. Radiotherapy and Oncology. 2002;62:37–49. [PubMed: 11830311]

[7]. Vinod SK, Jameson MG, Min M, Holloway LC. Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies. Radiotherapy and Oncology. 2016;121:169–79. [PubMed: 27729166]

[8]. Breunig J, Hernandez S, Lin J, Alsager S, Dumstorf C, Price J, et al. A System for Continual Quality Improvement of Normal Tissue Delineation for Radiation Therapy Treatment Planning. International Journal of Radiation Oncology Biology Physics. 2012;83:E703–E8.

[9]. Nelms BE, Tome WA, Robinson G, Wheeler J. Variations in the Contouring of Organs at Risk: Test Case from a Patient with Oropharyngeal Cancer. International Journal of Radiation Oncology Biology Physics. 2012;82:368–78.

[10]. Chen X, Pan L. A Survey of Graph Cuts/Graph Search Based Medical Image Segmentation. IEEE Reviews in Biomedical Engineering. 2018;11:112–24. [PubMed: 29994356]

[11]. Naqa IE, Yang D, Apte A, Khullar D, Mutic S, Zheng J, et al. Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning. Medical physics. 2007;34 12:4738–49. [PubMed: 18196801]

[12]. Pratondo A, Chui C, Ong S. Robust Edge-Stop Functions for Edge-Based Active Contour Models in Medical Image Segmentation. IEEE Signal Processing Letters. 2016;23:222–6.

[13]. Tsai A, Yezzi A, Wells W, Tempany C, Tucker D, Fan A, et al. A shape-based approach to the segmentation of medical imagery using level sets. IEEE Transactions on Medical Imaging. 2003;22:137–54. [PubMed: 12715991]

[14]. Dong X, Lei Y, Tian S, Liu Y, Wang T, Liu T, et al. Air, bone and soft-tissue Segmentation on 3D brain MRI Using Semantic Classification Random Forest with Auto-Context Model. ArXiv. 2019;doi: arXiv:1911.09264.

[15]. Isgum I, Staring M, Rutten A, Prokop M, Viergever MA, van Ginneken B. Multi-Atlas-Based Segmentation With Local Decision Fusion-Application to Cardiac and Aortic Segmentation in CT Scans. Ieee Transactions on Medical Imaging. 2009;28:1000–10. [PubMed: 19131298]

[16]. Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. Neuroimage. 2009;46:726–38. [PubMed: 19245840]

[17]. Iglesias JE, Sabuncu MR. Multi-atlas segmentation of biomedical images: A survey. Medical Image Analysis. 2015;24:205–19. [PubMed: 26201875]

[18]. Yang X, Rossi P, Ogunleye T, Marcus DM, Jani AB, Mao H, et al. Prostate CT segmentation method based on nonrigid registration in ultrasound-guided CT-based HDR prostate brachytherapy. Med Phys. 2014;41:111915. [PubMed: 25370648]

[19]. Ecabert O, Peters J, Schramm H, Lorenz C, von Berg J, Walker MJ, et al. Automatic model-based segmentation of the heart in CT images. Ieee Transactions on Medical Imaging. 2008;27:1189–201. [PubMed: 18753041]

[20]. Qazi AA, Pekar V, Kim J, Xie J, Breen SL, Jaffray DA. Auto-segmentation of normal and target structures in head and neck CT images: A feature-driven model-based approach. Medical Physics. 2011;38:6160–70. [PubMed: 22047381]

[21]. Sun SH, Bauer C, Beichel R. Automated 3-D Segmentation of Lungs With Lung Cancer in CT Data Using a Novel Robust Active Shape Model Approach. IEEE Transactions on Medical Imaging. 2012;31:449–60. [PubMed: 21997248]

[22]. Beam AL, Kohane IS. Translating Artificial Intelligence Into Clinical Care Jama-Journal of the American Medical Association. 2016;316:2368–9.

[23]. Pella A, Cambria R, Riboldi M, Jereczek-Fossa BA, Fodor C, Zerini D, et al. Use of machine learning methods for prediction of acute toxicity in organs at risk following prostate radiotherapy. Medical Physics. 2011;38:2859–67. [PubMed: 21815361]

[24]. Yang X, Wu N, Cheng G, Zhou Z, David SY, Beitler JJ, et al. Automated segmentation of the parotid gland based on atlas registration and machine learning: a longitudinal MRI study in head-and-neck radiation therapy. International Journal of Radiation Oncology* Biology* Physics. 2014;90:1225–33.

[25]. Bryce TJ, Dewhirst MW, Floyd CE, Hars V, Brizel DM. Artificial neural network model of survival in patients treated with irradiation with and without concurrent chemotherapy for advanced carcinoma of the head and neck. International Journal of Radiation Oncology Biology Physics. 1998;41:339–45.

[26]. Gulliford SL, Webb S, Rowbottom CG, Corne DW, Dearnaley DP. Use of artificial neural networks to predict biological outcomes for patients receiving radical radiotherapy of the prostate. Radiotherapy and Oncology. 2004;71:3–12. [PubMed: 15066290]

[27]. Tomatis S, Rancati T, Fiorino C, Vavassori V, Fellin G, Cagna E, et al. Late rectal bleeding after 3D-CRT for prostate cancer: development of a neural-network-based predictive model. Physics in Medicine and Biology. 2012;57:1399–412. [PubMed: 22349550]

[28]. Chen SF, Zhou SM, Zhang JN, Yin FF, Marks LB, Das SK. A neural network model to predict lung radiation-induced pneumonitis. Medical Physics. 2007;34:3420–7. [PubMed: 17926943]

[29]. Su M, Miften M, Whiddon C, Sun XJ, Light K, Marks L. An artificial neural network for predicting the incidence of radiation pneumonitis. Medical Physics. 2005;32:318–25. [PubMed: 15789575]

[30]. Ochi T, Murase K, Fujii T, Kawamura M, Ikezoe J. Survival prediction using artificial neural networks in patients with uterine cervical cancer treated by radiation therapy alone. Radiology. 2000;217:142-.

[31]. Boldrini L, Bibault JE, Masciocchi C, Shen YT, Bittner MI. Deep Learning: A Review for the Radiation Oncologist. Frontiers in Oncology. 2019;9. [PubMed: 30723704]

[32]. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436–44. [PubMed: 26017442]

[33]. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. Communications of the Acm. 2012;60:84–90.

[34]. Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y. OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks. CoRR. 2013;abs/1312.6229.

[35]. Shelhamer E, Long J, Darrell T. Fully Convolutional Networks for Semantic Segmentation. Ieee Transactions on Pattern Analysis and Machine Intelligence. 2017;39:640–51. [PubMed: 27244717]

[36]. Hesamian MH, Jia W, He XJ, Kennedy P. Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. Journal of Digital Imaging. 2019;32:582–96. [PubMed: 31144149]

[37]. Zhou T, Ruan S, Canu S. A review: Deep learning for medical image segmentation using multi-modality fusion. Array. 2019;3–4:100004.

[38]. Lei Y, Wang T, Tian S, Dong X, Jani AB, Schuster D, et al. Male pelvic multi-organ segmentation aided by CBCT-based synthetic MRI. Phys Med Biol. 2019;in press, doi: 10.1088/1361-6560/ab63bb.

[39]. Dong X, Lei Y, Tian S, Wang T, Patel P, Curran WJ, et al. Synthetic MRI-aided multi-organ segmentation on male pelvic CT using cycle consistent deep attention network. Radiother Oncol. 2019;141:192–9. [PubMed: 31630868]

[40]. Lei Y, Dong X, Tian Z, Liu Y, Tian S, Wang T, et al. CT prostate segmentation based on synthetic MRI-aided deep attention fully convolution network. Med Phys. 2019;in press, doi: 10.1002/mp.13933.

[41]. Lei Y, Dong X, Wang T, Higgins K, Liu T, Curran WJ, et al. Whole-body PET estimation from low count statistics using cycle-consistent generative adversarial networks. Phys Med Biol. 2019;64:215017. [PubMed: 31561244]

[42]. van der Heyden B, Wohlfahrt P, Eekers DBP, Richter C, Terhaag K, Troost EGC, et al. Dual-energy CT for automatic organs-at-risk segmentation in brain-tumor patients using a multi-atlas and deep-learning approach. Scientific Reports. 2019;9.

[43]. Lei Y, Tian S, He X, Wang T, Wang B, Patel P, et al. Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net. Med Phys. 2019;46:3194–206. [PubMed: 31074513]

[44]. Wang T, Lei Y, Tian S, Jiang X, Zhou J, Liu T, et al. Learning-based automatic segmentation of arteriovenous malformations on contrast CT images in brain stereotactic radiosurgery. Med Phys. 2019;46:3133–41. [PubMed: 31050804]

[45]. Dong X, Lei Y, Wang T, Thomas M, Tang L, Curran WJ, et al. Automatic multiorgan segmentation in thorax CT images using U-net-GAN. Med Phys. 2019;46:2157–68. [PubMed: 30810231]

[46]. Wang B, Lei Y, Tian S, Wang T, Liu Y, Patel P, et al. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. Med Phys. 2019;46:1707–18. [PubMed: 30702759]

[47]. Lei Y, Wang T, Wang B, He X, Tian S, Jani AB, et al. Ultrasound prostate segmentation based on 3D V-Net with deep supervision. SPIE Medical Imaging2019.

[48]. Wang B, Lei Y, Wang T, Dong X, Tian S, Jiang X, et al. Automated prostate segmentation of volumetric CT images using 3D deeply supervised dilated FCN. SPIE Medical Imaging2019.

[49]. Wang T, Lei Y, Tang H, Harms J, Wang C, Liu T, et al. A learning-based automatic segmentation method on left ventricle in SPECT imaging. SPIE Medical Imaging2019.

[50]. Wang B, Lei Y, Jeong JJ, Wang T, Liu Y, Tian S, et al. Automatic MRI prostate segmentation using 3D deeply supervised FCN with concatenated atrous convolution. SPIE Medical Imaging2019.

[51]. Lei Y, Liu Y, Dong X, Tian S, Wang T, Jiang X, et al. Automatic multi-organ segmentation in thorax CT images using U-Net-GAN. SPIE Medical Imaging2019.

[52]. Wang T, Lei Y, Shafai-Erfani G, Jiang X, Dong X, Zhou J, et al. Learning-based automatic segmentation on arteriovenous malformations from contract-enhanced CT images. SPIE Medical Imaging2019.

[53]. Wang T, Lei Y, Tang H, He Z, Castillo R, Wang C, et al. A learning-based automatic segmentation and quantification method on left ventricle in gated myocardial perfusion SPECT imaging: A feasibility study. J Nucl Cardiol. 2019;In press, doi: 10.1007/s12350-019-01594-2. .

[54]. Wu J, Xin J, Yang X, Sun J, Xu D, Zheng N, et al. Deep morphology aided diagnosis network for segmentation of carotid artery vessel wall and diagnosis of carotid atherosclerosis on black-blood vessel wall MRI. Med Phys. 2019;46:5544–61. [PubMed: 31356693]

[55]. Fu Y, Lei Y, Wang T, Tian S, Patel P, Jani AB, et al. Pelvic multi-organ segmentation on cone-beam CT for prostate adaptive radiotherapy. Med Phys. 2020;47:3415–22. [PubMed: 32323330]

[56]. Jun Guo B, He X, Lei Y, Harms J, Wang T, Curran WJ, et al. Automated left ventricular myocardium segmentation using 3D deeply supervised attention U-net for coronary computed tomography angiography; CT myocardium segmentation. Med Phys. 2020;47:1775–85. [PubMed: 32017118]

[57]. Lei Y, Dong X, Tian Z, Liu Y, Tian S, Wang T, et al. CT prostate segmentation based on synthetic MRI-aided deep attention fully convolution network. Med Phys. 2020;47:530–40. [PubMed: 31745995]

[58]. Lei Y, Fu Y, Roper J, Higgins K, Bradley JD, Curran WJ, et al. Echocardiographic Image Multi-Structure Segmentation using Cardiac-SegNet. Med Phys. 2021;(in press), doi: 10.1002/mp.14818.

[59]. Lei Y, He X, Yao J, Wang T, Wang L, Li W, et al. Breast tumor segmentation in 3D automatic breast ultrasound using Mask scoring R-CNN. Med Phys. 2021;48:204–14. [PubMed: 33128230]

[60]. Liu Y, Lei Y, Fu Y, Wang T, Zhou J, Jiang X, et al. Head and neck multi-organ auto-segmentation on CT images aided by synthetic MRI. Med Phys. 2020;47:4294–302. [PubMed: 32648602]

[61]. Fu YB, Mazur TR, Wu X, Liu S, Chang X, Lu YG, et al. A novel MRI segmentation method using CNN-based correction network for MRI-guided adaptive radiotherapy. Medical Physics. 2018;45:5129–37. [PubMed: 30269345]

[62]. Liu Y, Lei Y, Fu Y, Wang T, Tang X, Jiang X, et al. CT-based multi-organ segmentation using a 3D self-attention U-net network for pancreatic radiotherapy. Medical Physics. 2020;47:4316–24. [PubMed: 32654153]

[63]. He X, Guo BJ, Lei Y, Wang T, Fu Y, Curran WJ, et al. Automatic segmentation and quantification of epicardial adipose tissue from coronary computed tomography angiography. Phys Med Biol. 2020;65:095012. [PubMed: 32182595]

[64]. Jeong J, Lei Y, Kahn S, Liu T, Curran WJ, Shu HK, et al. Brain tumor segmentation using 3D Mask R-CNN for dynamic susceptibility contrast enhanced perfusion imaging. Phys Med Biol. 2020;65:185009. [PubMed: 32674075]

[65]. Lei Y, Wang T, Tian S, Dong X, Jani AB, Schuster D, et al. Male pelvic multi-organ segmentation aided by CBCT-based synthetic MRI. Phys Med Biol. 2020;65:035013. [PubMed: 31851956]

[66]. Fu YB, Liu S, Li HH, Yang DS. Automatic and hierarchical segmentation of the human skeleton in CT images. Physics in Medicine and Biology. 2017;62:2812–33. [PubMed: 28195561]

[67]. He X, Guo B, Lei Y, Tian S, Wang T, Curran WJ, et al. Thyroid gland delineation in noncontrast-enhanced CT using deep convolutional neural networks. Phys Med Biol. 2020;DOI: 10.1088/1361-6560/abc5a6.

[68]. Zhang Y, He X, Tian Z, Jeong JJ, Lei Y, Wang T, et al. Multi-Needle Detection in 3D Ultrasound Images Using Unsupervised Order-Graph Regularized Sparse Dictionary Learning. IEEE Transactions on Medical Imaging. 2020;39:2302–15. [PubMed: 31985414]

[69]. Harms J, Lei Y, Tian S, McCall NS, Higgins K, Bradley JD, et al. Automatic Delineation of Cardiac Substructures using a Region-Based Fully Convolutional Network. Med Phys. 2021;(in press) DOI: 10.1002/mp.14810.

[70]. Zhang Y, Tian Z, Lei Y, Wang T, Patel P, Jani AB, et al. Automatic multi-needle localization in ultrasound images using large margin mask RCNN for ultrasound-guided prostate brachytherapy. Phys Med Biol. 2020;65:205003. [PubMed: 32640435]

[71]. Dai X, Lei Y, Zhang Y, Qiu RLJ, Wang T, Dresser SA, et al. Automatic multi-catheter detection using deeply supervised convolutional neural network in MRI-guided HDR prostate brachytherapy. Medical Physics. 2020;47:4115–24. [PubMed: 32484573]

[72]. Zhang Y, Lei Y, Qiu RLJ, Wang T, Wang H, Jani AB, et al. Multi-needle Localization with Attention U-Net in US-guided HDR Prostate Brachytherapy. Med Phys. 2020;47:2735–45. [PubMed: 32155666]

[73]. Dong X, Wang T, Lei Y, Higgins K, Liu T, Curran WJ, et al. Synthetic CT generation from non-attenuation corrected PET images for whole-body PET imaging. Phys Med Biol. 2019;64:215016. [PubMed: 31622962]

[74]. Liu Y, Lei Y, Wang Y, Shafai-Erfani G, Wang T, Tian S, et al. Evaluation of a deep learning-based pelvic synthetic CT generation technique for MRI-based prostate proton treatment planning. Phys Med Biol. 2019;64:205022. [PubMed: 31487698]

[75]. Lei Y, Wang T, Harms J, Fu Y, Dong X, Curran WJ, et al. CBCT-Based Synthetic MRI Generation for CBCT-Guided Adaptive Radiotherapy. In: Nguyen D, Xing L, Jiang S, editors.

Artificial Intelligence in Radiation Therapy. Cham: Springer International Publishing; 2019. p. 154–61.

[76]. Shafai-Erfani G, Lei Y, Liu Y, Wang Y, Wang T, Zhong J, et al. MRI-Based Proton Treatment Planning for Base of Skull Tumors. International Journal of Particle Therapy. 2019;6:12–25. [PubMed: 31998817]

[77]. Yang X, Lei Y, Wang T, Liu Y, Tian S, Dong X, et al. CBCT-guided Prostate Adaptive Radiotherapy with CBCT-based Synthetic MRI and CT. International Journal of Radiation Oncology • Biology • Physics ASTRO: Elsevier; 2019. p. S250.

[78]. Yang X, Liu Y, Lei Y, Wang Y, Shafai-Erfani G, Wang T, et al. MRI-based Proton Radiotherapy for Prostate Cancer Using Deep Convolutional Neural Networks. International Journal of Radiation Oncology • Biology • Physics ASTRO: Elsevier; 2019. p. S200.

[79]. Wang T, Manohar N, Lei Y, Dhabaan A, Shu HK, Liu T, et al. MRI-based treatment planning for brain stereotactic radiosurgery: Dosimetric validation of a learning-based pseudo-CT generation method. Med Dosim. 2019;44:199–204. [PubMed: 30115539]

[80]. Liu Y, Lei Y, Wang Y, Wang T, Ren L, Lin L, et al. MRI-based treatment planning for proton radiotherapy: dosimetric validation of a deep learning-based liver synthetic CT generation method. Phys Med Biol. 2019;64:145015. [PubMed: 31146267]

[81]. Liu Y, Lei Y, Wang T, Kayode O, Tian S, Liu T, et al. MRI-based treatment planning for liver stereotactic body radiotherapy: validation of a deep learning-based synthetic CT generation method. Br J Radiol. 2019;92:20190067. [PubMed: 31192695]

[82]. Lei Y, Harms J, Wang T, Liu Y, Shu HK, Jani AB, et al. MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks. Med Phys. 2019;46:3565–81. [PubMed: 31112304]

[83]. Lei Y, Wang T, Liu Y, Higgins K, Tian S, Liu T, et al. MRI-based synthetic CT generation using deep convolutional neural network. SPIE Medical Imaging2019.

[84]. Shafai-Erfani G, Wang T, Lei Y, Tian S, Patel P, Jani AB, et al. Dose evaluation of MRI-based synthetic CT generated using a machine learning method for prostate cancer radiotherapy. Med Dosim. 2019;44:e64–e70. [PubMed: 30713000]

[85]. Yang X, Lei Y, Wang T, Patel PR, Jiang X, Liu T, et al. MRI-Based Synthetic CT for Radiation Treatment of Prostate Cancer. International Journal of Radiation Oncology • Biology • Physics ASTRO: Elsevier; 2018. p. S193–S4.

[86]. Charyyev S, Wang T, Lei Y, Ghavidel B, Beitler J, McDonald M, et al. Learning-Based Synthetic Dual Energy CT Imaging from Single Energy CT for Stopping Power Ratio Calculation in Proton Radiation Therapy. arXiv: Medical Physics. 2020;DOI: arXiv:2005.12908

[87]. Wang T, Lei Y, Fu Y, Curran W, Liu T, Yang X. Medical Imaging Synthesis using Deep Learning and its Clinical Applications: A Review. arXiv: Medical Physics. 2020;DOI: arXiv:2004.10322.

[88]. Wang T, Lei Y, Fu Y, Wynne JF, Curran WJ, Liu T, et al. A review on medical imaging synthesis using deep learning and its clinical applications. J Appl Clin Med Phys. 2021;22:11–36.

[89]. Liu Y, Lei Y, Wang T, Fu Y, Tang X, Curran WJ, et al. CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy. Med Phys. 2020;47:2472–83. [PubMed: 32141618]

[90]. Dai X, Lei Y, Fu Y, Curran WJ, Liu T, Mao H, et al. Multimodal MRI synthesis using unified generative adversarial networks. Medical Physics. 2020;DOI: 10.1002/mp.14539.

[91]. Liu R, Lei Y, Wang T, Zhou J, Roper J, Lin L, et al. Synthetic dual-energy CT for MRI-only based proton therapy treatment planning using label-GAN. Phys Med Biol. 2021;(in press) DOI: 10.1088/1361-6560/abe736.

[92]. Lei Y, Tian Z, Wang T, Higgins K, Bradley JD, Curran WJ, et al. Deep learning-based real-time volumetric imaging for lung stereotactic body radiation therapy: a proof of concept study. Physics in Medicine & Biology. 2020;65:235003. [PubMed: 33080578]

[93]. Harms J, Lei Y, Wang T, McDonald M, Ghavidel B, Stokes W, et al. Cone-beam CT-derived relative stopping power map generation via deep learning for proton radiotherapy. Med Phys. 2020;47:4416–27. [PubMed: 32579710]

[94]. Charyyev S, Lei Y, Harms J, Eaton B, McDonald M, Curran WJ, et al. High quality proton portal imaging using deep learning for proton radiation therapy: a phantom study. Biomedical Physics & Engineering Express. 2020;6:035029. [PubMed: 33438674]

[95]. Wang T, Lei Y, Harms J, Ghavidel B, Lin L, Beitler J, et al. Learning-Based Stopping Power Mapping on Dual Energy CT for Proton Radiation Therapy. arXiv: Medical Physics. 2020;DOI: arXiv:2005.12908.

[96]. Wang T, Lei Y, Fu Y, Curran W, Liu T, Yang X. Machine Learning in Quantitative PET Imaging. ArXiv. 2020;DOI: abs/2001.06597.

[97]. Dong X, Lei Y, Wang T, Higgins K, Liu T, Curran WJ, et al. Deep learning-based attenuation correction in the absence of structural information for whole-body PET imaging. Phys Med Biol. 2019;in press, doi: 10.1088/1361-6560/ab652c.

[98]. Wang T, Lei Y, Tian Z, Dong X, Liu Y, Jiang X, et al. Deep learning-based image quality improvement for low-dose computed tomography simulation in radiation therapy. J Med Imaging. 2019;6:043504.

[99]. Harms J, Lei Y, Wang T, Zhang R, Zhou J, Tang X, et al. Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. Med Phys. 2019;46:3998–4009. [PubMed: 31206709]

[100]. Wang T, Lei Y, Manohar N, Tian S, Jani AB, Shu HK, et al. Dosimetric study on learning-based cone-beam CT correction in adaptive radiation therapy. Med Dosim. 2019;44:e71–e9. [PubMed: 30948341]

[101]. Lei Y, Wang T, Harms J, Shafai-Erfani G, Dong X, Zhou J, et al. Image quality improvement in cone-beam CT using deep learning. SPIE Medical Imaging2019.

[102]. Yang X, Wang T, Lei Y, Jiang X, Jani A, Patel PR, et al. A Learning-Based Method to Improve Pelvis Cone Beam CT Image Quality for Prostate Cancer Radiation Therapy. International Journal of Radiation Oncology • Biology • Physics ASTRO: Elsevier; 2018. p. E377–E8.

[103]. Lei Y, Xu D, Zhou Z, Higgins K, Dong X, Liu T, et al. High-resolution CT image retrieval using sparse convolutional neural network. SPIE Medical Imaging2018.

[104]. Yang X, Lei Y, Dong X, Wang T, Higgins K, Liu T, et al. Attenuation and Scatter Correction for Whole-body PET Using 3D Generative Adversarial Networks. J Nucl Med2019. p. 174.

[105]. Yang X, Lei Y, Wang T, Dong X, Higgins K, Curran WJ, et al. Whole-body PET Estimation from Ultra-short Scan Durations using 3D Cycle-Consistent Generative Adversarial Networks. J Nucl Med2019. p. 247.

[106]. Wang T, Lei Y, Fu Y, Curran WJ, Liu T, Nye JA, et al. Machine learning in quantitative PET: A review of attenuation correction and low-count image reconstruction methods. Phys Med. 2020;76:294–306. [PubMed: 32738777]

[107]. Dai X, Lei Y, Liu Y, Wang T, Ren L, Curran WJ, et al. Intensity non-uniformity correction in MR imaging using residual cycle generative adversarial network. Physics in Medicine & Biology. 2020;65:215025. [PubMed: 33245059]

[108]. Fu Y, Lei Y, Wang T, Curran WJ, Liu TJ, Yang XJA. Deep Learning in Medical Image Registration: A Review. ArXiv. 2019;abs/1912.12318.

[109]. Haskins G, Kruger U, Yan P. Deep Learning in Medical Image Registration: A Survey. ArXiv. 2019;abs/1903.02026.

[110]. Lei Y, Fu Y, Harms J, Wang T, Curran WJ, Liu T, et al. 4D-CT Deformable Image Registration Using an Unsupervised Deep Convolutional Neural Network. In: Nguyen D, Xing L, Jiang S, editors. Artificial Intelligence in Radiation Therapy. Cham: Springer International Publishing; 2019. p. 26–33.

[111]. Yang X, Zeng Q, Lei Y, Tian S, Wang T, Dong X, et al. MRI-US Registration Using Label-driven Weakly-supervised Learning for Multiparametric MRI-guided HDR Prostate Brachytherapy. International Journal of Radiation Oncology • Biology • Physics ASTRO: Elsevier; 2019. p. E727.

[112]. Li HM, Fan Y. Non-Rigid Image Registration Using Self-Supervised Fully Convolutional Networks without Training Data 2018 Ieee 15th International Symposium on Biomedical Imaging (Isbi 2018). 2018:1075–8.

[113]. Fu Y, Lei Y, Wang T, Higgins K, Bradley J, Curran W, et al. LungRegNet: an unsupervised deformable image registration method for 4D-CT lung. Medical physics. 2020.

[114]. Fu Y, Lei Y, Wang T, Patel P, Jani A, Mao H, et al. Biomechanically constrained non-rigid MR-TRUS prostate registration using deep learning based 3D point cloud matching. Medical image analysis. 2020;67:101845. [PubMed: 33129147]

[115]. Fu Y, Wang T, Lei Y, Patel P, Jani A, Curran W, et al. Deformable MR-CBCT Prostate Registration using Biomechanically Constrained Deep Learning Networks. Medical physics. 2020.

[116]. Lei Y, Fu Y, Wang T, Liu Y, Patel P, Curran W, et al. 4D-CT deformable image registration using multiscale unsupervised deep learning. Physics in medicine and biology. 2020.

[117]. Fu Y, Lei Y, Wang T, Patel P, Jani AB, Mao H, et al. Biomechanically constrained non-rigid MR-TRUS prostate registration using deep learning based 3D point cloud matching. Med Image Anal. 2021;67:101845. [PubMed: 33129147]

[118]. Fu Y, Lei Y, Wang T, Higgins K, Bradley JD, Curran WJ, et al. LungRegNet: An unsupervised deformable image registration method for 4D-CT lung. Med Phys. 2020;47:1763–74. [PubMed: 32017141]

[119]. Fu Y, Wang T, Lei Y, Patel P, Jani AB, Curran WJ, et al. Deformable MR-CBCT prostate registration using biomechanically constrained deep learning networks. Med Phys. 2020;DOI: 10.1002/mp.14584.

[120]. Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X. Deep learning in medical image registration: a review. Phys Med Biol. 2020;65:20TR01.

[121]. Lei Y, Fu Y, Wang T, Liu Y, Patel P, Curran WJ, et al. 4D-CT deformable image registration using multiscale unsupervised deep learning. Phys Med Biol. 2020;65:085003. [PubMed: 32097902]

[122]. Zeng Q, Fu Y, Tian Z, Lei Y, Zhang Y, Wang T, et al. Label-driven magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy. Phys Med Biol. 2020;65:135002. [PubMed: 32330922]

[123]. Tong N, Gou S, Yang S, Ruan D, Sheng K. Fully automatic multi-organ segmentation for head and neck cancer radiotherapy using shape representation model constrained fully convolutional neural networks. Med Phys. 2018;45:4558–67. [PubMed: 30136285]

[124]. Men K, Chen XY, Zhang Y, Zhang T, Dai JR, Yi JL, et al. Deep Deconvolutional Neural Network for Target Segmentation of Nasopharyngeal Cancer in Planning Computed Tomography Images. Frontiers in Oncology. 2017;7.

[125]. Kazemifar S, Balagopal A, Nguyen D, McGuire S, Hannan R, Jiang S, et al. Segmentation of the prostate and organs at risk in male pelvic CT images using deep learning. Biomedical Physics & Engineering Express. 2018;4.

[126]. Javaid U, Dasnoy D, Lee JA. Multi-organ Segmentation of Chest CT Images in Radiation Oncology: Comparison of Standard and Dilated UNet. Advanced Concepts for Intelligent Vision Systems, Acivs 2018. 2018;11182:188–99.

[127]. Elguindi S, Zelefsky MJ, Jiang J, Veeraraghavan H, Deasy JO, Hunt MA, et al. Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy. Physics and Imaging in Radiation Oncology. 2019;12:80–6. [PubMed: 32355894]

[128]. Lamb J, Cao M, Kishan A, Agazaryan N, Thomas DH, Shaverdian N, et al. Online Adaptive Radiation Therapy: Implementation of a New Process of Care. Cureus. 2017;9:e1618. [PubMed: 29104835]

[129]. Fu Y, Mazur TR, Wu X, Liu S, Chang X, Lu Y, et al. A novel MRI segmentation method using CNN-based correction network for MRI-guided adaptive radiotherapy. Med Phys. 2018;45:5129–37. [PubMed: 30269345]

[130]. Liu Y, Lei Y, Fu Y, Wang T, Tang X, Jiang X, et al. CT-based Multi-organ Segmentation using a 3D Self-attention U-Net Network for Pancreatic Radiotherapy. Medical physics. 2020.

[131]. Fu Y, Lei Y, Wang T, Tian S, Patel P, Jani A, et al. Pelvic Multi-organ Segmentation on CBCT for Prostate Adaptive Radiotherapy Medical physics. 2020.

[132]. Tappeiner E, Proll S, Honig M, Raudaschl PF, Zaffino P, Spadea MF, et al. Multi-organ segmentation of the head and neck area: an efficient hierarchical neural networks approach. Int J Comput Assist Radiol Surg. 2019;14:745–54. [PubMed: 30847761]

[133]. Liu Y, Lei Y, Fu Y, Wang T, Zhou J, Jiang X, et al. Head and Neck Multi-Organ Auto-Segmentation on CT Images Aided by Synthetic MRI. Medical physics. 2020.

[134]. Hu P, Wu F, Peng J, Bao Y, Chen F, Kong D. Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets. International Journal of Computer Assisted Radiology and Surgery. 2017;12:399–411. [PubMed: 27885540]

[135]. Gibson E, Giganti F, Hu Y, Bonmati E, Bandula S, Gurusamy K, et al. Automatic Multi-Organ Segmentation on Abdominal CT With Dense V-Networks. IEEE Trans Med Imaging. 2018;37:1822–34. [PubMed: 29994628]

[136]. Chen S, Zhong X, Hu S, Dorn S, Kachelriess M, Lell M, et al. Automatic multi-organ segmentation in dual-energy CT (DECT) with dedicated 3D fully convolutional DECT networks. Med Phys. 2019.

[137]. Wang T, Lei Y, Fu Y, Curran W, Liu T, Yang X. Machine Learning in Quantitative PET Imaging. arXiv e-Print. 2020;arXiv:2001.06597.

[138]. Shin HC, Orton MR, Collins DJ, Doran SJ, Leach MO. Stacked Autoencoders for Unsupervised Feature Learning and Multiple Organ Detection in a Pilot Study Using 4D Patient Data. Ieee Transactions on Pattern Analysis and Machine Intelligence. 2013;35:1930–43. [PubMed: 23787345]

[139]. Alex V, Vaidhya K, Thirunavukkarasu S, Kesavadas C, Krishnamurthi G. Semisupervised learning using denoising autoencoders for brain lesion detection and segmentation. J Med Imaging (Bellingham). 2017;4:041311. [PubMed: 29285516]

[140]. Vaidhya K, Thirunavukkarasu S, Varghese A, Krishnamurthi G. Multi-modal Brain Tumor Segmentation Using Stacked Denoising Autoencoders. Brainles@MICCAI2015.

[141]. Ahmad M, Yang J, Ai D, Qadri SF, Wang Y. Deep-Stacked Auto Encoder for Liver Segmentation. In: Wang Y, Wang S, Liu Y, Yang J, Yuan X, He R, et al., editors. Advances in Image and Graphics Technologies. Singapore: Springer Singapore; 2018. p. 243–51.

[142]. Qadri SF, Zhao Z, Ai D, Ahmad M, Wang Y. Vertebrae segmentation via stacked sparse autoencoder from computed tomography images: SPIE; 2019.

[143]. Wang C, Elazab A, Jia F, Wu J, Hu Q. Automated chest screening based on a hybrid model of transfer learning and convolutional sparse denoising autoencoder. Biomed Eng Online. 2018;17:63. [PubMed: 29792208]

[144]. Wang X, Zhai S, Niu Y. Automatic Vertebrae Localization and Identification by Combining Deep SSAE Contextual Features and Structured Regression Forest. J Digit Imaging. 2019;32:336–48. [PubMed: 30631979]

[145]. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Trans Med Imaging. 2015;34:1993–2024. [PubMed: 25494501]

[146]. He KM, Zhang XY, Ren SQ, Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. 2015 Ieee International Conference on Computer Vision (Iccv). 2015:1026–34.

[147]. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE. 1998;86:2278–324.

[148]. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision. 2015;115:211–52.

[149]. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, editors. ECCV. Cham: Springer International Publishing; 2014. p. 818–33.

[150]. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR. 2014;abs/1409.1556.

[151]. Szegedy C, Wei L, Yangqing J, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)2015. p. 1–9.

[152]. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)2016. p. 770–8.

[153]. Huang G, Liu Z, Maaten Lvd, Weinberger KQ. Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)2017. p. 2261–9.

[154]. Roth HR, Lu L, Farag A, Shin HC, Liu JM, Turkbey EB, et al. DeepOrgan: Multi-level Deep Convolutional Networks for Automated Pancreas Segmentation. Medical Image Computing and Computer-Assisted Intervention - Miccai 2015, Pt I. 2015;9349:556–64.

[155]. Hamidian S, Sahiner B, Petrick N, Pezeshk A. 3D Convolutional Neural Network for Automatic Detection of Lung Nodules in Chest CT. Proc SPIE Int Soc Opt Eng. 2017;10134.

[156]. Armato Iii SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. Medical Physics. 2011;38:915–31. [PubMed: 21452728]

[157]. Zhong T, Huang X, Tang F, Liang SJ, Deng XG, Zhang Y. Boosting-based cascaded convolutional neural networks for the segmentation of CT organs-at-risk in nasopharyngeal carcinoma. Medical Physics. 2019.

[158]. Harten LDv Noothout JMH, Verhoeff J Wolterink JM, Išgum I Automatic Segmentation of Organs at Risk in Thoracic CT scans by Combining 2D and 3D Convolutional Neural Networks. SegTHOR@ISBI2019.

[159]. Johnson JM, Alahi A, Fei-Fei L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. ECCV2016.

[160]. Wolterink JM, Leiner T, Viergever MA, Išgum I. Dilated Convolutional Neural Networks for Cardiovascular MR Segmentation in Congenital Heart Disease. In: Zuluaga MA, Bhatia K, Kainz B, Moghari MH, Pace DF, editors. Reconstruction, Segmentation, and Analysis of Medical Images. Cham: Springer International Publishing; 2017. p. 95–102.

[161]. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Medical Image Analysis. 2017;36:61–78. [PubMed: 27865153]

[162]. Roth HR, Lu L, Farag A, Sohn A, Summers RM. Spatial Aggregation of Holistically-Nested Networks for Automated Pancreas Segmentation. In: Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Cham: Springer International Publishing; 2016. p. 451–9.

[163]. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. Medical Physics. 2017;44:547–57. [PubMed: 28205307]

[164]. K JD, R G, A M. Fuzzy-C-Means Clustering Based Segmentation and CNN-Classification for Accurate Segmentation of Lung Nodules. Asian Pac J Cancer Prev. 2017;18:1869–74. [PubMed: 28749127]

[165]. Zhou XR, Takayama R, Wang S, Zhou XX, Hara T, Fujita H. Automated segmentation of 3D anatomical structures on CT images by using a deep convolutional network based on end-to-end learning approach. Medical Imaging 2017: Image Processing. 2017;10133.

[166]. Bae HJ, Kim CW, Kim N, Park B, Kim N, Seo JB, et al. A Perlin Noise-Based Augmentation Strategy for Deep Learning with Small Data Samples of HRCT Images. Sci Rep. 2018;8:17687. [PubMed: 30523268]

[167]. Chmelik J, Jakubicek R, Walek P, Jan J, Ourednicek P, Lambert L, et al. Deep convolutional neural network-based segmentation and classification of difficult to define metastatic spinal lesions in 3D CT data. Med Image Anal. 2018;49:76–88. [PubMed: 30114549]

[168]. Gudmundsson E, Straus CM, Armato SG 3rd. Deep convolutional neural networks for the automated segmentation of malignant pleural mesothelioma on computed tomography scans. J Med Imaging (Bellingham). 2018;5:034503. [PubMed: 30840717]

[169]. Nardelli P, Jimenez-Carretero D, Bermejo-Pelaez D, Washko GR, Rahaghi FN, Ledesma-Carbayo MJ, et al. Pulmonary Artery-Vein Classification in CT Images Using Deep Learning. IEEE Trans Med Imaging. 2018;37:2428–40. [PubMed: 29993996]

[170]. Thyreau B, Sato K, Fukuda H, Taki Y. Segmentation of the hippocampus by transferring algorithmic knowledge for large cohort processing. Med Image Anal. 2018;43:214–28. [PubMed: 29156419]

[171]. Wang GT, Li WQ, Zuluaga MA, Pratt R, Patel PA, Aertsen M, et al. Interactive Medical Image Segmentation Using Deep Learning With Image-Specific Fine Tuning. Ieee Transactions on Medical Imaging. 2018;37:1562–73. [PubMed: 29969407]

[172]. Zhou XR, Yamada K, Kojima T, Takayama R, Wang S, Zhou XX, et al. Performance evaluation of 2D and 3D deep learning approaches for automatic segmentation of multiple organs on CT images. Medical Imaging 2018: Computer-Aided Diagnosis. 2018;10575.

[173]. Liu H, Wang L, Nan Y, Jin F, Wang Q, Pu J. SDFN: Segmentation-based deep fusion network for thoracic disease classification in chest X-ray images. Comput Med Imaging Graph. 2019;75:66–73. [PubMed: 31174100]

[174]. Tang YC, Huo YK, Xiong YX, Moon H, Assad A, Moyo TK, et al. Improving Splenomegaly Segmentation by Learning from Heterogeneous Multi-Source Labels. Medical Imaging 2019: Image Processing. 2019;10949.

[175]. Yun J, Park J, Yu D, Yi J, Lee M, Park HJ, et al. Improvement of fully automated airway segmentation on volumetric computed tomographic images using a 2.5 dimensional convolutional neural net. Med Image Anal. 2019;51:13–20. [PubMed: 30388500]

[176]. Zhu J, Zhang J, Qiu B, Liu Y, Liu X, Chen L. Comparison of the automatic segmentation of multiple organs at risk in CT images of lung cancer between deep convolutional neural network-based and atlas-based techniques. Acta Oncol. 2019;58:257–64. [PubMed: 30398090]

[177]. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention, Pt Iii. 2015;9351:234–41.

[178]. Milletari F, Navab N, Ahmadi S-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. Fourth International Conference on 3D Vision. 2016:565–71.

[179]. Schlemper J, Oktay O, Schaap M, Heinrich M, Kainz B, Glocker B, et al. Attention gated networks: Learning to leverage salient regions in medical images. Med Image Anal. 2019;53:197–207. [PubMed: 30802813]

[180]. Zhou XR, Takayama R, Wang S, Hara T, Fujita H. Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method. Medical Physics. 2017;44:5221–33. [PubMed: 28730602]

[181]. Binder T, Tantaoui E, Pati P, Catena R, Set-Aghayan A, Gabrani M. Multi-Organ Gland Segmentation Using Deep Learning. Frontiers in Medicine. 2019;6.

[182]. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. MICCAI2016.

[183]. Men K, Dai JR, Li YX. Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. Medical Physics. 2017;44:6377–89. [PubMed: 28963779]

[184]. Oda M, Shimizu N, Roth HR, Karasawa K, Kitasaka T, Misawa K, et al. 3D FCN Feature Driven Regression Forest-Based Pancreas Localization and Segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. 2017;10553:222–30.

[185]. Brosch T, Saalbach A. Foveal Fully Convolutional Nets for Multi-Organ Segmentation. Medical Imaging 2018: Image Processing. 2018;10574.

[186]. Chen L, Bentley P, Mori K, Misawa K, Fujiwara M, Rueckert D. DRINet for Medical Image Segmentation. IEEE Trans Med Imaging. 2018;37:2453–62. [PubMed: 29993738]

[187]. Ad Gelder, Huisman HJA. Autoencoders for Multi-Label Prostate MR Segmentation. ArXiv. 2018;abs/1806.08216.

[188]. Gibson E, Li WQ, Sudre C, Fidon L, Shakir DI, Wang GT, et al. NiftyNet: a deep-learning platform for medical imaging. Computer Methods and Programs in Biomedicine. 2018;158:113–22. [PubMed: 29544777]

[189]. Gonzalez G, Washko GR, Estepar RS. Multi-structure Segmentation from Partially Labeled Datasets. Application to Body Composition Measurements on CT Scans. Image Analysis for Moving Organ, Breast, and Thoracic Images. 2018;11040:215–24.

[190]. Kakeya H, Okada T, Oshiro Y. 3D U-JAPA-Net: Mixture of Convolutional Networks for Abdominal Multi-organ CT Segmentation. Medical Image Computing and Computer Assisted Intervention - Miccai 2018, Pt Iv. 2018;11073:426–33.

[191]. Roth HR, Shen C, Oda H, Sugino T, Oda M, Hayashi Y, et al. A Multi-scale Pyramid of 3D Fully Convolutional Networks for Abdominal Multi-organ Segmentation. Medical Image Computing and Computer Assisted Intervention - Miccai 2018, Pt Iv. 2018;11073:417–25.

[192]. Zhou SH, Nie D, Adeli E, Gao YZ, Wang L, Yin JP, et al. Fine-Grained Segmentation Using Hierarchical Dilated Neural Networks. Medical Image Computing and Computer Assisted Intervention - Miccai 2018, Pt Iv. 2018;11073:488–96.

[193]. Vandewinckele L, Willems S, Robben D, Veen JVD, Crijns W, Nuyts S, et al. Segmentation of head-and-neck organs-at-risk in longitudinal CT scans combining deformable registrations and convolutional neural networks. Computer methods in biomechanics and biomedical engineering Imaging & visualization. 2019:1–10.

[194]. Anthimopoulos M, Christodoulidis S, Ebner L, Geiser T, Christe A, Mougiakakou S. Semantic Segmentation of Pathological Lung Tissue With Dilated Fully Convolutional Networks. IEEE J Biomed Health Inform. 2019;23:714–22. [PubMed: 29993791]

[195]. Chen G, Zhang J, Zhuo D, Pan Y, Pang C. Identification of pulmonary nodules via CT images with hierarchical fully convolutional networks. Med Biol Eng Comput. 2019;57:1567–80. [PubMed: 31025248]

[196]. Gu X, Wang J, Zhao J, Li Q. Segmentation and suppression of pulmonary vessels in low-dose chest CT scans. Med Phys. 2019;46:3603–14. [PubMed: 31240721]

[197]. Li XL, Wang YY, Tang QS, Fan Z, Yu JH. Dual U-Net for the Segmentation of Overlapping Glioma Nuclei. IEEE Access. 2019;7:84040–52.

[198]. Nguyen NQ, Lee SW. Robust Boundary Segmentation in Medical Images Using a Consecutive Deep Encoder-Decoder Network. Ieee Access. 2019;7:33795–808.

[199]. Park B, Park H, Lee SM, Seo JB, Kim N. Lung Segmentation on HRCT and Volumetric CT for Diffuse Interstitial Lung Disease Using Deep Convolutional Neural Networks. J Digit Imaging. 2019;32:1019–26. [PubMed: 31396776]

[200]. Park J, Yun J, Kim N, Park B, Cho Y, Park HJ, et al. Fully Automated Lung Lobe Segmentation in Volumetric Chest CT with 3D U-Net: Validation with Intra- and Extra-Datasets. J Digit Imaging. 2019.

[201]. Xu XAN, Zhou FG, Liu B, Bai XZ. Multiple Organ Localization in CT Image Using Triple-Branch Fully Convolutional Networks. IEEE Access. 2019;7:98083–93.

[202]. van Rooij W, Dahele M, Brandao HR, Delaney AR, Slotman BJ, Verbakel WF. Deep Learning-Based Delineation of Head and Neck Organs at Risk: Geometric and Dosimetric Evaluation. International Journal of Radiation Oncology Biology Physics. 2019;104:677–84.

[203]. Heinrich MP, Oktay O, Bouteldja N. OBELISK-Net: Fewer layers to solve 3D multi-organ segmentation with sparse deformable convolutions. Med Image Anal. 2019;54:1–9. [PubMed: 30807894]

[204]. Wang Y, Zhou Y, Shen W, Park S, Fishman E, Yuille A. Abdominal multi-organ segmentation with organ-attention networks and statistical fusion. Medical image analysis. 2019;55:88–102. [PubMed: 31035060]

[205]. Chan J, Kearney V, Haaf S, Wu S, Bogdanov M, Reddick M, et al. A convolutional neural network algorithm for automatic segmentation of head and neck organs at risk using deep lifelong learning. Medical physics. 2019;46 5:2204–13. [PubMed: 30887523]

[206]. Zhou Y, Li Z, Bai S, Wang C, Chen X, Han M, et al. Prior-Aware Neural Network for Partially-Supervised Multi-Organ Segmentation. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). 2019:10671–80.

[207]. Kim H, Jung J, Kim J, Cho B, Kwak J, Jang JY, et al. Abdominal multi-organ auto-segmentation using 3D-patch-based deep convolutional neural network. Scientific Reports 2020;10.

[208]. Fu Y, Ippolito J, Ludwig DR, Nizamuddin R, Li H, Yang D. Automatic segmentation of CT images for ventral body composition analysis. Medical physics. 2020.

[209]. Peng Z, Fang X, Yan P, Shan H, Liu T, Pei X, et al. A Method of Rapid Quantification of Patient-Specific Organ Doses for CT Using Deep-Learning based Multi-Organ Segmentation and GPU-accelerated Monte Carlo Dose Computing. Medical physics. 2020.

[210]. Gou S-p, Tong N, Qi S, Yang S, Chin R, Sheng K. Self-channel-and-spatial-attention neural network for automated multi-organ segmentation on head and neck CT images. Physics in medicine and biology. 2020.

[211]. Girshick RB, Donahue J, Darrell T, Malik J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. IEEE Conference on Computer Vision Pattern Recognition. 2013:580–7.

[212]. Uijlings JRR, Sande Kvd, Gevers T, Smeulders AWM. Selective Search for Object Recognition. International Journal of Computer Vision. 2013;104:154–71.

[213]. Girshick RB. Fast R-CNN. IEEE International Conference on Computer Vision. 2015:1440–8.

[214]. Ren S, He K, Girshick RB, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis Machine Intelligence. 2015;39:1137–49.

[215]. He K, Gkioxari G, Dollár P, Girshick RB. Mask R-CNN. IEEE International Conference on Computer Vision. 2017:2980–8.

[216]. Christ PF, Elshaer MEA, Ettlinger F, Tatavarty S, Bickel M, Bilic P, et al. Automatic Liver and Lesion Segmentation in CT Using Cascaded Fully Convolutional Neural Networks and 3D Conditional Random Fields. MICCAI2016.

[217]. Wu L, Xin Y, Li S, Wang T, Heng P, Ni D. Cascaded Fully Convolutional Networks for automatic prenatal ultrasound image segmentation. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)2017. p. 663–6.

[218]. Trullo R, Petitjean C, Nie D, Shen DG, Ruan S. Joint Segmentation of Multiple Thoracic Organs in CT Images with Two Collaborative Deep Architectures. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. 2017;10553:21–9. [PubMed: 29707697]

[219]. Xu X, Zhou F, Liu B, Fu D, Bai X. Efficient Multiple Organ Localization in CT Image using 3D Region Proposal Network. IEEE Trans Med Imaging. 2019.

[220]. Xu Z, Wu Z, Feng J. CFUN: Combining Faster R-CNN and U-net Network for Efficient Whole Heart Segmentation. ArXiv. 2018;abs/1812.04914.

[221]. Bouget D, Jorgensen A, Kiss G, Leira HO, Lango T. Semantic segmentation and detection of mediastinal lymph nodes and anatomical structures in CT data for lung cancer staging. Int J Comput Assist Radiol Surg. 2019;14:977–86. [PubMed: 30891655]

[222]. Li Y, Zhang L, Chen H, Yang N. Lung Nodule Detection With Deep Learning in 3D Thoracic MR Images. IEEE Access. 2019;7:37822–32.

[223]. Huang X, Sun W, Tseng TB, Li C, Qian W. Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks. Comput Med Imaging Graph. 2019;74:25–36. [PubMed: 30954678]

[224]. Chen S, Roth H, Dorn S, May M, Cavallaro A, Lell M, et al. Towards Automatic Abdominal Multi-Organ Segmentation in Dual Energy CT using Cascaded 3D Fully Convolutional Network. ArXiv. 2017;abs/1710.05379.

[225]. Roth HR, Oda H, Zhou XR, Shimizu N, Yang Y, Hayashi Y, et al. An application of cascaded 3D fully convolutional networks for medical image segmentation. Computerized Medical Imaging and Graphics. 2018;66:90–9. [PubMed: 29573583]

[226]. Liu M, Dong J, Dong X, Yu H, Qi L, Technology. Segmentation of Lung Nodule in CT Images Based on Mask R-CNN. 9th International Conference on Awareness Science. 2018:1–6.

[227]. He K, Cao X, Shi Y, Nie D, Gao Y, Shen D. Pelvic Organ Segmentation Using Distinctive Curve Guided Fully Convolutional Networks. IEEE Transactions on Medical Imaging. 2019;38:585–95. [PubMed: 30176583]

[228]. Kopelowitz E, Engelhard G. Lung Nodules Detection and Segmentation Using 3D Mask-RCNN. ArXiv. 2019;abs/1907.07676.

[229]. Wessel J, Heinrich MP, Berg Jv, Franz A, Saalbach A. Sequential Rib Labeling and Segmentation in Chest X-Ray using Mask R-CNN. ArXiv. 2019.

[230]. Zhang R, Cheng C, Zhao X, Li X. Multiscale Mask R-CNN-Based Lung Tumor Detection Using PET Imaging. Mol Imaging. 2019;18:1536012119863531. [PubMed: 31364467]

[231]. Wang S, He K, Nie D, Zhou S, Gao Y, Shen D. CT male pelvic organ segmentation using fully convolutional networks with boundary sensitive representation. Medical Image Analysis. 2019;54:16X 178.

[232]. Zhang L, Zhang J, Shen P, Zhu G, Li P, Lu X, et al. Block Level Skip Connections Across Cascaded V-Net for Multi-Organ Segmentation. IEEE Transactions on Medical Imaging. 2020;39:2782–93. [PubMed: 32091995]

[233]. Liang S, Thung K-H, Nie D, Zhang Y, Shen D. Multi-View Spatial Aggregation Framework for Joint Localization and Segmentation of Organs at Risk in Head and Neck CT Images. IEEE Transactions on Medical Imaging. 2020;39:2794–805. [PubMed: 32091997]

[234]. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets. NIPS2014.

[235]. Yi X, Walia E, Babyn P. Generative Adversarial Network in Medical Imaging: A Review. Medical image analysis. 2019;58:101552. [PubMed: 31521965]

[236]. Ying X, Guo H, Ma K, Wu JY, Weng Z, Zheng Y. X2CT-GAN: Reconstructing CT from Biplanar X-Rays with Generative Adversarial Networks. CVPR2019.

[237]. Dai W, Dong N, Wang Z, Liang X, Zhang H, Xing EP. SCAN: Structure Correcting Adversarial Network for Organ Segmentation in Chest X-Rays. DLMIA/ML-CDS@MICCAI2017.

[238]. Zhang Q, Wang H, Lu H, Won D, Yoon SW. Medical Image Synthesis with Generative Adversarial Networks for Tissue Recognition. 2018 IEEE International Conference on Healthcare Informatics (ICHI)2018. p. 199–207.

[239]. Han C, Murao K, Satoh Si, Nakayama H. Learning More with Less: GAN-based Medical Image Augmentation. ArXiv. 2019;abs/1904.00838.

[240]. Yang D, Xu D, Zhou SK, Georgescu B, Chen M, Grbic S, et al. Automatic Liver Segmentation Using an Adversarial Image-to-Image Network MICCAI2017.

[241]. Mondal AK, Dolz J, Desrosiers C. Few-shot 3D Multi-modal Medical Image Segmentation using Generative Adversarial Learning. ArXiv. 2018;abs/1810.12241.

[242]. Kamnitsas K, Baumgartner C, Ledig C, Newcombe V, Simpson J, Kane A, et al. Unsupervised Domain Adaptation in Brain Lesion Segmentation with Adversarial Networks. In: Niethammer M, Styner M, Aylward S, Zhu H, Oguz I, Yap P-T, et al., editors. Information Processing in Medical Imaging. Cham: Springer International Publishing; 2017. p. 597–609.

[243]. Moeskops P, Veta M, Lafarge MW, Eppenhof KAJ, Pluim JPW. Adversarial Training and Dilated Convolutions for Brain MRI Segmentation. In: Cardoso MJ, Arbel T, Carneiro G, Syeda-Mahmood T, Tavares JMRS, Moradi M, et al., editors. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer International Publishing; 2017. p. 56–64.

[244]. Rezaei M, Harmuth K, Gierke W, Kellermeier T, Fischer M, Yang H, et al. A Conditional Adversarial Network for Semantic Segmentation of Brain Tumor. ArXiv. 2017;abs/1708.05227.

[245]. Son J, Park SJ, Jung K-H. Retinal Vessel Segmentation in Fundoscopic Images with Generative Adversarial Networks. ArXiv. 2017;abs/1706.09318.

[246]. Zhu W, Xiang X, Tran TD, Hager GD, Xie X. Adversarial deep structured nets for mass segmentation from mammograms. IEEE 15th International Symposium on Biomedical Imaging. 2017:847–50.

[247]. Li Z, Wang Y, Yu J. Brain Tumor Segmentation Using an Adversarial Network. In: Crimi A, Bakas S, Kuijf H, Menze B, Reyes M, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Cham: Springer International Publishing; 2018. p. 123–32.

[248]. Rezaei M, Yang H, Meinel C. Whole Heart and Great Vessel Segmentation with Context-aware of Generative Adversarial Networks. In: Maier A, Deserno TM, Handels H, Maier-Hein KH, Palm C, Tolxdorff T, editors. Bildverarbeitung für die Medizin 2018. Berlin, Heidelberg: Springer Berlin Heidelberg; 2018. p. 353–8.

[249]. Rezaei M, Yang H, Meinel C. Conditional Generative Refinement Adversarial Networks for Unbalanced Medical Image Semantic Segmentation. ArXiv. 2018;abs/1810.03871.

[250]. Xue Y, Xu T, Zhang H, Long LR, Huang X. SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation. Neuroinformatics. 2018;16:383–92. [PubMed: 29725916]

[251]. Zhang L, Pereañez M, Piechnik SK, Neubauer S, Petersen SE, Frangi AF. Multi-Input and Dataset-Invariant Adversarial Learning (MDAL) for Left and Right-Ventricular Coverage Estimation in Cardiac MRI. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Cham: Springer International Publishing; 2018. p. 481–9.

[252]. Zhang Y, Miao S, Mansi T, Liao R. Task Driven Generative Modeling for Unsupervised Domain Adaptation: Application to X-ray Image Segmentation. Medical Image Computing and Computer Assisted Intervention - Miccai 2018, Pt Ii. 2018;11071:599–607.

[253]. Mahmood F, Borders D, Chen R, McKay GN, Salimian KJ, Baras A, et al. Deep Adversarial Training for Multi-Organ Nuclei Segmentation in Histopathology Images. IEEE Trans Med Imaging. 2019.

[254]. Trullo R, Petitjean C, Dubray B, Ruan S. Multiorgan segmentation using distance-aware adversarial networks. Journal of Medical Imaging. 2019;6.

[255]. Tong N, Gou S, Yang S, Cao M, Sheng K. Shape constrained fully convolutional DenseNet with adversarial training for multiorgan segmentation on head and neck CT and low-field MR images. Medical physics. 2019;46 6:2669–82. [PubMed: 31002188]

[256]. Cai J, Xia Y, Yang D, Xu D, Yang L, Roth H. End-to-End Adversarial Shape Learning for Abdomen Organ Deep Segmentation. MLMI@MICCAI2019.

[257]. Bnouni N, Rekik I, Rhim MS, Amara NB. Context-Aware Synergetic Multiplex Network for Multi-organ Segmentation of Cervical Cancer MRI. PRIME@MICCAI2020.

[258]. Arjovsky M, Chintala S, Bottou L. Wasserstein Generative Adversarial Networks. ICML2017.

[259]. Dai X, Lei Y, Wang T, Dhabaan A, McDonald M, Beitler J, et al. Synthetic MRI-aided Head-and-Neck Organs-at-Risk Auto-Delineation for CBCT-guided Adaptive Radiotherapy. arXiv: Medical Physics. 2020.

[260]. Lei Y, Wang T, Harms J, Shafai-Erfani G, Tian S, Higgins K, et al. MRI-based pseudo CT generation using classification and regression random forest. SPIE Medical Imaging2019.

[261]. Lei Y, Harms J, Wang T, Tian S, Zhou J, Shu HK, et al. MRI-based synthetic CT generation using semantic random forest with iterative refinement. Phys Med Biol. 2019;64:085001. [PubMed: 30818292]

[262]. Lei Y, Jeong JJ, Wang T, Shu HK, Patel P, Tian S, et al. MRI-based pseudo CT synthesis using anatomical signature and alternating random forest with iterative refinement model. J Med Imaging. 2018;5:043504.

[263]. Yang J, Veeraraghavan H, Armato SG 3rd, Farahani K, Kirby JS, Kalpathy-Kramer J, et al. Autosegmentation for thoracic radiation treatment planning: A grand challenge at AAPM 2017. Med Phys. 2018;45:4568–81. [PubMed: 30144101]

[264]. Raudaschl PF, Zaffino P, Sharp GC, Spadea MF, Chen A, Dawant BM, et al. Evaluation of segmentation methods on head and neck CT: Auto-segmentation challenge 2015. Med Phys. 2017;44:2020–36. [PubMed: 28273355]

[265]. Feng X, Qing K, Tustison NJ, Meyer CH, Chen Q. Deep convolutional neural network for segmentation of thoracic organs-at-risk using cropped 3D images. Medical Physics. 2019;46:2169–80. [PubMed: 30830685]

[266]. Wang Y, Zhao L, Wang M, Song Z. Organ at Risk Segmentation in Head and Neck CT Images Using a Two-Stage Segmentation Framework Based on 3D U-Net. IEEE Access. 2018;7:144591–602.

[267]. Zhu W, Huang Y, Zeng L, Chen X, Liu Y, Qian Z, et al. AnatomyNet: Deep learning for fast and fully automated whole - volume segmentation of head and neck anatomy. Medical Physics. 2019;46:576–89. [PubMed: 30480818]

[268]. Tang H, Chen X, Liu Y, Lu Z, You J, Yang M, et al. Clinically applicable deep learning framework for organs at risk delineation in CT images. 2019.

[269]. Ren X, Xiang L, Nie D, Shao Y, Zhang H, Shen D, et al. Interleaved 3D-CNNs for joint segmentation of small-volume structures in head and neck CT images. Med Phys. 2018;45:2063–75. [PubMed: 29480928]

[270]. Lei Y, Tang X, Higgins K, Lin J, Jeong J, Liu T, et al. Learning-based CBCT correction using alternating random forest based on auto-context model. Med Phys. 2019;46:601–18. [PubMed: 30471129]

[271]. Yang X, Wang T, Lei Y, Higgins K, Liu T, Shim H, et al. MRI-based attenuation correction for brain PET/MRI based on anatomic signature and machine learning. Phys Med Biol. 2019;64:025001. [PubMed: 30524027]

[272]. Zhou X-Y, Yang G-Z. Normalization in Training U-Net for 2-D Biomedical Semantic Segmentation. IEEE Robotics Automation Letters. 2018;4:1792–9.

[273]. Taghanaki SA, Zheng Y, Zhou S, Georgescu B, Sharma P, Xu D, et al. Combo Loss: Handling Input and Output Imbalance in Multi-Organ Segmentation. Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society. 2019;75:24–33. [PubMed: 31129477]

[274]. Johnson J, Khoshgoftaar T. Survey on deep learning with class imbalance. Journal of Big Data. 2019;6:1–54.

1. Comprehensive review of deep learning-based multi-organ segmentation
2. Categorization of pixel-wise classification and end-to-end segmentation
3. Pixel-wise classification includes AE and CNN
4. End-to-end segmentation includes FCN, R-FCN, GAN and synthetic image-aided
5. Benchmark of algorithms' performances for thoracic and head-neck CT segmentation

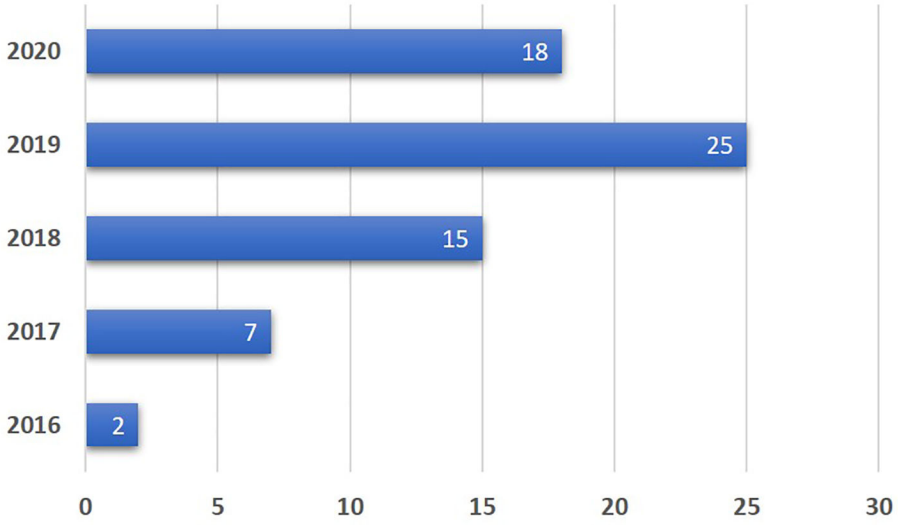

Fig. 1.
The number of publications for DL-based multi-organ segmentation (till October 2020).

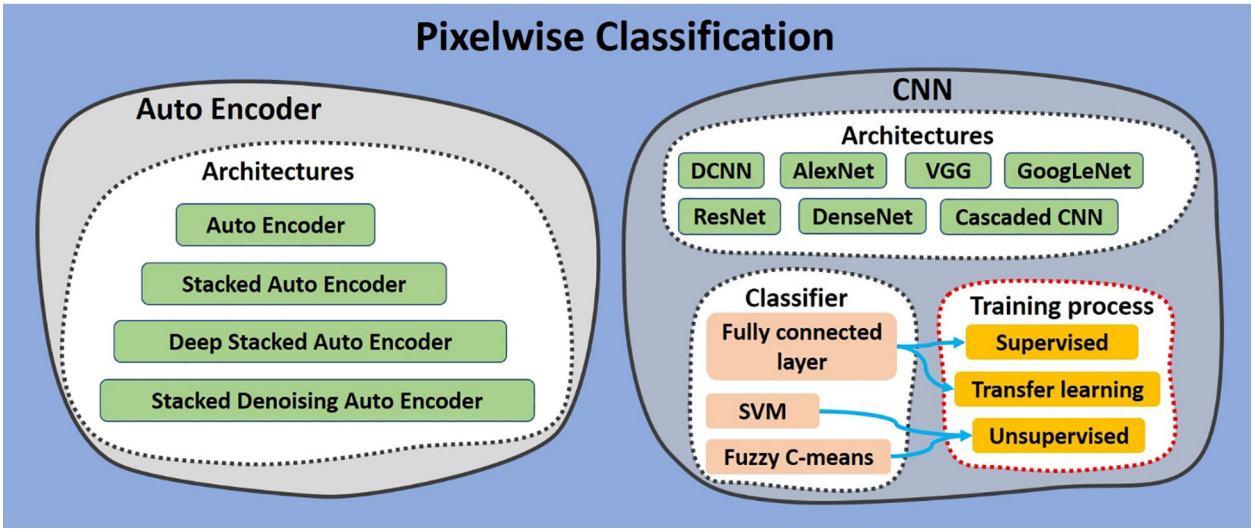

Fig. 2.
The network components of the pixelwise classification methods.

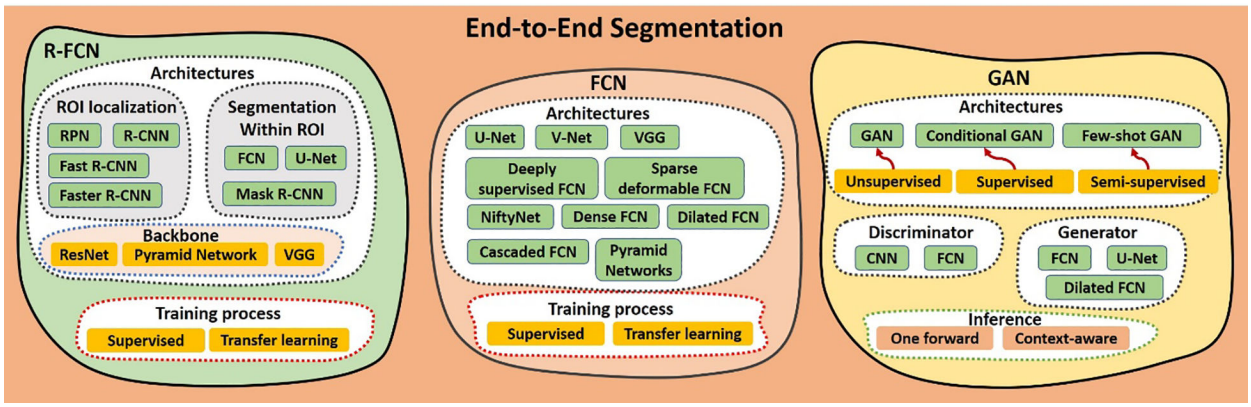

Fig. 3.

The network components of the end-to-end segmentation methods.

Table 1

Overview of AE methods

Ref.	Year	Network	Supervision	Dimension	Site	Modality
[138]	2013	SAE	Weakly supervised	3D patch	Abdomen	4D DCE-MRI
[140]	2015	SDAE	Supervised	3D patch	Brain Gliomas	MRI
[139]	2017	SDAE	Semi-supervised	2D patch	Brain lesion	MRI
[141]	2017	SAE	Transfer learning	2D slice	Liver	CT
[143]	2018	CSDAE	Transfer learning	2D slice	Thoracic	chest X-rays
[142]	2019	SSAE	Unsupervised	2D patch	Vertebrae	CT
[144]	2019	SSAE	Unsupervised	2D patch	Vertebrae	CT
[132]	2019	Hierarchical 3D AE	Supervised	N.A.	Head & Neck	CT

Table 2

Overview of CNN methods

Ref.	Year	Network	Dimension	Site	Modality
[124]	2017	Deep deconvolutional neural network (DDNN)	2D slice	Brain	CT
[161]	2017	3D CNN	3D patch	Brain lesion	MRI
[154]	2015	Multi-level DCNN	2D patch	Pancreas	CT
[162]	2016	Holistically Nested CNN	2D patch	Pancreas	CT
[155]	2017	3D CNN	3D patch	Chest	CT
[134]	2017	3D DCNN	Not specified	Abdomen	CT
[163]	2017	CNN	3D patch	Head & Neck	CT
[164]	2017	Fuzzy-C-Means CNN	3D patch	Lung nodule	CT
[165]	2017	DCNN	2D Slice	Body, Chest, Abdomen	CT
[166]	2018	Fusion Net	2D patch	100 ROIs	HRCT
[167]	2018	DCNN	2D patch	Spinal lesion	CT
[168]	2018	DCNN	2D slice	Malignant pleural mesothelioma	CT
[169]	2018	2D and 3D CNN	2D slice, 3D volume	Artery / vein	CT
[170]	2018	3D ConvNets	3D volume	Brain	MRI
[171]	2018	CNN with specific fine-tuning	2D slice, 3D volume	Brain, abdomen	Fetal MRI
[172]	2018	2D and 3D DCNN	2D slice, 3D volume	Whole body	CT
[173]	2019	Deep fusion Network	2D slice	Chest	CXR
[174]	2019	DCNN	2D slice	Abdomen	CT
[175]	2019	2.5D CNN	2.5D patch	Thorax	CT
[157]	2019	Cascaded CNN	2D slice	Head & Neck	CT
[158]	2019	2D and 3D CNN	2D slice, 3D volume	Thorax	CT
[176]	2019	U-Net Neural Network	3D patch	Lung	CT

Table 3

Overview of FCN methods

Ref.	Year	Network	Dimension	Site	Modality
[177]	2015	U-Net	2D slice	Neuronal structure	Electron microscopic
[182]	2016	3D U-Net	3D volume	Kidney	Microscopic
[183]	2017	Dilated FCN	2D slice	Abdomen	CT
[184]	2017	3D FCN Feature Driven Regression Forest	3D patch	Pancreas	CT
[180]	2017	2D FCN	2.5D slices	Whole body	CT
[185]	2018	Foveal Fully Convolutional Nets	N.A.*	Whole body	CT
[186]	2018	DRINet	2D slice	Brain, abdomen	CT
[187]	2018	3D U-Net	3D volume	Prostate	MRI
[135]	2018	Dense V-Net	3D volume	Abdomen	CT
[188]	2018	NiftyNet	3D volume	Abdomen	CT
[189]	2018	PU-Net, CU-Net	2D slice	Pelvis	CT
[126]	2018	Dilated U-Net	2D slice	Chest	CT
[190]	2018	3D U-JAPA-Net	3D volume	Abdomen	CT
[125]	2018	U-Net	2D slice	Pelvis	CT
[191]	2018	Multi-scale Pyramid of 3D FCN	3D patch	Abdomen	CT
[123]	2018	Shape representation model constrained FCN	3D volume	Head & Neck	CT
[192]	2018	Hierarchical Dilated Neural Networks	2D slice	Pelvis	CT
[129]	2018	Dense 3D FCN	3D volume	Abdomen	MRI
[193]	2018	3D FCN	3D patch	Head & Neck	CT
[194]	2019	Dilated FCN	2D slice	Lung	CT
[181]	2019	Dense-U-Net	2D slice	Head & Neck	Stained colon adenocarcinoma dataset
[195]	2019	2D and 3D FCNs	2D slice and 3D volume	Pulmonary nodule	CT
[136]	2019	Dedicated 3D FCN	3D patch	Thorax/abdomen	DECT
[127]	2019	2D FCN (DeepLabV3+)	2D slice	Pelvis	MRI
[196]	2019	2D FCN	2D patch	Pulmonary vessels	CT
[197]	2019	Dual U-Net	2D slice	Glioma Nuclei	Hematoxylin and eosin (H&E)-stained histopathological image
[198]	2019	Consecutive deep encoder-decoder Network	2D slice	Skin lesion	CT
[199]	2019	U-Net	2D slice	Lung	HRCT
[200]	2019	3D U-Net	3D volume	Chest	CT
[42]	2019	3D U-Net with Multi-atlas	3D volume	Brain tumor	Dual-energy CT
[201]	2019	Triple-Branch FCN	Not specified	Abdomen/torso	CT
[43]	2019	2.5D Deeply supervised V-Net	2.5 patch	Prostate	Ultrasound
[46]	2019	Group dilated deeply supervised FCN	3D volume	Prostate	MRI
[44]	2019	3D FCN	3D volume	Arteriovenous malformations	Contract-enhanced CT
[53]	2019	3D FCN	3D volume	Left ventricle	SPECT

Ref.	Year	Network	Dimension	Site	Modality
[54]	2019	DeepMAD	2.5D patch	Vessel wall	MRI
[202]	2019	3D U-Net	3D volume	Head & Neck	CT
[203]	2019	OBELISK-Net	3D volume	Abdomen	CT
[204]	2019	OAN-RC	2D slice	Abdomen	CT
[205]	2019	Multi-stage 3D FCN	3D volume	Head & Neck	CT
[206]	2019	2D/3D FCN	3D patch	Abdomen	CT
[207]	2020	U-Net	3D patch	Abdomen	CT
[208]	2020	2.5D U-Net	2.5D patch	Body	CT
[130]	2020	3D Attention U-Net	3D patch	Pancreas/Abdomen	CT
[209]	2020	3D U-Net	3D patch	Thoracic/Abdomen	CT
[210]	2020	3D U-Net	3D volume	Head & Neck	CT

Table 4

Overview of Region-based FCN methods

Ref.	Year	Network	Dimension	Site	Modality
[216]	2016	Cascaded FCN	3D volume	Liver and lesion	CT
[224]	2017	3D Cascaded U-Net	3D volume	Abdomen	DECT
[225]	2018	Cascade 3D FCN	3D patch	Abdomen	CT
[226]	2018	Mask R-CNN	2D slice	Lung nodule	CT
[227]	2019	3D FCN	3D patch	Pelvic organs	CT
[220]	2018	Combination of Faster R-CNN and U-Net (CFUN)	3D volume	Cardiac	CT
[221]	2019	Combination of U-Net and Mask R-CNN	2D slice	Chest	CT
[223]	2019	Faster R-CNN	2D slice	Thorax/pulmonary nodule	CT
[228]	2019	3D Mask R-CNN	3D volume	Lung nodule	CT
[222]	2019	3D Faster R-CNN	3D volume	Thorax/ lung nodule	MRI
[229]	2019	Mask R-CNN	N.A.*	Chest	X-Ray
[219]	2019	3D RPN	3D volume	Whole body	CT
[230]	2019	Multiscale Mask R-CNN	2D slice	Lung tumor	PET
[231]	2019	2.5D U-Net	3D patch	Pelvic organs	CT
[232]	2020	3D Dense V-Net	3D volume	Thorax/Abdomen	CT
[233]	2020	2.5D CNN	2.5D	Head & Neck	CT

*N.A.: not available, i.e. not explicitly indicated in the publication

Table 5

Overview of GAN methods

Ref.	Year	Network	Dimension	Site	Modality
[237]	2015	SCAN	2D slice	Chest	X-rays
[242]	2017	Multi-connected adversarial networks	2D slice	Brain	Multi-modality MRI
[243]	2017	Dilated GAN	2D slice	Brain	MRI
[244]	2017	Conditional GAN	2D slice	Brain tumor	MRI
[245]	2017	GAN	2D patch	Retinal Vessel	Fundoscopic
[240]	2017	Adversarial Image-to-Image Network	3D volume	Liver	CT
[246]	2017	Adversarial FCN-CRF Nets	2D slice	Mass	Mammograms
[247]	2018	GAN	Not specified	Brain tumor	MRI
[241]	2018	Few-shot GAN	3D patch	Brain	MRI
[248]	2018	Context-aware GAN	2D cropped slices	Cardiac	MRI
[249]	2018	Conditional Generative Refinement Adversarial Networks	2D slice	Brain	MRI
[250]	2018	SegAN	2D slice	Brain	MRI
[251]	2018	MDAL	2D slice	Left and Right-Ventricular	Cardiac MRI
[252]	2018	TD-GAN	2D slice	Whole body	X-ray
[45]	2019	U-Net-GAN	3D volume	Thorax	CT
[253]	2019	Conditional GAN	2D slice	Nuclei	Histopathology Images
[254]	2019	Distance-aware GAN	2D slice	Chest	CT
[255]	2019	Shape Constraint GAN	3D volume	Head & Neck	CT/MRI
[256]	2019	Shape Constraint GAN	3D volume	Abdomen	CT
[257]	2020	CycleGAN	2D slice	Pelvic organs	MRI

Table 6

Overview of synthetic image-aided image segmentation

Ref.	Year	Network	Dimension	Site	Modality
[39, 40]	2019	Synthetic MRI-aided	2.5D patch	Pelvic	CT
[38, 131]	2019	Synthetic MRI-aided	3D volume	Pelvic	CBCT
[133, 259]	2020	Synthetic MRI-aided	3D volume	Head-and-Neck	CT/CBCT

Table 7

DL-based methods using the 2017 AAPM Thoracic Auto-segmentation Challenge datasets.

Metric	Method	Esophagus	Heart	Left Lung	Right Lung	Spinal Cord
DSC	DCNN Team Elekta *	0.72±0.10	**0.93±0.02**	0.97±0.02	0.97±0.02	0.88±0.037
	3D U-Net [265]	0.72±0.10	**0.93±0.02**	0.97±0.02	0.97±0.02	0.89±0.04
	Multi-class CNN Team Mirada *	0.71±0.12	0.91±0.02	**0.98±0.02**	0.97±0.02	0.87±0.110
	2D ResNet Team Beaumont *	0.61±0.11	0.92±0.02	0.96±0.03	0.95±0.05	0.85±0.035
	3D and 2D U-Net Team WUSTL *	0.55±0.20	0.85±0.04	0.95±0.03	0.96±0.02	0.83±0.080
	U-Net-GAN [45]	**0.75±0.08**	0.87±0.05	0.97±0.01	**0.97±0.01**	**0.90±0.04**
MSD (mm)	DCNN Team Elekta *	2.23±2.82	2.05±0.62	0.74±0.31	1.08±0.54	0.73±0.21
	3D U-Net [265]	2.34±2.38	2.30±0.49	**0.59±0.29**	0.93±0.57	0.66±0.25
	Multi-class CNN Team Mirada *	2.08±1.94	2.98±0.93	0.62±0.35	0.91±0.52	0.76±0.60
	2D ResNet Team Beaumont *	2.48±1.15	2.61±0.69	2.90±6.94	2.70±4.84	1.03±0.84
	3D and 2D U-Net Team WUSTL *	13.10±10.39	4.55±1.59	1.22±0.61	1.13±0.49	2.10±2.49
	U-Net-GAN [45]	**1.05±0.66**	**1.49±0.85**	0.61±0.73	**0.65±0.53**	**0.38±0.27**
HD95 (mm)	DCNN Team Elekta *	7.3+10.31	5.8±1.98	2.9±1.32	4.7±2.50	2.0±0.37
	3D U-Net [265]	8.71+10.59	6.57±1.50	2.10±0.94	3.96±2.85	1.89±0.63
	Multi-class CNN Team Mirada *	7.8±8.17	9.0±4.29	2.3±1.30	3.7±2.08	2.0±1.15
	2D ResNet Team Beaumont *	8.0±3.80	8.8±5.31	7.8±19.13	14.5±34.4	2.3±0.50
	3D and 2D U-Net Team WUSTL *	37.0±26.88	13.8±5.49	4.4±3.41	4.1±2.11	8.10±10.72
	U-Net-GAN [45]	**4.52±3.81**	**4.58±3.67**	**2.07±1.93**	**2.50±3.34**	**1.19±0.46**

* Note: Participating methods of the AAPM thorax challenge [263].

Table 8

DL-based methods using the 2015 MICCAI Head and Neck Auto-segmentation Challenge datasets.

Metric	Organs	Shape model constrained FCN [264]	Two-stage U-Net [266]	AnatomyNet [267]	DL-based [268]	Synthetic MRI-aided [133]	3D U-Net [182]	3D-CNN [269]
DSC	**Brain Stem**	0.87 ± 0.03	0.88 ± 0.02	0.87 ± 0.02	0.87 ± 0.03	**0.91 ± 0.02**	0.80 ± 0.08	N.A.
	Chiasm	0.58 ± 0.1	0.45 ± 0.17	0.53 ± 0.15	0.62 ± 0.1	**0.73 ± 0.11**	N.A.	0.58 ± 0.17
	Mandible	0.87 ± 0.03	0.93 ± 0.02	0.93 ± 0.02	0.95 ± 0.01	**0.96 ± 0.01**	0.94 ± 0.02	N.A.
	Left Optic Nerve	0.65 ± 0.05	0.74 ± 0.15	0.72 ± 0.06	0.75 ± 0.07	**0.78 ± 0.09**	0.72 ± 0.06	0.72 ± 0.08
	Right Optic Nerve	0.69 ± 0.5	0.74 ± 0.09	0.71 ± 0.1	0.72 ± 0.06	**0.78 ± 0.11**	0.70 ± 0.07	0.70 ± 0.09
	Left Parotid	0.84 ± 0.02	0.86 ± 0.02	0.88 ± 0.02	**0.89 ± 0.02**	0.88 ± 0.04	0.87 ± 0.03	N.A.
	Right Parotid	0.83 ± 0.02	0.85 ± 0.07	0.87 ± 0.04	**0.88 ± 0.05**	0.88 ± 0.06	0.85 ± 0.07	N.A.
	Left Submandibular	0.76 ± 0.06	0.76 ± 0.15	0.81 ± 0.04	0.82 ± 0.05	**0.86 ± 0.08**	0.76 ± 0.09	N.A.
	Right Submandibular	0.81 ± 0.06	0.73 ± 0.01	0.81 ± 0.04	0.82 ± 0.05	**0.85 ± 0.10**	0.78 ± 0.07	N.A.
HD95 (mm)	**Brain Stem**	4.01 ± 0.93	**2.01 ± 0.33**	N.A.	N.A.	N.A.	N.A.	N.A.
	Chiasm	**2.17 ± 1.04**	2.83 ± 1.42	N.A.	N.A.	N.A.	N.A.	2.81 ± 1.56
	Mandible	1.50 ± 0.32	**1.26 ± 0.50**	N.A.	N.A.	N.A.	N.A.	N.A.
	Left Optic Nerve	2.52 ± 1.04	2.53 ± 2.34	N.A.	N.A.	N.A.	N.A.	**2.33 ± 0.84**
	Right Optic Nerve	2.90 ± 1.88	2.13 ± 2.45	N.A.	N.A.	N.A.	N.A.	**2.13 ± 0.96**
	Left Parotid	3.97 ± 2.15	**2.41 ± 0.54**	N.A.	N.A.	N.A.	N.A.	N.A.
	Right Parotid	4.20 ± 1.27	**2.93 ± 1.48**	N.A.	N.A.	N.A.	N.A.	N.A.
	Left Submandibular	5.59 ± 3.93	**2.86 ± 1.60**	N.A.	N.A.	N.A.	N.A.	N.A.
	Right Submandibular	4.84 ± 1.67	**3.44 ± 1.55**	N.A.	N.A.	N.A.	N.A.	N.A.