



A Systematic Review of the Techniques for the Automatic Segmentation of Organs-at-Risk in Thoracic Computed Tomography Images

Malvika Ashok¹ · Abhishek Gupta¹

Received: 17 February 2020 / Accepted: 13 September 2020 / Published online: 23 September 2020
© CIMNE, Barcelona, Spain 2020

Abstract

The standard treatment for the cancer is the radiotherapy where the organs nearby the target volumes get affected during treatment called the Organs-at-risk. Segmentation of Organs-at-risk is crucial but important for the proper planning of radiotherapy treatment. Manual segmentation is time consuming and tedious in regular practices and results may vary from experts to experts. The automatic segmentation will produce robust results with precise accuracy. The aim of this systematic review is to study various techniques for the automatic segmentation of organs-at-risk in thoracic computed tomography images and to discuss the best technique which give the higher accuracy in terms of segmentation among all other techniques proposed in the literature. PRISMA guidelines had been used to conduct this systematic review. Three online databases had been used for the identification of the related papers and a query had been formed for the search purpose. The papers were shortlisted based on the various inclusion and exclusion criteria. Four research questions had been designed and answers of those were explored. After reviewing all the techniques, the best technique had been selected and discussed in detail which gave the precise accuracy based on Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD). Both DSC and HD were used in the literature to evaluate the performance of their proposed technique for the automatic segmentation of four organs (esophagus, heart, trachea and aorta). However, the value of these parameters vary as per the validation sample size. Consequently, various challenges faced by the researchers had been listed. This paper includes the summary of the various automatic segmentation techniques for the Organs-at-risk in thoracic computed tomography images in terms of four research questions. Different techniques, Datasets, Performance accuracy and various challenges had been discussed.

1 Introduction

Organ-at-risk (OARs) mostly refer to healthy organs which are located near the target volume and may get affected by the radiation exposure during radiation therapy treatment. (OARs) [1–3] are the healthy organs that may be harmed during radiotherapy treatment. The radiotherapy treatment is given to the tumorous organ called target volumes [4]. The primary aim of radiotherapy [5] is to deliver a prescribed dose [6] to the target volume while protecting normal tissue

including OARs [7]. The radiation dose given to the target volume kills the cells of the tumour so that it cannot grow further. The high radiation dosages (> 50 Gy) are utilized for the treatment of malignant tissues in radiotherapy (RT). However, in the case of benign tissues, low to intermediate dose (3–50 Gy) is used to control the growth of tissues effectively [8]. The risk of damage relies upon the size, number and recurrence of radiation divisions, irradiated tissue volume, the span of treatment, and strategy for radiation conveyance [9]. Irradiation can cause pathological changes in the ordinary tissue of OAR with irreversible functional outcomes. Therefore, the amount and plan of radiation dose [10, 11] along with precise delivery must be weighted in the treatment planning [12]. It is fundamental to incorporate the resistance dose of the organs-at-risk in the treatment technique [13]. Organs with a low resistance to radiations must be delineated regardless of whether they are not situated in the prompt region of the target volume.

✉ Abhishek Gupta
abhishekgupta10@yahoo.co.in
Malvika Ashok
malvika.ashok1@gmail.com

¹ School of Computer Science and Engineering, Shri Mata Vaishno Devi University, Kakryal, Katra, Jammu and Kashmir 182320, India

Lung, breast or esophagus cancer [14–21] affects the Organs-at-Risk in the thorax region most frequently [22, 23]. Table 1 is showing the different cancerous organs in human body with affected Organs at risk during radiotherapy treatment. It is vital to detect the organs which are at a risk in the human body. Computed Tomography (CT) scan images are used to detect the malignant organs due to the uncontrolled growth of abnormal tissue, which affects the normal working of the human body. Segmentation of Organs-at-risk is crucial but important for the proper planning of radiotherapy treatment [24–26]. In routine clinical practices, segmentation is performed by the experts manually for the planning of radiotherapy treatment. In doing so, reliability, reproducibility, and repeatability of the manual segmentation results depend on the process of learning or the experience of the operator. Also, this manual process is likewise time-consuming and tedious in regular practices [27]. Inter and intra-observer variability is a common issue encountered in manual segmentation [28].

Automatic segmentation techniques provide efficient and precise results that can be used [29]. It will help the expert to exploit the anatomy in lesser time with more accuracy. The experts can effectively analyze the amount of dose needed to be given to the target volumes during radiotherapy treatment so that the nearby organs should be less affected by the radiations [30–33].

The automatic segmentation will produce more accurate and robust results. But achieving precise accuracy in automatic segmentation of organs through CT images is a quite challenging task [34] because of a few factors such as soft-tissue organ images having low contrast, variable size of organs from patient to patient, the likeness between the shapes of organs and overfitting [35] towards organs with high intensity or better-structured organs [36]. So, considering all the factors, various techniques have been reported in the literature [37–39] that automatically segment the

organs-at-risk in thoracic images. However, these techniques were implemented on different datasets of different challenges of the same organs. There are several variabilities in all the experiments conducted by the researchers which are based on the techniques, datasets, performance parameters, results, and challenges. The objective of the systematic review was to analyze the different techniques along with their datasets, performance parameters, results and challenges for the automatic segmentation of Organ-at-Risk in thoracic CT images. Accordingly, recommendations are provided for potential researchers in the field for the further development of the improved techniques.

2 Materials and Methods

This systematic review was conducted under the guidelines of PRISMA [40]. For the identification of concerned papers, a detailed search has been carried out on three different electronic databases i.e. Google scholar, PubMed, and Jurn during 23–25 August 2019. Google Scholar is an openly available web index that lists the full content or metadata of academic writing over a variety of distributed configurations and controls. PubMed is a free internet searcher getting to essentially the MEDLINE database of references and digests on life sciences and biomedical points. Jurn is a free online quest device for the finding and downloading of free full-message insightful works. It was set up in an open online open beta form in February 2009, at first for discovering open-access electronic diary articles in expressions of the human experience and humanities. A hand search of the related papers is also included in this study.

Search was carried out by using the query “(((Segmentation) or (automatic segmentation)) AND ((Organ-at-risk) or (thoracic organ-at-risk) or (heart organ-at-risk) or (esophagus organ-at-risk) or (trachea organ-at-risk) or (aorta organ-at-risk) or (multiple organ-at-risk)) AND ((CT)))”. Papers related to the automatic segmentation of Organs-at-risk in thoracic CT images only were included. The total number of searched papers was accepted and rejected based on the various inclusion and exclusion principles as given in Table 2. The detailed screening of the collected papers is carried out according to the PRISMA flowchart given in Fig. 1 at different stages. A total of 370 papers was searched in three search engines including hand search also.

At the first stage of screening, all the duplicate papers and papers published before 2011 were eliminated and 263 papers were selected for the second stage of screening. In the second stage, 227 papers were eliminated based on titles and abstract reading. 36 papers were left out after the second screening stage. In the eligibility stage further 12 papers were eliminated based on full-text reading and finding various reasons for the exclusion of those reasons as mentioned

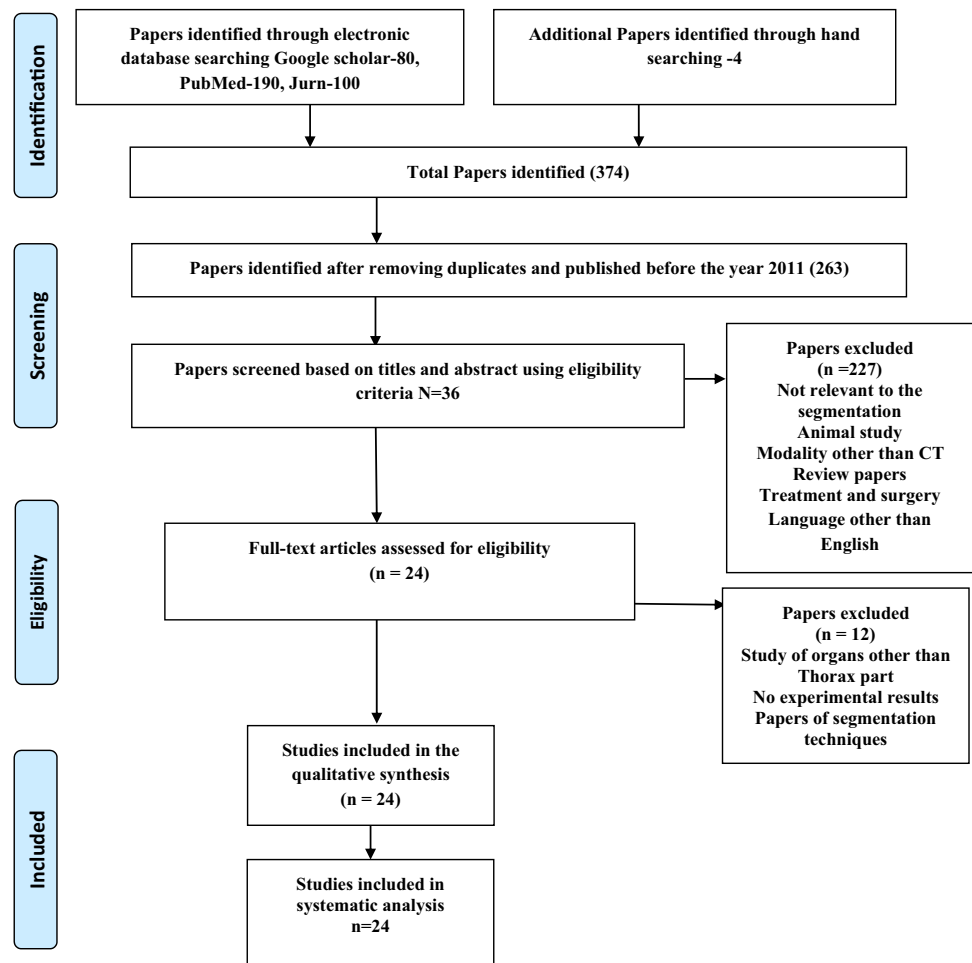
Table 1 Cancerous organs affecting organs-at-risk in different parts of the human body during radiotherapy treatment

S. no	Cancerous organs	Organ-at-risk
1.	Lung cancer	Heart, trachea, aorta, esophagus
2.	Breast cancer	Lungs, heart, trachea, aorta, esophagus, liver
3.	Stomach cancer	Livers, kidneys, spleen, prostate, rectum
4.	Brain cancer	The brain stem, eyes, mandible, temporomandibular joints, parotid glands, spinal cord, and skin
5.	Cervical cancer	Lungs, liver, bladder, vagina, and rectum
6.	Gallbladder cancer (uncommon)	Eyes, tongue and skin
7.	Mouth cancer	Jaws, throat, and ears
8.	Ovarian cancer	Intestines, liver, and stomach

Table 2 Inclusion and exclusion criterion used for shortlisting articles

S. no.	Parameters	Inclusion criteria	Exclusion criteria
1.	Time period	Papers published from the year 2011–2019	Papers published before 2011
2.	Interventions/ investigation	Studies including the segmentation of the human body	Studies including the segmentation of the animal’s body
3.	Comparator	Studies concentrating on the segmentation of the thorax part (heart, throat, trachea, aorta) of the human body	Studies including the segmentation of other than thorax part
4.	Modality type	Studies including CT modality segmentation	Studies including other than CT segmentation
5.	Study design	Studies involving experimental results	Studies including treatment and surgery of different organs Reviews, case study, Patent Language other than English

Fig. 1 Flowchart of the search process along with inclusion and exclusion criteria



in the PRISMA flowchart in Fig. 1. In the inclusion and final stage, 24 papers were finally selected for this systematic review considering all the inclusion criteria.

2.1 Quality Assessment

All the 24 papers were selected considering the various quality parameters. The relevancy to the topic was evaluated based on the inclusion and exclusion criteria.

Each shortlisted paper comprises a technique for the automatic segmentation of Organ-at-risk using CT images. All papers included in the review have an empirical study and their results are shown in the next sections in tabular form based on different parameters as per the description of designed research questions.

2.2 Literature Review

Various techniques have been developed using different frameworks (Deep learning-based [41–44], Model-based [45–47], Atlas-based [48–51]), etc. to segment the OARs efficiently. This systematic review aims to find out the answers to the following research questions.

RQ1 What are the various automatic segmentation techniques used for organ-at-risk in thoracic CT images?

RQ2 Which automatic technique gives the precise accuracy or improved performance in organ-at-risk segmentation?

RQ3 What are the minimum and maximum size of the dataset used for the organ-at-risk segmentation?

RQ4 What are the various challenges faced by the authors in organ-at-risk segmentation?

Tables 3 and 4 demonstrated the analysis of all the short-listed 24 papers accepted for the review. Table 4 contains 12 articles that were based on the SegTHOR competition plus 1 paper of the organizing members while Table 3 contains the other articles.

The authors had proposed different methods to segment the different organs in the thorax part as shown in Table 3. In every paper at least one of the four organs (heart, trachea, aorta, and esophagus) was segmented automatically using a different technique. The esophagus organ in a CT image is the most difficult to segment due to its low contrast [63] compared to other organs in the image. Kurugol et al. [52] proposed a model-based 3D level set technique for esophagus segmentation where PCT (Principle Curve Tracing) algorithm was used to detect the centerline of the esophagus. Segmentation error was reduced from 2.6 ± 2.1 mm to 2.1 ± 1.9 mm using PCT. Meyer et al. [53], in the same year, proposed another method using a model-based segmentation framework to segment the three organs trachea, esophagus, and heart. The authors used the prior shape knowledge [64], which was an advantage for segmenting the low contrast esophagus. Grosgeorge et al. [54], proposed a method to segment the esophagus using two approaches: the first step includes the segmentation using graph cut with prior knowledge of skeleton [65] and the second step includes 2D propagation [66]. Comparing with the manual ground truth, the results improved remarkably by achieving accuracy of 0.61 ± 0.06 using Dice Metric (DM) as a comparison parameter. Fechter et al. [57] used a deep learning architecture i.e. 3D fully convolution Neural Network to segment the esophagus from CT images. CNN [67] was used to develop a feature map followed by the Active Contour Model which was used to find the location of the esophagus and the result was computed through the implementation of both the techniques sequentially. Using a deformable registration

algorithm in the proposed multi-atlas segmentation [68, 69] with an online atlas selection approach [58], the esophagus was automatically segmented from CT scans of 21 head and neck patients and 15 thoracic cancer patients. The segmentation accuracy achieved on the 56% of the test set had a DSC value equal or greater to 0.7 and HD value equal or greater to 0.3 mm in 86% of the test set. Using data augmentation and sharp mask architecture in FCN, Trullo et al. [59] achieved 0.72 ± 0.07 segmentation accuracy for the esophagus measured using DSC with standard deviation.

For the segmentation of the heart, Larrey-Ruiz et al. [55], proposed a method using image-based segmentation. But the technique can segment only the left heart being its limitation. The precise segmentation of the left heart is performed employing the Isodata algorithm [70] which take the advantage of histogram containing two main clusters of gray levels, one for oxygenated blood and other for non-oxygenated contained in the resulting images after pre-processing. Schreiber et al. [56] segment the organs in thoracic part by matching the suitable atlases with the patient's datasets using deformable registration [71], deduced a most probable shape using STAPLE algorithm [72] and using customized level set-based algorithms [73] to produce a highly accurate segmentation. Dong et al. [61], proposed a novel technique to automatically segment the esophagus and heart from CT scan. They used U-Net-generative adversarial network (U-Net-GAN) to efficiently train the deep neural network to accurately perform the segmentation of the esophagus and heart from the given dataset. GAN-based segmentation model consists of a generator network to produce the segmentation maps of the images of multiple organs and a discriminator network as FCN to discriminate between the ground truth and the actual segmented OARs generated by the generator. In the year 2019, another novel technique based on deep CNN was proposed by Feng et al. [62], to segment the thoracic organs-at-risk in CT images. Firstly, the 3D thoracic images were cropped to patches of smaller sizes each containing only one organ. Then the organs were segmented from each cropped patch using a separate CNN network. The final results were produced by resampling and merging of the individual segmented organs. Noothout et al. [60], performed automatic segmentation of the ascending aorta, aortic arms and descending aorta in low dose chest CT scan using dilated CNN and two-fold cross-validation was performed to evaluate the performance of the proposed technique. Employment of dilated CNN enables the large receptive field analysis while keeping the network parameters low. The accurate detection of the aorta is allowed by large receptive fields based on context information. Also, the network can analyze images of variable sizes because of its purely convolutional nature.

The automatic segmentation of Organ-at-risk is crucial and important for proper planning of Radiotherapy treatment

Table 3 A review of the shortlisted articles containing different automatic techniques to segment the thoracic organ-at-risk

S. no.	References	Objectives	Sample size	Features extraction	Techniques	Results	Conclusion	Remarks
1.	Kurugol et al. [52]	To automatically segment low contrast esophagus in CT scan	CT scan of 8 patients, resolution $0.98 \times 0.98 \times 3.75 \text{ mm}^3$	Landmarks, contrast regions	Model-based with (PCT) Principle Curve Tracing algorithm	Point-wise mean error: $2.6 \pm 2.1 \text{ mm}$ and maximum error: 17.6 mm to $2.1 \pm 1.9 \text{ mm}$ and the maximum error of 15.1 mm	The algorithm using prior models achieved a smooth segmentation result	The technique is tested on small no of the dataset
2.	Meyer et al. [53]	Fully automatic segmentation of Complex organs (heart, trachea, esophagus)	Database of 23 cardiovascular CT angiographic volumes of different patients, each slice of thickness $0.4\text{--}1 \text{ mm}$ and 512×512 pixel matrix	The boundary of the objects	Shape constrained model-based segmentation	Model: trachea only* trachea $-\epsilon_{\text{mean}} - 1.92$ Heart + trachea: trachea $-\epsilon_{\text{mean}} - 1.48$ Esophagus only*: esophagus $-\epsilon_{\text{mean}} - 4.01$ (ϵ : Euclidean distance)	Extension of automatic model-based segmentation is used for segmenting the complex organs. The combined segmentation method gives efficient result for esophagus and trachea	Combined segmentation framework helps in locating the low contrast organs esophagus
3.	Grosgeorge et al. [54]	Esophagus segmentation from 3D CT data	Dataset of 6 patients, each CT volume includes between 73 and 108 slices. Voxelsizeis $0.98 \times 0.98 \times 2.0 \text{ mm}$ or $0.98 \times 0.98 \times 2.5 \text{ mm}$	Construction of the skeleton model	Graph cut with Skeleton-shape model	$3D + 2D\text{Seg}(\text{total})$ $DM-0.61 \pm 0.06$ $P (\%) = 88.9 \pm 11.9$	The use of 2D segmentation allows avoiding overestimation given by 3D graph method in slices much more difficult to segment and to obtain better results	The technique is tested on small no of the dataset
4.	Larrey-Ruiz et al. [55]	Image-based segmentation of the heart	32 Clinical Datasets of varying numbers, each slice of size 512×512 pixels. The resolution of the slice varies from 0.488 to 0.781 mm	Intensity	Multi-thresholding, image cleaning, mathematical morphology, and image filtering method	Left heart: Mean error— 1.22% Stdev— 0.68% Mean correlation ratio: 94.42%	The proposed segmentation technique is significantly faster and automatically delineates all the contours	The proposed method can segment only the left part of the heart
5.	Schreibmann et al. [56]	Multiatlas segmentation of thoracic and abdominal anatomy	Datasets of both abdominal and thoracic scanning protocols produce images of 2.5 mm voxel size, with a typical scan having $100\text{--}200$ slices	Rigid matching, the center of mass	Atlas registration algorithm STAPLE algorithm	Dice coefficient: 0.971 for the aorta, 0.869 for the trachea, 0.958 for the lungs, 0.788 for the heart	The method gives accurate segmentation even with the variability in the size of the image	The method is extendable to 4D images also with further modifications

Table 3 (continued)

S. no.	References	Objectives	Sample size	Features extraction	Techniques	Results	Conclusion	Remarks
6.	Fechter et al. [57]	Esophagus segmentation in CT	20 clinical datasets, voxel resolutions from 0.78×0.78 mm to 1.17×1.17 mm in axial plan and 30 publicly-available CTs taken with a resolution ranging from 0.67×0.67 mm to 0.98×0.98 mm	Spatial overlap	Fully Convolution Neural Network (CNN) & Active Contour Model (ACM)	DSC: 0.76 ASSD: 1.36 HD: 11.6	The technique accurately segments the low contrast esophagus and can be used on a large number of datasets. It can also be used for the segmentation of other organs	Performing 3D convolution exploits the 3D spatial context and volume-wise prediction also performed efficiently
7.	Yang et al. [58]	Automatic segmentation of esophagus from CT scan	Internal: 15 thoracic CTs External: 3D CT scans of 30 patients, slice resolution ranges from 0.57 to 0.96 mm	Local anatomy similarity in atlas selection	Multi-atlas segmentation (MAS)	Proposed MAS-AS For DSC: 0.73 ± 0.07 MSD (mm): 2.2 ± 0.8 HD (mm): 16.9 ± 8.9	The robustness of MAS is improved by using online atlas selection approach In the proposed method	The major obstacle in atlas-based segmentation is the presence of air bubbles
8.	Trullo et al. [59]	To present extensions of the FCN methods to segment the esophagus with more accuracy and efficiency	30 thoracic CT scans, a slice resolution of $0.98 \times 0.98 \times 2.5$ mm ³ for a size of $512 \times 512 \times (150-284)$ voxels	Patch match	Sharp mask architecture and conditional random fields	Dice score of proposed technique: (esophagus) 0.72 ± 0.07	Dice scores obtained on a dataset of 30 patients are very encouraging and show valuable results	Extension in FCNN effectively combine low-level features with the high-level feature to improve the accuracy
9.	Noothout et al. [60]	To segment Thoracic aorta in low dose chest CT	24 low dose chest CT scans, in-plane resolution varying b/w 0.46 and 0.86 mm, a slice thickness varying b/w 1.25 and 4.00 mm	Voxel	CNN with dilated convolutions	Multi-class Dice coefficients- Thoracic aorta: 0.89 ± 0.05 ASSD: Thoracic aorta- 1.67 ± 1.02	The proposed network accurately segments the aorta. The results of the multiclass network are similar to the two-class network	The result of non-contrast images needs to improve further
10.	Dong et al. [61]	Automatic Multiorgan segmentation in thorax CT images	35 sets of thoracic CT images used, obtained from 2017 AAPM Thoracic Auto-segmentation challenges	Textural information Structural information	U-NET Generative adversarial network (GAN)	Esophagus DSC: 0.75 ± 0.08 MSD: 1.05 ± 0.66 HD: 4.52 ± 3.81 Heart DSC: 0.87 ± 0.05 MSD: 1.49 ± 0.85 HD: 4.58 ± 3.67	This novel technique is very useful in lung radiotherapy treatment planning	The proposed technique does not work well for the variable data

Table 3 (continued)

S. no.	References	Objectives	Sample size	Features extraction	Techniques	Results	Conclusion	Remarks
11.	Feng et al. [62]	Segmentation of CT thoracic organs	60 Thoracic CT scans, dividing into three parts: 36 training cases, 12 offline test cases, 12 online test cases	Voxel Spacing, intensity	Deep CNN using cropped 3D images	Dice Score Esophagus-0.726 ± 0.094 Heart-0.925 ± 0.015 Hausdorff Distance Esophagus-8.714 ± 10.588 Heart-6.570 ± 1.501	A novel technique was developed to efficiently segment the organ at risk in thorax part by achieving good performance and less cost	Low-quality images suffer poor segmentation

of cancer in the thorax part. Trullo et al. [37], used a CNN [84] in combination with sharp mask [85, 86] and conditional random field, implemented as recurrent neural network architecture for the automatic segmentation of Organs-at-risk in Thoracic CT images. In the April month of the year 2019, a competition SegTHOR was organized by Trullo et al. on the topic “Automatic segmentation of Organ-at-risk in Thoracic CT images”. The provided dataset consists of CT scans of a total of 60 patients, 40 for training purposes and 20 for testing purposes. Different novel techniques had been proposed in the SegTHOR for the automatic segment the Organ-at-risk in Thoracic CT images. The performance of the techniques was measured using two parameters, Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD). This systematic review included all the papers published in the current year on the same topic. 12 papers published with different techniques mostly based on deep learning frameworks [87] for the automatic segmentation of OARs had been found which participated in the competition including 1 paper of the organizing team as shown in Table 4. All the authors used the same dataset and the same parameters Dice coefficient (DSC) and Hausdorff Distance (HD) were calculated to evaluate the accuracy of their proposed technique. Satya et al. [36], proposed a technique based on Dilated U-Net with 14 layers for the automatic segmentation of OARs. To overcome the overlapping of organ regions in the output results, the authors applied post-processing integrated with the model which further improved the overall accuracy of segmentation results. Another technique using the 2D dilated residual Network was proposed by Vesal et al. [75] to precisely perform the segmentation of OARs. Using the dilated convolutions, the receptive field was expanded in the lowest level of the network enabling the technique to use both local and global information without any increase in the network complexity. Louis et al. [74], proposed their technique which was the combination of 2D and 3D Convolutional Neural Network (CNN) to produce the segmentation results. The 3D network was an FCNN [88, 89] using residual blocks allowing the network to focus on detailed local information whereas the 2D network contained dilated convolutions that allow large receptive fields while inputting the images.

A two-stage cascaded network was used by Kim et al. [76] for the segmentation of thoracic organs. The two-step included: first the selection of slices and the second was the segmentation using the segmentation network with ensemble technique. But this technique gave the least accuracy among all other proposed techniques especially for the esophagus because of its small size. A 3D 3-stage multi-scale network was used by Wang et al. [39] to segment the four thoracic organs and their results show improvement for all organs. U-Net inspired network comprises a context pathway and a localization pathway was used by Chan et al. [77], to

Table 4 Demonstration of studies (13) performed under SegTHOR competition

S. no.	References	Objectives	Sample Size	Techniques	Results	Conclusion	Remarks
1.	Trullo et al. [37]	To propose a novel framework for automatic segmentation of Organ-at-risk	30 CT scans, each scan of size 512×512 and resolution of each is $0.98 \times 0.98 \times 2.5 \text{ mm}^3$	Sharp Mask architecture and Conditional random field	DR \pm stdev esophagus: 0.69 ± 0.05 heart: 0.90 ± 0.03 trachea: 0.87 ± 0.02 aorta: 0.89 ± 0.04	Collaborative deep architecture helps in learning the spatial relationship between the organs which further improve the performance by refining the segmentation	Results need to be improved further
2.	Manoj et al. [36]	To segment the OAR in thorax part by avoiding coarse output from Dilated U-Net	60 CT scans, each scan of size 512×512 and resolution of each is $0.98 \times 0.98 \times 2.5 \text{ mm}^3$	Dilated U-Net	Dice ratio: Heart 0.8595, Aorta 0.8537, Esophagus—0.6425, Trachea—0.4694 Hausdorff distance: Heart—0.8930, Aorta, 1.4495, Esophagus—2.8883, Trachea—2.7342	To separate from the background all the four organs are segmented individually and augmented to train the model. Post-processing further improved the results	Using pooling after two consecutive layers saves spatial resolution
3.	van Harten et al. [74]	To present a combination of 2D and 3D CNN for automatic segmentation of Thoracic organ-at-risk	3D: $64 \times 64 \times 64$ voxel sub-images 2D: $40 \times 256 \times 256$ pixel sub-images	Combination of 2D and 3D CNN	Dice coefficients of 0.84 ± 0.05 , 0.94 ± 0.02 , 0.91 ± 0.02 , and 0.93 ± 0.01 (esophagus, heart, trachea) and aorta Hausdorff distance: Eso— 3.4 ± 2.3 , Heart— 2.0 ± 1.1 , Trachea— 2.1 ± 1.0 , Aorta— 2.7 ± 3.6	The analysis showed that a combination of both the networks 2D and #D CNN's produce the best results	Increasing the size of the networks will increase the computation time and cost
4.	Vesal et al. [75]	Multiorgan Segmentation in Thoracic CT images using dilated residual 2D U-NET	CT scans of 60 patients of size 512×512 pixels, with a resolution of $0.98 \times 0.98 \times 2.5 \text{ mm}^3$	2D dilated Residual U-Net	Metrics DSC (%) Esophagus 0.858 Heart 0.941 Trachea 0.926 Aorta 0.938 HD (mm) Esophagus 0.331 Heart 0.226 Trachea 0.193 Aorta 0.297	The proposed technique provides accurate and reproducible segmentation which improves consistency and robustness in the delineation of OARs	The segmentation of the organ-at-risk with poorly-visible boundaries such as the esophagus is still a challenging task

Table 4 (continued)

S. no.	References	Objectives	Sample Size	Techniques	Results	Conclusion	Remarks
5.	Kim et al. [76]	Segmentation of thoracic organs	CT slices of 60 patients, each slice of 512×512 pixels	Encoder-Decoder based CNN	DSC of 0.7518, 0.9328, 0.8885, 0.8919 and Hausdorff of 0.9267, 0.2184, 0.6164, and 1.1300 for esophagus, heart, trachea, aorta	CNN with cascaded two-step approach shows comparable performance with other networks	The proposed method provides a poor result for the low contrast organ esophagus
6.	Wang et al. [39]	To segment the organ-at-risk in Thoracic CT images	60 CT scans, size of CT scans is 512×512 with a resolution of 0.90–1.37 mm. z-resolution varies from 0.90 to 1.37 mm	3D enhanced multi-scale network (EMSN)	Dice coefficient: Esophagus 0.8597 Heart 0.9459 Trachea 0.9217 Aorta 0.9433 Hausdorff distance: Esophagus 0.2883 Heart 0.1594 Trachea 0.2045 Aorta 0.1551	Compared with the previous results, the proposed technique improved the results up to a great extent especially the low contrast organ esophagus	Training on (1600) dataset provides the reliability of the proposed technique that it can be applied on a large dataset and will provide notable results
7.	Chan et al. [77]	Segmentation of OARs in CT images of the thorax	60 CT scans, each scan of size 512×512 and resolution of each is $0.98 \times 0.98 \times 2.5$ mm ³	Extension of 3D U-shaped Neural Network	Dice coefficient: Esophagus 0.8166 Heart 0.9329 Trachea 0.8910 Aorta 0.9232 Hausdorff distance: Esophagus 0.4914 Heart 0.2417 Trachea 0.2746 Aorta 0.3081	Post-processing after segmentation restores the original image	Dice loss function reduces the problem of unbalanced class in the data
8.	He et al. [78]	Multi-Task learning for the segmentation of thoracic organ-at-risk in CT images	60 abdominal CT scans, each scan of size 512×512 and resolution of each is $0.98 \times 0.98 \times 2.5$ mm ³	U-Net with extensions	Dice coefficient: Esophagus—0.8594 Heart—0.9500 Trachea—0.9201 Aorta—0.9484 Hausdorff distance: Esophagus—0.2743 Heart—0.1383 Trachea—0.1824 Aorta—0.1129	Enabling different network with U-like architecture is very helpful in discovering organ dependence	Multi-task learning using U-Net improves the segmentation

Table 4 (continued)

S. no.	References	Objectives	Sample Size	Techniques	Results	Conclusion	Remarks
9.	Kondratenko et al. [79]	Segmentation of thoracic OAR in CT images using Localization and organ-specific CNN	Abdominal CT scans, each scan of size 512×512 pixels and resolution of each is $0.98 \times 0.98 \times 2.5 \text{ mm}^3$	T-Net architecture in CNN	Dice coefficient: Esophagus—0.80 Heart—0.93 Trachea—0.89 Aorta—0.92 Hausdorff: Esophagus—0.62 Heart—0.30 Trachea—0.81 Aorta—0.27	The approach helps train the complex model with large no of datasets	Results are not uniform in all the images
10.	Zhang et al. [80]	Segmentation of thoracic organ-at-risk in CT images combining coarse and fine network	CT scans have the resolution 512×512	Combination of Coarse and Fine segmentation Network	Dice: Esophagus—0.7732 Heart—0.9384 Trachea—0.8939 Aorta—0.9232 Hausdorff: Esophagus—1.6774 Heart—0.2089 Trachea—0.2741 Aorta—0.3081	Coarse segmentation helped in finding the exact location of organs and features of various organs are obtained in fine segmentation	Memory constraint is there in using the proposed technique on high-quality images
11.	Lachinonet al. [81]	Segmentation of thoracic organs using pixel shuffle	60 CT scans, resolution 512×512 , voxel size varying between 0.90 and 1.37 mm per pixel	Deep learning with pixel shuffle	Dice Score: Eso—0.8303 Heart—0.9381 Trachea—0.9088 Aorta—0.9353	The proposed method gives the notable performance with the dice score values without any further post-processing	Group normalization in deep networks reduces the computational cost
12.	Feng et al. [82]	Multi-organ segmentation	60 CT scans, resolution 512×512 , voxel size varying between 0.90 and 1.37 mm per pixel	SIMPLIFIEDDENSEV-NET	Dice: Eso—0.773450 Heart—0.941403 Trachea—0.892730 Aorta—0.923325 Hausdorff: Eso—0.640093 Heart—0.182138 Trachea—0.307711 Aorta—0.235788	Variety of optimization techniques improve the stability and performance of the model	The accuracy with the proposed technique may vary with the increase in the no of datasets

Table 4 (continued)

S. no.	References	Objectives	Sample Size	Techniques	Results	Conclusion	Remarks
13.	Han et al. [83]	Segmentation of CT thoracic organs	60 CT scans, resolution 512 × 512, voxel size varying between 0.90 and 1.37 mm per pixel	Multi-resolution VB-nets	Dice metric Ensemble: Eso—0.8651, Heart—0.9536, Trachea—0.9276, Aorta—0.9464 Hausdorff distance Ensemble: Eso—0.2590, Heart—0.1272, Trachea—0.1453, Aorta—0.1209	The proposed technique reduces the cost of the GPU memory and also maintain the high accuracy	Efficient in terms of memory and time consumption

perform the segmentation. A multi-test step was added further on the esophagus to improve its performance. Another U-like architecture was abstracted from the U-Net by He et al. [78], used a uniform encoder-decoder architecture, involving Res-Net by eliminating the linearly connected layers. Due to the dependent organ occurrence, they used the multi-task learning for the thoracic organ’s segmentation to give the best result and achieved the second rank in the competition. Kondratenko et al. [79], used 2d T-Net architecture for the organ segmentation. Random slices were selected to train the model using batch size 8. Then 5-cross validation was performed on size 22. Different mean dice score values were obtained on the test set collected from the training data. Using the proposed technique, they stood 24th in the competition. Another competitor Zhang et al. [80], proposed a deep neural network combining coarse and fine networks. The coarse network was used to find a region of interest localization, then the features were used to get fine segmentation. A fully convolution Neural Network architecture based on U-Net was used by Lachinonet et al. [81], to complete the task by exploring two concepts: attention mechanism and pixel shuffle as an upsampling operator. The no of feature channels used at the starting was equal to 16. Their proposed method produced notable results for the segmentation. A simplified Dense V-Net with input size 128³ was used by Feng et al. [82] to achieve a better result comparing with the baseline [37].The final competitor and winner of the competition are the Han et al. [83]. They proposed a multi-resolution VB-Net framework to achieve the best results and ranked 1st in the competition. Their proposed method reduced the computation cost while maintaining high segmentation accuracy for the esophagus, heart, trachea and aorta organs.

Various techniques have been mentioned in Table 5 that had been used in the different papers on different datasets to automatically segment the four thoracic organs (esophagus, heart, aorta, and trachea). Convolutional Neural Network-based techniques were used in 8 papers and the accuracy of the technique of these papers was compared using the same parameter i.e. Dice coefficient.

2.2.1 Automatic Segmentation Techniques Based on Different Frameworks Proposed in the Literature

Deep learning-based networks like CNN had been implemented in the literature as shown in Table 5 and it has outperformed in the zone of medical image segmentation. Among all the CNN based proposed techniques, van Harten et al. [74], achieved the highest value for DSC and HD for the esophagus, heart, trachea, and aorta with the combination of 2D and 3D CNN. Both the network contains different architectures. The 3D network performed the multi-label segmentation whereas the 2D network performed the

Table 5 Different Automatic Segmentation Techniques based on different frameworks proposed by the various papers and showing the best accuracy among four OAR taking Dice Score Coefficient as the parameter

S. no.	Networks	Techniques	References	Accuracy (DSC)	Remarks
1.	Convolutional Neural Network (CNN)	Fully Convolution Neural Network (FCNN) and Active Contour Model (ACM)	Fechter et al. [57]	Esophagus: 76%	
		Sharp Mask architecture	Trullo et al. [59]	Esophagus: 72%	
		CNN with dilated convolutions	Noothout et al. [60]	Thoracic aorta: 89%	
		Deep CNN using cropped 3D images	Feng et al. [62]	Esophagus: 72.6% Heart: 92.5%	
		Sharp Mask architecture and Conditional random field	Trullo et al. [37]	Esophagus: 67% Heart: 90% Trachea: 82% Aorta: 86%	
		Combination of 2D and 3D CNN	van Harten et al. [74]	Esophagus: 84% Heart: 94% Trachea: 91% Aorta: 93%	The best accuracy for the esophagus, heart, trachea and aorta
		Encoder-Decoder based CNN	Kim et al. [76]	Esophagus: 75.1% Heart: 93.2% Trachea: 88.8% Aorta: 89.1%	
2.	U-Net based framework	T-Net architecture in CNN	Kondratenko et al. [79]	Esophagus: 80% Heart: 93% Trachea: 89% Aorta: 92%	
		U-NET Generative adversarial network (GAN)	Dong et al. [61]	Esophagus DSC: 75% Heart DSC: 87% ± 0.05	
		Dilated U-Net	Manoj et al. [36]	Esophagus—64.25% Heart—85.95%, Aorta 85.37% Trachea 46.94%	
		2D dilated Residual U-Net	Vesal et al. [75]	Esophagus 85.8% Heart 94.1% Trachea 92.6% Aorta 93.8%	Best accuracy for the Trachea
		Extension of 3D U-shaped Neural Network	Chan et al. [77]	Esophagus 81.66% Heart 93.29% Trachea 89.10% Aorta 92.32%	
		U-Net with extensions	He et al. [78]	Esophagus—85.94% Heart—95.00% Trachea—92.01% Aorta—94.84%	Best accuracy for the esophagus& Heart& Aorta
		Deep learning with pixel shuffle	Lachinov et al. [81]	Eso—83.03% Heart—93.81% Trachea—90.88% Aorta—93.53%	

Table 5 (continued)

S. no.	Networks	Techniques	References	Accuracy (DSC)	Remarks
3.	Model-based segmentation	Principle Curve Tracing algorithm	Kurugol et al. [52]	Esophagus: Point-wise mean error: 2.6 ± 2.1 mm	
		Shape-constrained model-based segmentation	Meyer et al. [53]	Model: trachea only* trachea — $\epsilon_{\text{mean}} - 1.92$ heart + trachea: trachea — $\epsilon_{\text{mean}} - 1.48$ Esophagus only*: esophagus — $\epsilon_{\text{mean}} - 4.01$	
4.	Atlas-based segmentation	Atlas registration algorithm	Schreibmann et al. [56]	Heart 78.8% Trachea 86.9% Aorta 97.1%	Best and highest accuracy for Aorta
		STAPLE algorithm			
		Multi-atlas segmentation (MAS-AS)	Yang et al. [58]	Heart 73%	
5.	Image-based segmentation	Multi-thresholding, image cleaning, mathematical morphology, and image filtering method	Larrey-Ruiz et al. [55]	Heart: Mean error—1.22% Stdev—0.68%	
6.	Skelton-shape model	Graph cut method	Grosgeorge et al. [54]	Esophagus 61%	
7.	V-Net framework	SIMPLIFIEDDENSEVNET	Feng et al. [82]	Esophagus—85.97% Heart—94.59% Trachea—92.17% Aorta—94.33%	
		3D enhanced multi-scale network (EMSN)	Wang et al. [39]	Esophagus 77.34% Heart 94.14% Trachea 89.27% Aorta 92.33%	
		Multi-resolution VB-nets	Han et al. [83]	Eso—86.51% Heart—95.36% Trachea—92.76 Aorta—94.64%	Best accuracy for esophagus, heart, trachea and aorta
8.	Coarse and Fine segmentation	—	Zhang et al. [80]	Esophagus—77.32% Heart—93.84% Trachea—89.39% Aorta—92.32%	

multi-class segmentation in their respective output layer and this distinction in the network promotes the additional diversity. Dilated Convolution [90] in the 2D network allows large receptive fields as the input to layers that help the network to extract high-level features. The advantage of using both the network is that it does not increase the computational cost. So, it can be used for performing the larger segmentation.

Another deep learning-based network was the U-Net based techniques that were used in 6 papers to perform automatic segmentation of thoracic organs. U-Net contains the up-sampling layers instead of pooling which increases the resolution in the output layer. The precise accuracy in terms of Dice Similarity Coefficient using U-Net based techniques was achieved by He et al. [78] for the esophagus, heart, and aorta. The multi-task learning scheme used in the network helps discover the dependencies among organs and the transfer learning in training boosted the performance of the network. Vesal et al. [75], achieved the highest value

of DSC for the aorta due to its high contrast, regular shape and large size compared to other organs using dilated U-Net based method. Dilated convolution expanded the receptive fields to efficiently extract the features of the objects without increasing the network complexity. Residual connections in U-Net incorporated global context and dense information which was the reason for the precise segmentation of the organ among all other U-Net based papers.

Model-based segmentation is characterized as the assignment of labels to pixels or voxels by matching the a priori known object model to the image data. Labels may have probabilities expressing their uncertainty [45]. Techniques based on this type of segmentation was used in two paper Kurugol et al. [52] and Meyer et al. [53] to segment esophagus, heart, and trachea. But the parameters on which the accuracy of their proposed techniques was measured were different. Schreibmann et al. [56] and J. Yang et al. [58] used Atlas-based segmentation for the segmentation of the

heart, trachea, and aorta. The highest accuracy in terms of DSC value for the aorta was achieved by [56]. Atlas registration algorithm [91–93] and STAPLE algorithm [94] were used to achieve accurate segmentation. To exclude unnecessary regions from the metric calculation, a mask of the atlas segmentation with a margin is used for the moving image to calculate the deformation only in voxels within defined organs. To adjust for variability among individual atlas results, STAPLE algorithm was used.

Larrey-Ruiz et al. [55] used an image-based segmentation method to segment the heart and used stdev and mean error as the parameters to calculate the accuracy of their proposed technique.

Feng et al. [82], Wang et al. [39] and Han et al. [83] used V-Net framework to segment the four organs. The accuracy was measured as dice coefficient value. Among the discussed three studies, precise accuracy for the esophagus, heart, trachea & aorta was achieved by [83]. Coarse and Fine segmentation was used by Zhang et al. [80] only and computed dice coefficient accuracy for the four organs. The distribution of different techniques used in the literature also shown in the Fig. 2 with the Pie-chart representation.

2.2.2 Precise Accuracy Achieved Among All the Proposed Techniques for the Automatic Segmentation of the Esophagus, Heart, Trachea, and Aorta

The Common parameters that were used in the literature are the Dice similarity coefficient (DSC) and the Hausdorff Distance. They were used to calculate the accuracy of the segmented part by comparing it with the ground truth

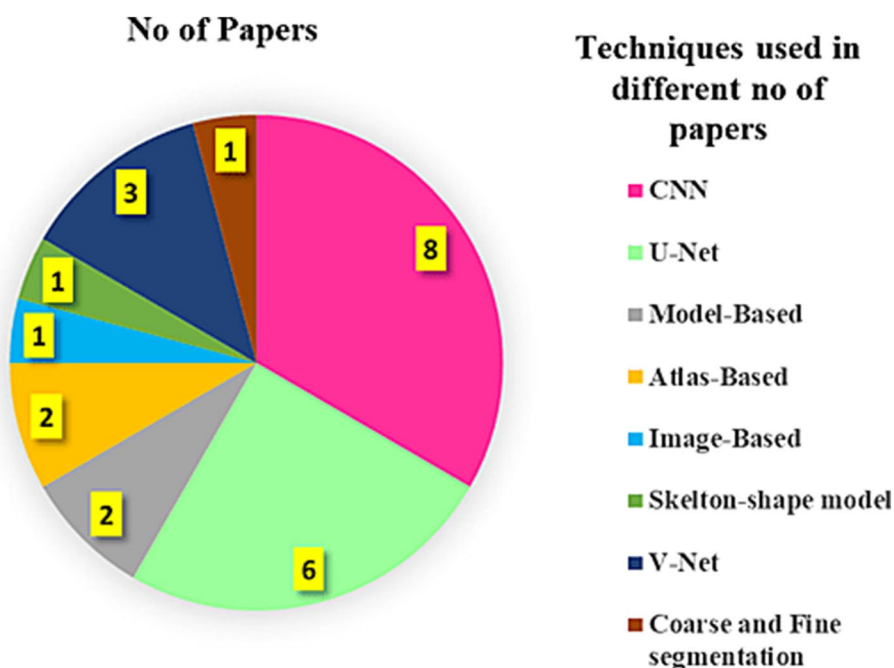
values. Dice Score coefficients and Hausdorff Distance (HD) calculated for the four thoracic OARs by the different proposed techniques in the literature are represented in the form of graphs given in Figs. 3 and 4. In the case of the Dice similarity coefficient (DSC), the more the value of DSC, the more accurate the technique would be whereas, in the case of Hausdorff Distance, the less the value of HD, the technique would be more efficient.

The overall Precise accuracy for the four thoracic organs (esophagus, heart, aorta, and trachea) of the proposed techniques among 24 papers using the parameter Dice coefficient and Hausdorff Distance was achieved by Han et al. [83] who were also the winner of the competition sector. The result can be visualized from the graph in Figs. 3 and 4. They introduced a multi-resolution VB-Net framework to achieve the 1strank and accurate segmentation for the four organs. They trained the VB-Net in the coarse resolution to roughly locate the region of interest (ROI) for each organ and in the fine resolution, the organ boundaries were correctly delineated within the marked ROI.

Multi-resolution VB-NET:

Han et al. [83] proposed a new architecture using the V-Net framework and they called it VB-Net. B stands for a bottleneck in the network. The bottleneck structure which replaces the conventional layers in the existing network inside the down block and up blocks. The network was divided into two main parts. The left side of the network called the contraction path is used to extract the high-level features of the image by performing convolution and downsampling.

Fig. 2 Distribution of the different techniques used among the literature



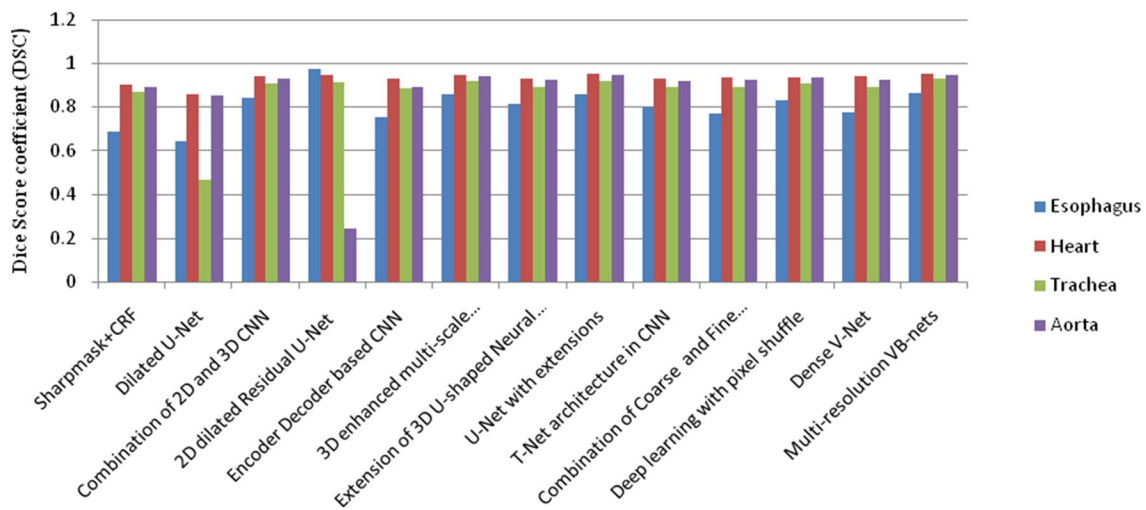
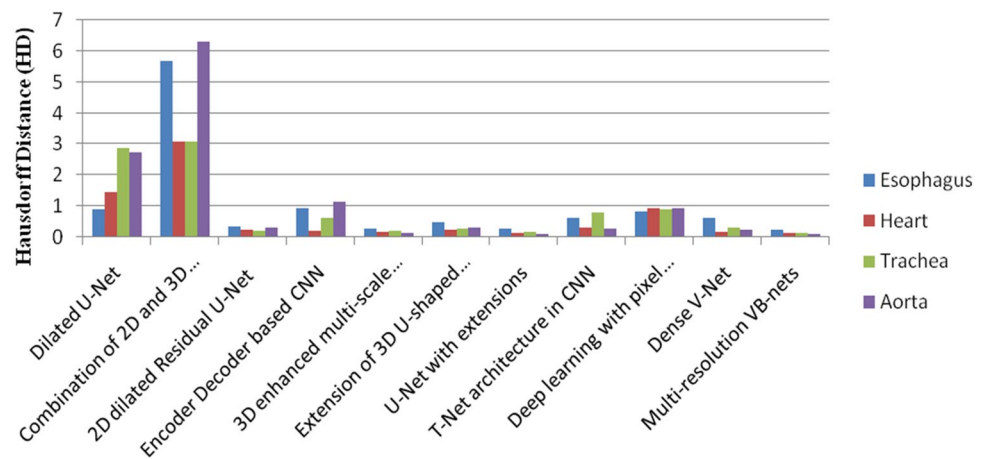


Fig. 3 Dice similarity coefficient (DSC) calculated by different techniques for the four Thoracic organ-at-risk

Fig. 4 Hausdorff Distance (HD) calculated by different techniques for the four Thoracic organ-at-risk



The right side of the network called the expansion path is used to expand the feature map to gather and assemble necessary information to output to five-channel volumetric segmentation.

There is a horizontal connection between the two paths that are used to transfer the features extracted on the left side of the network to the right side of the network. It will help in providing location information to the right path and improve the quality of the final output.

The authors gave the network name as a bottleneck because in the bottleneck structure the size of the input block is reduced first to lower down the computational cost and again the size of the block is restored to its original size.

Multi-resolution Strategy:

As dealing with 3D medical images that are of large size (512 × 512 × 300), processing it on CNN will consume a lot of GPU memory and also computation time. The advantage of using VB-Net was to perform the

segmentation task and to solve the above problem. The authors used two VB-Nets to achieve the precise accuracy for segmenting the OARs (Esophagus, Heart, Trachea and Aorta).

First VB-Net was used to get the Volume of interest (VOI) from each image by reducing the size of the input for the second network. The second VB-Net was used to accurately delineate the organ boundary from the detected VOI.

In the proposed architecture the downsampling is performed using different no of filters in the different layers of the VB-Net starting from the 16 filters in the first layer to 256 filters in the last layer of the left side of the network. As the no of filters increases in the network down the layers, the no of feature maps also increases with reduced size of the output blocks. The upsampling was performed by taking the no of filters in the reverse order. The upsampling layer increases the size of the input channel without adding extra information into it. The final segmented result was in the

same size as of the input image after performing the upsampling on the right side of the network.

2.2.3 Minimum and Maximum Size of the Dataset

A varying number of Datasets was being used in the selected 24 papers shown in Table 6 and all of them were captured through the same type of modality i.e. Computed Tomography. The Table 6 is arranged in the ascending order (minimum to maximum) of the no of the datasets used in different papers.

All the given datasets in Table 6 are of the thorax region. Most of the images of CT scans have a size of 512x512 pixels, slice thickness varying from 0 to 2.5 mm. It contained the images of patients from different planes: axial, coronal and sagittal. The dataset used by the SegTHOR competitors was downloaded from the Coda Lab website <https://codalab.org/> which is the largest number of datasets used for the segmentation purpose among all the selected papers. The minimum no of the dataset used was a CT scan of 6 patients [54]. The Table 7 listed Datasets downloaded from various sources or sites used in the literature. In a research study, it has been found that training on a large amount of dataset generate more accurate and efficient results [62, 78, 83] and also provide € validation of the developed techniques. Pre-and post-processing of the datasets had been performed according to the need for the different techniques being implemented for the segmentation which was the key step in the network performance. Using a dataset of bigger size in deep learning frameworks increases the computational cost and memory utilization. So, dimension reduction in the dataset was performed by many researchers before giving the image as input. Although the reduction in pixel dimension could also lead to a loss of information that happened in deep learning networks. That's why post-processing was being done after the final segmentation outputs to enhance the results. The Fig. 5 shows the Pie-chart representing the different Dataset utilization among different papers.

2.2.4 Various Challenges Faced by the Authors While Segmenting the Organs in CT Images

Segmentation is the crucial and important step in radiotherapy treatment planning as said earlier. Performing the segmentation with the respective proposed techniques by the various authors, several challenges are encountered.

1. *Low contrast images* As we have considered the study of Computed Tomography images only so the very first challenge that all the authors faced is the dataset containing low contrast images or organs. Low contrast in CT usually refers to 4–10 HU difference indicate that the objects are barely differentiable from the background.

So, it becomes difficult for the doctors, researchers or clinicians to accurately identify object boundaries or to separate object and background. The CT contains high contrast for hard tissues like bone and low contrast for soft tissues. The automatic segmentation of the esophagus is exceptionally challenging: the boundaries in CT images are almost invisible due to its low contrast. The shape is also irregular.

2. *Computational time* Most of the proposed techniques are based on deep learning as it has performed tremendous work in medical image segmentation than other techniques. But using deep learning-based frameworks like CNN or RCNN takes much time and memory in processing of the high-resolution medical images. Depending upon the no of layers being used in the Convolutional Neural Networks, the complexity of the technique is measured.
3. *Reduction of the spatial resolution of feature maps* Methods involving deep learning framework gives better segmentation accuracy. But as the no of layers increases in CNN to get more feature maps, the size of the input image decreases resulting in low spatial resolution. It's the common challenge faced by many authors in using deep learning-based architecture. Although many authors have efficiently handled this challenge by making changes in the existing architecture and by adding additional features in it like up-sampling, increasing the stride no in the kernel.
4. *Small no. of sample size* Taking small no of the dataset for training and testing does not validate the efficient working of the proposed techniques. Training on small and large datasets by the same technique.
5. *Variability in patients organs* Organ's shape and position differ greatly between patients. It is difficult for the automatic algorithms to attain accurate and consistent segmentation results. They can give varying accuracy for different patients.
6. There is no standardized technique developed that will give higher accuracy for the Organs at Risk in Thoracic CT images as the size of the organs varies from patient to patient so there will always be a scope of getting less accuracy for any organ with the already developed techniques.
7. Feature extraction does not ensure the correct training of the model. Therefore, more work is needed to enhance the technique required for the extraction of features.

3 Discussion

Following the PRISMA guidelines, a total of 24 research papers were selected for the literature review. Figure 6 represents the year-wise distribution of the 24 papers focusing

Table 6 List of different numbers of Datasets used in various papers to perform automatic segmentation of thoracic organs

S. no	Datasets	Authors	References	Results	
				Dice coefficient	Hausdorff distance (mm)
1.	6 patients	Damien Grosgeorge et al.	[54]	DM-0.61 ± 0.06	Not reported
2.	8 patients	S. Kurugol et al.	[52]	Mean error: 2.6 ± 2.1 mm	
3.	23 cardiovascular CT angiographic volumes	Meyer et al.	[53]	Model trachea only: trachea - $\epsilon_{\text{mean}} - 1.92$ heart + trachea: trachea - $\epsilon_{\text{mean}} - 1.48$ esophagus only: esophagus - $\epsilon_{\text{mean}} - 4.01$	
4.	24 low dose chest CT scans Training: 10 Test: 10	J. Noothout et al.	[60]	Thoracic aorta 0.89 ± 0.05	Not reported
5.	30 thoracic CT scans Training: 25 Test: 5	Trullo, Roger et al. Trullo et al.	[59] [37]	DSC: 0.72 ± 0.07 DR ± stdev esophagus: 0.69 ± 0.05 heart: 0.90 ± 0.03 trachea: 0.87 ± 0.02 aorta: 0.89 ± 0.04	Not reported Not reported
6.	32 Clinical Datasets	J. Larrey-Ruiz et al.	[55]	Left heart: Mean error—1.22% Stdev—0.68%	
7.	35 sets of thoracic CT images Training:34 sets Test:1 set	Xue Dong et al.	[61]	Esophagus DSC: 0.75 ± 0.08 MSD: 1.05 ± 0.66 Heart DSC: 0.87 ± 0.05 MSD: 1.49 ± 0.85	Esophagus 4.52 ± 3.81 Heart 4.58 ± 3.67
8.	45 thoracic CT scans	J. Yang et al.	[58]	DSC: 0.73 ± 0.07 MSD (mm): 2.2 ± 0.8	16.9 ± 8.9
9.	20 clinical datasets, 30 publicly-available CTs	Tobias Fechter et al. (2019)	[57]	DSC: 0.76 ASSD: 1.36	HD: 11.6
10.	60 Thoracic CT scans Training:36 Test:24	Xue Feng et al. (2019)	[62]	Esophagus—0.726 ± 0.094 Heart-0.925 ± 0.015	Esophagus-8.714 ± 10.588 Heart-6.570 ± 1.501
11.	60 Thoracic CT scans Training:40 Test:20	Manojet al. (2019)	[36]	Esophagus—0.6425 Heart—0.8595 Trachea—0.4694 Aorta 0.8537	Esophagus 2.8883 Heart—0.8930 Trachea—0.8930 Aorta, 1.4495
		Louis D. van Harten et al. (2019)	[74]	Eso—0.84 ± 0.05 Heart—0.94 ± 0.02 Trachea—0.91 ± 0.02 Aorta—0.93 ± 0.01	Eso—3.4 ± 2.3 Heart—2.0 ± 1.1 Trachea—2.1 ± 1.0 Aorta-2.7 ± 3.6
		Vesal et al. (2019)	[75]	Esophagus—0.858 Heart—0.941 Trachea—0.926 Aorta—0.938	Esophagus—0.331 Heart—0.226 Trachea—0.193 Aorta—0.297
		S. Kim et al. (2019)	[76]	Eso—0.7518 Heart—0.9328 Trachea—0.8885 Aorta—0.8919	Eso—0.9267 Heart—0.2184 Trachea—0.6164 Aorta—1.1300
		Q. Wang et al. (2019)	[39]	Esophagus—0.8597 Heart 0.9459 Trachea 0.9217 Aorta 0.9433	Esophagus—0.2883 Heart—0.1594 Trachea—0.2045 Aorta—0.1551

Table 6 (continued)

S. no	Datasets	Authors	References	Results	
				Dice coefficient	Hausdorff distance (mm)
		Pan Chan et al. (2019)	[77]	Esophagus: 0.8166 Heart 0.9329 Trachea 0.8910 Aorta 0.9232	Esophagus—0.4914 Heart—0.2417 Trachea—0.2746 Aorta—0.3081
		Tao He et al. (2019)	[78]	Esophagus—0.8594 Heart—0.9500 Trachea—0.9201 Aorta—0.9484	Esophagus—0.2743 Heart—0.1383 Trachea—0.1824 Aorta—0.1129
		V. Kondratenko et al. (2019)	[79]	Esophagus—0.80 Heart—0.93 Trachea—0.89 Aorta—0.92	Esophagus—0.62 Heart—0.30 Trachea—0.81 Aorta—0.27
		Li Zhang et al. (2019)	[80]	Esophagus—0.7732 Heart—0.9384 Trachea—0.8939 Aorta—0.9232	Esophagus—1.6774 Heart—0.2089 Trachea—0.2741 Aorta—0.3081
		D. Lachinov et al. (2019)	[81]	Eso—0.8303 Heart—0.9381 Trachea—0.9088 Aorta—0.9353	Not reported
		Ming Feng et al. (2019)	[82]	Eso—0.773450, Heart—0.941403 Trachea—0.892730 Aorta—0.923325	Eso—0.640093 Heart—0.182138 Trachea—0.307711 Aorta—0.235788
		Miaofei Han et al. (2019)	[83]	Eso—0.8651 Heart—0.9536 Trachea—0.9276 Aorta—0.9464	Eso—0.2590 Heart—0.1272 Trachea—0.1453 Aorta—0.1209

on the automatic segmentation of OARs (esophagus, heart, trachea, and aorta). The period selected was 2011–2019. Maximum papers related to the topic fall in the year 2019 and that was 13 whereas we could not find any relevant paper from the year 2015 and 2016.

As this review focused on the automatic segmentation of four thoracic organs, the Fig. 7 shows the organ-wise total no of papers selected from the year 2011–2019 containing the segmentation of the respective organs solely or combined. We did not find any paper of the year 2015–2016 which

could match the eligibility criteria of paper inclusion. As depicted in the Fig. 7, most of the research on the topic “Automatic Segmentation of OARs in Thoracic CT images” was done in the year 2019. The interval 2019 contains a total of 14 papers containing the segmentation of the esophagus, heart, trachea, and aorta. Due to the SegTHOR competition organized by one of the researchers, 12 papers were of those researchers who took part in that competition.

The performance evaluation [99] of the different techniques proposed in different papers was done using different

Table 7 Datasets downloaded from various sources or sites used in different papers for training and test purposes

S. no.	Dataset	Source/site
1.	23 cardiovascular CT angiographic volumes [53]	Acquired from 16 to 64 slice Clinical CT scanner
2.	32 Clinical Datasets [55]	Acquired on two different CT scanners (Siemens Sensation 64 and Toshiba Aquilion)
3.	Datasets of the both abdominal and thoracic part [56]	Acquired from the institution scanner (LightSpeed RT 16; GE Medical Systems, Waukesha, WI)
4.	24 low dose chest CT scans [60]	Acquired in the National Lung Screening Trial (NLST) [95]
5.	35 sets of thoracic CT images [61]	Taken from 2017 AAPM Thoracic Auto-segmentation Challenges [96–98]
6.	50 clinical datasets [57]	20: acquired on a Philips GEMINI TF Big Bore or a Siemens Emotion 30: “Multi-Atlas Labeling Beyond the Cranial Vault - Workshop and Challenge” https://www.synapse.org/#!/Synapse:syn3193805/wiki/89480
7.	60 Thoracic CT scans [62]	Obtained from three clinical sites: MD Anderson Cancer Center, Memorial Sloan-Kettering Cancer Center and MAASTRO clinic
8.	60 Thoracic CT scans [37, 39, 74–83]	https://codalab.org/

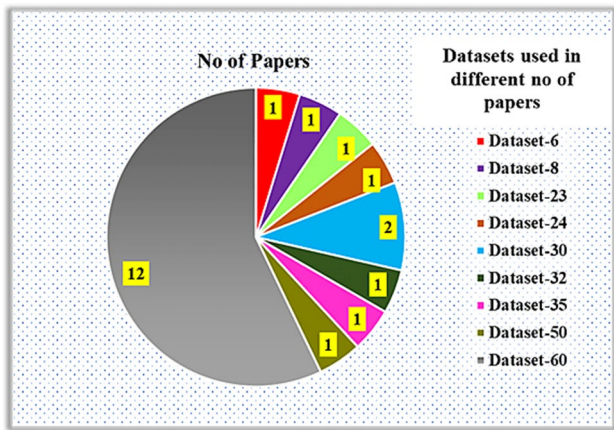


Fig. 5 Distribution of Datasets used in different papers

parameters [100, 101]. Table 8 contains the different parameters that were used by the researchers in calculating the accuracy of their technique for the automatic segmentation of organs focusing (esophagus, heart, trachea, and aorta).

4 Conclusion

This systematic review summarized the literature of 24 papers focusing on the automatic segmentation of Thoracic OARs in CT images. The paper contained the overview of the various techniques being proposed for the segmentation of four organs (Esophagus, heart, trachea and aorta), the dataset used by the authors, the accuracy achieved by them using different parameters. The various challenges faced by the authors in segmenting the organs had been discussed. Most of the techniques proposed in the different papers were based on the deep learning framework and provided excellent segmentation results. The highest accuracy in terms of Dice coefficient and Hausdorff distance

Fig. 6 Year-wise topic related selected papers

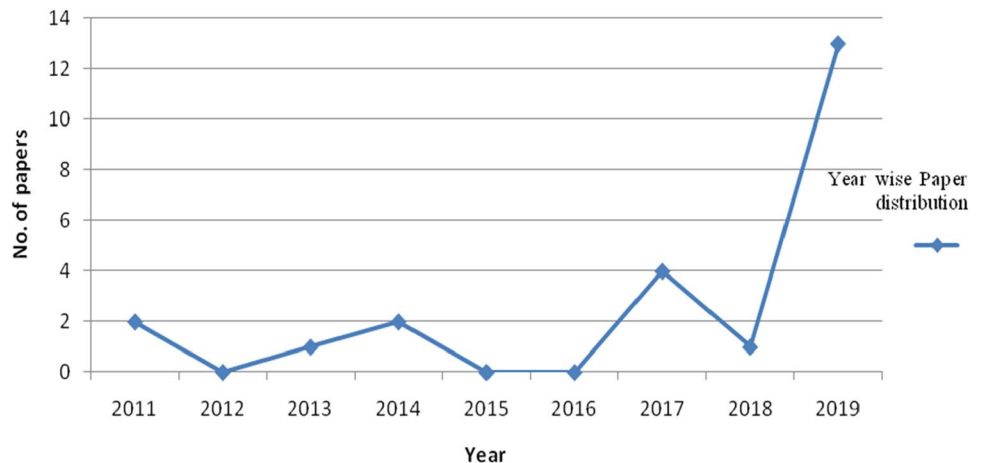


Fig. 7 Year-wise classification of the papers based on the four organs

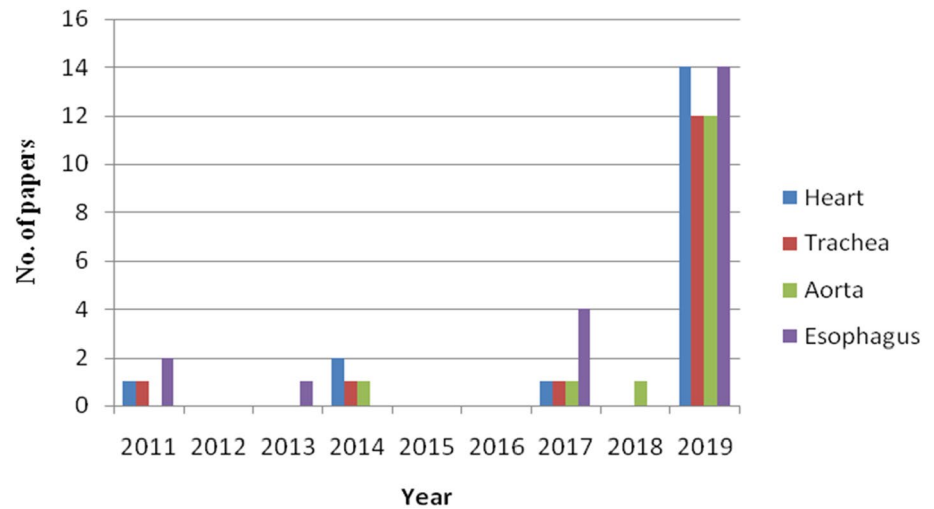


Table 8 Parameters used in different papers for calculating the accuracy of proposed automatic segmentation techniques [102, 103]

S. no	Parameter	Papers	Description	References
1.	Dice similarity coefficient (DSC)	21	It is also known as the Dice index, Dice score, etc. It is calculated as twice the intersection of the automatic and ground truth value to the total value of the automatic and ground truth $DSC = 2(A \cap G / A + G)$	[36, 37, 39, 54, 56–62, 74–83]
2.	Hausdorff Distance (HD)	15	It is defined as the maximum distance from a point in Automatic segmentation to the minimum point in the Ground Truth or vice versa $HD = \max_a \min_b d(a, b)$	[36, 39, 57, 58, 61, 62, 74–80, 82, 83]
3.	Average Symmetrical Surface Distance (ASSD's)	2	The distance b/w the boundary pixels of the segmented part and the ground truth. The average of these boundary pixel distances gives the Average Symmetrical Surface Distance	[57, 60]
4.	Correlation ratio	1	It is defined as the overlapping volume between the segmentation part and the ground truth	[55]
5.	Standard deviation	3	It is defined as the change in the segmented value and the ground truth	[21, 25, 55]
6.	Mean surface distance (MSD)	2	It is the distance of the surface's b/w the automated segmentation and the ground truth	[58, 61]
7.	Dice loss function	2	It is the pixel-wise cross-entropy used to examine each pixel separately comparing the class predictions	[30, 31]
8.	Mean error (ME)	2	It is the term used to calculate the average of all the means b/w segmented and ground truth value	[52, 55]
9.	Euclidean distance mean	1	It is used to find whether the two points are similar or not b/w which the distance is calculated $E_d = \sqrt{(a_2 - a_1)^2 + (b_2 - b_1)^2}$	[53]

was achieved by Han et al. [83] for the four organs. They proposed a multi-resolution VB-Net framework to achieve the accurate segmentation and ranked 1st in the SegTHOR competition. The researchers who want to work or working in the same field can explore the technique for further modification so that the results of segmentation can be enhanced.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Grosu A-L, Sprague LD, Molls M (2006) Definition of target volume and organs at risk. *Biological target volume*. *New Technol Radiat Oncol* 167:167–177
- Kazemifar S et al (2018) Segmentation of the prostate and organs at risk in male pelvic CT images using deep learning. *Biomed Phys Eng Express* 4(5):55003
- T. Morbidity (2016) 6. Organs at risk and morbidity-related concepts and volumes. *J. ICRU vol 13, no 1–2 Rep.* 89. Oxford Univ. Press
- Shirato H et al (2018) Selection of external beam radiotherapy approaches for precise and accurate cancer treatment. *J Radiat Res* 59(suppl_1):i2–i10
- Sharma N et al (2010) Automated medical image segmentation techniques. *J Med Phys* 35(1):3
- Baskar R, Lee KA, Yeo R, Yeoh KW (2012) Cancer and radiation therapy: current advances and future directions. *Int J Med Sci* 9(3):193–199
- Astaraki M et al (2018) Evaluation of localized region-based segmentation algorithms for CT-based delineation of organs at risk in radiotherapy. *Phys Imaging Radiat Oncol* 5:52–57
- McKeown SR, Hatfield P, Prestwich RJD, Shaffer RE, Taylor RE (2015) Radiotherapy for benign disease; assessing the risk of radiation-induced cancer following exposure to intermediate dose radiation. *Br J Radiol* 88(1056):20150405
- Iyer R, Jhingran A (2006) Radiation injury: imaging findings in the chest, abdomen and pelvis after therapeutic radiation. *Cancer Imaging* 6:S31
- Yao J et al (2016) A prospective study on radiation doses to organs at risk (OARs) during intensity-modulated radiotherapy for nasopharyngeal carcinoma patients Classification based on GTV. *Oncotarget* 7(16):21742
- Dische S, Saunders MI, Williams C, Hopkins A, Aird E (1993) Precision in reporting the dose given in a course of radiotherapy. *Radiat Oncol* 29(3):287–293
- Yartsev S, Muren LP, Thwaites DI (2013) Treatment planning studies in radiotherapy. *Radiat Oncol* 109(3):342–343
- Nazemi-Gelyan H et al (2015) Evaluation of organs at risk's dose in external radiotherapy of brain tumors. *Iran J Cancer Prev* 8(1):47–52
- Hammerschmidt S, Wirtz H (2009) Lung cancer : current diagnosis and treatment. *Dtsch Arzteblatt Int* 106(49):809–821
- Aggarwal A et al (2016) The state of lung cancer research: a global analysis. *J Thorac Oncol* 11(7):1040–1050
- Worley S (2014) Lung cancer research is taking on new challenges knowledge of tumors' molecular diversity is opening new pathways to treatment. *Pharm Ther* 39(10):698–704
- Palani D, Venkatalakshmi K (2019) An IoT based predictive modelling for predicting lung cancer using fuzzy cluster based segmentation and classification. *J Med Syst* 43(2):21
- Tong CWS, Wu M, Cho WCS, To KKW (2018) Recent advances in the treatment of breast cancer. *Front Oncol* 8:227
- Caballo M, Boone JM, Mann R, Sechopoulos I (2018) An unsupervised automatic segmentation algorithm for breast tissue classification of dedicated breast computed tomography images. *Med Phys* 45(6):2542–2559
- Lin J, Tsai J, Chang C, Jen Y, Li M, Liu W (2015) Comparing treatment plan in all locations of esophageal cancer. *Medicine* 94(17):1–9
- Münch S, Oechsner M, Combs SE, Habermehl D (2017) DVH- and NTCP-based dosimetric comparison of different longitudinal margins for VMAT-IMRT of esophageal cancer. *Radiat Oncol* 12(1):128
- Schena M, Battaglia AF, Munoz F (2017) Esophageal cancer developed in a radiated field: can we reduce the risk of a poor prognosis cancer? *J Thorac Dis* 9(7):1767–1771
- Ayata HB, Güden M, Ceylan C, Küçük N, Engin K (2011) Original article Comparison of dose distributions and organs at risk (OAR) doses in conventional tangential technique (CTT) and IMRT plans with different numbers of beam in left-sided breast cancer. *Rep Pract Oncol Radiother* 16(3):95–102
- Machiels M et al (2019) Reduced inter-observer and intra-observer delineation variation in esophageal cancer radiotherapy by use of fiducial markers. *Acta Oncol (Madr)* 58(6):943–950
- Lv M, Li Y, Kou B, Zhou Z (2017) Integer programming for improving radiotherapy treatment efficiency. *PLoS ONE* 12(7):1–9
- Nielsen MH et al (2013) Delineation of target volumes and organs at risk in adjuvant radiotherapy of early breast cancer: national guidelines and contouring atlas by the Danish Breast Cancer Cooperative Group. *Acta Oncol (Madr)* 52(4):703–710
- Meyer-Baese A, Schmid V (2014) Computer-aided diagnosis for diagnostically challenging breast lesions in DCE-MRI. *Pattern Recognit Signal Anal Med Imaging* 2:391–420
- Thomson D et al (2014) Evaluation of an automatic segmentation algorithm for definition of head and neck organs at risk. *Radiat Oncol* 9(1):1–12
- Stoloiu-Crişan C, Holban Ş (2013) A Comparison of X-ray image segmentation techniques. *Adv Electr Comput Eng* 13(3):85–92
- Fietkau R (2017) Which fractionation of radiotherapy is best for limited-stage small-cell lung cancer? *Lancet Oncol* 18(8):994–995
- Newhauser WD (2016) A review of radiotherapy-induced late effects research after advanced technology treatments. *Front Oncol* 6:1–11
- Basu T, Bhaskar N (2019) Overview of important 'Organs at Risk' (OAR) in modern radiotherapy for head and neck cancer (HNC). In: Afroz D (ed) *Cancer survivorship*, IntechOpen. <https://doi.org/10.5772/intechopen.80606>
- Whitfield GA, Price P, Price GJ, Moore CJ (2013) Automated delineation of radiotherapy volumes: are we going in the right direction? *Br J Radiol* 86(1021):1–9
- Khalifa F, Beache GM, Gimel G, Suri JS, El-baz A (2011) State-of-the-art medical image registration methodologies: a survey. In: *Multi modality state-of-the-art medical image segmentation and registration methodologies*. Springer, Boston, MA, pp 235–280
- Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. *J Big Data* 6:60
- Satya M, Gali K, Garg N, Vasamsetti S (2015) Dilated U-Net based segmentation of organs at risk in thoracic CT images. In: *SegTHOR@ ISBI*, pp 2–5
- Trullo R, Petitjean C, Ruan S, Dubray B, Nie D, Shen D (2017) Segmentation of organs at risk in thoracic CT images using a SharpMask architecture and conditional random fields. In: *Proceedings of international symposium on biomedical imaging*, pp 1003–1006
- Han M, Ma J, Li Y, Li M, Song Y, Li Q (2015) Segmentation of organs at risk in CT volumes of head, thorax, abdomen, and pelvis. In: *Medical Imaging 2015 Image Processing*, vol 9413, no 2258, p 94133J
- Wang Q et al (2019) 3D enhanced multi-scale network for thoracic organs segmentation. In: *SegTHOR@ ISBI*, no 3, pp 1–5
- Moher D et al (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 6(7):e1000097

41. Ghosh S, Das N, Das I, Maulik U (2019) Understanding deep learning techniques for image segmentation. *ACM Comput Surv* 52(4):1–58
42. Razzak MI, Naz S, Zaib A (2018) Deep learning for medical image processing: overview, challenges and the future. *Lect Notes Comput Vis Biomech* 26:323–350
43. Zhou T, Ruan S, Canu S (2019) A review: deep learning for medical image segmentation using multi-modality fusion. *Array* 3–4:100004
44. An F (2019) Medical image segmentation algorithm based on feedback mechanism CNN. *Contrast Media Mol Imaging* 2019:1–14
45. Suetens P, Verbeek R, Delaere D (1991) Model-based image segmentation : methods and applications. In: *AIME 91*. Springer, Berlin, pp 3–24
46. Kaus MR, McNutt T, Shoenbill J (2006) Model-based segmentation for treatment planning with Pinnacle3. philips white paper. Technical report. Philips Healthcare, Andover, pp 1–4
47. Freedman D et al (2005) Model-based segmentation of medical imagery by matching distributions. *IEEE Trans Med Imaging* 24(3):281–292
48. Paragios N, Duncan J, Ayache N (2015) *Handbook of biomedical imaging: methodologies and clinical research*. Springer, Boston, pp 1–511
49. Duay V, Houhou N, Thiran JP (2005) Atlas-based segmentation of medical images locally constrained by level sets. In: *Proceedings of International Conference on Image Processing (ICIP)*, vol 2, pp 1286–1289
50. Wirth L (1958) Use of chlorpromazine in cough, with particular reference to whooping cough. *Mil Med* 122(3):195–196
51. Pardeshi A. A survey on atlas-based segmentation of medical imaging. *Int J Res Eng Appl Manag* 2(8):1–7. ISSN 2494-9150
52. Kurugol S, Bas E, Erdogmus D, Dy JG, Sharp GC, Brooks DH (2011) Centerline extraction with principal curve tracing to improve 3D level set esophagus segmentation in CT images. In: *Conference Proceedings of the IEEE Engineering in Medicine and Biology Society*, vol 2011, pp 3403–3406
53. Meyer C, Peters J, Weese J (2011) Fully automatic segmentation of complex organ systems: example of trachea, esophagus and heart segmentation in CT images. In: *Medical Imaging 2011 Image Processing*, vol 7962, p 796216
54. Grosgeorge D, Petitjean C, Dubray B, Ruan S (2013) Esophagus segmentation from 3D CT data using skeleton prior-based graph cut. *Comput Math Methods Med* 2013:2–7
55. Larrey-Ruiz J, Morales-Sánchez J, Bastida-Jumilla MC, Menchón-Lara RM, Verdú-Monedero R, Sancho-Gómez JL (2014) Automatic image-based segmentation of the heart from CT scans. *Eurasip J Image Video Process* 2014(1):1–13
56. Schreibmann E, Marcus DM, Fox T (2014) Multiatlas segmentation of thoracic and abdominal anatomy with level set-based local search. *J Appl Clin Med Phys* 15(4):22–38
57. Fechter T et al (2017) Esophagus segmentation in CT via 3D fully convolutional neural network and random walk. *Med Phys* 44(12):6341–6352
58. Yang J et al (2017) Atlas ranking and selection for automatic segmentation of the esophagus from CT scans. *Phys Med Biol* 62(23):9140–9158
59. Trullo R, Petitjean C, Nie D, Shen D, Ruan S (2017) Fully automated esophagus segmentation with a hierarchical deep learning approach. In: *Proceedings of 2017 IEEE International Conference on Signal Image Processing and Applications, ICSIPA 2017*, pp 503–506
60. Noothout J, de Vos B, Wolterink J, Isgum I (2018) Automatic segmentation of thoracic aorta segments in low-dose chest CT. In: *Medical Imaging 2018 Image Processing*, vol 10574. International Society for Optics and Photonics, p 63
61. Dong X et al (2019) Automatic multiorgan segmentation in thorax CT images using U-net-GAN. *Med Phys* 46(5):2157–2168
62. Feng X, Qing K, Tustison NJ, Meyer CH, Chen Q (2019) Deep convolutional neural network for segmentation of thoracic organs-at-risk using cropped 3D images. *Med Phys* 46(5):2169–2180
63. Patil V, Rudrakshi S (2013) Enhancement of medical images using image processing in Matlab. *Int J Eng Res Technol* 2(4):2359–2364
64. Ecabert O et al (2008) Automatic model-based segmentation of the heart in CT images. *IEEE Trans Med Imaging* 27(9):1189–1202
65. Freedman D, Zhang T (20005) Interactive graph cut based segmentation with shape priors. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol 1, pp 755–762
66. Waggoner J et al (2013) 3D materials image segmentation by 2D propagation: a graph-cut approach considering homomorphism. *IEEE Trans Image Process* 22(12):5282–5293
67. Kayalibay B, Jensen G, van der Smagt P (2017) CNN-based segmentation of medical imaging data. *arXiv Preprint arXiv:1701.03056*
68. Iglesias JE, Sabuncu MR (2015) Multi-atlas segmentation of biomedical images: a survey. *Med Image Anal* 24(1):205–219
69. Kirisli HA et al (2010) Fully automatic cardiac segmentation from 3D CTA data: a multi-atlas based approach. In: *Medical Imaging 2010 Image Processing*, vol 7623, May 2014, p 762305
70. Ridler TW, Calvard S (1978) Picture thresholding using an iterative selection method. *IEEE Trans Syst Man Cybern smc-8(8):630–632*
71. Schreibmann E, Yang Y, Boyer A, Li T, Xing L (2005) SU-FF-J-21: image interpolation in 4D CT using a BSpline deformable registration model. *Med Phys* 32(6):1924
72. Warfield SK, Zou KH, Wells WM (2004) Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans Med Imaging* 23(7):903–921
73. Andrew AM (2019) Book reviews “Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science”, in Printed in the United Kingdom © 2000 Cambridge University Press book reviews *LEVEL*, 18, 2000, 2019, pp 89–92
74. Van Harten LD, Noothout JMH, Verhoeff JJC, Wolterink JM, II (2019) Automatic segmentation of organs at risk in thoracic ct scans by combining 2D and 3D convolutional neural networks. In: *SegTHOR@ ISBI*, pp 3–6
75. Vesal S, Ravikumar N, Maier A (2019) A 2D dilated residual U-net for multi-organ segmentation in thoracic CT. In: *CEUR Workshop Proceedings*, vol 2349, pp 2–5
76. Kim S, Jang Y, Han K, Shim H, Chang HJ (2019) A cascaded two-step approach for segmentation of thoracic organs. In: *CEUR Workshop Proceedings*, vol 2349, no. c, pp 3–6
77. Sun F, Chen P, Xu C, Li X, Ma Y (2019) Two-stage network for OAR segmentation, pp 3–4
78. Xu X, Yi Z, He T, Guo J, Wang J (2019) Multi-task learning for the segmentation of thoracic organs at risk in CT images. In: *SegTHOR@ ISBI*, pp 10–13
79. Kondratenko V, Denisenko D, Pimkin A (2012) Segmentation of thoracic organs at risk in CT images using localization and organ-specific CNN. In: *SegTHOR@ ISBI*, pp 4–7
80. Chen H, Zhang L, Wang L, Huang Y (2019) Segmentation of thoracic organs at risk in CT images combining coarse and fine network. In: *SegTHOR@ ISBI*, pp 2–4
81. Lachinov D, Intel (2019) Segmentation of thoracic organs using pixel shuffle. In: *SegTHOR@ ISBI*, pp 1–4

82. Xie Y, Feng M, Huang W, Wang Y (2019) Multi-organ segmentation using simplified dense V-NET with post processing. In: SegTHOR@ ISBI, pp 1–4
83. Han M et al (2019) Segmentation of CT thoracic organs by multi-resolution VB-nets. In: CEUR Workshop Proceedings, vol 2349, pp 1–4
84. Tajbakhsh N et al (2016) Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans Med Imaging* 35(5):1299–1312
85. Pinheiro PO, Lin T, Collobert R, Doll P (2016) Learning to refine object segments. [arXiv:1603.08695v2](https://arxiv.org/abs/1603.08695v2) [cs.CV], pp 1–18
86. Kaur G (2013) An enhancement of classical unsharp mask filter for contrast and edge preservation. *Int J Eng Sci Res Technol* 2(8):2073–2079
87. Hesamian MH, Jia W, He X, Kennedy P (2019) Deep learning techniques for medical image segmentation: achievements and challenges. *J Digit Imaging* 3:582–596
88. Nekrasov V, Ju J, Choi J (2016) Global deconvolutional networks for semantic segmentation. In: British Machine Vision Conference 2016, BMVC 2016, vol 2016-September, pp 124.1–124.14
89. Sakinis T et al (2019) Interactive segmentation of medical images through fully convolutional neural networks. *arXiv Preprint arXiv:1903.08205*, pp 442–448
90. Lei X, Pan H, Huang X (2019) A dilated CNN model for image classification. *IEEE Access* PP:1
91. Young AV, Wortham A, Wernick I, Evans A, Ennis RD (2011) Atlas-based segmentation improves consistency and decreases time required for contouring postoperative endometrial cancer nodal volumes. *Int J Radiat Oncol Biol Phys* 79(3):943–947
92. Li Z, Hoffman EA, Reinhardt JM (2006) Atlas-driven lung lobe segmentation in volumetric X-ray CT images. *IEEE Trans Med Imaging* 25(1):1–16
93. Reed VK et al (2009) Automatic segmentation of whole breast using atlas approach and deformable image registration. *Int J Radiat Oncol Biol Phys* 73(5):1493–1500
94. Mitchell H (2010) STAPLE: simultaneous truth and performance. *IEEE Trans Med Imaging* 23(7):903–921
95. Garc JC, Vaca-boh ML (2011) The national lung screening trial: overview and study design. *Radiology* 258:243–253
96. Jegou S, Drozdal M, Vazquez D, Romero A, Bengio Y (2017) The one hundred layers tiramisu: fully convolutional DenseNets for semantic segmentation. In: IEEE computer society on conference on computer vision and pattern recognition workshops, vol 2017-July, pp 1175–1183
97. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of 30th IEEE conference on Computer Vision and Pattern Recognition, CVPR 2017, vol 2017-January, pp 2261–2269
98. Veit A, Wilber M, Belongie S (2016) Residual networks behave like ensembles of relatively shallow networks. In: Advances in neural information processing systems, pp 550–558
99. Ferrarese FP, Menegaz G (2013) Performance evaluation in medical image segmentation. *Curr Med Imaging Rev* 9:7–17
100. Yeghiazaryan V, Voiculescu I, Yeghiazaryan V, Voiculescu I (2015) An overview of current evaluation methods used in medical image segmentation. Thesis, pp 1–21
101. Prabha DS, Kumar JS (2016) Performance evaluation of image segmentation using objective methods. *Indian J Sci Technol* 9(8):1–8
102. Taha AA, Hanbury A (2015) Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging* 15(1):29
103. Kumar SN, Fred AL, Kumar HA, Varghese PS (2018) Performance metric evaluation of segmentation algorithms for gold standard medical images. Springer, Singapore

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.