European Radiology (2019) 29:1961–1967
https://doi.org/10.1007/s00330-018-5748-9

HEAD AND NECK

Deep-learning-based detection and segmentation of organs at risk in nasopharyngeal carcinoma computed tomographic images for radiotherapy planning

Shujun Liang[1] · Fan Tang[1,2] · Xia Huang[1] · Kaifan Yang[3] · Tao Zhong[1] · Runyue Hu[1] · Shangqing Liu[1] · Xinrui Yuan[1] · Yu Zhang[1]

Received: 4 June 2018 / Revised: 16 August 2018 / Accepted: 10 September 2018 / Published online: 9 October 2018

Abstract

Objective Accurate detection and segmentation of organs at risks (OARs) in CT image is the key step for efficient planning of radiation therapy for nasopharyngeal carcinoma (NPC) treatment. We develop a fully automated deep-learning-based method (termed organs-at-risk detection and segmentation network (ODS net)) on CT images and investigate ODS net performance in automated detection and segmentation of OARs.

Methods The ODS net consists of two convolutional neural networks (CNNs). The first CNN proposes organ bounding boxes along with their scores, and then a second CNN utilizes the proposed bounding boxes to predict segmentation masks for each organ. A total of 185 subjects were included in this study for statistical comparison. Sensitivity and specificity were performed to determine the performance of the detection and the Dice coefficient was used to quantitatively measure the overlap between automated segmentation results and manual segmentation. Paired samples *t* tests and analysis of variance were employed for statistical analysis.

Results ODS net provides an accurate detection result with a sensitivity of 0.997 to 1 for most organs and a specificity of 0.983 to 0.999. Furthermore, segmentation results from ODS net correlated strongly with manual segmentation with a Dice coefficient of more than 0.85 in most organs. A significantly higher Dice coefficient for all organs together ($p = 0.0003 < 0.01$) was obtained in ODS net (0.861 ± 0.07) than in fully convolutional neural network (FCN) (0.8 ± 0.07). The Dice coefficients of each OAR did not differ significantly between different T-staging patients.

Conclusion The ODS net yielded accurate automated detection and segmentation of OARs in CT images and thereby may improve and facilitate radiotherapy planning for NPC.

Key Points

- *A fully automated deep-learning method* (*ODS net*) *is developed to detect and segment OARs in clinical CT images.*
- *This deep-learning-based framework produces reliable detection and segmentation results and thus can be useful in delineating OARs in NPC radiotherapy planning.*
- *This deep-learning-based framework delineating a single image requires approximately 30 s*, *which is suitable for clinical workflows.*

Keywords Image processing · Tomography, x-ray computed · Head and neck neoplasms · Organs at risk · Radiotherapy

✉ Yu Zhang
yuzhang@smu.edu.cn

[1] Guangdong Provincial Key Laboratory of Medical Image Processing, School of Biomedical Engineering, Southern Medical University, No. 1838 Guangzhou Northern Avenue, Baiyun District, Guangzhou 510515, Guangdong, China

[2] Department of Radiation Oncology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, China

[3] Department of Medical Imaging Center, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, China

Abbreviations

CNN	Convolutional neural network
FCN	Fully convolutional neural network
GPU	Graphics processing unit
NPC	Nasopharyngeal carcinoma
OARs	Organs at risk
ODS net	Organs-at-risk detection and segmentation network

Introduction

Nasopharyngeal carcinoma (NPC) is a cancer of the head and neck that occurs with cancer of the throat in the region between the throat and nasal passages [1]. External beam radiotherapy is one of the most effective treatments for NPC. In the treatment planning stage, radiation oncologists or radiologists manually segment organs at risk (OARs) in non-contrast CT (in the following sections, we briefly use CT instead of non-contrast CT) data primarily. After that, radiation physicist calculates the dose and distribution of irradiation on these CT data according to the manual segmentation and formulate treatment plan based on the target of precision radiotherapy that radiation fields are optimized to deliver the prescribed dose to a tumor volume in NPC while sparing the nearby healthy organs. Therefore, segmentation of organs at risk (OARs) is a crucial task for radiation oncologists or radiologists. Manual segmentation is normally adopted in most clinical practices, but this approach is time-consuming and labor intensive. Moreover, manual segmentation often suffers from large inter-operator variability [2], which directly influences the quality of treatment plan and especially the dose distribution for OARs [3]. Accordingly, it is clinically desirable to develop a robust, accurate, and automatic algorithm for the segmentation of OARs in NPC from CT images.

Some of the auto-segmentation methods including model-based segmentations [4–6], atlas-based auto-segmentation [7–9], and hybrid approaches [10–12] have been discussed. Due to the low soft-tissue contrast in CT images and image artifacts (e.g., caused by dental implants), the use of fully automated segmentation methods in radiation therapy planning remains a challenge.

Deep-learning techniques involving convolutional neural network (CNN) are currently considered state-of-the-art approaches in detection, classification, and segmentation of images; this approach arises from the recent success in ImageNet Large Scale Visual Recognition Competition and PASCAL VOC detection and segmentation challenges [13, 14]. The application of deep learning in medical imaging receives interest because of the recent success of deep learning and its promising results. The applications of deep learning include detection of pleural diffusion and cardiomegaly at chest radiography; detection of mediastinal lymph nodes and lung nodules at CT; segmentation of pancreas, brain structures, brain tumors, cardiac structures, and bone and cartilage at MR; and segmentation of cells at microscopic images [15–17]. The first deep-learning-based method used in head and neck CT image segmentation has segmented the spinal cord, mandible, parotid, submandibular glands, larynx, pharynx, eye globes, optic nerves, and optic chiasm, and the results vary from a 37.4% Dice coefficient for chiasm to an 89.5% Dice coefficient for mandible in different datasets [18]. However, the segmentation always fails due to the interaction of organs as revealed by low soft-tissue contrast and blurred boundary. To increase the accuracy of OAR segmentation, we developed and evaluated a fully automated deep-learning-based detection and segmentation method for simultaneous detection and segmentation of OARs in NPC radiotherapy planning by using CT images. In this approach, segmentation was focused on the detection area, thereby reducing the influence of similar or neighboring structures on the results.

Materials and methods

This study was approved by the Institutional Review Board, and written informed consent requirement was waived.

Patients

Two hundred and eight patients from January 2014 to December 2016 for scheduled treatment of nasopharyngeal carcinoma were included in this retrospective study. The inclusion criteria of patients were as follows: (1) newly diagnosed histologically proven nasopharyngeal carcinoma and (2) not given prior treatment of any form. All diagnoses were confirmed histologically according to the guidelines of the TNM staging system. Patients who lost histological confirmation were excluded ($n = 22$). All the patients underwent CT scanning ($n = 186$). One patient was excluded due to the strong artifacts caused by dental implants. One hundred and eighty-five patients fulfilled all criteria and were included in this study. Detailed clinical and demographics of the patients are summarized in Table 1.

CT scanning

All scans were conducted for clinical indications. The CT acquisitions were performed on a Brilliance Big Bore CT scanner (Philips) with the following settings: 120-kVp tube voltage, 275- to 375-mA tube current, 3-mm slice thickness, and 1024 × 1024 × (80–145) imaging matrix that results in an in-plane resolution of 0.363–0.639 × 0.363–0.639.

Table 1 Clinical and demographic data for all patients

N		185
Sex	Male	129
	Female	56
Ages (years)		48.49 ± 11.89 (14–77)
T-staging	T1	27
	T2	42
	T3	84
	T4	32
N-staging	N0	5
	N1	42
	N2	115
	N3	23
M-staging	M0	173
	M1	12

The values in the "Ages" row represented as mean ± standard deviation (minimum and maximum values)

Manual segmentation

Manual segmentation was performed by a radiologist (with 10 years of experience) using ITK-SNAP [19]. For all patients, brain stem, right/left eye balls, right/left lens, right/left optic nerves, right/left parotids, right/left mastoids, right/left mandibles, right/left t-m joints, spinal cord, larynx, and oral cavity were annotated in the CT scanning data. Inclusion of certain OARs into the treatment planning procedure depends on the position of the tumor and image field of view. The OARs mentioned above are mostly included in NPC treatment planning procedure and can be clearly identified by the radiologist in the CT images.

ODS net architecture

Deep-learning models can learn a hierarchy of features by building high-level features from low-level features. The convolutional neural networks (CNNs) are one of the most popular types of deep-learning models, in which the trainable filters and neighborhood pooling operations are applied in an alternating sequence, starting with the input images. The convolution operation sequence can be described as a set of filtered operations of the input images and the hierarchical features. The last layer of network is usually a softmax layer that is actually composed of cross entropy loss. The loss is back propagated to the whole network. The deep-learning-based framework (organs-at-risk detection and segmentation network (ODS net)) was built on the basis of network structures that have been used in studies for natural image object recognition [20] and for image segmentation [21]. An illustration of the framework is shown in Fig. 1. ODS net involves two stages, namely, detection and segmentation. In the detection stage, a CNN proposes organ bounding boxes along with their scores, i.e., probability of them being a certain organ. In the segmentation stage, a second CNN utilizes the proposed bounding boxes to predict segmentation masks for the OARs. The ODS net iteratively estimates the location of target organs and outputs segmented masks and then compares the output masks with reference masks. The reference masks ensure that the network learns the relationship between CT image contexts and these reference masks. Given a new CT data, the trained ODS net can output the location of the target organs and the corresponding segmented masks.

ODS net implementation details

The training data for ODS net consist of non-contrast CT images as input and the manual segmentation as the reference masks. For the training procedure, 3D CT data were encoded into the model as a stack of 2D axial images. All 2D input images were resampled to a 400 × 400 matrix size by using bilinear interpolation before being used as input to the ODS net. Network weights were initialized by using a pre-trained VGG-16 model [22] as adopted in [23] and updated by using stochastic gradient descent [24]. The initial learning rate in the

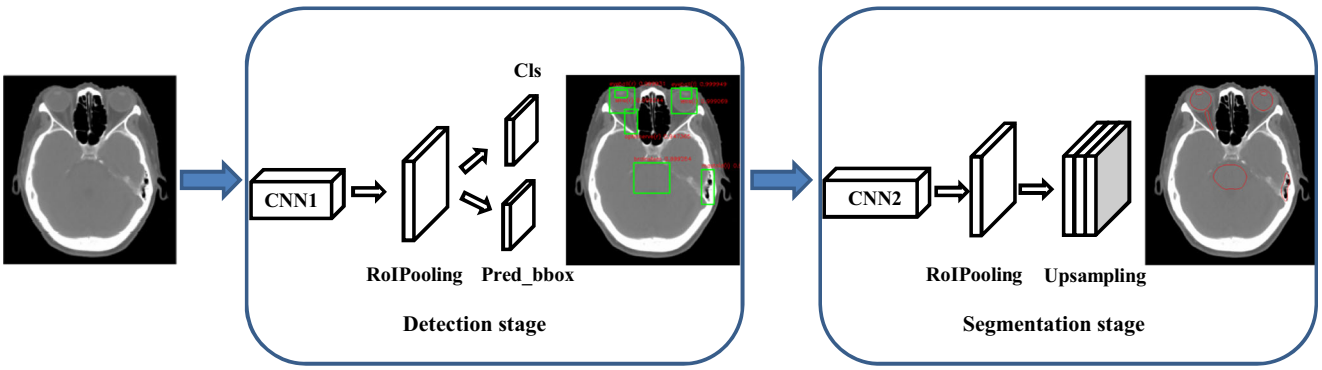

Fig. 1 Flowchart for the construction of the ODS net. The network consists of two stages, detection and segmentation. Detection: a CNN1 proposes organ bounding boxes along with their scores, Segmentation: a second CNN2 utilizes the proposed bounding boxes to predict segmentation masks for the OARs. *CNN* convolutional neural network, *Cls* classification, *Pred_bbox* predicted bounding box

detection stage is 0.001, which was reduced during training. In the segmentation stage, the learning rate is 1e−10, and the momentum is 0.9. ODS net training was performed using 70,000 iteration steps, which correspond to 28 epochs in our training data, to achieve training loss convergence.

In this study, considering the standard cross-validation strategy, we split the dataset into four fixed folds that individually contain approximately the same number of samples. We applied cross-validation, i.e., training the model on three out of four subsets and testing it on the remaining subset. The ODS framework was implemented in a hybrid computing environment involving Python, MATLAB, and C/C++. The ODS network was modified and fine-tuned based on Caffe implementation with graphics processing unit (GPU) parallel computing support [25]. All training and testing procedures were performed on a computer running a 64-bit Linux operating system (Ubuntu 14.04) with an Intel i5 3570-K 3.4-gHz processor, 4 TB of hard disk space, 128 GB of RAM, and a CUDA-enabled Nvidia Titan X 12 GB GPU.

Statistical and data analysis

Sensitivity and specificity were performed to determine the performance of the detection stage of the ODS net. The adjusted Wald method was used to determine the 95% confidence intervals of the sensitivity and specificity [26]. The Dice coefficient, which is a similarity measure ranging from 0 to 1 and describes the overlap between two labels, was calculated for OARs for the segmentation performance. The labels segmented from the ODS net and the ground truth (delineated by an experienced radiologist) were compared. A high Dice coefficient indicates good labeling accuracy. Paired samples *t* tests were used in pairwise comparison between ODS net and fully convolutional neural network (FCN). Analysis of variance was used to compare the Dice coefficients of each OAR between different T-staging patients. Statistical analysis was performed using Matlab version 2012a (MathWorks), and significant difference was defined by $p < 0.01$.

Results

The training stage of ODS net necessitated approximately 18 h, whereas labeling a single input image by using the trained model required approximately 30 s (0.16 s for detection and 30 s for segmentation, a suitable condition for clinical workflows). The training stage of FCN necessitated approximately 24 h, and labeling a single input image by using the trained model required approximately 52 s.

Detection

Sensitivity and specificity values were obtained from the detection results (Table 2). The sensitivity and specificity were as follows: 1 and 0.995 for brainstem, 1 and 0.998 for eyeballs, 0.818 and 0.998 for lens, 1 and 0.999 for larynx, 0.999 and 0.996 for mandible, 1 and 0.997 for oral cavity, 1 and 0.996 for mastoids, 1 and 0.983 for spinal cord, 1 and 0.999 for parotids, 1 and 0.994 for t-m joints, and 1 and 0.997 for optic nerves. Table 2 also shows the 95% confidence intervals on the sensitivity and specificity for each OAR.

Segmentation

Examples of CT images, manual segmentation by an experienced radiologist, and results of the ODS net and FCN are shown in Fig. 2. Manual segmentation and automated segmentation from the ODS net correlated well regarding the automated OAR segmented results. Table 3 shows the mean Dice coefficients of the ODS net. The mean Dice coefficient was high for brainstem (mean ± standard deviation, 0.896 ± 0.03), eyeballs (0.934 ± 0.04), lens (0.836 ± 0.07), larynx (0.87 ± 0.04), mandible (0.913 ± 0.04), oral cavity (0.928 ± 0.03), mastoids (0.823 ± 0.06), spinal cord (0.884 ± 0.07), parotids (0.851 ± 0.05), t-m joints (0.845 ± 0.05), and optic nerves (0.689 ± 0.1). As shown in Fig. 2 and Table 3, ODS net performed better than FCN. The mean Dice coefficients of OARs in ODS net and FCN significantly differed ($p = 0.0003 < 0.01$). Paired samples *t* test of each OAR showed that ODS net provided the significantly higher

Table 2 Performance of the detection stage of ODS net

	Sen (95% CI)	Spe (95% CI)
Brain stem	1.000 (0.990, 1.000)	0.995 (0.99, 0.998)
Eye ball, L	1.000 (0.988, 1.000)	0.997 (0.992, 0.999)
Eye ball, R	1.000 (0.998, 1.000)	0.998 (0.994, 0.999)
Lens, L	0.786 (0.704, 0.851)	0.998 (0.995, 0.999)
Lens, R	0.850 (0.772, 0.904)	0.998 (0.994, 0.999)
Larynx	1.000 (0.973, 1.000)	0.999 (0.997, 1.000)
Mandible, L	1.000 (0.989, 1.000)	0.996 (0.991, 0.998)
Mandible, R	0.997 (0.983, 1.000)	0.995 (0.989, 0.997)
Oral cavity	1.000 (0.989, 1.000)	0.997 (0.993, 0.999)
Mastoid, L	1.000 (0.980, 1.000)	0.995 (0.990, 0.998)
Mastoid, R	1.000 (0.981, 1.000)	0.996 (0.991, 0.998)
Spinal cord	1.000 (0.994, 1.000)	0.983 (0.974, 0.989)
Parotid, L	1.000 (0.967, 1.000)	0.999 (0.996, 1.000)
Parotid, R	1.000 (0.968, 1.000)	0.998 (0.994, 0.999)
T-M joint, L	1.000 (0.957, 1.000)	0.994 (0.989, 0.996)
T-M joint, R	1.000 (0.961, 1.000)	0.994 (0.988, 0.996)
Optic nerve, L	1.000 (0.947, 1.000)	0.996 (0.992, 0.998)
Optic nerve, R	1.000 (0.952, 1.000)	0.997 (0.993, 0.999)

Sen sensitivity, *Spe* specificity, *CI* confidence interval

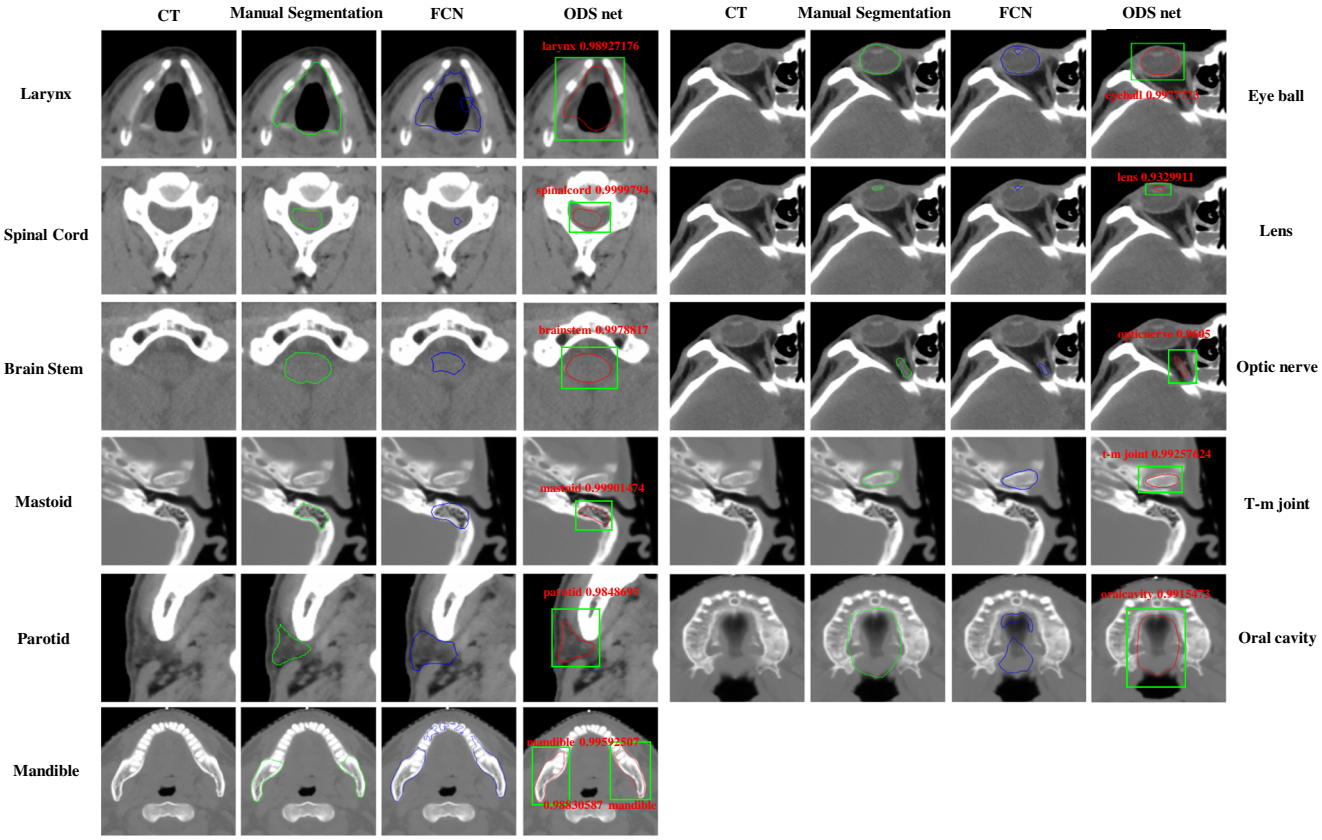

Fig. 2 Comparison of results by using ODS net and FCN. Green line is the manual segmentation and blue line is the segmentation result of FCN. The results of ODS net contain detection results and segmentation results: green bounding boxes is the detection result along with their score (in red), and the red line is the segmentation result. ODS net organs-at-risk detection and segmentation network, FCN fully convolutional neural network

Table 3 Comparisons of the Dice values with ODS net and FCN between different organs at risk

	ODS net			FCN	*t* stat	*t* critical	*p* value
	Left	Right	Average				
Brain stem	–	–	0.896 ± 0.03	0.837 ± 0.14	5.39	2.61	< 0.00001*
Eye balls	0.932 ± 0.04	0.936 ± 0.03	0.934 ± 0.04	0.88 ± 0.12	4.59	2.59	< 0.00001*
Lens	0.83 ± 0.07	0.842 ± 0.07	0.836 ± 0.07	0.77 ± 0.2	3.13	2.63	0.002*
Larynx	–	–	0.87 ± 0.04	0.814 ± 0.14	3.69	2.61	0.0003*
Mandible	0.914 ± 0.04	0.912 ± 0.03	0.913 ± 0.04	0.80 ± 0.11	10.2	2.59	< 0.00001*
Oral cavity	–	–	0.928 ± 0.03	0.9 ± 0.07	3.45	2.6	0.0006*
Mastoids	0.821 ± 0.05	0.824 ± 0.06	0.823 ± 0.06	0.74 ± 0.17	6.26	2.61	< 0.00001*
Spinal cord	–	–	0.884 ± 0.07	0.771 ± 0.22	9.55	2.58	< 0.00001*
Parotids	0.852 ± 0.05	0.85 ± 0.05	0.851 ± 0.05	0.821 ± 0.08	2.38	2.68	0.02
T-m joints	0.846 ± 0.04	0.844 ± 0.06	0.845 ± 0.05	0.828 ± 0.17	2.83	2.65	0.006*
Optic nerves	0.661 ± 0.1	0.717 ± 0.1	0.689 ± 0.1	0.642 ± 0.12	2.93	2.69	0.005*
Overall			0.861 ± 0.07	0.8 ± 0.07	5.71	3.25	0.0003*

The values were the Dice values, represented as mean ± standard deviation

ODS net organs-at-risk detection and segmentation network, *FCN* fully convolutional neural network

* $p < 0.01$ and t stat > t critical was considered significant

Table 4 Segmentation results of different T-staging patients

	T1	T2	T3	T4
Brain stem	0.89 ± 0.04	0.88 ± 0.03	0.89 ± 0.04	0.9 ± 0.03
Eye ball	0.936 ± 0.05	0.926 ± 0.02	0.931 ± 0.03	0.93 ± 0.04
Lens	0.84 ± 0.08	0.824 ± 0.1	0.831 ± 0.07	0.83 ± 0.07
Larynx	0.88 ± 0.03	0.88 ± 0.05	0.85 ± 0.05	0.87 ± 0.05
Mandible	0.919 ± 0.03	0.906 ± 0.03	0.907 ± 0.04	0.92 ± 0.03
Oral cavity	0.935 ± 0.02	0.922 ± 0.03	0.927 ± 0.03	0.92 ± 0.03
Mastoid	0.82 ± 0.04	0.827 ± 0.06	0.819 ± 0.06	0.828 ± 0.06
Spinal cord	0.876 ± 0.08	0.88 ± 0.07	0.89 ± 0.05	0.87 ± 0.09
Parotid	0.837 ± 0.06	0.844 ± 0.05	0.864 ± 0.05	0.85 ± 0.04
T-M joint	0.87 ± 0.02	0.85 ± 0.06	0.833 ± 0.04	0.83 ± 0.06
Optic nerve	0.68 ± 0.12	0.69 ± 0.07	0.68 ± 0.09	0.67 ± 0.09

The values were the Dice values, represented as mean ± standard deviation
T T-staging

Dice coefficients than FCN for 10 out of 11 OARs with $p < 0.01$. Moreover, Table 3 also shows the t and p values of paired samples t test of each OAR. Further, the Dice coefficients of each OAR did not differ significantly between different T-staging patients (Table 4).

Discussion

This study developed and investigated automated detection and segmentation of OARs in NPC patients by a novel deep-learning-based approach (ODS net). The method proved to allow for accurate detection and segmentation in CT images. Detection accuracy was high. Automated segmentation correlated well with manual segmentations and the reliable segmentation results were shown in different T-staging patients. Overlap measured by the Dice coefficients was high for each OAR. Compared with FCN which directly segments a whole image without the detection operation, ODS net provided the significantly higher Dice coefficients.

This achievement of detection and segmentation can be explained by the fact the general appearance of OARs and surrounding tissues remains similar among CT images of different patients and OARs have a very relatively stable position in the image that can be accurately modeled by deep-learning approaches to gain the high performance of detection and segmentation. The added detection operation plays an important role in segmentation operation that can only focus on a certain region containing an organ and then segment this organ from the background, thus avoiding the influence of other organs.

The accurate automated detection and segmentation offer improved approaches to clinical assessment in the imaging routine, as these may allow for a more precise therapy planning. Further, manual segmentation in clinical routine is time-consuming; an automated segmentation is therefore warranted. What is more, one advantage of deep leaning is its transferable learning performance. Transferable learning refers to the transfer of trained model parameters to a new model to help train the new model. That is, our ODS net is easily trained to be a new model in other clinical applications by using new medical images.

Limitation

This study has several limitations. First, evaluation of the proposed techniques relied on manual segmentation that was delineated by a radiologist. Thus, all manual segmentation of different subjects as delineated by a radiologist must be consistent. When a manual segmentation is conducted by another radiologist to evaluate the proposed techniques, the accuracy of the result is affected. Although the reduction in accuracy of the result is minimal, the manual segmentation conducted by a certain radiologist possibly does not reflect the true standard of the reference used in segmentation. Second, the proposed technique was evaluated in subjects undergoing a clinical CT scan (Brilliance Big Bore, Philips) but not evaluated in subjects examined by other CT scanners. Further evaluation of ODS net in subjects examined by other CT scanners is necessary to determine its robustness to CT imaging devices. Third, in our study, since the treatment plan is conducted on the non-contrast CT scanning, the ODS net is trained only on non-contrast CT data and this may limit the accuracy of detection and segmentation. Nevertheless, we plan to adapt the ODS net with combination of the non-contrast CT images and contrast CT images, even the MRI scanning after registration as training data to yield improved results on the non-contrast CT images in future studies. At last, we only consider the detection and segmentation of serval anatomical structures of NPC patients. We plan to include much more anatomical structures, such as optic chiasm, pituitary, and thyroid, even the NPC tumor volume to transfer the ODS net for the entire purpose of planning stereotactic radiation therapy.

Conclusion

Fully automated detection and segmentation of OARs in clinical non-contrast CT images of NPC patients based on deep-learning method (ODS net) were accurate and reliable. The accurate detection and segmentation may improve and facilitate therapy planning for NPC patients.

Acknowledgements The author(s) would like to thank the reviewers for their fruitful comments.

Funding This study has received funding by the National Natural Science Foundation of China under Grant No. 61671230 and No.31271067, the Science and Technology Program of Guangdong Province under Grant No. 2017A020211012, the Guangdong Provincial

Key Laboratory of Medical Image Processing under Grant No.2014B030301042, and the Science and Technology Program of Guangzhou under Grant No. 201607010097.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Yu Zhang.

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

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- retrospective
- experimental
- performed at one institution

References

1. Mohammed MA, Ghani MKA, Hamed RI, Ibrahim DA (2017) Review on nasopharyngeal carcinoma: concepts, methods of analysis, segmentation, classification, prediction and impact: a review of the research literature. J Comput Sci 21:283–298
2. Brouwer CL, Steenbakkers RJ, Heuvel EVD et al (2012) 3D variation in delineation of head and neck organs at risk. Radiat Oncol 7(1):1–10
3. Weiss E, Hess CF (2003) The impact of gross tumor volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences. Strahlenther Onkol 179(1):21–30
4. Faggiano E, Fiorino C, Scalco E et al (2011) An automatic contour propagation method to follow parotid gland deformation during head-and-neck cancer tomotherapy. Phys Med Biol 56(3):775–791
5. Rueckert D, Frangi AF, Schnabel JA (2003) Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. IEEE Trans Med Imaging 22(8):1014–1025
6. Fritscher KD, Grünerbl A, Schubert R (2007) 3D image segmentation using combined shape-intensity prior models. Int J Comput Assist Radiol Surg 1(6):341–350
7. Han X, Hibbard LS, O'Connell NP, Willcut V (2010) Automatic segmentation of parotids in head and neck CT images using multiatlas fusion. In: van Ginneken B, Murphy K, Heimann T, Pekar V, Deng X (eds.) Med Image Analysis for the Clinic:A Grand Challenge, Beijing, 297–304
8. Han X, Hibbard LS, O'Connell NP (2009) Automatic segmentation of head and neck CT images by GPU-accelerated multi-atlas fusion. On 3D Segmentation. Retrieved from http://www.midasjournal.org/handle/10380/3111
9. Daisne JF, Blumhofer A (2013) Atlas-based automatic segmentation of head and neck organs at risk and nodal target volumes: a clinical validation. Radiat Oncol 8(1):154
10. Fritscher KD, Peroni M, Zaffino P, Spadea MF, Schubert R, Sharp G (2014) Automatic segmentation of head and neck CT images for radiotherapy treatment planning using multiple atlases, statistical appearance models, and geodesic active contours. Med Phys 41(5):051910
11. Gorthi S, Duay V, Houhou N et al (2009) Segmentation of head and neck lymph node regions for radiotherapy planning using active contour-based atlas registration. IEEE Journal of Selected Topics in Signal Processing 3(1):135–147
12. Qazi AA, Pekar V, Kim J, Xie J, Breen SL, Jaffray DA (2011) Auto-segmentation of normal and target structures in head and neck CT images: a feature-driven model-based approach. Med Phys 38(11):6160–6170
13. Russakovsky O, Deng J, Su H et al (2014) ImageNet large scale visual recognition challenge. Int J Comput Vis 115(3):211–252
14. Girshick R, Donahue J, Darrell T, Malik J (2016) Region-based convolutional networks for accurate object detection and segmentation. IEEE Trans Pattern Anal Mach Intell 38(1):142–158
15. Lakhani P, Sundaram B (2017) Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 284(2):574–582
16. Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB (2017) Deep learning MR imaging-based attenuation correction for PET/MR imaging. Radiology 286(2):170700
17. Akram SU, Kannala J, Eklund L, Heikkilä J (2016) Cell Segmentation proposal network for microscopy image analysis. IEEE International Conference on Image Processing
18. Ibragimov B, Xing L (2017) Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. Med Phys 44(2):547
19. Yushkevich PA, Piven J, Hazlett HC et al (2006) User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31(3):1116–1128
20. Ren S, He K, Girshick R, Sun J (2017) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149
21. Long J, Shelhamer E, Darrell T (2017) Fully convolutional networks for semantic segmentation. IEEE Trans Pattern Anal Mach Intell 39(4):640–651
22. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. ArXiv: 1409.1556 [cs.CV]
23. Zhang L, Lin L, Liang X, He K (2016) Is faster R-CNN doing well for pedestrian detection? In: Leibe B, Matas J, Sebe N, Welling M (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, vol 9906. Springer, Cham pp 443–457
24. Bottou L (2010) Large-scale machine learning with stochastic gradient descent. In: Lechevallier Y, Saporta G (eds) Proceedings of COMPSTAT'2010. Physica-Verlag HD 177–186
25. Jia Y, Shelhamer E, Donahue J et al (2014) Caffe: convolutional architecture for fast feature embedding. arXiv:1408.5093v1 [cs.CV]
26. Agresti A, Coull BA (1998) Approximate is better than "exact" for interval estimation of binomial proportions. Am Stat 52(2):119–126