Deep Probability Contour Framework for Tumour Segmentation and Dose Painting in PET Images

Wenhui Zhang(✉) and Surajit Ray

University of Glasgow, Glasgow G12 8QQ, UK
W.Zhang.2@research.gla.ac.uk, Surajit.Ray@glasgow.ac.uk

Abstract. The use of functional imaging such as PET in radiotherapy (RT) is rapidly expanding with new cancer treatment techniques. A fundamental step in RT planning is the accurate segmentation of tumours based on clinical diagnosis. Furthermore, recent tumour control techniques such as intensity modulated radiation therapy (IMRT) dose painting requires the accurate calculation of multiple nested contours of intensity values to optimise dose distribution across the tumour. Recently, convolutional neural networks (CNNs) have achieved tremendous success in image segmentation tasks, most of which present the output map at a pixel-wise level. However, its ability to accurately recognize precise object boundaries is limited by the loss of information in the successive downsampling layers. In addition, for the dose painting strategy, there is a need to develop image segmentation approaches that reproducibly and accurately identify the high recurrent-risk contours. To address these issues, we propose a novel hybrid-CNN that integrates a kernel smoothing-based probability contour approach (KsPC) to produce contour-based segmentation maps, which mimic expert behaviours and provide accurate probability contours designed to optimise dose painting/IMRT strategies. Instead of user-supplied tuning parameters, our final model, named KsPC-Net, applies a CNN backbone to automatically learn the parameters and leverages the advantage of KsPC to simultaneously identify object boundaries and provide probability contour accordingly. The proposed model demonstrated promising performance in comparison to state-of-the-art models on the MICCAI 2021 challenge dataset (HECKTOR).

Keywords: Image Segmentation · PET imaging · Probability Contour · Dose painting · Deep learning

1 Introduction

Fluorodeoxyglucose Positron Emission Tomography (PET) is widely recognized as an essential tool in oncology [10], playing an important role in the stag-

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43901-8_51.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14223, pp. 534–543, 2023.
https://doi.org/10.1007/978-3-031-43901-8_51

ing, monitoring, and follow-up radiotherapy (RT) planning [2,19]. Delineation of Region of Interest (ROI) is a crucial step in RT planning. It enables the extraction of semi-quantitative metrics such as standardized uptake values (SUVs), which normalize pixel intensities based on patient weight and radiotracer dose [20]. Manual delineation is a time-consuming and laborious task that is prone to poor reproducibility in medical imaging, and this is particularly true for PET, due to its low signal-to-noise ratio and limited spatial resolution [10]. In addition, manual delineation depends heavily on the expert's prior knowledge, which often leads to large inter-observer and intra-observer variations [8]. Therefore, there is an urgent need for developing accurate automatic segmentation algorithms in PET images which will reduce expert workload, speed up RT planning while reducing intra-observer variability.

In the last decade, CNNs have demonstrated remarkable achievements in medical image segmentation tasks. This is primarily due to their ability to learn informative hierarchical features directly from data. However, as illustrated in [9,23], it is rather difficult for CNNs to recognize the object boundary precisely due to the information loss in the successive downsampling layers. Despite the headway made in using CNNs, their applications have been restricted to the generation of pixel-wise segmentation maps instead of smooth contour. Although CNNs may yield satisfactory segmentation results, low values of the loss function may not always indicate a meaningful segmentation. For instance, a noisy result can create incorrect background contours and blurry object boundaries near the edge pixels [6]. To address this, a kernel smoothing-based probability contour (KsPC) approach was proposed in our previous work [22]. Instead of a pixel-wise analysis, we assume that the true SUVs come from a smooth underlying spatial process that can be modelled by kernel estimates. The KsPC provides a surface over images that naturally produces contour-based results rather than pixel-wise results, thus mimicking experts' hand segmentation. However, the performance of KsPC depends heavily on the tuning parameters of bandwidth and threshold in the model, and it lacks information from other patients.

Beyond tumour delineation, another important use of functional images, such as PET images is their use for designing IMRT dose painting (DP). In particular, dose painting uses functional images to paint optimised dose prescriptions based on the spatially varying radiation sensitivities of tumours, thus enhancing the efficacy of tumour control [14,18]. One of the popular DP strategies is dose painting by contours (DPBC), which assigns a homogeneous boost dose to the subregions defined by SUV thresholds. However, there is an urgent need to develop image segmentation approaches that reproducibly and accurately identify the high recurrent-risk contours [18]. Our previously proposed KsPC provides a clear framework to calculate the probability contours of the SUV values and can readily be used to define an objective strategy for segmenting tumours into subregions based on metabolic activities, which in turn can be used to design the IMRT DP strategy.

To address both tumour delineation and corresponding dose painting challenges, we propose to combine the expressiveness of deep CNNs with the versa-

tility of KsPC in a unified framework, which we call KsPC-Net. In the proposed KsPC-Net, a CNN is employed to learn directly from the data to produce the pixel-wise bandwidth feature map and initial segmentation map, which are used to define the tuning parameters in the KsPC module. Our framework is completely automatic and differentiable. More specifically, we use the classic UNet [17] as the CNN backbone and evaluate our KsPC-Net on the publicly available MICCAI HECKTOR (HEad and neCK TumOR segmentation) challenge 2021 dataset. Our proposed KsPC-Net yields superior results in terms of both Dice similarity scores and Hausdorff distance compared to state-of-art models. Moreover, it can produce contour-based segmentation results which provide a more accurate delineation of object edges and provide probability contours as a byproduct, which can readily be used for DP planning.

2 Methods

2.1 Kernel Smoothing Based Probability Contour

Kernel-based method and follow up approach of modal clustering [13,16] have been used to cluster high-dimensional random variables and natural-scene image segmentation. In this work, we propose to model the pixel-specific SUV as a discretized version of the underlying unknown smooth process of "metabolic activity". The smooth process can then be estimated as kernel smoothed surface of the SUVs over the domain of the entire slice. In particular, let $Y = (y_1, y_2, ..., y_N)$ denote N pixel's SUV in a 2D PET image sequentially, and $\mathbf{x_i} = (x_{i1}, x_{i2}), i = 1, ..., N$ denote position vector with x_{i1} and x_{i2} being the position in 2D respectively. Note that $\mathbf{x_i} \in \mathbb{R}^d$ and $d = 2$ in our case. We assume that for each position vector $\mathbf{x}$, the SUV represents the frequency of $\mathbf{x}$ appearing in the corresponding grid. The SUV surface can therefore be modelled as kernel density estimate (KDE) [3,15] of an estimated point $\mathbf{x}$, which is defined generally as

$$\hat{f}(\mathbf{x}; \mathbf{H}) = \left(\sum_{i=1}^{N} y_i \right)^{-1} \sum_{t=1}^{y_1+...+y_N} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_t), \tag{1}$$

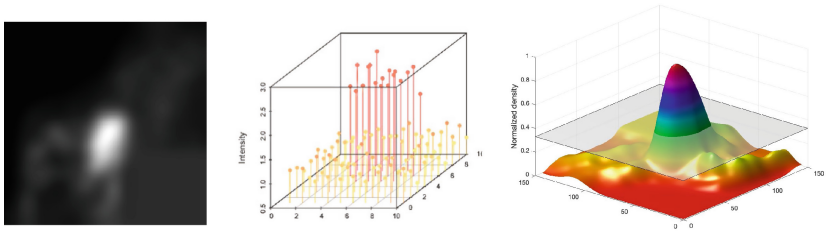
where K is a kernel function and $\mathbf{H}$ is a symmetric, positive definite, $d \times d$ matrix of smoothing tuning parameters, called bandwidth which controls the orientation and amount of smoothing via the scaled kernel $K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K(|\mathbf{H}|^{-1/2}\,\mathbf{x})$. On the other hand, since $\mathbf{x}_t$ is counted y_i times at the same position, Eq. 1 can be further simplified as

$$\hat{f}(\mathbf{x}; \mathbf{H}) = \left(\sum_{i=1}^{N} y_i \right)^{-1} \sum_{i=1}^{N} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i) y_i. \tag{2}$$

A scaled kernel is positioned so that its mode coincides with each data point $\mathbf{x}_i$ which is expressed mathematically as $K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i)$. In this paper, we have used a Guassian kernel which is denoted as:

$$K_{\mathbf{H}}(\mathbf{x} - \mathbf{x_i}) = (2\pi)^{-1/2} |\mathbf{H}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x_i})^T \mathbf{H}^{-1}(\mathbf{x} - \mathbf{x_i})\right),$$

which is a normal distribution with mean $\mathbf{x}_i$ and variance-covariance matrix $\mathbf{H}$. Therefore, we can interpret $\hat{f}$ in Eq. (2) as the probability mass of the data point $\mathbf{x}$ which is estimated by smoothing the SUVs of the local neighbourhood using the Gaussian kernel. The resulting surface built by the KDE process can be visualized in Fig. 1(c). By placing a threshold plane, a contour-based segmentation map can naturally be obtained. Note that one can obtain a pixel-based segmentation map, by thresholding the surface at the observed grid points.



(a) (b) (c)

Fig. 1. A visualization example of how KsPC works: (a) An example of a PET image (b) Grid-level intensity values as observations (c) The resulting smoothed surface built by KsPC with a threshold plane.

After delineating the gross tumour volume, a follow-up application of the kernel smoothed surface is to construct probability contours. Mathematically, a $100\,\omega\%$ region of a density f is defined as the level set $\mathcal{L}(f_\omega) = \{f(\mathbf{x}) \geq f_\omega\}$ with its corresponding contour level f_ω such that $\mathcal{P}(\mathbf{x} \in \mathcal{L}(f_\omega) = 1 - \omega$, where $\mathbf{x}$ is a random variable and $\mathcal{L}(f_\omega)$ has a minimal hypervolume [11]. In other words, for any $\omega \in (0, 1)$, the $100\,\omega\%$ contour refers to the region with the smallest area which encompasses $100\,\omega\%$ of the probability mass of the density function [11]. In practice, f_ω can be estimated using the following result.

Result. The estimated probability contour level f_ω can be computed as the ω-th quantile of $\hat{f}_\omega$ of $\hat{f}(\mathbf{x}_1; \mathbf{H}), ..., \hat{f}(\mathbf{x}_n; \mathbf{H})$ (Proof in supplementary materials).

The primary advantage of utilizing probability contours is their ability to assign a clear probabilistic interpretation on the defined contours, which are scale-invariant [5]. This provides a robust definition of probability under the perturbation of the input data. In addition, these contours can be mapped to the IMRT dose painting contours, thus providing an alternative prescription strategy for IMRT. Examples of the application of probability contours will be demonstrated and explained in Sect. 4.2.

2.2 The KsPC-Net Architecture

In the KsPC module, the model performance heavily depends on the bandwidth matrix $\mathbf{H}$ and it is often assumed that each kernel shares the same scalar bandwidth parameter. However, one may want to use different amounts of smoothing

in the kernel at different grid positions. The commonly used approach for bandwidth selection is cross-validation [4], which is rather time-consuming even in the simpler scalar situation. In this paper, we instead use the classic 2D-Unet [17] as our CNN backbone to compute the pixel-level bandwidth feature map, which informs the KsPc bandwidth. Additionally, we obtain the optimal threshold for constructing the KsPC contour from the initial segmentation map. As shown in Fig. 2 the proposed KsPC-Net integrates the KsPC approach with a CNN backbone (UNet) in an end-to-end differentiable manner. First, the initial segmentation map and pixel-level bandwidth parameter map $h(x_{i1}, x_{i2})$ of KsPC are learned from data by the CNN backbone. Then the KsPC module obtains the quantile threshold value for each image by identifying the quantile corresponding to the minimum SUV of the tumour class in the initial segmentation map. The next step involves transmitting the bandwidth map, quantile threshold, and raw image to KsPC module to generate the segmentation map and its corresponding probability contours. The resulting output from KsPC is then compared to experts' labels using a Dice similarity loss function, referred to KsPC loss. Additionally, the initial Unet segmentation can produce another loss function, called CNN loss, which serves as an auxiliary supervision for the CNN backbone. The final loss can then be constructed as the weighted sum of CNN loss and KsPC loss. By minimizing the final loss, the error can be back-propagated through the entire KsPC architecture to guide the weights updating the CNN backbone.

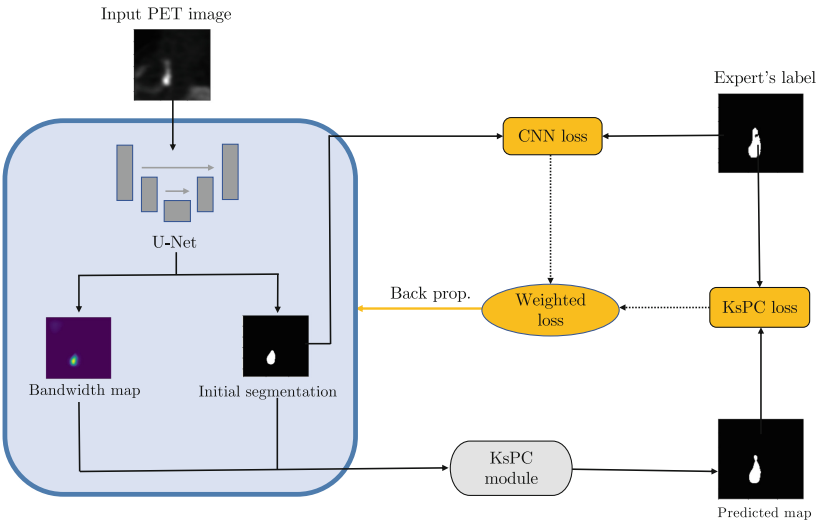


Fig. 2. The architecture of KsPC-Net. KsPC-Net is an end-to-end trainable framework with KsPC module. In contrast to the traditional kernel smoothing process, bandwidth parameters learned by the CNN backbone are pixel-wise functions $h(x_{i1}, x_{i2})$ rather than scalars.

2.3 Loss Function

The Dice similarity coefficient is widely employed to evaluate segmentation models. We utilize the Dice loss function to optimize the model performance during training, which is defined as:

$$\mathcal{L}_{dice}(y, \hat{y}) = 1 - \frac{2\sum_i^N y_i \hat{y}_i}{\sum_i^N y_i + \sum_i^N \hat{y}_i + \epsilon},$$

where y_i is the label from experts and $\hat{y_i}$ is the predicted label of i-th pixel. N is the total number of pixels and ϵ is a small constant in case of zero division. As shown in Fig. 2, we construct the weighted Dice loss to train the model as follows:

$$\mathcal{L}_{final} = \alpha * \mathcal{L}_{CNN} + (1 - \alpha) * \mathcal{L}_{KsPC},$$

where $\mathcal{L}_{final}$ denotes the weighed dice loss while $\mathcal{L}_{CNN}$ and $\mathcal{L}_{KsPC}$ denotes the CNN loss and KsPC loss, respectively. In addition, α is a balancing parameter and is set to be 0.01 in this work.

3 Experiments

3.1 Dataset

The dataset is from the HECKTOR challenge in MICCAI 2021 (HEad and neCK TumOR segmentation challenge). The HECKTOR training dataset consists of 224 patients diagnosed with oropharyngeal cancer [1]. For each patient, FDG-PET input images and corresponding labels in binary description (0 s and 1 s) for the primary gross tumour volume are provided and co-registered to a size of $144 \times 144 \times 144$ using bounding box information encompassing the tumour. Five-fold cross-validation is used to generalize the performance of models.

3.2 Implementation Details

We used Python and a trained network on a NVIDIA Dual Quadro RTX machine with 64 GB RAM using the PyTorch package. We applied a batch size of 12 and the Adam algorithm [12] with default parameters to minimize the dice loss function. All models were trained for 300 epochs. Each convolutional layer is followed by RELU activation and batch normalization.

4 Results

4.1 Results on HECKTOR 2021 Dataset

To evaluate the performance of our KsPC-Net, we compared it with the results of 5-fold cross-validation against three widely-used models namely, the standard 2D Unet, the 2D residual Unet and the 3D Unet. Additionally, we compare our

performance against newly developed approaches MSA-Net [7] and CCUT-Net [21] which were reported in the HECKTOR 2021 challenges [1]. To quantify the performance, we report several metrics including Dice similarity scores, Precision, Recall, and Hausdorff distance. Table 1 shows the quantitative comparison of different approaches on HECKTOR dataset. It is worth mentioning that since our KsPC-Net is in a 2D Unet structure, the Hausdorff distance here was calculated on slice averages to use a uniform metric across all 2D and 3D segmentation models. However, the results of 2D Hausdorff distances of MSA-Net and CCUT-Net are not available and therefore they are omitted in the table of comparison.

Table 1. Mean segmentation results of different models and our proposed model. The model with best performance for each metric is indicated in **bold***.

Method	Dice Score	Hausdorff Dist	Precision	Recall
2D-Unet	0.740	0.561	0.797	0.873
Res-Unet	0.680	0.611	0.740	0.841
3D-Unet	0.764	0.546	**0.839***	0.797
MSA-Net	0.757	-	0.788	0.785
CCUT-Net	0.750	-	0.776	0.804
KsPC-Net(Ours)	**0.768***	**0.521***	0.793	**0.911***

The results clearly demonstrate that the proposed KsPC-Net is effective in segmenting H&N tumours, achieving a mean Dice score of 0.768. This represents a substantial improvement over alternative approaches, including 2D-UNet (0.740), 3D U-Net (0.764), Residual-Unet (0.680), MSA-Net (0.757) and CCUT-Net (0.750). While we acknowledge that there was no statistically significant improvement compared to other SOTA models, it is important to note that our main goal is to showcase the ability to obtain probability contours as a natural byproduct while preserving state-of-the-art accuracy levels. On the other hand, in comparison to the baseline 2D-Unet model, KsPC-Net yields a higher Recall (0.911) with a significant improvement (4.35%), indicating that KsPC-Net generates fewer false negatives (FN). Although the Precision of KsPC-Net is slightly lower than the best-performing method (3D Unet), it achieves a relatively high value of 0.793. In addition, the proposed KsPC-Net achieves the best performance on Hausdorff distance among the three commonly used Unet models (2D-Unet, Res-Unet and 3D-Unet), which indicates that KsPC-Net exhibits a stronger capacity for accurately localizing the boundaries of objects. This is consistent with the mechanisms of KsPC, which leverages neighbouring weights to yield outputs with enhanced smoothness.

4.2 Probability Contours

One of the byproducts of using the kernel-smoothed densities to model the SUVs is the associated probability contours, which can be readily used to develop a

comprehensive inferential framework and uncertainty quantification. For example, Fig. 3 provides two examples of PET image segmentation maps by KsPC-Net and their corresponding probability contours in the last column. There are 5 contours in each case which is linear in probability space, in the sense that each contour encloses 10%, 30%, 50%, 70% and 90% probability mass respectively (from inner to outer), thus dividing the density surface into subregions with attached probability mass.

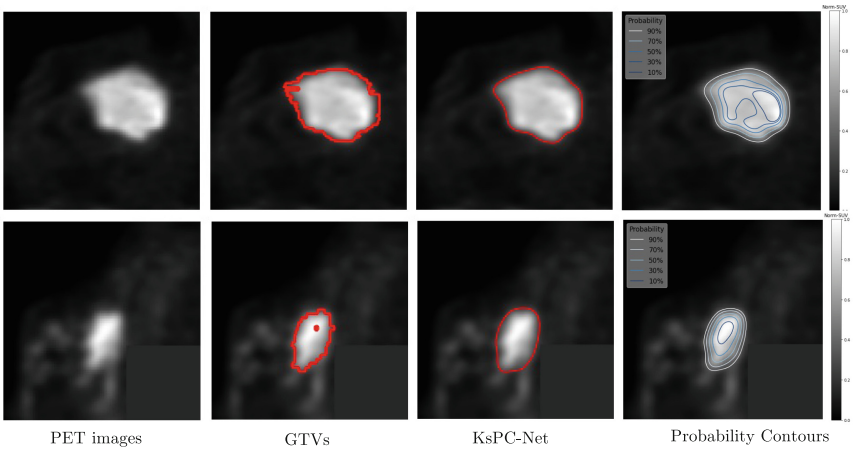


Fig. 3. Illustrations of the segmentation results and probability contours on two examples. The four columns are original PET images, ground truth provided by experts, segmentation maps from KsPC-Net and its probability contours (in 10%, 30%, 50%, 70%, 90% respectively).

These probability contours can provide a rigorous framework for designing the number and magnitude of SUV thresholds for designing optimal DP strategies. Since the SUVs are smoothed by the kernel density heights, the inner 10% probability contour corresponds to the subregion with relatively higher SUVs. In other words, there is an inverse mapping between the probability contours and the amount of dose boost assigned to subvolumes.

5 Conclusion

In this paper, we present a novel network, KsPC-Net, for the segmentation in 2D PET images, which integrates KsPC into the UNet architecture in an end-to-end differential manner. The KsPC-Net utilizes the benefits of KsPC to deliver both contour-based and grid-based segmentation outcomes, leading to improved precision in the segmentation of contours. Promising performance was achieved by our proposed KsPC-Net compared to the state-of-the-art approaches on the MICCAI 2021 challenge dataset (HECKTOR). It is worth mentioning that the

architecture of our KsPC-Net is not limited to Head & Neck cancer type and can be broadcast to different cancer types. Additionally, a byproduct application of our KsPC-Net is to construct probability contours, which enables probabilistic interpretation of contours. The subregions created by probability contours allow for a strategy planning for the assigned dose boosts, which is a necessity for the treatment planning of radiation therapy for cancers.

Acknowledgement. This work was supported by the Carnegie Trust of Scotland PhD Scholarships Fund.

References

1. Andrearczyk, V., Oreiller, V., Hatt, M., Depeursinge, A.: Head and Neck Tumor Segmentation and Outcome Prediction. LNCS, vol. 13209. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98253-9
2. Bai, B., Bading, J., Conti, P.S.: Tumor quantification in clinical positron emission tomography. Theranostics **3**(10), 787 (2013)
3. Bowman, A., Foster, P.: Density based exploration of bivariate data. Stat. Comput. **3**, 171–177 (1993)
4. Bowman, A.W.: An alternative method of cross-validation for the smoothing of density estimates. Biometrika **71**(2), 353–360 (1984)
5. Chacón, J.E., Duong, T.: Multivariate Kernel Smoothing and Its Applications. Chapman and Hall/CRC, Boca Raton (2018)
6. Chen, X., Williams, B.M., Vallabhaneni, S.R., Czanner, G., Williams, R., Zheng, Y.: Learning active contour models for medical image segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11632–11640 (2019)
7. Cho, M., Choi, Y., Hwang, D., Yie, S.Y., Kim, H., Lee, J.S.: Multimodal spatial attention network for automatic head and neck tumor segmentation in FDG-PET and CT images. In: Andrearczyk, V., Oreiller, V., Hatt, M., Depeursinge, A. (eds.) HECKTOR 2021. LNCS, vol. 13209, pp. 75–82. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98253-9_6
8. Gudi, S., et al.: Interobserver variability in the delineation of gross tumour volume and specified organs-at-risk during IMRT for head and neck cancers and the impact of FDG-PET/CT on such variability at the primary site. J. Med. Imaging Radiat. Sci. **48**(2), 184–192 (2017)
9. Hatamizadeh, A., et al.: Deep active lesion segmentation. In: Suk, H.-I., Liu, M., Yan, P., Lian, C. (eds.) MLMI 2019. LNCS, vol. 11861, pp. 98–105. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32692-0_12
10. Hatt, M., et al.: The first MICCAI challenge on pet tumor segmentation. Med. Image Anal. **44**, 177–195 (2018)
11. Hyndman, R.J.: Computing and graphing highest density regions. Am. Stat. **50**(2), 120–126 (1996)
12. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
13. Li, J., Ray, S., Lindsay, B.G.: A nonparametric statistical approach to clustering via mode identification. J. Mach. Learn. Res. **8**(8) (2007)

14. Ling, C.C., et al.: Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. Int. J. Radiat. Oncol. Biol. Phys. **47**(3), 551–560 (2000)
15. Parzen, E.: On estimation of a probability density function and mode. Ann. Math. Stat. **33**(3), 1065–1076 (1962)
16. Ray, S., Lindsay, B.G.: The topography of multivariate normal mixtures. Ann. Stat. **33**(5) (2005). https://doi.org/10.1214/009053605000000417
17. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
18. Shi, X., Meng, X., Sun, X., Xing, L., Yu, J.: PET/CT imaging-guided dose painting in radiation therapy. Cancer Lett. **355**(2), 169–175 (2014)
19. Vallieres, M., et al.: Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. Sci. Rep. **7**(1), 10117 (2017)
20. Visser, E.P., Boerman, O.C., Oyen, W.J.: SUV: from silly useless value to smart uptake value. J. Nucl. Med. **51**(2), 173–175 (2010)
21. Wang, J., Peng, Y., Guo, Y., Li, D., Sun, J.: CCUT-net: pixel-wise global context channel attention UT-net for head and neck tumor segmentation. In: Andrearczyk, V., Oreiller, V., Hatt, M., Depeursinge, A. (eds.) HECKTOR 2021. LNCS, vol. 13209, pp. 38–49. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98253-9_2
22. Zhang, W., Ray, S.: Kernel smoothing-based probability contours for tumour segmentation. In: 26th UK Conference on Medical Image Understanding and Analysis. Springer, Cham (2022)
23. Zhang, Y., Chung, A.C.S.: Deep supervision with additional labels for retinal vessel segmentation task. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11071, pp. 83–91. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00934-2_10