Deep Interactive Segmentation of Uncertain Regions with Shadowed Sets

Haiyan Zheng
College of Electronics and Information Engineering, Tongji University, Shanghai, China, 200092
18800205138@163.com

Yufei Chen*
College of Electronics and Information Engineering, Tongji University, Shanghai, China, 200092
april337@163.com

Xiaodong Yue
School of Computer Engineering and Science, Shanghai University, Shanghai, China, 200444
yswantfly@shu.edu.cn

Chao Ma
Department of Radiology, Changhai Hospital of Shanghai, Second Military Medical University, Shanghai, China, 200433
machaohdsd@126.com

ABSTRACT
Pancreas segmentation is a challenging task in medical image analysis because of its large variations in texture, location, shape and size and the high similarity to the surrounding tissues especially around the boundary regions, which leads to the high segmentation uncertainty and makes the results inaccurate. Existing fully automatic segmentation methods rarely achieve sufficiently accurate and robust results. To tackle this problem, we propose a deep learning based interactive uncertain segmentation method which can involve the domain knowledge in the process of segmentation in an interactive and iterative way. Specially, the proposed method describes the uncertain regions of pancreatic CT images based on shadowed sets theory which are further corrected through interaction. The proposed method is evaluated on a challenging 3D pancreatic CT images dataset collected from the Changhai Hospital. The experimental results demonstrate that our proposed method outperforms the existing methods in terms of both the Dice similarity coefficient of 78% and the pixel-wise accuracy of 96%, which reveals the effectiveness and the potential of our method in clinical settings.

CCS Concepts
• **Computing methodologies → Image segmentation**

Keywords
Pancreas segmentation; Shadowed sets; Uncertainty; Interaction

1. INTRODUCTION
Recently, due to the great development in deep neural network and increasing medical needs, the analysis of medical image in Computer-Aided Diagnosis (CAD) system has become a new fashion, in which segmentation of target organs is fundamentally important. However, among different abdominal organs, the segmentation of the pancreas is especially challenging, as the target suffers from high variability in shape, size and location [11], while occupying only a very small fraction (e.g., < 0.5%) of the entire CT volume. Besides, because of a relatively poor contrast and similar neighboring abdominal structures, deep neural networks can be disrupted by a large fraction of background. Consequently, the segmentation result becomes inaccurate especially around the boundary areas and fully automatic segmentation methods rarely achieve sufficiently accurate and robust results to be clinically useful [13].

To alleviate these limitations, interactive segmentation methods are desirable due to higher accuracy and robustness. There are many kinds of user interactions, such as click-based [6], contour-based [15] and bounding box-based methods [12]. Drawing scribbles is user-friendly and particularly popular, e.g., in Graph Cuts [2], GeoS [1], [4], and Random Walks [5]. However, most of these methods rely on low-level features and require a relatively large amount of user interactions to deal with images with low contrast and ambiguous boundaries [14]. In this paper, the interactive method can generate the flexible results with a few scribbles or dots of domain experts' interactions. In order to obtain the robust correction, the discriminative features that can represent mid-level cues beyond image intensity or gradient are adaptively extracted from ROI according to both the training set and the interaction [9].

In addition, both traditional and existing deep learning methods are limited to segmentation methods which based on probability and statistics theory. Therefore, the uncertainty of the organ segmentation of the boundaries which are easily confused with the surrounding tissues is neglected. To tackle this problem, this paper proposes a new interactive method to determine the uncertain regions of pancreatic CT image based on shadowed sets theory, which not only reduces the scope of interaction, but also considers the segmentation uncertainty of pancreatic boundaries.

In proposed method, firstly we train a 3D U-net to obtain the initial segmentation results. According to the interactions provided by doctor based on the initial segmentation results, the region of interest (ROI) is obtained. Then, the probability map that each voxel belongs to foreground and background is obtained as the membership function. The separation threshold determined by the shadowed sets is used to divide voxels into three classes: confident foreground, confident background and uncertain regions. Finally, the doctor corrects the segmentation results by second labeling of the uncertain regions to obtain the final segmentation

ISICDM 2019, August 24–26, 2019, Xi'an, China.

ACM ISBN 978-1-4503-7262-6/19/08…$15.00
DOI: https://doi.org/10.1145/3364836.3364885

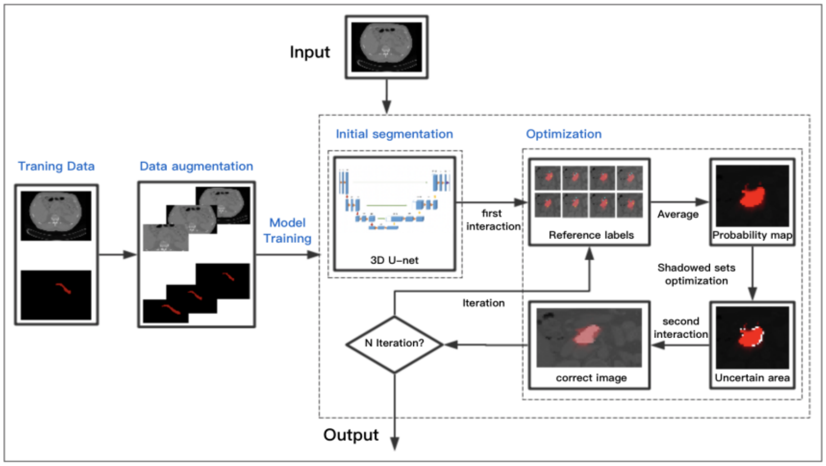


Figure 1. Overview of the proposed interactive segmentation framework.

results. The contributions of this work are two-fold: 1) We propose a new method to describe the uncertain region of pancreatic CT images based on shadowed sets theory. 2) We demonstrate that our interactive uncertain framework leads to higher accuracy and less interactions for pancreas segmentation, which incorporates the knowledge of domain experts and the uncertainty of pancreatic segmentation.

2. INTERACTIVE SEGMENTATION OF UNCERTAIN REGIONS FRAMEWORK

The proposed framework of interactive segmentation in uncertain regions is illustrated in Figure 1, which is an iterative process when clinicians insert interactions into the wrong parts of the previous segmentation L_{t-1} to correct the errors in segmentation where t is the iteration number. The processes of corrections are repeated until a satisfactory segmentation result is obtained. As shown in Figure 1, an initial segmentation is obtained at first and then the interactive corrections are conducted on the initial segmentation which includes the following procedures. (1) In the local area near the interaction, the reference labels, which are similar to both the current interaction area and previous segmentation, are selected from the training data. (2) Based on the selected reference labels, a probability map is generated in which the elements denote the probability that the corresponding voxel in image belongs to pancreas. Then a separation threshold is calculated based on shadowed sets theory [10] according to which the voxels are separated into confident pancreas and background, and uncertain regions. (3) The uncertain regions are then given to the clinicians for second corrections and the final segmentation result is obtained after the second correction. The details are described in the following sub-sections.

2.1 Initial Segmentation

An initial pancreas segmentation L_0 is obtained at first such that clinicians' interaction can be inserted into the wrong part of the segmentation and the corrections can be conducted repeatedly. The initial pancreas segmentation L_0 can be obtained by any kind of automatic or interactive segmentation methods.

In this paper, we build up a deep convolutional neural network to obtain the initial segmentation of pancreas motivated by 3D U-net [3], which has been demonstrated to be powerful in image segmentation. In our work, we train the 3D U-net on preprocessed pancreas CT images and use the network to generate the initial segmentation L_0. The process of interactive correction starts from the next sub-section.

2.2 Determination of ROI from User Interactions

Based on the initial segmentation, the clinicians can start the corrections to obtain a satisfactory segmentation result in an interactive and iterative way. Our proposed framework accepts two kinds of interactions, points or scribbles, which represents the foreground and background. After each interaction, the corrections are performed on the local regions near the interaction points or scribbles, which is based on the idea that segmentation errors exist in the regions near the interaction area, while the rest regions of L_{t-1} are assumed right temporarily. If there are errors on the rest regions, they can be corrected sequentially in subsequent interactions.

The local region of interest (ROI) Φ_t serves as a bounding box that includes the interaction area and a margin which is thick enough to cover possible local variations. Considering the large thickness of CT image and the difference in size of the pancreas, we set the edge as 9×9×2 voxels. In each interaction, a voxel $v_i(i = 1,2,\cdots,N)$ on Φ_t is labeled as foreground($U_t(v_i) = 1$), background ($U_t(v_i) = -1$) or unlabeled($U_t(v_i) = 0$) where U_t is the label image of t^{th} round interaction.

2.3 Generation of Probability Map

In order to generate the probability map which estimates the probability that the corresponding voxels belong to pancreas, the appropriate training data should be selected according to interactions. We assume that the suitable training data have label M which is both well matched with U_t on the interaction voxels and well matched with L_{t-1} on the other voxels. To select the appropriate training data, we evaluate the similarity between M and both U_t and L^{t-1} and then choose the k most similar M as training data. The similarity cost $S(M, U^t, L^{t-1})$ proposed in [9] is used in our work

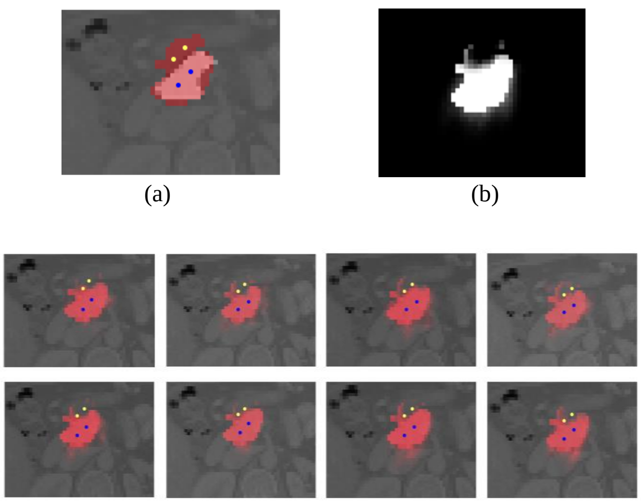
$$S\left(M,U^t,L^{t-1}\right)=\sum_{v\in\Phi^t,U^t(v)\neq 0}\left(\delta\left(M(v)-U^t(v)\right)\right) + w_U\sum_{v\in\Phi^t,U^t(v)=0}\left(\delta\left(M(v)-L^{t-1}(v)\right)\right) \quad (1)$$

where δ is the Kronecker delta. The first term denotes the number of voxels with the same label in M and U_t and the second term denotes the number of voxels with the same label in M and L_{t-1}. The higher the consistency of M with U_t and L_{t-1}, the higher the value of similarity cost function. w_U is the parameter used to balance the two terms. Since the number of voxels in the interaction($U^t(v_i) \neq 0$) is much smaller than that of the unlabeled voxels($U^t(v_i) = 0$), w_U is set to be a smaller value (i.e. w_U= 0.01). After each interaction, the similarity of all training data will be calculated, and the top k training labels with the highest similarity cost will be selected as the reference training data. An example of a group of reference label($k = 8$) is shown in Figure 2 (c).

To estimate the confident region of pancreas, the probability value in the probability map is calculated by averaging the value of the reference labels obtained before, which is shown in Figure 2(b). Then the probability value $p(v_i)$ for voxels is formulated as:

$$p(v_i) = \frac{\sum_{1}^{k} M^k(v_i)}{k} \tag{2}$$

Simple voting does not reflect the true label of the complex pancreas's borders with great uncertainty. Inspired by the shadowed sets theory, the probability value of voxels in probability map can be regarded as the membership function such that the voxels can be separated into confident regions and uncertain regions. For the uncertain regions, the further interactive correction can be conducted and therefore the regions with great uncertainty in pancreas can be segmented accurately.



(a) (b)

(c)

Figure 2. (a) Initial segmentation (red) and ground truth (white), foreground interaction (blue) and background interaction (yellow) respectively. (c) selected reference labels (red), and (b) probabilistic label constructed by averaging the reference labels.

3. UNCERTAIN REGIONS SEGMENTATION WITH SHADOWED SETS

The high variability in shape, size and location of pancreas [16] brings the high uncertainty in pancreas segmentation. Traditional deep learning pancreas segmentation methods are short of the ability to handle the uncertainty which makes the segmentation results inaccurate, especially around the boundary areas of pancreas. The shadowed sets theory proposed by Pedrycz[10] provides an alternate mechanism for handling uncertainty, aiming at ambiguity demarcation especially boundary regions. Therefore, to solve the problem of uncertainty in pancreas segmentation, we introduce the shadowed sets theory into the segmentation process and implement the interactive segmentation in the uncertain regions in pancreas.

3.1 Shadowed Sets

Shadowed sets provide a meaningful description of information granules by abstracting corresponding fuzzy sets into three categories: full acceptance, full rejection and uncertain. Suppose a continuous fuzzy membership function $x \to f(x)$, $f(x) \in [0,1]$, the elimination of uncertainty and formed shadows is illustrated in Figure 3 and quantified as follows:

$$\begin{aligned} V(\alpha) &= |\Omega_1 + \Omega_2 - \Omega_3| \\ &= \left| \int_{x_i f(x) \le \alpha} f(x)dx + \int_{x: f(x) > 1-\alpha} (1 - f(x))dx - \int_{x: \alpha < f(x) < 1-\alpha} dx \right| \end{aligned} \tag{3}$$

The optimal partition threshold α is achieved based on the principle of uncertainty balance, through minimizing the following objective function:

$$\alpha_{opt} = \arg\min V(\alpha) = \operatorname{argmin} |\Omega_1 + \Omega_2 - \Omega_3| \tag{4}$$

where $\alpha \epsilon [0,0.5)$. The discrete version of formula (4) can be formed as follows

$$V(\alpha) = \left| \sum_{u_i < \alpha} u_i + \sum_{u_i > 1-\alpha} (1 - u_i) - \text{card}_{\alpha < u_i < 1-\alpha}(\{u_i\}) \right| \tag{5}$$

where u_i is the fuzzy membership value and $card(X)$ stands for the cardinality of the set X.

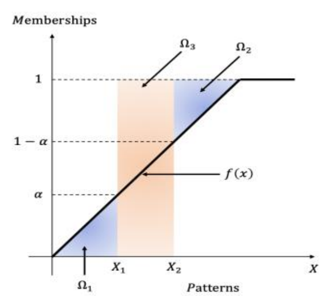


Figure 3. The schematic diagram of shadowed sets.

3.2 Determination of Uncertain Regions

Based on the probability value $p(v_i)$ obtained in section 2.3, we suppose a discrete fuzzy set $A = \{(v_i, u_i)\}(i = 1,2,\cdots,N)$, $u_i = f(p(v_i))$. Since the probability function constructed by voxels is discrete, we determine the separation threshold α^* through minimizing the following objective function:

$$V(\alpha) = \left| \sum_{u_i < \alpha} u_i + \sum_{u_i > 1-\alpha} (1 - u_i) - \text{card}\{u_i \mid \alpha \le u_i \le 1-\alpha\} \right| \tag{6}$$

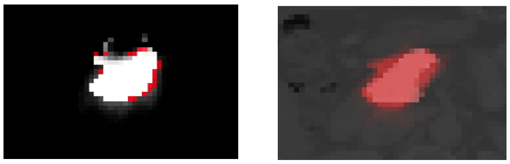
According to the obtained separation threshold α^*, items in fuzzy set A can be partitioned into three mutually disjoint categories: the foreground region(confident pancreas regions)($y(v_i) = 1$), the confident background region($y(v_i) = -1$) and the uncertain region($y(v_i) = 0$). The formulas are given as follows:

$$y(v_i) = \begin{cases} 1, & u_i > 1-\alpha^* \\ -1, & u_i < \alpha^* \\ 0, & \alpha^* < u_i < 1-\alpha^* \end{cases} \tag{7}$$

Thus, the uncertain regions in the pancreatic CT image have been obtained. Most existing segmentation methods are inaccurate in boundary regions. It can be shown in Figure 4(a) that the uncertain regions detected by shadowed sets are basically the boundary of pancreas which is meet expectations. The following step is to conduct the second interaction on the uncertain region and the interactive corrections are detailed in the next sub-section.

3.3 Uncertain Segmentation of Pancreas with Interactive Corrections

The second interaction asks the clinicians for the interactions on the uncertain regions obtained with shadowed sets theory, which involves the domain knowledge into the segmentation process in the uncertain regions of pancreas. With the information given in the second interaction, the uncertain regions on previous segmentation can be further portioned into pancreas regions or background.



(a) (b)

Figure 4. (a) uncertain region (red) obtained by separation thresholds α^*, and (b) corrected final segmentation results (red) and ground truth (white).

For a good interactive segmentation method, as few as user interactions are required to obtain accurate results, which makes the interaction method efficient. Therefore, to implement an efficient interaction process, our proposed method only requests points or scribbles representing foreground and background respectively on the uncertain regions, which greatly reduces the time-consuming and laborious manual interactions of clinicians.

In order to make full use of the information of interactions, the confidence of the foreground points should be increased while the confidence of the background points should be reduced. To implement this, we adopted the following Gaussian distribution our work:

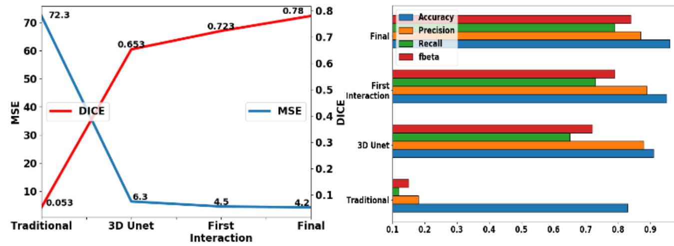
$$N(\bar{x} \mid \bar{u}, \Sigma) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left[-\frac{1}{2} (\bar{x} - \bar{u})^T \Sigma^{-1} (\bar{x} - \bar{u}) \right] \quad (8)$$

where $\bar{\boldsymbol{u}}$ is the coordinate of an interactive point, and Σ is a covariance matrix used to control the density of Gaussian distribution. It is noted that since the image is a set of discrete points and the distribution generated by the Gaussian distribution produces continuous results, the coordinates should be rounded after obtaining random data according to the distribution. The proposed correction for uncertain region on pancreas is depicted in Figure. 4 (b), which is also the final result of the proposed model.

4. EXPERIMENTS

To validate the proposed interactive segmentation method for the uncertain regions in pancreas, we evaluate our method on a 3D pancreas CT image dataset collected from the Changhai Hospital which consists of 20 images of different patients with pancreatic endocrine neoplasms. The image size is 512×512×(36 ~ 74) with 0.66×0.66×1.5 mm^3 voxel spacing. The dataset includes various subjects with large shape and appearance variations. The boundary of pancreas in each image was manually delineated by human experts according to domain knowledge and used as the ground truth for measuring the segmentation performance. The Dice similarity coefficient (DSC) and Mean Square Error (MSE) are adopted to measure the segmentation performance. Considering the segmentation problem as a binary classification problem which classifies the voxel into foreground or background, we also adopted accuracy, precision, recall and F1 Score to evaluate the classification results. We perform the 4-fold cross validation and take the average as the experimental results.

We implement our proposed interactive segmentation framework based on a deep convolutional model 3D U-net. To train the model, we first preprocess both the original CT image and the ground truth images. The images are resized to be 256×256 and data augmentation are conducted on the images such as translation transform and scale transform referring to [7] to enlarge the training dataset and prevent the deep learning model overfitting. For the optimization of separation threshold $\boldsymbol{\alpha}^*$, since there are no closed-form solutions to minimize the optimization objective function (6), the optimal value is computed to be 0.375 in our work.



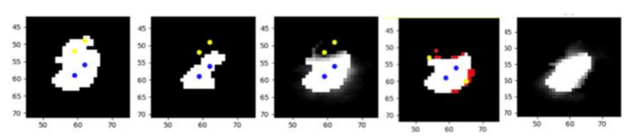
(a) (b)

Figure 5. Segmentation results of different methods.

Table 1. MSE and DICE of different methods

	MSE	ΔMSE	DICE	ΔDICE
Traditional	72.3	--	0.05	--
3D U-net	6.3	-66.0	0.65	+0.60
First Interaction	4.5	-1.8	0.72	+0.07
Final	4.2	-0.3	0.78	+0.06

We compare our model with two elegant segmentation methods, including the traditional method of threshold segmentation [8], 3D U-net [3]. Besides, we also compare the segmentation results of the first interaction and the second interaction which is the final result of our proposed method. Figure 5(a) illustrates the DSC and MSE of different methods and Table 1 presents the details. It is clear that our method gets the better performance. Especially, when involving the second interaction focusing on the uncertain regions of pancreas, the performance of segmentation can be improved, which verifies the effectiveness of our uncertain segmentation method. Figure 5(b) presents the accuracy, precision, recall and F1 Score of different methods when the segmentation problem is considered as the classification problem for voxels. It can be found that our method achieves the most precise and stable classification results.



(a) (b) (c) (d) (e)

Figure 6. Schematic diagram of segmentation process

Figure. 6 shows a case of segmentation by our model, (a) is the initial segmentation given by 3D U-net and (b) is the ground truth. (c) is the result after first interaction. We can see that segmentation errors are mostly corrected, making (c) closer to the ground truth, which proves the effectiveness of our first interaction. In (d), the clinicians will further make interaction on the uncertain region (red) with the background interaction (yellow) and the foreground interaction (blue). After the second correction, (e) is the final segmentation result, which is the closest to the ground truth. This gradual procedure proves the effectiveness of the determination of uncertain region by shadowed sets and two interactions. In this case, our method achieves Dice score 82%, while only 66% compared with 3D U-net.

Figure 7 presents the pancreas segmentation results of patients selected randomly with traditional methods, 3D U-net methods, the first and second correction on uncertain region. It can be clear found that our model obtains the most similar results to the ground truth and achieves the highest DICE (82%, 78%, 80%, 72%).

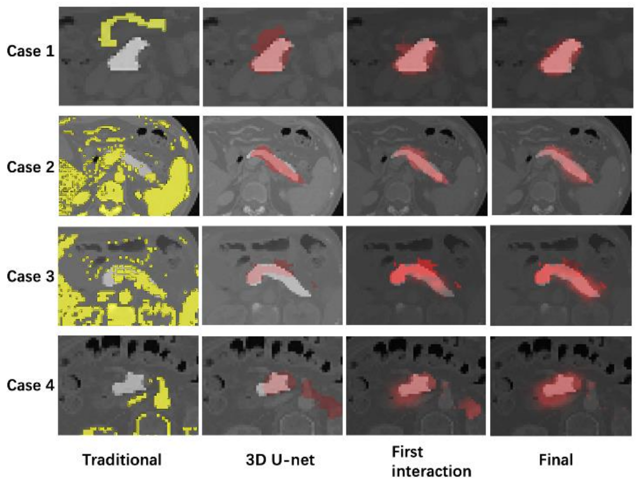


Figure 7. The white color is the ground truth given by the doctor, and the red color is the segmentation result of the model.

5. CONCLUSION

In this paper, we propose an interactive pancreas segmentation framework, leading to higher accuracy and less interactions, which incorporates the knowledge of domain experts and the uncertainty of pancreatic segmentation. The method gracefully addresses the problems that requiring laborious, time-consuming manual interactions and ignoring the uncertainty of segmentation. Our results indicate that uncertain segmentation greatly improves performance for pancreas segmentation. Moreover, uncertain interactive method provides clinicians with information permitting them to quickly analyze uncertain organ boundaries which further make more accurate diagnosis and treatment.

6. ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. 61573235), the Shanghai Innovation Action Project of Science and Technology (No. 17411952200), and the Fundamental Research Funds for the Central Universities.

7. REFERENCES

[1] Bai, X., & Sapiro, G. (2007, October). A geodesic framework for fast interactive image and video segmentation and matting. In 2007 IEEE 11th International Conference on Computer Vision (pp. 1-8). IEEE.

[2] Boykov, Y. Y., & Jolly, M. P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. In Proceedings eighth IEEE international conference on computer vision. ICCV 2001 (Vol. 1, pp. 105-112). IEEE.

[3] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016, October). 3D U-Net: learning dense volumetric segmentation from sparse annotation. In International conference on medical image computing and computer-assisted intervention (pp. 424-432). Springer, Cham.

[4] Criminisi, A., Sharp, T., & Blake, A. (2008, October). Geos: Geodesic image segmentation. In European Conference on Computer Vision (pp. 99-112). Springer, Berlin, Heidelberg.

[5] Grady, L. (2006). Random walks for image segmentation. IEEE Transactions on Pattern Analysis & Machine Intelligence, (11), 1768-1783.

[6] Haider, S. A., Shafiee, M. J., Chung, A., Khalvati, F., Oikonomou, A., Wong, A., & Haider, M. A. (2015, April). Single-click, semi-automatic lung nodule contouring using hierarchical conditional random fields. In 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI) (pp. 1139-1142). IEEE.

[7] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[8] Otsu, N. (1979). A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics, 9(1), 62-66.

[9] Park, S. H., Gao, Y., Shi, Y., & Shen, D. (2014, September). Interactive prostate segmentation based on adaptive feature selection and manifold regularization. In International Workshop on Machine Learning in Medical Imaging (pp. 264-271). Springer, Cham.

[10] Pedrycz, W. (1998). Shadowed sets: representing and processing fuzzy sets. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 28(1), 103-109.

[11] Roth, H. R., Lu, L., Farag, A., Shin, H. C., Liu, J., Turkbey, E. B., & Summers, R. M. (2015, October). Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation. In International conference on medical image computing and computer-assisted intervention (pp. 556-564). Springer, Cham.

[12] Rother, C., Kolmogorov, V., & Blake, A. (2004, August). Grabcut: Interactive foreground extraction using iterated graph cuts. In ACM transactions on graphics (TOG) (Vol. 23, No. 3, pp. 309-314). ACM.

[13] Sharma, N., & Aggarwal, L. M. (2010). Automated medical image segmentation techniques. Journal of medical physics/Association of Medical Physicists of India, 35(1), 3.

[14] Wang, G., Zuluaga, M. A., Li, W., Pratt, R., Patel, P. A., Aertsen, M., ... & Vercauteren, T. (2018). DeepIGeoS: a deep interactive geodesic framework for medical image segmentation. IEEE transactions on pattern analysis and machine intelligence.

[15] Xu, C., & Prince, J. L. (1998). Snakes, shapes, and gradient vector flow. IEEE Transactions on image processing, 7(3), 359-369.

[16] Zhou, Y., Xie, L., Shen, W., Wang, Y., Fishman, E. K., & Yuille, A. L. (2017, September). A fixed-point model for pancreas segmentation in abdominal CT scans. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 693-701). Springer, Cham