MEDINFO 2023 — The Future Is Accessible
J. Bichel-Findlay et al. (Eds.)

doi:10.3233/SHTI231105

Position Prior Attention Network for Pancreas Tumor Segmentation

Kaiqi DONG[a], Peijun HU[b], Xiang LI[c], Yu TIAN[a], Yan ZHU[a], Xueli BAI[c], Tingbo LIANG[c] and Jingsong LI[a,b,1]

[a] *Engineering Research Center of EMR and Intelligent Expert System, Ministry of Education, College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China*

[b] *Research Center for Healthcare Data Science, Zhejiang Laboratory, Hangzhou, China*

[c] *Department of Hepatobiliary and Pancreatic Surgery, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China*

Abstract. Segmentation of pancreatic tumors on CT images is essential for the diagnosis and treatment of pancreatic cancer. However, low contrast between the pancreas and the tumor, as well as variable tumor shape and position, makes segmentation challenging. To solve the problem, we propose a Position Prior Attention Network (PPANet) with a pseudo segmentation generation module (PSGM) and a position prior attention module (PPAM). PSGM and PPAM maps pancreatic and tumor pseudo segmentation to latent space to generate position prior attention map and supervises location classification. The proposed method is evaluated on pancreatic patient data collected from local hospital and the experimental results demonstrate that our method can significantly improve the tumor segmentation results by introducing the position information in the training phase.

Keywords. Pancreas tumor segmentation, position prior knowledge, variational AutoEncoders

1. Introduction

Automatic and accurate segmentation of pancreatic tumors is essential for computer-assisted diagnosis and prognosis of pancreatic cancer. In recent years, many approaches have been proposed for tackling the task. The nnUNet [1] achieved the segmentation of pancreas and pancreatic tumor by using cascade network approach. C2FNAS [2] utilized huge computing power for neural architecture search to optimize and improve the structure of 3D UNet to improve the segmentation results of pancreatic tumors.

However, the precision of pancreatic tumor segmentation is currently inadequate. Incorporating prior knowledge into the model has the potential to enhance its segmentation abilities. For example, in the study [3], the shape prior of the pancreas is learned unsupervised using a Variational AutoEncoder (VAE) network. Wang et al. [4]

[1] Corresponding Author: Jingsong Li, Tel: +86-571-87951564, Fax: +86-571-87951564, email: ljs@zju.edu.cn.

introduced the breast cancer classification information as a prior knowledge into the segmentation of breast tumors and achieved the improvement of segmentation accuracy.

In this paper, we labeled the CT images with pancreatic cancer according to the anatomical location of the tumor in the pancreas (head, body, tail) as prior knowledge to guide the segmentation. As far as we know, we are the first to leverage tumor position information as prior knowledge to assist pancreatic tumor segmentation.

2. Methods

The overview of the image processing is illustrated in Figure 1. First, a 3D UNet network is trained as a pancreas segmentation network on the Medical Segmentation Decathlon (MSD) dataset [5], and then the model infers the pancreas pseudo segmentation on the First Affiliated Hospital, Zhejiang University School of Medicine (FAH) dataset. Second, we train a 3D position prior attention network (PPANet) to segment tumors on the FAH dataset with the pancreas pseudo segmentation and tumor position information.

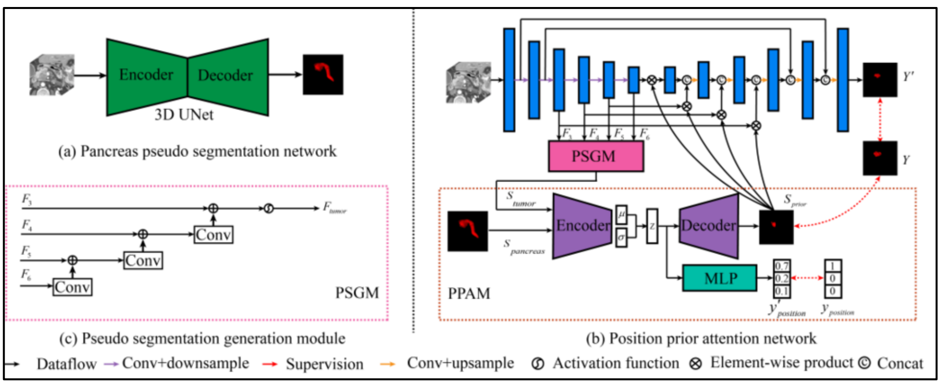

Figure 1. Overview of our proposed method.

2.1. Pancreas Segmentation Network

As shown in Figure 1 (a), we trained a 3D UNet network on the MSD dataset, referred as $f_{pancreas}(\cdot)$, to segment the pancreas from CT images. The pixel-level segmentation loss combines the cross-entropy loss and Dice loss:

$$L_{pancreas} = L_{CE}(S'_{pancreas}, Y_{MSD}) + \beta L_{DC}(S'_{pancreas}, Y_{MSD}), \quad (1)$$

where $S'_{pancreas}$ is the pancreas prediction of MSD dataset, Y_{MSD} is the ground truth of input image, and β is a balance factor (β =1).

Subsequently, $f_{pancreas}(\cdot)$ is used to obtain pancreas pseudo segmentation $S_{pancreas}$ from the CT images on the FAH dataset.

2.2. Position Prior Attention Network

PPANet embeds a pseudo segmentation generation module (PSGM) and a position prior attention module (PPAM) on a 3D UNet like network, as shown in Figure 1 (b) PSGM is used to obtain tumor pseudo segmentation and PPAM introduce position prior attention map, respectively.

2.2.1 Pseudo segmentation generation module

To obtain tumor pseudo segmentation, a pseudo segmentation generation module (PSGM) is designed with reference to the decoder network of FPN[6], as detailed in Figure 1(c). PSGM combines the deepest and shallowest features, aligns feature maps by trilinear interpolation, and fuses multiscale features by element-wise addition. Using $F_s, s \in (3,4,5)$ to represent the multiscale feature maps, feature fusion process can be expressed as:

$$F_s' = f_{up}(W_{s+1}F_{s+1}) \oplus F_s, \tag{2}$$

where W_{s+1} denotes a $1 \times 1 \times 1$ kernel convolution, $f_{up}(\cdot)$ denotes the trilinear interpolation. Next, Tanh activation function is employed to compress the values of the feature map F_3' to [0,1] to generate the tumor pseudo segmentation S_{tumor}:

$$S_{tumor} = Tanh(W_{out}F_3') \tag{3}$$

2.2.2 Position prior attention module

PPAM comprises a Variational AutoEncoders (VAE) [7] module and a classification branch to introduce position prior knowledge by supervised position classification, as shown in Figure 1 (b). First, PPAM maps tumor pseudo segmentation and pancreas pseudo segmentation to the latent space z via VAE encoder $q_\theta(\cdot)$. z follows the Gaussian distribution. The mean of the Gaussian distribution μ and the Standard Deviation σ are obtained from the VAE encoder $q_\theta(\cdot)$. Secondly, the features are sampled from z to predict the tumor position label $y_{position}'$ through a multilayer perception (MLP) sub-network and used to generate the position prior attention map S_{prior} through VAE decoder. The loss function is defined:

$$\begin{aligned} L_{prior} &= L_{CE}(S_{prior}, Y) \\ &+ D_{KL}(q_\theta(z|S_{tumor}, S_{pancreas})||p_\phi(z)) \\ &+ L_{CE}\left(y_{position}', y_{position}\right), \end{aligned} \tag{4}$$

where the first two terms are the typical ELBO (Evidence Lower Bound) VAE loss function. $p_\phi(\cdot)$ is the reparameterization method in VAE. Y is the ground truth segmentation. $y_{position} \in \{0,1,2\}$ is true label for the tumor position, where 0,1,2 indicate that the tumor occurred in the head, body and tail of the pancreas respectively. Then, S_{prior} highlight the important region in F_s with element-wise multiplication.

2.2.3 Loss Function

The total segmentation loss of PPANet consists of two components: L_{seg} segmentation loss for supervised training of tumor segmentation prediction and L_{prior} position prior loss for PPAM parameter optimization. L_{seg} consists of the the cross-entropy loss and Dice loss. The total segmentation loss is:

$$L_{total} = L_{seg} + \alpha L_{prior} = L_{CE}(Y', Y) + L_{DC}(Y', Y) + \alpha L_{prior}, \tag{5}$$

where Y' is the prediction of tumor segmentation and α is a balance factor ($\alpha = 0.5$).

3. Results

3.1 Dataset

We collected 71 CT images from the First Affiliated Hospital (FAH), Zhejiang University School of Medicine. Manual annotations and position labels of the tumors are performed by experienced radiologists. The region of tumorigenesis is divided into head, body and tail (head 40, body 19, tail 12). The quality of the model segmentation was evaluated using Dice, Precision and Recall through a 4-fold cross-validation.

3.2 Ablation Results

To evaluate the rationality of the design, we conduct two experiments: i) To verify the effectiveness of the VAE module. We used a multi-task learning framework for both the classification task and the segmentation task, i.e., the encoder and MLP were retained and the other structures of the VAE module were removed. ii) To verify the mapping of tumor pseudo segmentation to latent space, a convolutional auto-encode (CAE) is constructed to replace VAE.

Table 1 reveals that our methods can effectively improve the segmentation ability of the model (Dice improved by a maximum of 2.23% and recall improved by a maximum of 3.1%), although it is lower than the model with CAE in terms of precision.

Table 1. Effectiveness of Position prior attention module.

Method	Dice (%)	Precision (%)	Recall (%)
W/O VAE	61.26	74.70	59.18
CAE	60.57	**76.59**	57.83
VAE	**62.80**	75.24	**60.93**

3.3 Comparison with existing methods

Table 2 compares the segmentation results of our proposed method with existing methods, i.e. 2D segmentation networks UNet [8], Attention UNet [9], and a cascaded segmentation network nnUNet [1]. The results indicate that our proposed framework with position prior knowledge is competitive.

Table 2. Comparison of existing methods.

Method	Dice	Precision	Recall
2D UNet [8]	45.11	73.80	40.37
2D Attention UNet [9]	52.32	69.81	49.93
3D nnUNet [1]	62.07	73.05	**64.98**
Ours	**62.80**	**75.24**	60.93

4. Discussion

Our study demonstrates the effectiveness of incorporating tumor position information as prior knowledge to improve pancreatic tumor segmentation. The ablation studies validate the rationality of the sub-module design. However, the dataset used for validation contains only 71 cases. Large-scale datasets could further verify the method's robustness. Additionally, only tumor position is explored as prior knowledge. Incorporating other prior knowledge like tumor shape and texture may further improve performance.

5. Conclusions

In this paper, we firstly design a position prior attention network (PPANet) to introduce position prior information to assist in pancreas tumor segmentation. In the future, we will explore integrating other prior knowledge and apply the framework to various segmentation tasks.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 12101571, No. 82172069), the Key Research Project of Zhejiang Laboratory (No. 2022ND0AC01), and the Zhejiang Provincial Key Research and Development Program (No. 2020C03117).

References

[1] F. Isensee, P.F. Jaeger, S.A. Kohl, J. Petersen, and K.H. Maier-Hein, nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, *Nature Methods* **18** (2021), 203-211.

[2] Q. Yu, D. Yang, H. Roth, Y. Bai, Y. Zhang, A.L. Yuille, and D. Xu, C2FNAS: Coarse-to-Fine Neural Architecture Search for 3D Medical Image Segmentation, in, 2019, p. arXiv:1912.09628.

[3] Y. Yao, F. Liu, Z. Zhou, Y. Wang, W. Shen, A.L. Yuille, and Y. Lu, Unsupervised Domain Adaptation through Shape Modeling for Medical Image Segmentation, in, 2022, pp. 1444-1458.

[4] J. Wang, Y. Zheng, J. Ma, X. Li, C. Wang, J. Gee, H. Wang, and W. Huang, Information bottleneck-based interpretable multitask network for breast cancer classification and segmentation, *Medical Image Analysis* **83** (2023), 102687.

[5] M. Antonelli, A. Reinke, S. Bakas, K. Farahani, AnnetteKopp-Schneider, B.A. Landman, G. Litjens, B. Menze, O. Ronneberger, R.M. Summers, B. van Ginneken, M. Bilello, P. Bilic, P.F. Christ, R.K.G. Do, M.J. Gollub, S.H. Heckers, H. Huisman, W.R. Jarnagin, M.K. McHugo, S. Napel, J.S. Goli Pernicka, K. Rhode, C. Tobon-Gomez, E. Vorontsov, H. Huisman, J.A. Meakin, S. Ourselin, M. Wiesenfarth, P. Arbelaez, B. Bae, S. Chen, L. Daza, J. Feng, B. He, F. Isensee, Y. Ji, F. Jia, N. Kim, I. Kim, D. Merhof, A. Pai, B. Park, M. Perslev, R. Rezaiifar, O. Rippel, I. Sarasua, W. Shen, J. Son, C. Wachinger, L. Wang, Y. Wang, Y. Xia, D. Xu, Z. Xu, Y. Zheng, A.L. Simpson, L. Maier-Hein, and M.J. Cardoso, The Medical Segmentation Decathlon, in, 2021, p. arXiv:2106.05735.

[6] T.-Y. Lin, P. Dollár, R.B. Girshick, K. He, B. Hariharan, and S.J. Belongie, Feature Pyramid Networks for Object Detection, in, 2017, pp. 936-944.

[7] Q. Zhao, E. Adeli, N. Honnorat, T. Leng, and K.M. Pohl, Variational AutoEncoder for Regression: Application to Brain Aging Analysis, in: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019,* D. Shen, T. Liu, T.M. Peters, L.H. Staib, C. Essert, S. Zhou, P.-T. Yap, and A. Khan, eds., Springer International Publishing, Cham, 2019, pp. 823-831.

[8] T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, A. Dovzhenko, O. Tietz, C. Dal Bosco, S. Walsh, D. Saltukoglu, T.L. Tay, M. Prinz, K. Palme, M. Simons, I. Diester, T. Brox, and O. Ronneberger, U-Net: deep learning for cell counting, detection, and morphometry, *Nature Methods* **16** (2019), 67-70.

[9] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y Hammerla, B. Kainz, B. Glocker, and D. Rueckert, Attention U-Net: Learning Where to Look for the Pancreas, in: *Conference on Medical Imaging with Deep Learning*, 2018.