2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

3D Dense Volumetric Network for Accurate Automated Pancreas Segmentation

Ruijie Wang[1], Xudong Liu[1], Huikai Shao[1], Qiling Li[2], and Dexing Zhong[1,3,4*]

Abstract— **Pancreatic cancer poses a great threat to our health with an overall five-year survival rate of 8%. Automatic and accurate segmentation of pancreas plays an important and prerequisite role in computer-assisted diagnosis and treatment. Due to the ambiguous pancreas borders and intertwined surrounding tissues, it is a challenging task. In this paper, we propose a novel 3D Dense Volumetric Network (3D^{2}VNet) to improve the segmentation accuracy of pancreas organ. Firstly, 3D fully convolutional architecture is applied to effectively incorporate the 3D pancreas and geometric cues for volume-to-volume segmentation. Then, dense connectivity is introduced to preserve the maximum information flow between layers and reduce the overfitting on limited training data. In addition, a auxiliary side path is constructed to help the gradient propagation to stabilize the training process. Adequate experiments are conducted on a challenging pancreas dataset in Medical Segmentation Decathlon challenge. The results demonstrate our method can outperform other comparison methods on the task of automated pancreas segmentation using limited data.**

Clinical relevance—**This paper proposes an accurate automated pancreas segmentation method, which can provide assistance to clinicians in the diagnosis and treatment of pancreatic cancer.**

I. INTRODUCTION

Pancreatic cancer has a high mortality rate with a low five-year survival rate of 8%, which is the 4th most challenging cancer of death [1]. The early diagnosis of pancreatic cancer is difficult. Only less than 30% of patients have the opportunity of surgery at the first diagnosis, and more than 70% of patients have local progress or distant metastasis, completely losing the opportunity of surgery. For clinicians, the key to the diagnosis and clinical operation of pancreatic cancer is its accurate segmentation and positioning. However, pancreatic organs have some characteristics that make the accurate segmentation challenging. 1) The boundary of pancreas is blurred, and it is difficult for clinicians to delineate its boundary clearly. For small pancreas tumors (diameter $<$ 1cm), although there are some signal changes in imaging, it is not easy to observe the subject with our naked eyes, which makes the diagnosis difficult. 2) Pancreatic tumor and the surrounding tissues are intertwined, and their three-dimensional structure is irregular. Only using common scanning and reconstruction methods cannot accurately assess the overall appearance of pancreatic tumor, which is the objective reason why it is difficult to be accurately segmented. Therefore, there is an urgent need for automated and accurate pancreas segmentation method to assist the clinicians in treatment and diagnosis, as done in other applications [2].

At present, deep learning has achieved great success in many difficult tasks, including medical image processing and computer-aided diagnosis [3]. For the task of automatic pancreas segmentation, there are also several deep learning-based methods, which obtain promising results, such as [4] and [5]. However, there are still challenges for applying existing segmentation algorithms to accurate automatic pancreas segmentation, especially when there is only limited training data.

To address these challenges, we propose a 3D Dense Volumetric Network (3D^{2}VNet) to improve the performance of automatic pancreas segmentation. Fig. 1 shows the overview of our method. Firstly, a 3D fully convolutional architecture is adopted to make full use of 3D volumetric information, which can realize effective volume-to-volume prediction. Then, the module of dense connectivity is applied to avoid the learning of redundant feature maps and speed up the training of the network. In addition, a short cut is constructed to provide auxiliary supervision for the network training. Compared with other 3D convolutional networks, our 3D^{2}VNet has fewer parameters and simpler structure, which can prevent the overfitting with limited training samples. Adequate experiments are conducted on a challenging pancreas dataset in Medical Segmentation Decathlon challenge [6]. The results demonstrate the superiority of our 3D^{2}VNet.

The rest of the paper is organized as follows: Section 2 mainly introduces the related work. Section 3 presents the our method in detail. The detailed experiments and result analysis are presented in section 4. Section 5 concludes the paper.

II. RELATED WORK

Pancreas segmentation is an important task in the diagnosis and treatment of pancreatic cancer for clinicians. Recently, several methods have been proposed to address this issue, and since deep learning was introduced to pancreas

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant xzy022020054 and in part by the Clinical Research Award of the First Affiliated Hospital of Xi'an Jiaotong University under Grant XJTU1AF-CRF-2019-002.

[1]R. Wang, X. Liu, and H. Shao are the with School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China

[2]Q. Li is with the Department of Obstetrics and Gynecology, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710061, China.

[1,3,4]D. Zhong is with the School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China, Pazhou Laboratory, Guangzhou 510335, China, and State Key Laboratory of Novel Software Technology, Nanjing University, Nanjing 210093, China. bell@xjtu.edu.cn

*Corresponding author

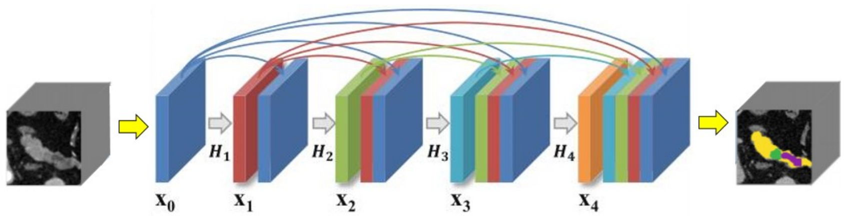

Fig. 1. The overview of our proposed 3D^{2}VNet. 3D pancreas CT volumes are processed to segment the tumor.

segmentation, its accuracy and efficiency have been greatly improved. In [7], in order to improve the segmentation accuracy of small pancreas organs, Zhou *et al.* adopted a predicted segmentation mask to shrink the input region and proposed a fixed-point model. They stated that a smaller input region can lead to more accurate segmentation. Roth *et al.* [8] proposed a two-stage cascaded approach for automatic organ segmentation from 3D computed tomography (CT) volumes, *i.e.*, pancreas localization and pancreas segmentation. Dmitriev *et al.* [9] proposed a semi-automatic segmentation algorithm for pancreas with cysts, and a combination of random walker and region growing method was adopted to delineate the boundaries of pancreas and cysts. Chen *et al.* [10] proposed a multi-scale feature fusion (MsFF) model to segment the pancreas from CT images, which was a well-recognized encoder–decoder framework. Zhou *et al.* [5] proposed a 3D fully convolution neural network for pancreas segmentation, called Hyper-Pairing Network (HPN), which integrated information from different phases. Jiang *et al.* [11] proposed a novel multi-phase and multi-level selective feature fusion network (MMNet) for automated pancreas segmentation, which contains a core adaptive cross refinement (ACR) module. In order to reduce the requirements of deep learning based pancreas segmentation methods for training data, Wang *et al.* [12] introduced federated learning and quantitatively compared it with local training methods. Man *et al.* [13] proposed a deep Q network (DQN) driven method with deformable U-Net for accurate pancreas segmentation, which learned a context-adaptive localization policy and extracted anisotropic features from pancreas. For segmenting the pancreas in magnetic resonance imaging (MRI), Asaturyan *et al.* [14] proposed a dual-stage automatic segmentation method. In [15], Zhao *et al.* proposed a fully automated two stage framework for accurate pancreas segmentation and obtained promising results. Farag *et al.* [16] proposed a novel bottom-up approach for pancreas segmentation in abdominal CT scans, which generated a hierarchical cascade of information propagation and consisted of four steps.

Another line of related work is 3D convolutional neural networks (CNN), which is modified from 2D CNN to process 3D volumetric information. Çiçek *et al.* [17] replaced the 2D operations of previous 2D U-net architecture with 3D counterparts for dense 3D segmentation. Dou *et al.* [18] proposed an end-to-end learning and inference method, 3D deeply supervised network (3D DSN), for automatic liver segmentation from CT volumes. In order to improve the accuracy, Wu *et al.* [19] constructed a novel joint 3D+2D fully convolutional framework to segment the subcortical structures from MRIs. To mitigate the noise artifact in function magnetic resonance imaging (fMRI) data, Zhao *et al.* [20] proposed a fully data-driven 3D convolutional encapsulated Long Short-Term Memory (3DConv-LSTM) method based on adversarial network. Inspired by them, in this paper, we propose 3D^{2}VNet to further improve the accuracy of automatic segmentation.

III. METHOD

A. Dense Connectivity

Recently, CNN has outperformed the handcrafted approaches and obtained the state-of-the-art performance in the computer vision field. For a common CNN, there are several layers, which can be denoted as

$$x_l = H_l\left(x_{l-1}\right), \tag{1}$$

where x_l is the output of l^{th} layer and the x_{l-1} is the output of $(l-1)^{th}$ layer, which is also the input of l^{th} layer. H_l represents the non-linear transformation from the output of the previous layer, which can consist of rectified linear units (ReLU), convolution, Batch Normalization (BN), and pooling.

Then, in order to improve the performance, ResNet [21] constructs a skip-connection to boost the training against the vanishing gradients, which bypasses the non-linear transformations with an identity function:

$$x_l = H_l\left(x_{l-1}\right) + x_{l-1}. \tag{2}$$

It sums the identity function from later layers with the earlier layers, which may hinder the flow of information in the network.

In order to further improve the information flow between different layers, DenseNet [22] introduces an idea of dense connectivity, which implements the connections from any layer to all subsequent layers. It can be defined as:

$$x_l = H_l\left(\left[x_0, x_1, ..., x_{l-1}\right]\right), \tag{3}$$

where $[x_0, x_1, ..., x_{l-1}]$ represents the concatenation of different layers. The dense connectivity can improve the flow of information and gradients in the network to make it easy to train. In addition, there is a *growth rate* referred to *k* feature maps of the output in each layer, which can be set to a small number to reduce the overfitting with smaller training set sizes in medical image processing.

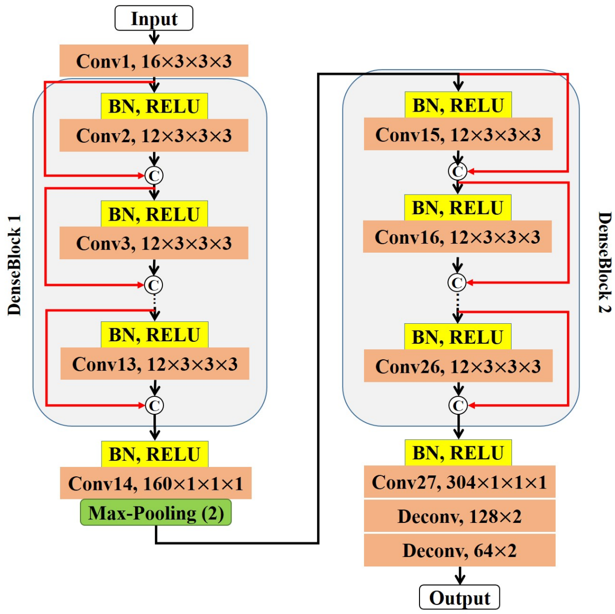

Fig. 2. The parameters of 3D²VNet proposed. The filter size, strides, windows, and activation functions are described.

B. The Architecture of 3D²VNet

According to the characteristic of 3D volumes, our method adopts 3D CNN architecture to learn volumetric feature representation. Dense connectivity is introduced to alleviate the vanishing-gradient problem and reduce the overfitting on tasks with fewer training samples. There are two DenseBlocks comprised of several layers with dense connections. Each layer consists of convolution, BN, ReLU, max-pooling and activation function. The details and parameters of each layer are list in Fig. 2

IV. EXPERIMENTS AND RESULTS

A. Dataset

The pancreas dataset in Medical Segmentation Decathlon challenge is adopted to evaluate our method [6]. There are 421 3D portal venous phase CT volumes collected by Memorial Sloan Kettering Cancer Center (New York, NY, USA) and comprised of patients undergoing resection of pancreatic masses. In each slice, the background, pancreatic parenchyma, and pancreatic mass (cyst or tumour) were manually segmented by an expert abdominal radiologist using the Scout application. In the official challenge, 421 3D volumes are divided into 282 volumes for training and 139 volumes for testing. In fact, the ground truth annotation of test set is not available. So in our experiments, only 282 pancreas CT volumes from official train set are adopted to evaluate our method, which are further divided into 225 training volumes and 57 testing volumes. Some typical samples are shown in Fig. 3

B. Implementation Details

The experiments are implemented using TensorFlow on a NVIDIA GPU GTX2080 Super and i9-3.6GHz processors. The base learning rate is set to 0.0001 and RMSProp [23] optimizer is adopted. A simple three dimensional data augmentation was adopted to leverage the limited training data, which includes image flipping and the rotation with 90, 180 and 270°. Each original slice is cut into small patches with the size of 32×32×32. Then 1,000 patches are randomly selected to training the model and 250 patches are used for testing. The Dice Similarity Coefficients (DSC) of background, pancreas, and tumor are adopted as the performance metric to evaluate the model.

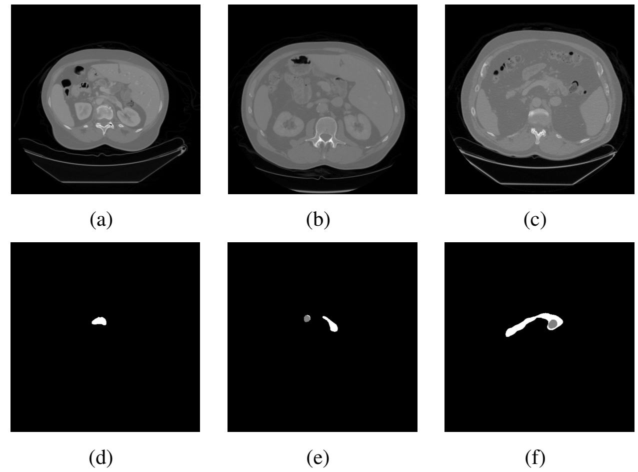

Fig. 3. Some typical samples of pancreas dataset. (a), (b), and (c) are original CT volumes; (d), (e), and (f) are their ground truth annotation.

TABLE I
THE PERFORMANCE OF DIFFERENT METHODS

Method	DSC (%)			
	Background	Pancreas	Tumor	Average
Med3D [24]	89.04	47.19	17.29	51.17
V-NET [25]	88.33	53.62	18.31	53.42
HC [26]	89.72	51.03	17.94	53.42
3D U-net [17]	90.88	53.85	12.14	52.90
3D²VNet (ours)	**93.18**	**54.84**	12.63	**53.55**

C. Results and Analyses

Table I shows the results of our 3D²VNet. The pancreas organ is small ans its boundary is blurred, so it is difficult to segment. From the results, it can be observed that our method can obtain promising results on the task of automated pancreas segmentation.

In order to evaluate the effectiveness of our 3D²VNet, we present the results of some baseline methods, as shown in Table I. Med3D [24] is a heterogeneous 3D network and aimed to provide a pre-trained model using 3D multi-domain medical data. V-NET [25] is trained using the Dice overlap coefficient between the predicted segmentation and the annotation. High-resolution compact (HC) network architecture [26] investigates the efficient and flexible dilated convolution and residual connection to improve the performance of volumetric image segmentation. 3D U-net [17] replaces all

2D operations in previous 2D U-net architecture with their 3D counterparts to learn 3D representation. Note that all the comparison methods adopt the same parameters and structures. It can be observed that our 3D^{2}VNet can obtain better segmentation results, suggesting that it can extract more discriminative features to better process the medical images.

V. Conclusion

In this paper, a novel 3D^{2}VNet is proposed to improve the performance of automated pancreas segmentation. A 3D fully convolutional network is firstly adopted to fully incorporate the 3D information of pancreas volumes for effective volume-to-volume prediction. In order to help the training of network and reduce the overfitting with limited training set sizes, the strategy of dense connectivity is further introduced. Adequate experiments are conducted on the pancreas dataset in Medical Segmentation Decathlon challenge and the results show that our 3D^{2}VNet can outperform other baseline models by a large margin.

References

[1] J. K. Bjerregaard, M. B. Mortensen, H. A. Jensen, M. Nielsen, and P. Pfeiffer, "Prognostic factors for survival and resection in patients with initial nonresectable locally advanced pancreatic cancer treated with chemoradiotherapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 83, no. 3, pp. 909–915, 2012.

[2] Y. Guo, Y. Gao, and D. Shen, "Deformable MR prostate segmentation via deep feature learning and sparse patch matching," *IEEE Trans. Medical Imaging*, vol. 35, no. 4, pp. 1077–1089, 2016.

[3] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Anal.*, vol. 42, pp. 60–88, 2017.

[4] H. R. Roth, L. Lu, A. Farag, A. Sohn, and R. M. Summers, "Spatial aggregation of holistically-nested networks for automated pancreas segmentation," in *19th International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016, Athens, Greece, October 17-21, 2016*, pp. 451–459.

[5] Y. Zhou, Y. Li, Z. Zhang, Y. Wang, A. Wang, E. K. Fishman, A. L. Yuille, and S. Park, "Hyper-pairing network for multi-phase pancreatic ductal adenocarcinoma segmentation," in *22nd International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI 2019, Shenzhen, China, October 13-17, 2019*, pp. 155–163.

[6] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features," *Scientific Data*, vol. 4, 2017.

[7] Y. Zhou, L. Xie, W. Shen, Y. Wang, E. K. Fishman, and A. L. Yuille, "A fixed-point model for pancreas segmentation in abdominal CT scans," in *20th International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI 2017, Quebec City, QC, Canada, September 11-13, 2017*, pp. 693–701.

[8] H. R. Roth, L. Lu, N. Lay, A. P. Harrison, A. Farag, A. Sohn, and R. M. Summers, "Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation," *Medical Image Anal.*, vol. 45, pp. 94–107, 2018.

[9] K. Dmitriev, I. Gutenko, S. Nadeem, and A. E. Kaufman, "Pancreas and cyst segmentation," in *Medical Imaging 2016: Image Processing, San Diego, California, USA, February 27, 2016*.

[10] Z. Chen, X. Wang, K. Yan, and J. Zheng, "Deep multi-scale feature fusion for pancreas segmentation from CT images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 3, pp. 415–423, 2020.

[11] X. Jiang, Q. Luo, Z. Wang, T. Mei, Y. Wen, X. Li, K. Cheng, and X. Yang, "Multi-phase and multi-level selective feature fusion for automated pancreas segmentation from CT images," in *23rd International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI 2020, Lima, Peru, October 4-8, 2020*, pp. 460–469.

[12] P. Wang, C. Shen, H. R. Roth, D. Yang, D. Xu, M. Oda, K. Misawa, P. Chen, K. Liu, W. Liao, W. Wang, and K. Mori, "Automated pancreas segmentation using multi-institutional collaborative deep learning," in *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning - Second MICCAI Workshop, DART 2020, Lima, Peru, October 4-8, 2020*, pp. 192–200.

[13] Y. Man, Y. Huang, J. Feng, X. Li, and F. Wu, "Deep Q learning driven CT pancreas segmentation with geometry-aware u-net," *IEEE Trans. Medical Imaging*, vol. 38, no. 8, pp. 1971–1980, 2019.

[14] H. Asaturyan, E. L. Thomas, J. Fitzpatrick, J. D. Bell, and B. Villarini, "Advancing pancreas segmentation in multi-protocol MRI volumes using hausdorff-sine loss function," in *10th International Workshop on Machine Learning in Medical Imaging, MLMI 2019, Shenzhen, China, October 13, 2019*, pp. 27–35.

[15] N. Zhao, N. Tong, D. Ruan, and K. Sheng, "Fully automated pancreas segmentation with two-stage 3d convolutional neural networks," in *22nd International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI 2019, Shenzhen, China, October 13-17, 2019*, pp. 201–209.

[16] A. Farag, L. Lu, H. R. Roth, J. Liu, E. Turkbey, and R. M. Summers, "A bottom-up approach for pancreas segmentation using cascaded superpixels and (deep) image patch labeling," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 386–399, 2017.

[17] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: Learning dense volumetric segmentation from sparse annotation," in *19th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2016, Athens, Greece, October 17-21, 2016*, pp. 424–432.

[18] Q. Dou, H. Chen, Y. Jin, L. Yu, J. Qin, and P. Heng, "3d deeply supervised network for automatic liver segmentation from CT volumes," in *19th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2016, Athens, Greece, October 17-21, 2016*, pp. 149–157.

[19] J. Wu, Y. Zhang, and X. Tang, "A joint 3d+2d fully convolutional framework for subcortical segmentation," in *22nd International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI 2019, Shenzhen, China, October 13-17, 2019*, pp. 301–309.

[20] C. Zhao, H. Li, Z. Jiao, T. Du, and Y. Fan, "A 3d convolutional encapsulated long short-term memory (3dconv-lstm) model for denoising fmri data," in *23rd International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI 2020, Lima, Peru, October 4-8, 2020*. Springer, pp. 479–488.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pp. 770–778.

[22] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pp. 2261–2269.

[23] M. Andrychowicz, M. Denil, S. G. Colmenarejo, M. W. Hoffman, D. Pfau, T. Schaul, and N. de Freitas, "Learning to learn by gradient descent by gradient descent," in *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems, 2016, Barcelona, Spain*, pp. 3981–3989.

[24] S. Chen, K. Ma, and Y. Zheng, "Med3d: Transfer learning for 3d medical image analysis," *CoRR*, vol. abs/1904.00625, 2019. [Online]. Available: http://arxiv.org/abs/1904.00625

[25] F. Milletari, N. Navab, and S. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *Fourth International Conference on 3D Vision, 3DV 2016, Stanford, CA, USA, October 25-28, 2016*, pp. 565–571.

[26] W. Li, G. Wang, L. Fidon, S. Ourselin, M. J. Cardoso, and T. Vercauteren, "On the compactness, efficiency, and representation of 3d convolutional networks: Brain parcellation as a pretext task," in *Information Processing in Medical Imaging - 25th International Conference, IPMI 2017, Boone, NC, USA, June 25-30, 2017*, pp. 348–360.