



OPEN

Automated pancreas segmentation and volumetry using deep neural network on computed tomography

Sang-Heon Lim^{1,2,4}, Young Jae Kim^{2,4}, Yeon-Ho Park³, Doojin Kim³, Kwang Gi Kim^{1,2}✉ & Doo-Ho Lee³✉

Pancreas segmentation is necessary for observing lesions, analyzing anatomical structures, and predicting patient prognosis. Therefore, various studies have designed segmentation models based on convolutional neural networks for pancreas segmentation. However, the deep learning approach is limited by a lack of data, and studies conducted on a large computed tomography dataset are scarce. Therefore, this study aims to perform deep-learning-based semantic segmentation on 1006 participants and evaluate the automatic segmentation performance of the pancreas via four individual three-dimensional segmentation networks. In this study, we performed internal validation with 1,006 patients and external validation using the cancer imaging archive pancreas dataset. We obtained mean precision, recall, and dice similarity coefficients of 0.869, 0.842, and 0.842, respectively, for internal validation via a relevant approach among the four deep learning networks. Using the external dataset, the deep learning network achieved mean precision, recall, and dice similarity coefficients of 0.779, 0.749, and 0.735, respectively. We expect that generalized deep-learning-based systems can assist clinical decisions by providing accurate pancreatic segmentation and quantitative information of the pancreas for abdominal computed tomography.

The detection rate of benign or malignant lesions of the pancreas, and subsequent surgery are gradually increasing, owing to the early diagnosis of pancreatic neoplasm, which is a result of the development of imaging modalities, an increase in the health screening program, and the aging of the population¹⁻⁴. In particular, cystic tumors that are inadvertently identified in the pancreas require continuous follow-up^{2,3}. This is typically followed by computed tomography (CT) scans of the abdomen to observe the increase in lesion size. Subsequently, resection of benign or malignant tumors on the endocrine and exocrine function of the pancreas is implemented in the long-term, which greatly affects the patient's quality of life⁴⁻⁶.

It is necessary to investigate the change in the volume of the pancreas after resection; however, this is difficult to apply in clinical practice because it is cumbersome and laborious to obtain the volume of the pancreas from abdominal CT using current technology^{5,6}. The quantitative pancreatic volume cannot be measured in all patients after resection of the pancreas because obtaining the volume of the pancreas is a long and time-consuming task. In addition, determining the volume of the pancreas by hand is error-prone for each examiner. Therefore, this situation necessitates a computer-aided diagnosis (CAD) system based on artificial intelligence.

Automatically obtaining the volume of the pancreas from abdominal CT scans based on artificial intelligence can assist in calculating the quantitative pancreatic volume and the patient's endocrine and exocrine functions, which enables a more scientific and objective treatment for the patients. Therefore, the current study develops a technique for calculating the volume of the pancreas based on deep learning technology using abdominal CT scan images.

Recently, deep learning (DL)-based semantic segmentation networks were considered more beneficial for medical image segmentation tasks compared with traditional image segmentation methods, such as the intensity-based threshold, morphology, and geometry⁷⁻¹⁰. However, accurate pancreas segmentation is a challenging task

¹Department of Health Sciences and Technology, Gachon Advanced Institute for Health Sciences and Technology (GAIHST), Gachon University, 1342, Seongnamdaero, Sujeong-gu, Seongnam-si 13120, Republic of Korea. ²Department of Biomedical Engineering, Gil Medical Center, Gachon University College of Medicine, 38-13, Dokjeon-ro 3beon-gil, Namdong-gu, Incheon 21565, Republic of Korea. ³Department of Surgery, Gil Medical Center, Gachon University College of Medicine, 774-21, Namdong-daero, Namdong-gu, Incheon 21565, Republic of Korea. ⁴These authors contributed equally: Sang-Heon Lim and Young Jae Kim. ✉email: kimkg@gachon.ac.kr; dooholeemd@gilhospital.com

because the pancreas is structurally diverse, it occupies a small region in the abdomen, and it is closely attached to other organs, such as the duodenum and gallbladder^{11,12}.

However, a convolutional neural network (CNN)-based method was proposed as a promising method for pancreas segmentation, owing to the powerful advantages of the DL method. Subsequently, several studies proposed state-of-the-art CNN-based pancreas segmentation approaches via either cascaded or coarse-to-fine segmentation networks. However, previous pancreas segmentation studies were performed on small study populations^{11,13,14} that comprised 82 participants from the National Institutes of Health (NIH) clinical center. Furthermore, DL methods are sensitive to the features of the data that are encoded in the network; therefore, clinical assessment of the pancreatic segmentation performance is necessary for varied and large datasets. However, to the best of our knowledge, DL studies on large CT datasets that contain various pancreatic volumes are scarce.

Therefore, we aim to evaluate the performance of four DL-based three-dimensional (3D) pancreas segmentation networks on 1006 healthy participants. In addition, we evaluate the reliability of the pancreatic-volume estimation task using DL-based approaches. In this study, we exploit four semantic segmentation networks based on a 3D u-net. One of the four networks is the basic 3D u-net, and the other three networks are configured with residual modules, dense modules, and residual dense modules. We assess DL networks using segmentation metrics (i.e., dice similarity coefficient, precision, and recall) as well as a regression plot and Bland–Altman plot for pancreatic volumetric evaluation.

Methods

Study populations. We acquired abdominal CT images from 1006 patients, who were examined at the Gil Medical Center. All patient records were confirmed and retrospectively reviewed based on a clinical diagnosis from 2016 to 2019. This study was conducted in accordance with the Declaration of Helsinki and written informed consent was obtained from all the participants (IRB number: GDIRB2020-121). This study was approved by the Institutional Review Board of the Gil Medical Center. The inclusion criteria for this study were as follows: (1) patient did not undergo pancreatic resection, and (2) patient has no benign or malignant tumor in the pancreas. The feature of the CT dataset has a slice thickness of 3–5 mm and a pixel spacing of 0.58–0.97 mm. We used a CT scanner (SOMATOM Definition Edge, Siemens, Germany), and images were acquired using a tube-voltage of 80–150 kVp and tube-current of 52–641 mA.

Preprocessing and experimental setup. Manual delineation was conducted in the 2D axial plane using ImageJ (ver. 1.52a, NIH, USA) to generate a gold standard. As the acquired CT volumes have a different voxel-spacing couple, we unified the voxel spacing of all the volume data; the slice thickness was regularized to 3 mm, and the pixel spacing to 1 mm ($z, y, x = 3, 1, 1$). Moreover, because the manual delineation was conducted before conducting volume reconstruction, we simultaneously reconstructed the CT and mask volumes. Based on the reconstructed mask volume, a specific margin was assigned to crop the region of the pancreas.

Owing to the irregular shape of the pancreas (x -, y -, z -axis), we cropped the image considering the ratio of the depth, width, and height of the pancreas ($z:y:x = 1:2:3$). Additionally, the volume of the cropped pancreatic region varies according to the patient; therefore, bilinear interpolation was applied to create a volume with a particular single-channel size (64, 128, 256, 1) (Fig. 1a). The pancreas cropping process was conducted based on manual delineation.

The pancreas is attached to other organs, such as the duodenum and gallbladder; therefore, contrast enhancement was applied to the input volume to increase the visibility of the pancreas. First, we adjusted the CT images using a window center (60) and window width (400) to clearly observe the region of the abdomen¹⁵. The final dataset was generated by applying contrast-limited adaptive histogram equalization (CLAHE) to enhance the contrast of the pancreatic region (Fig. 1b)^{16–19}.

The output images and resized images were restored using the same voxel spacing as the raw CT data during preprocessing, before input to the network, to evaluate the pancreatic segmentation performance of the network. Additionally, we conducted fourfold cross validation on binary images that restored the voxel spacing of raw CT data (Fig. 1c). External validation was performed using the Cancer Imaging Archive (TCIA) pancreas-CT dataset^{14,20,21}, which was provided by the NIH clinical center ($n = 82$). The TCIA dataset was split as the ratio of 10:5:5 for fourfold cross validation. The segmentation performance assessment was performed via a pixel-wise comparison between the gold standard and prediction results of the DL network (if probability > 0.5 , positive). As a result of the assessment, we obtained a confusion matrix (i.e., true positive, false positive, true negative, and false negative) from 3D binary volume images.

Network architecture. We exploited 3D u-net-based architectures with skip connections and batch normalization, which consisted of four resolution steps^{22–26}. All convolution blocks comprised a convolutional kernel size of $(3 \times 3 \times 3)$, dilation rate²⁷ of $(1 \times 1 \times 1)$, and rectified linear units (ReLUs). The hyper-parameter setting was set to the hyper-parameter that achieved the best performance in all baseline networks. In addition, we employed simple 3D upsample layers instead of transposed convolution layers for the decoding steps. Figure 2 shows the architecture of the residual dense u-net for pancreas segmentation^{28,29}. We performed deep learning analysis using four semantic segmentation networks that had the same width, depth, and filter size, except for specific blocks (i.e., dense blocks, residual blocks, and residual dense blocks). For a network comparison, we experimented by replacing the blocks in the residual dense blocks in Fig. 2 with the other specific modules.

Implementation details. This study conducted a deep learning analysis on a Tesla V100 (32 GB) graphics processing unit (GPU). The networks were trained using the Adam optimizer (learning rate 0.001) to minimize dice loss. We utilized the following frameworks using Python (ver. 3.6.12, Python Software Foundation, USA):

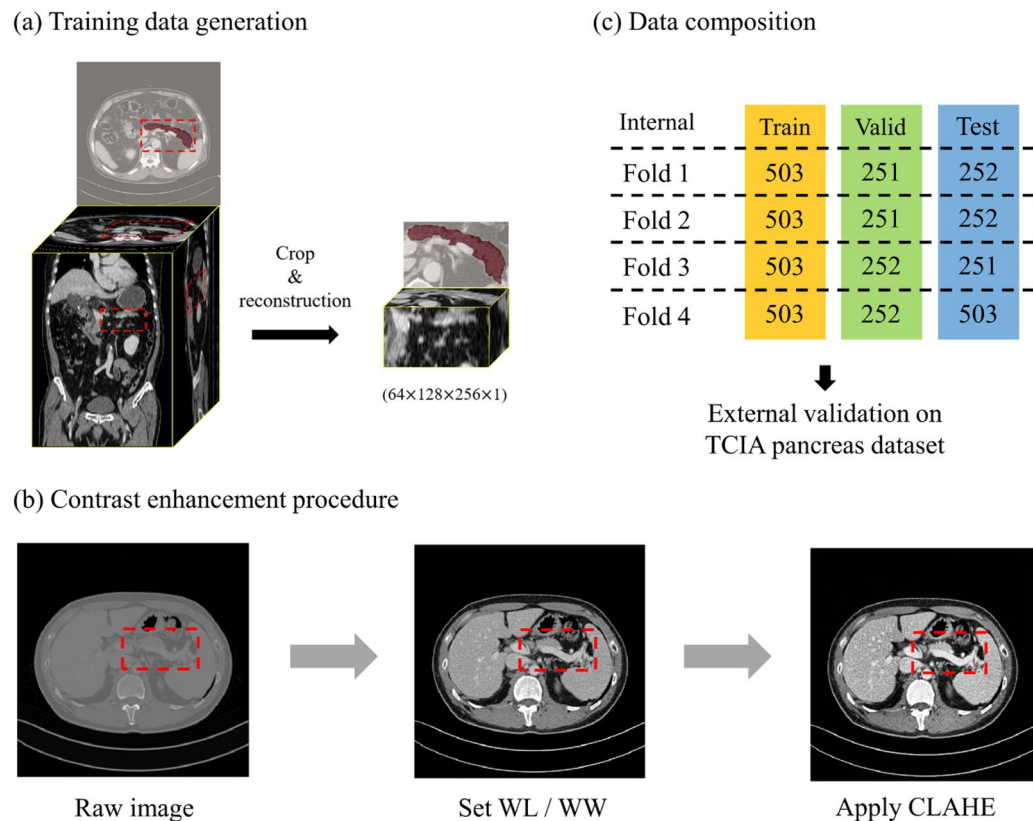


Figure 1. (a) Raw volume data were cropped and reconstructed for training data generation. (b) All data were divided into datasets that consisted of almost identical numbers of participants for cross validation. (c) A region of the pancreas was enhanced via CLAHE. Valid, validation; TCIA, the Cancer Imaging Archive; WL, window level; WW, window width; CLAHE, contrast-limited adaptive histogram equalization.

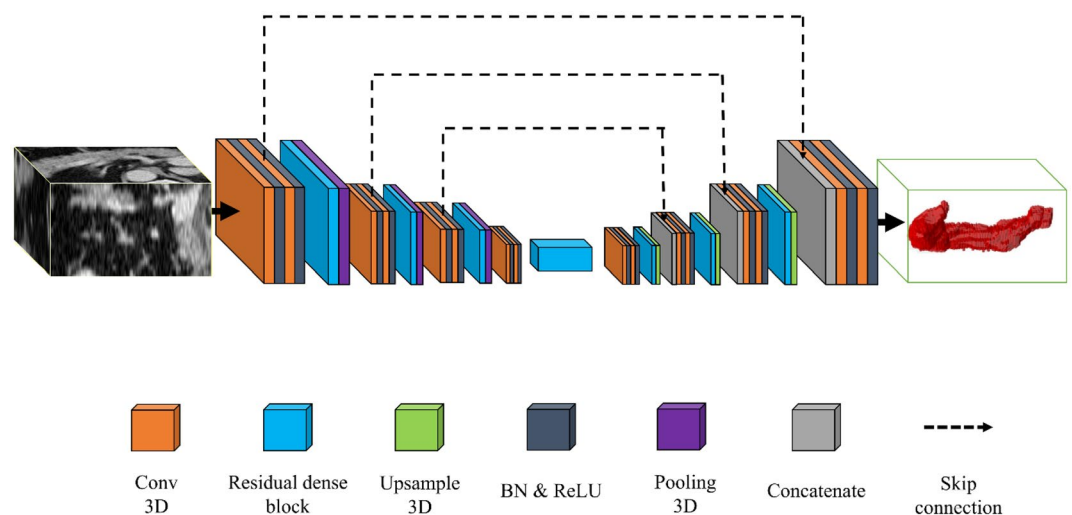


Figure 2. Architecture illustration of residual dense u-net. Conv, convolution; BN, batch normalization; ReLU, rectified linear unit.

Keras (ver. 2.2.5), TensorFlow-GPU (ver. 1.15.4). The training settings of all networks are as follows: batch size, 2; epoch, 500.

	Total	Men	Women	<i>p</i> value
Number	1006 (100)	530 (52.6)	476 (47.4)	
Age (years)	55.3 ± 15.6	55.6 ± 15.3	54.9 ± 15.9	0.508
Height (cm)	163.5 ± 8.9	169.3 ± 6.8	157.1 ± 6.2	<0.001
Weight (kg)	65.3 ± 12.1	70.3 ± 11.6	59.6 ± 10.0	<0.001
Body mass index (kg/m ²)	24.3 ± 3.5	24.4 ± 3.3	24.1 ± 3.7	0.165
Hypertension	364 (36.3)	205 (38.8)	159 (33.5)	0.078
Diabetes mellitus	190 (18.9)	108 (20.5)	82 (17.3)	0.198
Smoking	224 (22.3)	185 (35.0)	39 (8.2)	<0.001
Alcohol	330 (32.9)	246 (46.6)	84 (17.7)	<0.001
Volume of pancreas (cm ³)	62.6 ± 19.0	68.8 ± 19.5	55.8 ± 16.0	<0.001

Table 1. Demographics of participants. Values are expressed as *n* (%) or mean ± standard deviation, unless otherwise indicated.

	Precision	Recall	DSC	Trainable parameter
Basic U-net	<i>0.861 ± 0.468</i>	<i>0.816 ± 0.173</i>	<i>0.822 ± 0.143</i>	11,003,073
Dense U-net	<i>0.864 ± 0.114</i>	<i>0.828 ± 0.165</i>	<i>0.831 ± 0.134</i>	35,261,601
Residual U-net	<i>0.843 ± 0.127</i>	<i>0.810 ± 0.178</i>	<i>0.808 ± 0.146</i>	2,350,857
Residual Dense U-net	0.869 ± 0.110	0.842 ± 0.156	0.842 ± 0.128	47,074,657

Table 2. Evaluation metrics for four pancreas segmentation models. Results are indicated as mean ± standard deviation, and the best performances are indicated in bold. The results are highlighted in italics if the residual dense u-net performs significantly better than the corresponding method. We used a significance level of 0.05 and a paired *t* test for network comparison. DSC, dice similarity coefficient.

Results

Participant demographics. Table 1 shows the demographics of the participants who underwent abdominal CT for routine health check-ups via a health-care program. A total of 528 (52.6%) participants were men and 475 (47.4%) were women. The mean age was 55.3 years, and the mean body mass index was 24.3 kg/m². Significant differences were observed in height (men: 169.3 ± 6.8 vs. women: 157.1 ± 6.2 cm, *p* < 0.001), weight (70.3 ± 11.6 vs. 59.6 ± 10.0 kg, *p* < 0.001), smoking (*n* = 224, 22.3% vs. *n* = 39, 8.2%, *p* < 0.001), and alcohol consumption (*n* = 246, 46.6% vs. *n* = 84, 17.7%, *p* < 0.001) between men and women. The mean volume of the pancreas was 66.5 cm³, and a significant difference was observed between men and women (68.8 ± 19.5 vs. 55.8 ± 16.0, *p* < 0.001). No significant differences were observed in men and women based on age, body mass index, and the proportion of those with hypertension and diabetes mellitus.

Pancreas segmentation. Table 2 presents the evaluation results of the four 3D segmentation models. Networks using residual dense blocks achieved the highest precision, recall, and DSC, and they also exhibited the lowest standard deviation. Additionally, we performed paired *t* tests to verify the statistical significance between residual dense u-nets and other networks. The results showed that the residual dense u-net was promising and significantly different from the other three networks (all significance levels were *p* < 0.05). In contrast, the residual u-net achieved the lowest pancreas segmentation performance in terms of precision, recall, and DSC. The residual dense u-net obtained a mean precision, recall, and DSC of 0.779 ± 0.204, 0.749 ± 0.226, and 0.735 ± 0.197, respectively, on the NIH external dataset. Furthermore, the residual dense u-net achieved the highest mean DSC for every pancreas volume range: (1) 0–30 cm³, mean DSC of 0.808; (2) 30–60 cm³, mean DSC of 0.851; (3) 60–90 cm³, mean DSC of 0.872; (4) > 90 cm³, mean DSC of 0.870 on our dataset (Table 3). The statistical assessment was performed on 2D axial plane.

We visually assessed the four semantic segmentation models in the 2D axial plane and 3D volumes based on a single patient's CT (Fig. 3). 3D visualization was conducted via 3D volume rendering with a 3D Slicer (ver. 4.11.20200930; <http://www.slicer.org>).

Pancreas volume estimation. We evaluated the pancreas volume estimation performance of the residual dense u-net using the Bland–Altman plot and regression analysis (Fig. 4). Most of the estimation errors outside the coefficient of repeatability (± 1.96 SD) were underestimated (*n* = 32). In contrast, over-estimations did not occur often (*n* = 4). We performed correlation and intraclass correlation coefficient (ICC) analyses for pancreatic volumetry. For the internal validation, we obtained an R² score of 0.954 (*p* < 0.001) using the regression analysis, an R score of 0.977 (*p* < 0.001) using the correlation analysis, and an ICC score of 0.987. For the external validation, we obtained R², R, and ICC scores of 0.667 (*p* < 0.001), 0.817 (*p* < 0.001), and 0.894, respectively. We used MedCalc Statistical Software (ver. 14.8.1, <https://www.medcalc.org>) for the statistical analysis.

	DSC	*P-value
PV < 30 cm ³ (n = 54)		
Basic U-net	0.785 ± 0.100	< 0.001
Dense U-net	0.794 ± 0.089	0.013
Residual U-net	0.756 ± 0.111	< 0.001
Residual Dense U-net	0.808 ± 0.078	–
30 cm ³ ≤ PV < 60 cm ³ (n = 361)		
Basic U-net	0.834 ± 0.073	< 0.001
Dense U-net	0.842 ± 0.066	< 0.001
Residual U-net	0.815 ± 0.082	< 0.001
Residual Dense U-net	0.851 ± 0.060	–
60 cm ³ ≤ PV < 90 cm ³ (n = 441)		
Basic U-net	0.859 ± 0.047	< 0.001
Dense U-net	0.866 ± 0.039	< 0.001
Residual U-net	0.844 ± 0.053	< 0.001
Residual Dense U-net	0.872 ± 0.037	–
PV ≥ 90 cm ³ (n = 150)		
Basic U-net	0.852 ± 0.078	< 0.001
Dense U-net	0.857 ± 0.079	< 0.001
Residual U-net	0.836 ± 0.082	< 0.001
Residual Dense U-net	0.870 ± 0.074	–

Table 3. Comparison of pancreas segmentation performance according to pancreatic volumes using four independent 3D networks. Results are indicated as mean ± standard deviation, and the best performances are indicated in bold. PV, pancreatic volume; DSC, dice similarity coefficient. *We used a paired *t* test to compare the residual dense u-net with the corresponding network and used a significance level of 0.05.

Discussion

This study presents an automated deep learning method for pancreatic segmentation and volumetry using the abdominal CT images of 1006 participants who underwent a health checkup. Recently, various studies have suggested a promising DL network for pancreas segmentation. However, to the best of our knowledge, there is no existing study on a DL approach applied and evaluated on a large abdominal CT dataset of more than 1000 patients. DL-based medical image segmentation is highly dependent on the number of data points. However, previously presented DL-based pancreatic segmentation studies used the NIH pancreas-CT dataset (n = 82). Although the previously proposed DL networks achieved high performance for pancreas segmentation (mean DSC of 0.866¹¹, 0.854¹³ and 0.859³⁰), there is insufficient data to prove that those networks are reliable. Therefore, in this study, we presented a DL-based pancreas segmentation on a large dataset (i.e., 1,006 abdominal CT images) and conducted external validation on the NIH pancreas-CT dataset using four state-of-the-art 3D segmentation networks. We demonstrated that residual dense u-net enables accurate pancreas segmentation and volumetry: (1) mean precision, recall, and DSC of 0.869, 0.842, and 0.842 for internal validation; (2) mean precision, recall, and DSC of 0.779, 0.749, and 0.735 for external validation. We confirmed that the number of trainable parameters is proportional to the segmentation performance of the DL approaches. The segmentation performance on the external NIH pancreas CT dataset was significantly inferior to that of the internal dataset. We assume that these results were attributable to the different slice thicknesses of the CT images; the external dataset was acquired using a 1.5–2.5 mm slice thickness.

In this study, a DSC comparison was performed according to four pancreatic volume (PV) ranges in four networks used for pancreas 3D-segmentation: (1) PV < 30 cm³, n = 54; (2) 30 cm³ ≤ PV < 60 cm³, n = 361; (3) 60 cm³ ≤ PV < 90 cm³, n = 441; (4) PV > 90 cm³, n = 150. In the total volume range, the residual dense u-net achieved the highest mean DSCs (Fig. 3b; Table 3). The mean DSC had a positive correlation with the pancreas volume, and all networks achieved the highest mean DSC results in samples with a volume of 60–90 cm³. Generally, the network achieves a high DSC, which is proportional to the volume of the pancreas. However, we assumed that high segmentation performance was achieved for samples with a pancreatic volume of 60–90 cm³ owing to the high ratio of 60–90 cm³ samples in the dataset (43.84%). In contrast, all the networks achieved the lowest segmentation performance for abdominal CT for patients with a pancreatic volume of 0–30 cm³.

We assessed the residual dense u-net-based pancreatic volume measurements using the Bland–Altman plot and regression plots (Fig. 4). The agreement between the network pancreatic volume measurement and the manual measurements was high, and there were mean differences between DL-based and manual-based pancreatic volume estimation. For the internal validation, the mean difference was 1.67 cm³, and the mean difference of the external dataset was 2.34 cm³. Most pancreatic volume estimation results were reliable; however, a few underestimations (n = 32) and over-estimations (n = 4) existed in a total of 1006 datapoints. Most of the underestimation occurred for a pancreatic volume greater than 90 cm³, which is presumed to be owing to the blurred boundary or low density of soft tissue.

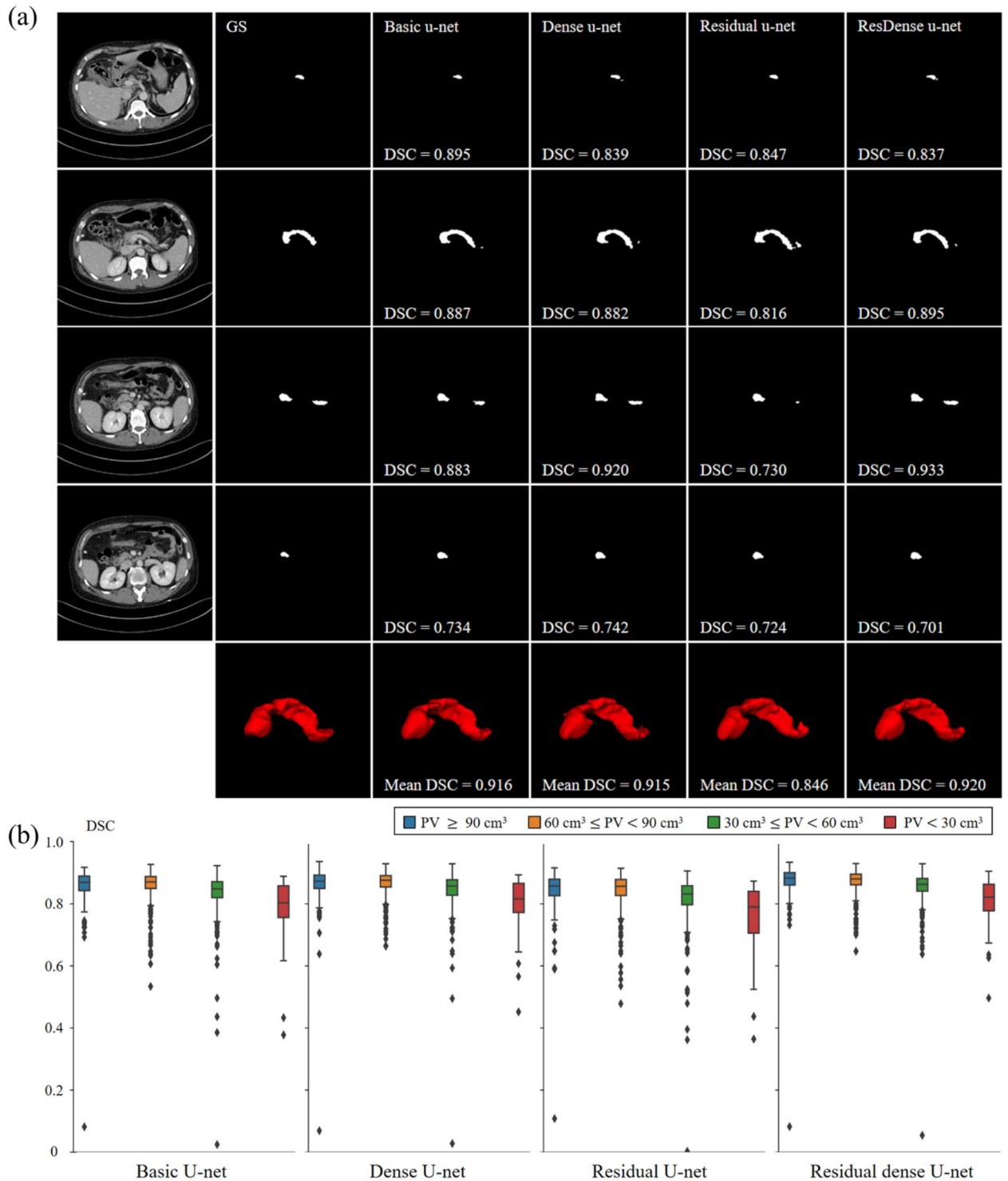


Figure 3. (a) Representative examples of pancreas segmentation in the 2D axial plane and 3D volume of one patient. (b) DSC metric in each deep learning model according to the volume of the pancreas. GS, gold standard; DSC, dice similarity coefficient; ResDense, residual dense; PV, pancreatic volume.

This study presented a semi-automated pancreas segmentation approach based on DL methods for 1006 participants. However, there are several limitations to our study. We manually cropped the volume of interest (region of pancreas) to train the DL networks, owing to a lack of random access memory and GPU memory. Accordingly, further study is necessary to achieve fully automated pancreas segmentation using two-stage methods, such as cascaded or coarse-to-fine networks. Moreover, we assumed that other segmentation methods^{31–34} may be appropriate for accurate pancreas segmentation, owing to the blurry boundaries of the pancreas. The

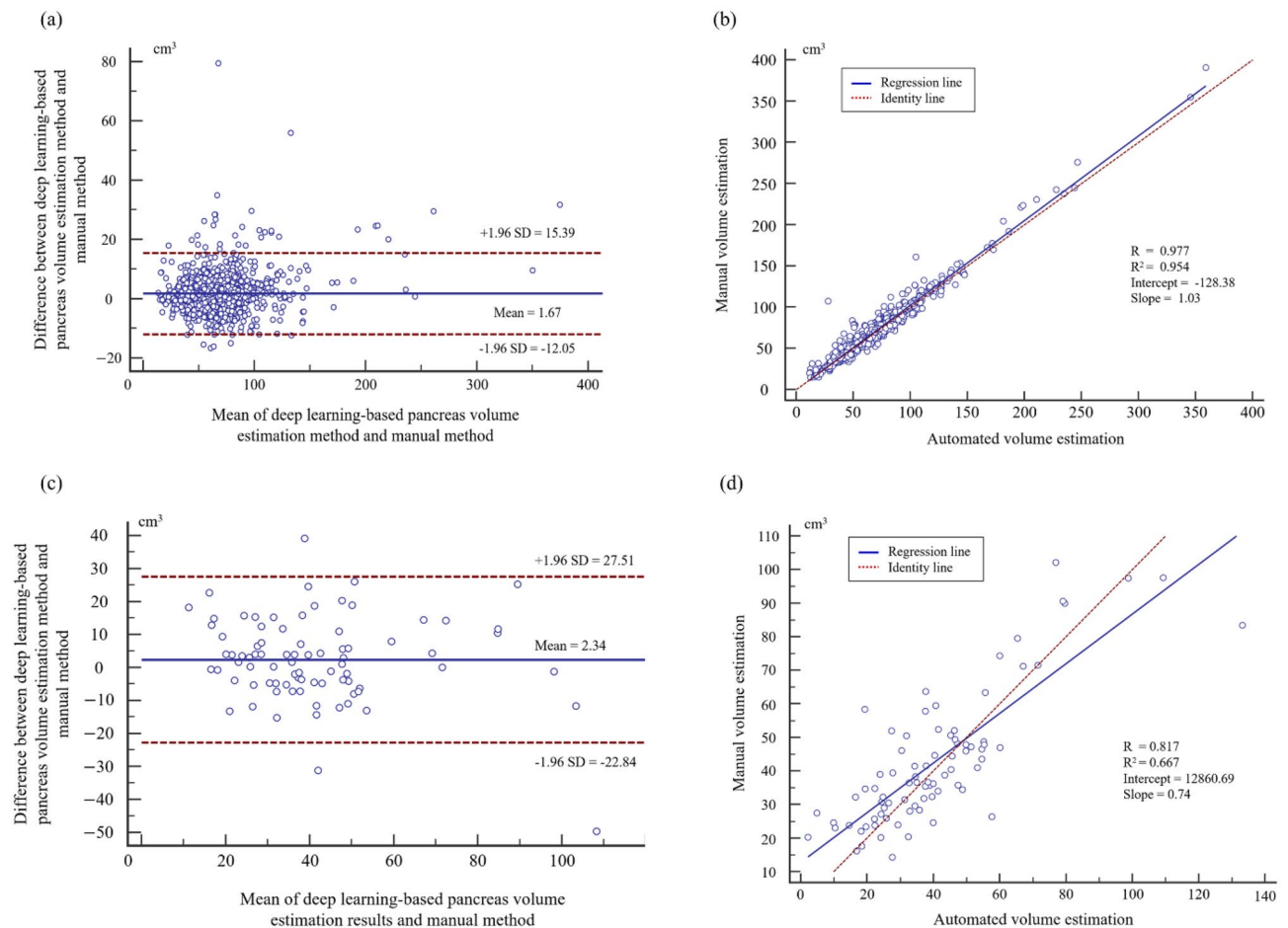


Figure 4. Estimation of pancreatic volume assessments using DL-prediction and manual pancreas segmentation. To validate the DL approaches, the (a) Bland–Altman plot and (b) regression plot were employed for internal validation and (c, d) external validation. SD, standard deviation.

study of two-stage networks and a state-of-the-art segmentation strategy for more accurate pancreas segmentation will be part of our future work.

Repeatable and reproducible pancreas segmentation is necessary for pancreatic volumetry, and CNNs may have broad applicability to this problem. Furthermore, automated abdominal organ segmentation^{35,36} and analysis applications can be used not only for CT but also for diverse modalities, such as magnetic resonance imaging and ultrasound. However, experiments using data that includes various races, ages, and pancreatic volumes are necessary to evaluate the applicability of these methods to clinical practice. Our study presented a DL-based semi-automated method on data from 1006 healthy Koreans; however, further DL-based studies on a dataset that includes various features are necessary to investigate reliable DL-based pancreas-segmentation strategies to aid clinicians.

Data availability

The datasets generated and/or analyzed during the current study are not publicly available because permission to share patient data was not granted by the institutional review board, but they are available from the corresponding author upon reasonable request.

Received: 30 September 2021; Accepted: 23 February 2022

Published online: 08 March 2022

References

- Lee, D. H. *et al.* Recent treatment patterns and survival outcomes in pancreatic cancer according to clinical stage based on single-center large-cohort data. *Ann. Hepatobiliary Pancreat. Surg.* **22**, 386–396. <https://doi.org/10.14701/ahbps.2018.22.4.386> (2018).
- Kim, J. R. *et al.* Clinicopathologic analysis of intraductal papillary neoplasm of bile duct: Korean multicenter cohort study. *HPB (Oxford)* <https://doi.org/10.1016/j.hpb.2019.11.007> (2019).
- Han, Y. *et al.* Progression of pancreatic branch duct intraductal papillary mucinous neoplasm associates with cyst size. *Gastroenterology* **154**, 576–584. <https://doi.org/10.1053/j.gastro.2017.10.013> (2018).
- Chang, Y. R. *et al.* Incidental pancreatic cystic neoplasms in an asymptomatic healthy population of 21,745 individuals Large-scale, single-center cohort study. *Medicine*. <https://doi.org/10.1097/MD.0000000000005535> (2016).

5. Lee, D. H. *et al.* Central pancreatectomy versus distal pancreatectomy and pancreaticoduodenectomy for benign and low-grade malignant neoplasms: A retrospective and propensity score-matched study with long-term functional outcomes and pancreas volumetry. *Ann. Surg. Oncol.* **27**, 1215–1224. <https://doi.org/10.1245/s10434-019-08095-z> (2020).
6. Shin, Y. C. *et al.* Comparison of long-term clinical outcomes of external and internal pancreatic stents in pancreaticoduodenectomy: Randomized controlled study. *HPB (Oxford)* **21**, 51–59. <https://doi.org/10.1016/j.hpb.2018.06.1795> (2019).
7. Valueva, M. V., Nagornov, N. N., Lyakhov, P. A., Valuev, G. V. & Chervyakov, N. I. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Math. Comput. Simul.* **177**, 232–243. <https://doi.org/10.1016/j.matcom.2020.04.031> (2020).
8. Fu, Y. *et al.* A review of deep learning based methods for medical image multi-organ segmentation. *Phys. Med.* **85**, 107–122. <https://doi.org/10.1016/j.ejmp.2021.05.003> (2021).
9. Karasawa, K. *et al.* Multi-atlas pancreas segmentation: Atlas selection based on vessel structure. *Med. Image Anal.* **39**, 18–28. <https://doi.org/10.1016/j.media.2017.03.006> (2017).
10. Lim, S. H. *et al.* Reproducibility of automated habenula segmentation via deep learning in major depressive disorder and normal controls with 7 Tesla MRI. *Sci Rep* **11**, 13445. <https://doi.org/10.1038/s41598-021-92952-z> (2021).
11. Yan, Y. & Zhang, D. Multi-scale U-like network with attention mechanism for automatic pancreas segmentation. *PLoS ONE* **16**, e0252287. <https://doi.org/10.1371/journal.pone.0252287> (2021).
12. Kumar, H., DeSouza, S. V. & Petrov, M. S. Automated pancreas segmentation from computed tomography and magnetic resonance images: A systematic review. *Comput. Methods Programs Biomed.* **178**, 319–328. <https://doi.org/10.1016/j.cmpb.2019.07.002> (2019).
13. Li, J., Lin, X., Che, H., Li, H. & Qian, X. Pancreas segmentation with probabilistic map guided bi-directional recurrent UNet. *Phys. Med. Biol.* <https://doi.org/10.1088/1361-6560/abfce3> (2021).
14. Roth, H. *et al.* DeepOrgan: Multi-level deep convolutional networks for automated pancreas segmentation. [arXiv:1506.06448](https://arxiv.org/abs/1506.06448) (2015).
15. Marin, D. *et al.* Detection of pancreatic tumors, image quality, and radiation dose during the pancreatic parenchymal phase: Effect of a low-tube-voltage, high-tube-current CT technique—preliminary results. *Radiology* **256**, 450–459. <https://doi.org/10.1148/radiol.10091819> (2010).
16. Singh, P., Mukundan, R. & De Ryke, R. Feature enhancement in medical ultrasound videos using contrast-limited adaptive histogram equalization. *J. Digit. Imaging* **33**, 273–285. <https://doi.org/10.1007/s10278-019-00211-5> (2020).
17. Anifah, L., Purnama, I. K., Hariadi, M. & Purnomo, M. H. Osteoarthritis classification using self organizing map based on gabor kernel and contrast-limited adaptive histogram equalization. *Open Biomed. Eng. J.* **7**, 18–28. <https://doi.org/10.2174/1874120701307010018> (2013).
18. Zhang, Y. D. *et al.* Smart detection of abnormal breasts in digital mammography based on contrast-limited adaptive histogram equalization and chaotic adaptive real-coded biogeography-based optimization. *Simul.-Trans. Soc. Mod. Simul.* **92**, 873–885. <https://doi.org/10.1177/0037549716667834> (2016).
19. Ravichandran, C. G. & Raja, J. B. A fast enhancement/thresholding based blood vessel segmentation for retinal image using contrast limited adaptive histogram equalization. *J. Med. Imag. Health Insur.* **4**, 567–575. <https://doi.org/10.1166/jmihi.2014.1289> (2014).
20. Clark, K. *et al.* The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. *J. Digit. Imaging* **26**, 1045–1057. <https://doi.org/10.1007/s10278-013-9622-7> (2013).
21. Roth, H. R. *et al.* Data from pancreas-CT. The Cancer Imaging Archive. <https://doi.org/10.7937/K9/TCIA.2016.tNB1kqBU> (2016).
22. Falk, T. *et al.* U-Net: Deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**, 67–70. <https://doi.org/10.1038/s41592-018-0261-2> (2019).
23. Nazem, F., Ghasemi, F., Fassihi, A. & Dehnavi, A. M. 3D U-Net: A voxel-based method in binding site prediction of protein structure. *J. Bioinform. Comput. Biol.* **19**, 2150006. <https://doi.org/10.1142/S0219720021500062> (2021).
24. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T. & Ronneberger, O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*. 424–432 (Springer).
25. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation, in *International conference on medical image computing and computer-assisted intervention*. 234–241 (Springer).
26. Ioffe, S. & Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift, in *International conference on machine learning*. 448–456 (PMLR).
27. Wu, H., Zhang, J., Huang, K., Liang, K. & Yu, Y. Fastfcn: Rethinking dilated convolution in the backbone for semantic segmentation. [arXiv:1903.11816](https://arxiv.org/abs/1903.11816) (2019).
28. Iandola, F. *et al.* Densenet: Implementing efficient convnet descriptor pyramids. [arXiv:1404.1869](https://arxiv.org/abs/1404.1869) (2014).
29. Ding, P. L. K., Li, Z., Zhou, Y. & Li, B. Deep residual dense U-Net for resolution enhancement in accelerated MRI acquisition, in *Medical Imaging 2019: Image Processing*. 109490F (International Society for Optics and Photonics).
30. Wang, W. *et al.* A Fully 3D cascaded framework for pancreas segmentation, in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. 207–211.
31. Gros, C., Lemay, A. & Cohen-Adad, J. SoftSeg: Advantages of soft versus binary training for image segmentation. *Med. Image Anal.* **71**, 102038. <https://doi.org/10.1016/j.media.2021.102038> (2021).
32. Zhang, D. *et al.* Automatic pancreas segmentation based on lightweight DCNN modules and spatial prior propagation. *Pattern Recogn.* **114**, 107762. <https://doi.org/10.1016/j.patcog.2020.107762> (2021).
33. Zhou, Y. *et al.* A fixed-point model for pancreas segmentation in abdominal CT scans. in *Medical image computing and computer assisted intervention—MICCAI 2017*. (eds Maxime Descoteaux *et al.*) 693–701 (Springer International Publishing).
34. Tang, Y. *et al.* High-resolution 3D abdominal segmentation with random patch network fusion. *Med. Image Anal.* **69**, 101894. <https://doi.org/10.1016/j.media.2020.101894> (2021).
35. Kim, H. *et al.* Abdominal multi-organ auto-segmentation using 3D-patch-based deep convolutional neural network. *Sci. Rep.* **10**, 6204. <https://doi.org/10.1038/s41598-020-63285-0> (2020).
36. Wang, Y. *et al.* Abdominal multi-organ segmentation with organ-attention networks and statistical fusion. *Med. Image Anal.* **55**, 88–102 (2019).

Acknowledgements

This work was supported by the Gachon University College of Medicine (grant No. 202107270001), and the Gachon University Gil Medical Center (No. FRD2021-20), and the GRRC program of Gyeonggi province (GRRC Gachon 2017-B01) and the Gachon Program (GCU-202008440010).

Author contributions

Conception and design: D.H.L., K.G.K. Provision of study materials or patients: Y.H.P., D.J.K., D.H.L. Collection and assembly of data: Y.H.P., D.J.K., D.H.L. Data analysis and interpretation: S.H.L., Y.J.K., D.H.L., K.G.K. Manuscript writing: S.H.L., Y.J.K., D.H.L. Final approval of manuscript: all authors.

Funding

Gyeonggi-do Regional Research Center (No. GRRC Gachon 2017-B01), Gachon University College of Medicine (No. 202107270001), Gachon University Gil Medical Center (No. FRD2021-20), Gachon Program (No. GCU-202008440010).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to K.G.K. or D.-H.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022