



OPEN

Deep 3D attention CLSTM U-Net based automated liver segmentation and volumetry for the liver transplantation in abdominal CT volumes

Jin Gyo Jeong^{1,5}, Sangtae Choi^{2,5}, Young Jae Kim³, Won-Suk Lee⁴✉ & Kwang Gi Kim^{1,3}✉

In living-donor liver transplantation, the safety of the donor is critical. In addition, accurately measuring the liver volume is significant as the amount that can be resected from living donors is limited. In this paper, we propose an automated segmentation and volume estimation method for the liver in computed tomography imaging based on a deep learning-based segmentation network. Our framework was trained using the data of 191 donors, achieved a dice similarity coefficient of 0.789, 0.869, 0.955, and 0.899, respectively, in the segmentation task for the left lobe, right lobe, caudate lobe, and whole liver. Moreover, the R² score reached 0.980, 0.996, 0.953, and 0.996 in the volume estimation task. We demonstrate that our approach provides accurate and quantitative liver segmentation results, reducing the error in liver volume estimation. Therefore, we expected to be used as an aid in estimating liver volume from CT volume data for living-donor liver transplantation.

Liver transplantation is the definitive treatment that increases the survival rate of patients with end-stage liver cancer and cirrhosis^{1,2}. Major hepatic resection is increasingly performed due to advances in surgical devices and preoperative diagnostic imaging that allows accurate delineation of anatomical variants and liver volume assessment in living liver transplantation³. Postoperative morbidity is mainly due to hepatic dysfunction and bile leaks following extended liver resection⁴. Thus, accurately estimating the remnant liver is critical to avoid liver failure as the most serious complication.

Liver volumetry is performed for the donor liver, to calculate the graft volume and remnant liver to avoid graft size mismatch⁵. The safety of the liver donor is a major concern in liver transplantation. Overestimation of the donor liver can result in excess hepatic resection, while underestimation of the recipient's standard liver volume can result in small-for-size graft syndrome^{5,6}. Therefore, to measure the liver volume accurately, it is necessary to perform an accurate liver segmentation.

In general, liver segmentation approaches include manual segmentation and automated segmentation. Manual segmentation is time-consuming, poorly reproducible, and results vary with operators, and it is a task that requires high-level technical skills. Automated segmentation may reduce errors from manual segmentation⁷.

Recently, convolutional neural networks (CNNs) have shown promise for performing various tasks in medical imaging^{8,9}. Traditional machine learning and image processing, in which limited features are extracted for segmentation, require cautious design by a researcher. In contrast, CNNs have the ability to learn relevant features, which define the anatomical structure and tissue of the liver, to segment the liver automatically in a training dataset of computed tomography (CT) volumes. For instance, U-Net is an encoder–decoder model that improves the performance by applying a skip connection to compensate for missing data due to convolution¹⁰. Recently, state-of-the-art approaches that apply an attention-mechanism (AM), deep supervision (DS), and nested architecture techniques, have improved segmentation performance^{11–14}.

¹Department of Health Sciences and Technology, GAIHST, Gachon University, Incheon 21999, Korea. ²Department of Surgery and Liver Transplantation, Gil Medical Center, Gachon University, Incheon 21565, Korea. ³Department of Biomedical Engineering, Gil Medical Center, College of Medicine, Gachon University, Incheon 21565, Korea. ⁴Department of Surgery, Gil Medical Center, College of Medicine, Gachon University, Incheon 21565, Korea. ⁵These authors contributed equally: Jin Gyo Jeong and Sangtae Choi. ✉email: lws@gilhospital.com; kimkg@gachon.ac.kr

	Females	Males	All
No. of patients	80	111	191
Age (years)	33.1 (10.7)	31.0 (12.1)	31.8 (11.5)
Body height (cm)	161.9 (5.5)	172.8 (5.8)	168.4 (7.8)
Body weight (kg)	58.9 (7.8)	70.8 (10.9)	65.9 (11.4)

Table 1. Characteristics of patients whose image data were used in this study. Data summarized as mean (\pm SD).

Imaging systems and parameters	Training/validation at our institution
Imaging system	Siemens
In-plane resolution (mm)	0.5–0.9
Section thickness (mm)	5.0
Collimation (mm)	NA
Tube current (mAs)	50–550
Tube voltage (kVp)	90–120

Table 2. Imaging systems and parameters.

An attention U-Net's AM approach demonstrated an increase in pancreas segmentation performance in TCIA Pancreas-CT Dataset¹¹. The pancreas is a small organ, with a shape that differs across patients. An attention U-Net improved performance of segmentation of pancreas of various shapes and small sizes, by using an AM with 1×1 convolution layer and a sigmoid activation function to reduce background weight and to preserve foreground weight.

An Attention U-Net⁺⁺ improved liver segmentation performance by applying an AM, DS, and nested architecture¹³. In order to integrate semantic information from feature maps at different semantic levels from output nodes, the convergence speed of the loss and the segmentation accuracy was improved by optimizing the losses obtained at each level. These approaches have the advantages of focusing on a small organ or tissue and accelerate optimal convergence^{11–14}.

Recently, several studies have been conducted for inter-frame and inter-slice learning from video data and three-dimensional (3D) volumetric data using convolutional long short-term memory (CLSTM)^{15–19}. LSTM is usually used in natural language processing and is mainly used for continuous data or for data with a temporal sequence. CLSTM was developed for CNN models^{20–22}. A previous study proposed a 3D segmentation approach that combined CLSTM in 3D fungus datasets. The proposed approach achieved higher performance than the generalized 3D CNN. The CLSTM approach demonstrated that it could learn efficiently in the context of inter-slice and 3D volume data.

In this study, we describe a method for automatic liver segmentation from abdominal CT volume data, in comparison with human liver volumetry conducted by a single liver transplant surgeon. First, we proposed a deep 3D CNN model comprised of AM and DS to achieve fast and accurate segmentation of CT volumetric data. Second, we used CLSTM in our network to learn the inter-slice context. Third, we evaluated and compared liver segmentation accuracy and liver volumetry with human-based volumetry data. Finally, we evaluated the CNN for left, right, and caudate lobe volumetry to show its clinical significance.

Methods

Data sources and patient demographics. In this study, our dataset was obtained 191 abdominal CT data from liver transplant adult donors (mean age, 31 years \pm 11 years; 58% male; mean weight, 65.9 \pm 11.4 kg) from 2005 to 2017 at Gachon University Gil Hospital. This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the Gil Medical Center (IRB No. GBIRB2021-229), and written informed consent was obtained from all the participants. All images were deidentified before inclusion in this study. The demographics and other characteristics of each cohort are summarized in Table 1. The imaging system is summarized in Table 2.

Datasets and data pre-processing. In this study, the window level and window width were set to clearly observe the liver on CT (window level: 50, window width: 180)²³. Figure 1 shows the result of the window setting. Furthermore, CT data had the same 512 \times 512 in-plane resolutions, but due to the computational limitation of the graphics card, the image was resized to a resolution of 256 \times 256. Finally, the image and ground truth data had a shape of 64 \times 256 \times 256 \times 1 and were divided into train: validation: test sets in the ratio of 70: 15: 15.

Liver segmentation using DALU-Net. The proposed model Deep Attention LSTM U-Net (DALU-Net) had an architecture similar to the standard U-Net, consisting of an encoder and a decoder¹⁰. The encoder could extract more complex hierarchical features and obtain contextual information. The decoder deconvolutes the

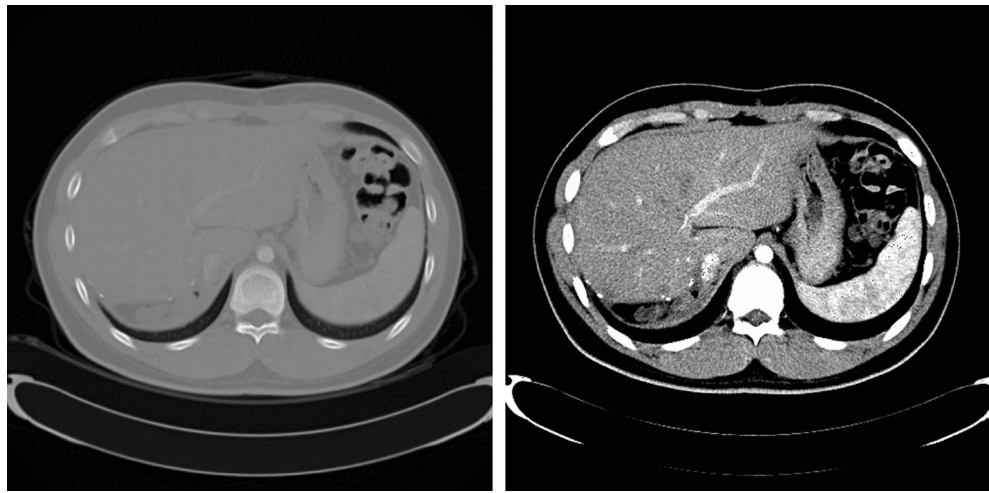


Figure 1. Comparison between the raw computed tomography axial view images (left) and with the window setting (right).

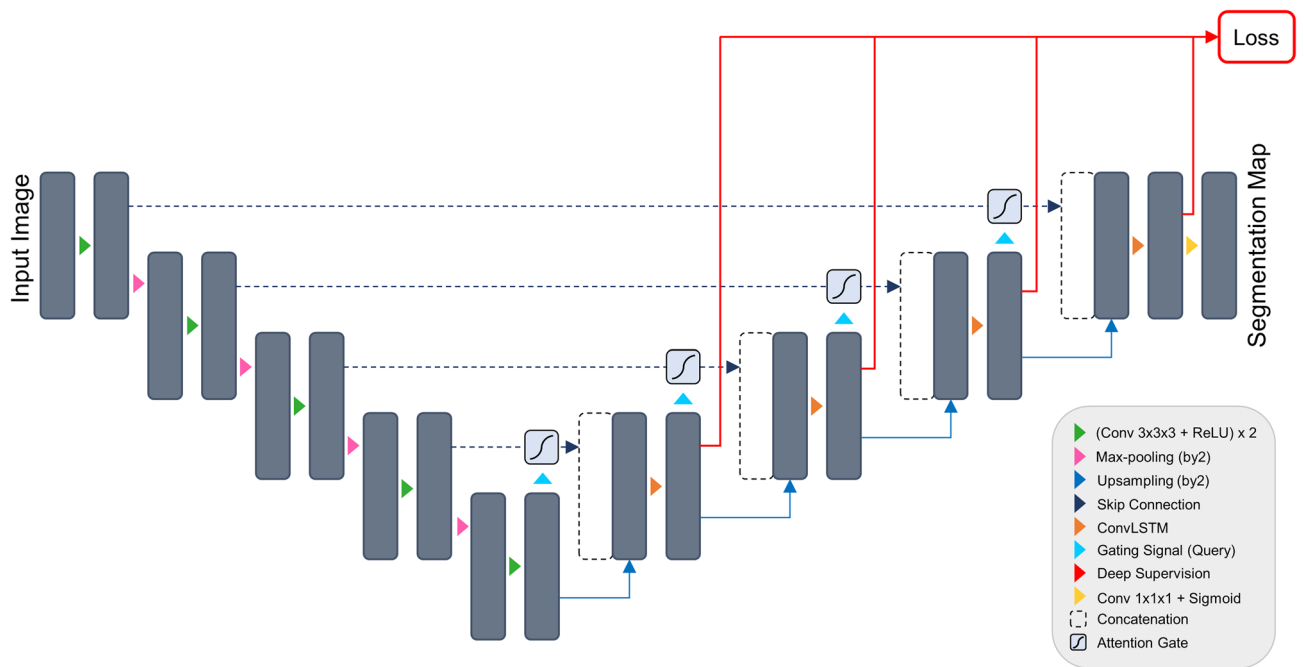


Figure 2. The proposed DALU-Net architecture.

features extracted by the encoder to reconstruct the size of the volume reduced by the convolution operation. In addition, it concatenates hierarchical features using the skip connection at every level of encoder and the decoder. Information about localization lost due to convolution and pooling layers in the encoder, can be corrected, and the network can segment objects more accurately. Figure 2 illustrated the DALU-Net architecture.

DALU-Net is a model that combines AM, DS, and CLSTM techniques. AM uses a module called attention gate (AG) to skip connections between the up-sampling layer and encoder. The CLSTM was used in the feature map from AG and was used only in the decoder. DS was used for fast convergence of the model, and the loss is calculated at every level of the decoder, and the final loss was calculated as the sum of each loss.

AM is commonly used for machine translation²⁴ and classification in natural language processing and graph neural networks^{25,26}. Recently, AM has been variously used in semantic segmentation and classification tasks²⁷. AM has demonstrated improved accuracy in the medical image^{11,13,28,29}. In the image segmentation and classification task, the AM was designed to generate an attention map by analyzing the gradient of the output class score for the input image, reduce the weight of the background by multiplying it with the input image, and focus on the object. The details of the AM are shown in Fig. 3.

DS supervises the CNN's hidden layer in the segmentation task, to speed up convergence and resolve the gradient-vanishing problem^{12,17,30}. The red arrow in Fig. 2, the loss is calculated using a sigmoid function for every

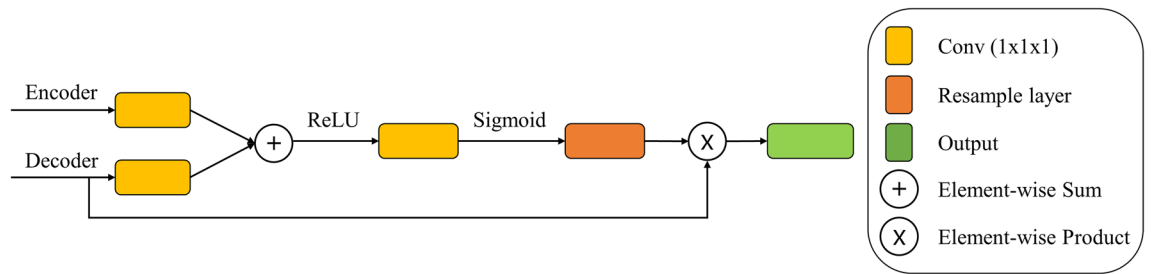


Figure 3. Schematic of attention mechanism.

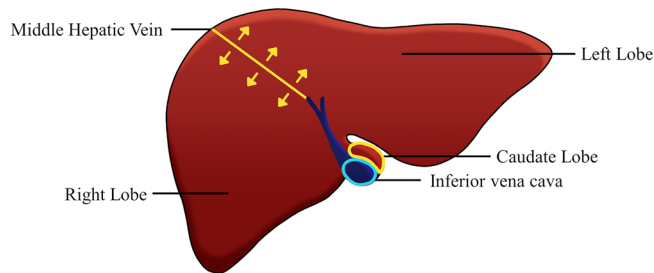


Figure 4. Illustration of the ground truth for the left lobe, right lobe, and caudate lobe at the axial view.

level of the decoder, and the final loss was determined by summing each loss. It has the advantage of allowing the final loss to converge rapidly by optimizing at each level. In this study, DS was applied to the output values from the attention gate of each level.

Liver volumetry. The segmented liver area was calculated using the proposed model on the axial image. The liver volume was calculated for each image by multiplying the liver area and the section thickness. Thus, the whole liver volume was estimated by adding the liver volume of all images.

Ground truth for liver segmentation and volumetry. The liver was manually labeled on all CT images under the supervision of a liver transplant surgeon with more than 5 years of experience in CT analysis related to liver transplantation. ImageJ software was used for manual segmentation (NIH, Bethesda, MD, USA). For liver volumetry, the calculated liver volume, based on manual segmentation, was used as the reference standard (as described in “Liver Volumetry” section).

Ground truth for left and right and caudate lobe. In this study, a comparative analysis was performed for segmentation and volume measurements for the left, right, and caudate lobe regions, according to the anatomical structure of the liver^{31,32}. The left lobe was defined as the upper region, above the middle hepatic vein, and the lower region was defined as the right lobe. The caudate lobe was located to the left of the inferior vena cava (IVC), without overlapping the left lobe in the coronal view^{31,32}. Figure 4 schematically illustrates the ground truth for the left lobe, right lobe, and caudate lobe.

Implementation details. The DALU-Net was implemented in Python 3.7, TensorFlow 2.1.3, Keras 2.3.1, and was run on four NVIDIA Tesla V100 GPU with 5120 cores and 32 GB of memory. Our networks were trained using the Adam optimizer to minimize the dice loss^{33,34}. When the loss was not minimized over 10 epochs, the learning rate was reduced by multiplying the initial learning rate by 0.1, with an initial value of 0.01. We terminated the process early when the loss did not improve for 30 epochs in the training procedure. Our network terminated the training procedure early in 90–150 epochs.

Evaluation metrics. We evaluated the performance of the proposed approach. The evaluation metrics included the dice similarity coefficient (DSC), intersection over union (IOU), and Hausdorff distance³⁵. The DSC was defined as the volume of overlap between the CNN and manual labeling segmentations divided by the average of the segmentation volume of the two methods. The IOU was defined as a mathematical indicator of how much the two objects’ positions coincide. The Hausdorff distance was defined as the difference measured between two subsets of metric space.

Informed consent. Informed consent was obtained from all subjects involved in the study.

Dataset	Model	Recall	Precision	DSC	IOU	HD (mm)
Left lobe	3D U-Net	0.838±0.124	0.803±0.081	0.771±0.082	0.783±0.127	6.895±1.794
	AU-Net	0.838±0.122	0.808±0.086	0.785±0.081	0.776±0.128	6.691±1.785
	AU-Net w/ DS	0.847±0.121	0.812±0.083	0.787±0.082	0.779±0.128	6.630±1.732
	DALU-Net	0.851±0.121	0.806±0.083	0.789±0.084	0.787±0.128	6.629±1.696
Right lobe	3D U-Net	0.966±0.069	0.786±0.067	0.855±0.065	0.821±0.138	5.144±1.635
	AU-Net	0.966±0.067	0.807±0.072	0.867±0.037	0.856±0.115	5.125±1.714
	AU-Net w/ DS	0.972±0.063	0.804±0.071	0.868±0.066	0.827±0.121	5.087±1.731
	DALU-Net	0.975±0.058	0.805±0.071	0.869±0.067	0.896±0.141	5.074±1.705
Caudate lobe	3D U-Net	0.825±0.326	0.793±0.316	0.849±0.307	0.819±0.316	2.133±1.476
	AU-Net	0.905±0.241	0.817±0.209	0.922±0.236	0.896±0.236	1.557±1.273
	AU-Net w/ DS	0.914±0.229	0.822±0.217	0.929±0.215	0.907±0.222	1.583±1.221
	DALU-Net	0.943±0.183	0.877±0.174	0.955±0.167	0.936±0.181	1.238±1.131
Whole liver	3D U-Net	0.899±0.155	0.850±0.287	0.822±0.295	0.770±0.288	3.236±1.256
	AU-Net	0.891±0.162	0.893±0.241	0.852±0.258	0.800±0.259	3.280±0.982
	AU-Net w/ DS	0.922±0.110	0.907±0.213	0.887±0.218	0.840±0.219	2.980±0.784
	DALU-Net	0.923±0.120	0.924±0.188	0.899±0.201	0.855±0.205	2.762±0.728

Table 3. Segmentation results of an ablation study of the proposed model on the four validation datasets.

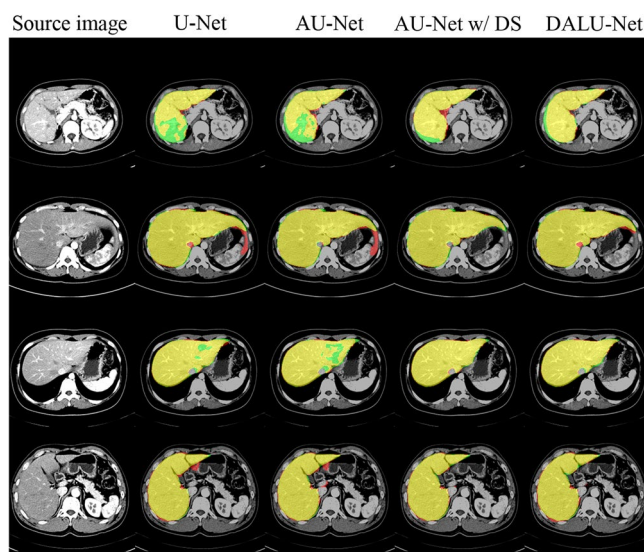


Figure 5. Examples of liver segmentation results at the axial view images for the various models. Segmentation results are color-coded: green is ground truth, red is automated segmentation, and yellow is the overlap between the ground truth and automated segmentation.

Results

Liver segmentation accuracy. We analyzed the accuracy of the liver segmentation network. To this end, we compared the 3D U-Net, Attention U-Net, Attention U-Net with DS, and DALU-Net. Table 3 shows the performance of our method in the four validation datasets: (a) left lobe, (b) right lobe, (c) caudate lobe, (d) whole liver. We used evaluation metrics to measure the accuracy of the segmentation results: the higher the value of the evaluation metrics, the better the segmentation results. On the other hand, the smaller the Hausdorff distance was, the better were the results. Our method performed better than 3D U-Net in terms of segmentation accuracy (DSC), with 7% and 10% improvement in the whole liver and caudate lobe datasets, respectively. Figure 5 shows the visual assessment of the automated liver segmentation.

Effect of layer on segmentation accuracy. To prove the effectiveness of our proposed method, we performed an ablation study on our dataset. We compared different configurations with the baseline of 3D U-Net. The result of the ablation study is summarized in Table 3. For the liver, the DSC increased from 0.822 to 0.852 with AG. Then, we evaluated the contribution of DS. By using DS, the DSC increased from 0.852 to 0.887. Finally, we analyzed the importance of CLSTM in our proposed model and showed that the DSC increased from 0.887 to 0.899.

Liver volumetry accuracy. The comparison of between human volumetry and automated liver volumetry is summarized in Fig. 6. Left lobe volumes obtained by manual segmentation ranged from 195 to 653 mL (mean volume, 379 ± 130 mL), and that obtained by automated segmentation ranged from 223 to 749 mL (mean volume, 431 ± 144 mL). Manual and automated left lobe volumetry correlated strongly ($R^2 = 0.980$, slope = 1.095, intercept = 15.841, Fig. 6a). When compared with the manual method, the automated method resulted in a slight underestimation of the left lobe volume (bias = 51.9 mL, $P < 0.001$), and the 95% limits of agreement (LoA) were 4.5 mL and 99.3 mL (Fig. 6b).

Right lobe volumes obtained by manual segmentation ranged from 562 to 1573 mL (mean volume, 1006 ± 290 mL), and that obtained by automated segmentation ranged from 651 to 1717 mL (mean volume, 1130 ± 316 mL). Manual and automated right lobe volumetry correlated strongly ($R^2 = 0.996$, slope = 1.087, intercept = 36.471, Fig. 6c). When compared with the manual method, the automated method resulted in a slight underestimation of right lobe volume (bias = 124.8 mL, $P < 0.001$), and the 95% LoA were 61.0 mL and 188.6 mL (Fig. 6d).

Caudate lobe volumes obtained by manual segmentation ranged from 3 to 31 mL (mean volume, 14 ± 6 mL), and that obtained by automated segmentation ranged from 3 to 30 mL (mean volume, 13 ± 5 mL). Manual and automated caudate lobe volumetry correlated strongly ($R^2 = 0.953$, slope = 0.862, intercept = 1.294, Fig. 6e). When compared with the manual method, the automated method resulted in a slight underestimation of the caudate lobe volume (bias = -0.7 mL, $P < 0.001$), and the 95% LoA were -3.7 mL and 2.3 mL (Fig. 6f).

Whole liver volumes obtained by manual segmentation ranged from 924 to 2100 mL (mean volume, 1400 ± 357 mL), and that obtained by automated segmentation ranged from 1072 to 2340 mL (mean volume, 1576 ± 388 mL). Manual and automated whole liver volumetry correlated strongly ($R^2 = 0.996$, slope = 1.084, intercept = 57.409, Fig. 6g). When compared with the manual method, the automated method resulted in a slight underestimation of the whole liver volume (bias = 176 mL, $P < 0.001$), and the 95% LoA were 101.3 mL and 250.8 mL (Fig. 6h).

Discussion

In this study, we showed that our proposed DALU-Net model, based on AG, DS, and CLSTM, can be trained to perform automated and accurate liver segmentation using abdominal CT data. We achieved high performance with a segmentation accuracy DSC of 0.899 for the whole liver. Compared to the reference model, 3D U-Net, the performance of the proposed model was improved by 7%. Table 3 presents the results of the AG, DS, and CLSTM models to the basic 3D U-Net model; whole liver segmentation accuracy improved by 3%, 3%, and 1%, respectively. Figure 5 presents the training results of the model. For the proposed DALU-Net model, segmentation results were more accurate than those of other models. We demonstrated that the segmentation accuracy increased with addition of each method.

Figure 6g, h show the results of the linear regression analysis between the human-based whole liver volumetry and the volume of the automatically segmented liver. The slope of the regression curve was 1.084 and R^2 was 0.996, indicating that the two values were similar. However, the regression line was located above the identity line. The automated method overestimated liver volume as compared to manual liver volumetry. Since the slice thickness was relatively large (5 mm), it is assumed that there were many false positives in the liver segmentation results, due to insufficient information for a 3D context. However, results of the Bland–Altman analysis showed that most values were within the 95% confidence interval. It can be confirmed that manual and automatic volume estimations are similar.

In previous liver segmentation studies, whole liver segmentation accuracy and volumetry were compared. In contrast, we further compared the segmentation accuracy and volumetry for the left, right, and caudate lobes. In the three-validation datasets, defined according to the anatomy of the liver, the DSC was 0.789, 0.869, and 0.955 for the left, right, and caudate lobes, respectively. There was no significant difference between the DSCs of the proposed model and those of the 3D U-Net model in the left and right lobes. In contrast, the DSC of the proposed model for the caudal lobe data was improved by 10% as compared to that of the 3D U-Net model. The caudate lobe is attached to the IVC. It is smaller than the left and right lobe, and with a 5-mm section thickness, it has a small number of slices in terms of CT volume data. Consequently, it was difficult to perform inter-slice learning with a general 3D convolution layer. The proposed model, using CLSTM, was assumed to be capable of inter-slice learning and achieved high accuracy.

Linear regression analysis in the three validation datasets yielded the slope of regression lines of 1.095, 1.087, and 0.862, which are close to 1, and R^2 of 0.980, 0.996, and 0.953, for the left, right, and caudate lobes, respectively. This indicated that manual and automated volumetry yielded similar results. In addition, using Bland–Altman analysis, we demonstrated that the gap between the manual liver volume and the CNN-based liver volume is small, as most values were within the 95% confidence interval.

We have several limitations to the dataset. The contrast agent injection amount and the scan time after injection are different, so it has various contrast in the dataset³⁶. In addition, there have several cases with low accuracy because the section thickness was 5 mm and there were not many slices overlapping the heart in the CT volume data, so it was not possible to distinguish between the myocardium and the liver. Additional work may be required to improve performance in these problems. More extensive testing helps to reduce the failed cases of CNN, so inform the choice of additional training data to further improve model performance.

In summary, the proposed DALU-Net, using AM, DS, and CLSTM, can perform liver segmentation and volumetry using abdominal CT volume data. Furthermore, the volume of the left, right, and caudate lobes can be measured, according to the anatomical structure of the liver. Thus, our approach can be used for estimating liver volumes from CT data to assist in planning of living-donor liver transplantation.

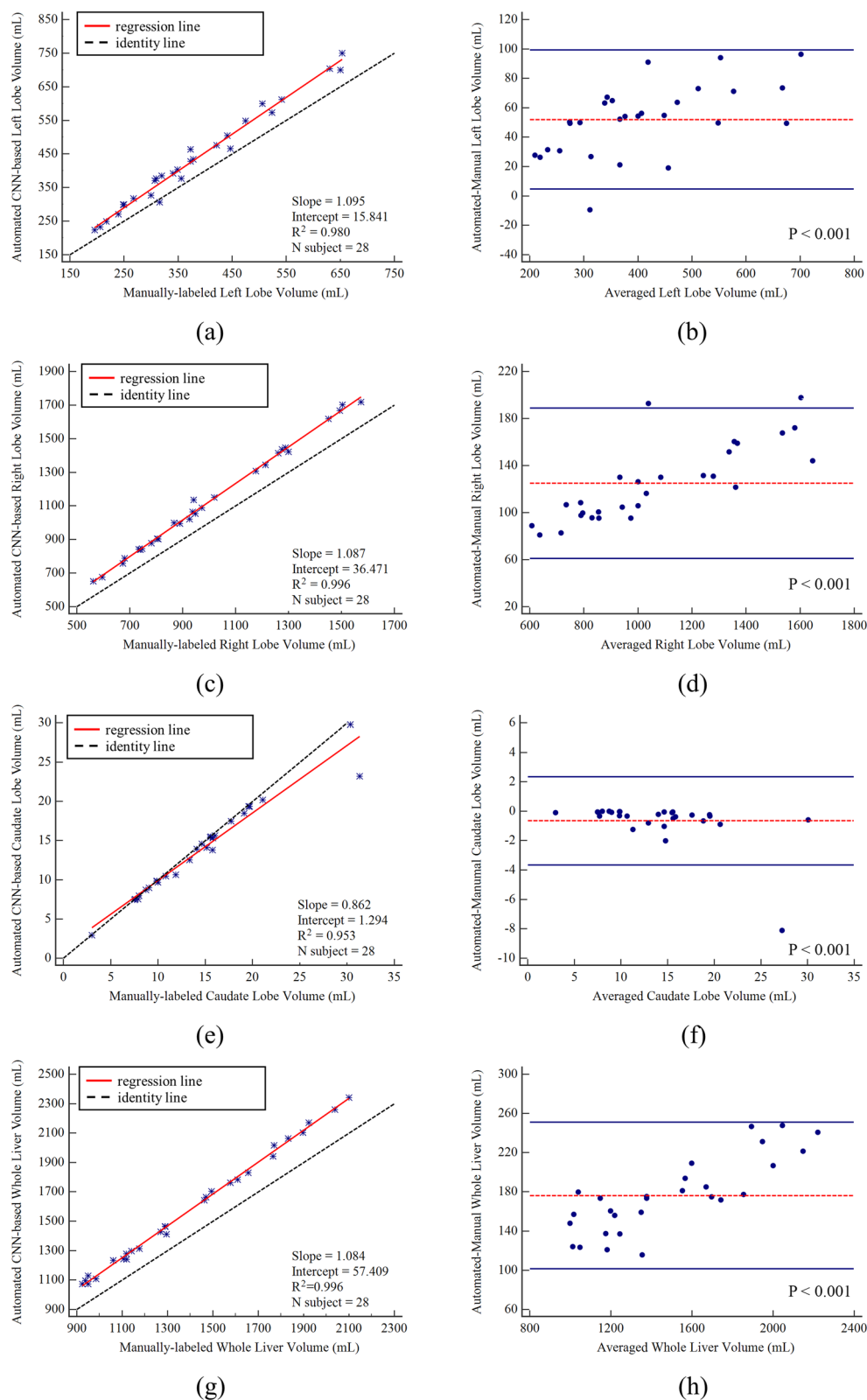


Figure 6. Agreement of liver volume assessments between convolutional neural network-predicted and manual liver segmentation. (a) Linear regression and (b) Bland–Altman analysis of left lobe volume assessments. (c) Linear regression and (d) Bland–Altman analysis of right lobe volume assessments. (e) Linear regression and (f) Bland–Altman analysis of caudate lobe volume assessments. (g) Linear regression and (h) Bland–Altman analysis of whole liver volume assessments.

Data availability

The CT volume data used to support the findings of this study are available upon request from the corresponding author.

Received: 14 September 2021; Accepted: 24 March 2022

Published online: 16 April 2022

References

1. Gotthardt, D. *et al.* Fulminant hepatic failure: Etiology and indications for liver transplantation. *Nephrol. Dial. Transplant.* **22**(Suppl 8), viii5–viii8. <https://doi.org/10.1093/ndt/gfm650> (2007).
2. Jain, A. *et al.* Long-term survival after liver transplantation in 4,000 consecutive patients at a single center. *Ann. Surg.* **232**, 490–500. <https://doi.org/10.1097/0000658-200010000-00004> (2000).
3. de Ridder, J. A. *et al.* Liver resection for metastatic disease; a population-based analysis of trends. *Dig. Surg.* **33**, 104–113. <https://doi.org/10.1159/000441802> (2016).
4. Guglielmi, A., Ruzzenente, A., Conci, S., Valdegamberi, A. & Iacono, C. How much remnant is enough in liver resection?. *Dig. Surg.* **29**, 6–17. <https://doi.org/10.1159/000335713> (2012).
5. Ma, K. W. *et al.* Impact of small-for-size liver grafts on medium-term and long-term graft survival in living donor liver transplantation: A meta-analysis. *World J. Gastroenterol.* **25**, 5559 (2019).
6. Fan, S. T. *et al.* Safety of donors in live donor liver transplantation using right lobe grafts. *Arch. Surg.* **135**, 336–340. <https://doi.org/10.1001/archsurg.135.3.336> (2000).
7. Gotra, A. *et al.* Liver segmentation: Indications, techniques and future directions. *Insights Imaging* **8**, 377–392. <https://doi.org/10.1007/s13244-017-0558-1> (2017).
8. Lee, J. H., Kim, Y. J. & Kim, K. G. Bone age estimation using deep learning and hand X-ray images. *Biomed. Eng. Lett.* **10**, 323–331. <https://doi.org/10.1007/s13534-020-00151-y> (2020).
9. Kim, D., Hwang, W., Bae, J., Park, H. & Kim, K. G. Video archiving and communication system (VACS): A progressive approach, design, implementation, and benefits for surgical videos. *Healthc. Inform. Res.* **27**, 162–167. <https://doi.org/10.4258/hir.2021.27.2.162> (2021).
10. Ronneberger, O., Fischer, P. & Brox, T. in *International Conference on Medical image computing and computer-assisted intervention*. 234–241 (Springer).
11. Oktay, O. *et al.* Attention u-net: Learning where to look for the pancreas. [arXiv:1804.03999](https://arxiv.org/abs/1804.03999) (2018).
12. Li, C. *et al.* Deep supervision with intermediate concepts. *IEEE Trans. Pattern Anal. Mach. Intell.* **41**, 1828–1843. <https://doi.org/10.1109/TPAMI.2018.2863285> (2019).
13. Li, C. *et al.* in *2020 IEEE International Conference on Image Processing (ICIP)*. 345–349 (IEEE).
14. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N. & Liang, J. UNet++: A nested U-Net architecture for medical image segmentation. *Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support* **2018**(11045), 3–11. https://doi.org/10.1007/978-3-030-00889-5_1 (2018).
15. Nwoye, C. I., Mutter, D., Marescaux, J. & Padoy, N. Weakly supervised convolutional LSTM approach for tool tracking in laparoscopic videos. *Int. J. Comput. Assist. Radiol. Surg.* **14**, 1059–1067 (2019).
16. Olulana, K., Owolawi, P., Tu, C. & Abe, B. in *International Symposium on Visual Computing*. 753–760 (Springer).
17. Wang, L., Lee, C.-Y., Tu, Z. & Lazebnik, S. Training deeper convolutional networks with deep supervision. [arXiv:1505.02496](https://arxiv.org/abs/1505.02496) (2015).
18. Rochan, M. Future semantic segmentation with convolutional lstm. [arXiv:1807.07946](https://arxiv.org/abs/1807.07946) (2018).
19. Chen, J., Yang, L., Zhang, Y., Alber, M. & Chen, D. Z. in *Advances in neural information processing systems*. 3036–3044.
20. Patraucean, V., Handa, A. & Cipolla, R. Spatio-temporal video autoencoder with differentiable memory. [arXiv:1511.06309](https://arxiv.org/abs/1511.06309) (2015).
21. Shi, X. *et al.* in *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1* 802–810 (MIT Press, Montreal, Canada, 2015).
22. Stollenga, M. F., Byeon, W., Liwicki, M. & Schmidhuber, J. in *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2* 2998–3006 (MIT Press, Montreal, Canada, 2015).
23. SABOURI, S., Khatami, A., Azadeh, P., Ghoroubi, J. & Azimi, G. Adding liver window setting to the standard abdominal CT scan protocol: Is it useful? (2008).
24. Bahdanau, D., Cho, K. & Bengio, Y. Neural machine translation by jointly learning to align and translate. [arXiv:1409.0473](https://arxiv.org/abs/1409.0473) (2014).
25. Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M. & Monfardini, G. The graph neural network model. *IEEE Trans. Neural Netw.* **20**, 61–80. <https://doi.org/10.1109/TNN.2008.2005605> (2009).
26. Veličković, P. *et al.* Graph attention networks. [arXiv:1710.10903](https://arxiv.org/abs/1710.10903) (2017).
27. Wang, F. *et al.* in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 3156–3164.
28. Jiang, H., Shi, T., Bai, Z. & Huang, L. Ahcnet: An application of attention mechanism and hybrid connection for liver tumor segmentation in ct volumes. *IEEE Access* **7**, 24898–24909 (2019).
29. Jin, Q., Meng, Z., Sun, C., Cui, H. & Su, R. RA-UNet: A hybrid deep attention-aware network to extract liver and tumor in CT scans. *Front. Bioeng. Biotechnol.* **8**, 605132. <https://doi.org/10.3389/fbioe.2020.605132> (2020).
30. Lee, C.-Y., Xie, S., Gallagher, P. W., Zhang, Z. & Tu, Z. Deeply-supervised nets. [arXiv:1409.5185](https://arxiv.org/abs/1409.5185) (2015).
31. Strasberg, S. M. Nomenclature of hepatic anatomy and resections: A review of the Brisbane 2000 system. *J. Hepatobiliary Pancreat. Surg.* **12**, 351–355. <https://doi.org/10.1007/s00534-005-0999-7> (2005).
32. Healey, J. E. & Schroy, P. C. Anatomy of the biliary ducts within the human liver: Analysis of the prevailing pattern of branchings and the major variations of the biliary ducts. *A.M.A. Arch. Surg.* **66**, 599–616 (1953).
33. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
34. Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S. & Jorge Cardoso, M. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. *Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support* **2017**(240–248), 2017. https://doi.org/10.1007/978-3-319-67558-9_28 (2017).
35. Karimi, D. & Salcudean, S. E. Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks. *IEEE Trans. Med. Imaging* **39**, 499–513. <https://doi.org/10.1109/TMI.2019.2930068> (2020).
36. Bae, K. T. Intravenous contrast medium administration and scan timing at CT: Considerations and approaches. *Radiology* **256**, 32–61. <https://doi.org/10.1148/radiol.10090908> (2010).

Acknowledgements

This study was approved by the Institutional Review Board (IRB) of Gil Medical Center, Gachon University, and the protocols used in the study were approved by the Committee of Human Subjects Protection of the Gil Medical Center, Gachon University, South Korea (IRB No. GBIRB2021-229). The study was approved by the Institutional Review Board (IRB) of Gil Medical Center, Gachon University. The protocols used in the study were approved by the Hospital's Protection of Human Subjects Committee.

Author contributions

J.G.J. conducted the deep learning analysis, drafted the manuscript, and performed statistical analyses. W.-S.L. and S.C. collected the dataset. Y.J.K., S.C., W.-S.L., and K.G.K. participated in the study design and revised the manuscript. All the authors read and approved the final manuscript.

Funding

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2021R1A5A2030333), by the Gachon Program (GCU-202008440010), and by the GRRC program of Gyeonggi province. [GRRC-Gachon2020(B01), AI-based Medical Image Analysis.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to W.-S.L. or K.G.K.

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