2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)
April 13-16, 2021, Nice, France

2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) | 978-1-6654-1246-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/ISBI48211.2021.9433929

Semantic Segmentation of Liver Tumor in Contrast-enhanced Hepatic CT by Using Deep Learning with Hessian-based Enhancer with Small Training Dataset Size

Muneyuki Sato[1,2], Ze Jin[1,2], Kenji Suzuki[1,2]

[1]Department of Information and Communications Engineering, School of Engineering, Tokyo Institute of Technology
[2]Laboratory for Future Interdisciplinary Research of Science and Technology, Institute of Innovative Research, Tokyo Institute of Technology

ABSTRACT

Deep learning requires a large dataset for training, but collecting and annotating such a large dataset are time-consuming and often very difficult. In this study, we proposed a 3D massive-training artificial neural network (MTANN) incorporated with a Hessian-based enhancer to achieve a high performance of our MTANN model with small training datasets. Contrast-enhanced CT scans of 42 patients with 194 liver tumors from the Liver Tumor Segmentation (LiTS) Benchmark database were used in this study. MTANN models were trained with: 14 and 28 tumors. The remaining 24 patients with 59 tumors were applied to the trained models for evaluation. Our method achieved a Dice of 0.703 with a training set of 14 tumors. The performance was comparable to the best performance in the MICCAI 2017 competition with less than one tenth of training cases. Our method would be essential in applications where a large training set is not available.

Index Terms— Deep learning, computer-aided segmentation, small training dataset size.

1. INTRODUCTION

Liver cancer is the sixth most commonly diagnosed cancer and the fourth leading cause of cancer-related death worldwide in 2018 [1]. In its diagnosis and therapy, accurate segmentation of liver tumors from CT volumes is an essential task. In clinical practice, tumor regions are generally delineated manually by radiologists. However, manual segmentation of tumor regions is very tedious and time consuming, which adds a heavy workload to radiologists [2]. In addition, there are inter-/intra observer variations [3], which result in unreliable, inconsistent diagnosis or treatment planning.

Recently, many automatic segmentation methods have been proposed to overcome the limitations of manual segmentation. Tian et al. and other 26 groups have submitted their methods to the Liver Tumor Segmentation (LiTS) Challenge in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) 2017 with a LiTS Benchmark dataset [4]. The database includes 131 patients with 908 tumors. Most of the submitted methods adopted fully automatic approaches and relied on supervised learning methods. In the competition, "U-Net derived architecture was overwhelmingly used" [5], two other automated methods were using a modified VGG-net and a k-CNN. Deep learning methods were dominant in the competition. However, deep learning requires a large dataset of cases for training. It is said that a deep learning model requires from 5,000 to 100,000 cases for adequate training. In real clinical practices, collecting and annotating a large number of cases are time-consuming and often very difficult especially at a single medical center even for a major cancer. Therefore, segmentation methods that can be trained with a small dataset are highly demanded.

In this study, we proposed a small-sample-size deep-learning technique for semantic segmentation of liver tumors in contrast-enhanced CT, based on a 3D massive-training artificial neural network (MTANN) combined with a Hessian-based ellipse enhancer. The 3D MTANN is a supervised volumetric deep-learning model consisting of convolution of neural network regression, which is trained with 3D input volumes and the corresponding "teaching" volume [6]. The 3D MTANN has a strong advantage of achieving high performance by training with a small number of cases [7]. The Hessian-based enhancer has the capability of enhancing ellipsoidal shapes. For evaluating whether our model could be trained with a small number of cases, we trained 2 proposed MTANN models with: 14 tumors from 12 patients and 28 tumors from 18 patients with the LiTS database, and compared the results with those achieved by the 27 methods in the MICCAI 2017 worldwide competition.

2. MATERIALS

The database used in this study contains dynamic contrast-enhanced CT scans of 42 patients with 194 liver tumors acquired from the LiTS database [4], which is a publicly available database used in the MICCAI 2017 competition. Each slice of the CT volumes in the dataset has a matrix size of 512 x 512 pixels, with in-plane pixel sizes of 0.60-1.0

Authorized licensed use limited to: Intel Corporation via the Virtual Library. Downloaded on February 05,2022 at 02:53:13 UTC from IEEE Xplore. Restrictions apply.

mm and slice thickness of 0.70-2.0 mm. Each liver tumor region in CT was manually segmented by trained radiologists in each clinical institution and then verified by three experienced radiologists as second-time review, along with a manually segmented liver region.

3. METHODS

3.1. Preprocessing of input segmented liver CT images and teaching images

Original liver CT images from the database were preprocessed to obtain input segmented liver CT images with manually segmented liver mask since we focused on the liver tumor segmentation within the liver. Firstly, cropped isotropic image of an in-plane matrix size of 512x512 with a voxel size of 0.60 x 0.60 x 0.60 mm was obtained in order to have the same physical scale on our spherical shape kernel of the 3D MTANN. Then, Z-score-normalization and an anisotropic diffusion filter were applied to a liver image masking by using a liver mask to get rid of non-liver information, reduce the noise, and unify complex histograms of tumors in different cases.

Since the border of manually-segmented region would contain uncertainties due to intra/inter reader variations, we applied a Gaussian filter to the binary manual segmentation image to obtain a map of the probability of being a tumor. The probability map was used as the teaching image.

3.2. Hessian-enhanced image

Sato et el. have proposed a three-dimensional (3D) filter based on the gradient vector and Hessian matrix to enhance specific 3D local intensity structures such as line-like, sheet-like, and blob-like structures [8]. Jin et al. modified the technique and selectively enhanced the ellipsoidal shaped brain aneurysm while suppressing the false positives by other ellipsoidal shaped strictures with the Hessian-based method [9]. Since the liver tumors generally have ellipsoidal shape, the liver tumor can be enhanced by the Hessian-based method and utilized in the training process to improve the performance.

To selectively enhance the ellipsoidal liver tumor concave region, the ratio of the semi-minor axis to the semi-major axis of the ellipsoid was used [9]. The ratio could be represented by the root of the ratio of the eigenvalues of the Hessian matrix:

$$\frac{semiminor}{semimajor} = \sqrt{\frac{|\lambda_3|}{|\lambda_1|}}, \quad (1)$$

where λ_1, λ_2 and λ_3 represents the first, second and third eigenvalues of the Hessian matrix and satisfying the condition $\lambda_1 > \lambda_2 > \lambda_3 \gg 0$. The Hessian matrix in the CT image is given by:

$$H = \begin{bmatrix} I_{\sigma_{xx}}(x,y,z) & I_{\sigma_{xy}}(x,y,z) & I_{\sigma_{xz}}(x,y,z) \\ I_{\sigma_{yx}}(x,y,z) & I_{\sigma_{yy}}(x,y,z) & I_{\sigma_{yz}}(x,y,z) \\ I_{\sigma_{zx}}(x,y,z) & I_{\sigma_{zy}}(x,y,z) & I_{\sigma_{zz}}(x,y,z) \end{bmatrix}, \quad (2)$$

where $I_\sigma(x,y,z)$ represents the intensity value of the original CT image applied with a Gaussian kernel at voxel (x,y,z) with a standard deviation σ.

3.3. Concept of our semantic segmentation technique with 3D MTANN deep learning

Suzuki et al. invented supervised image-processing techniques based on convolution of an artificial neural network, called "neural filters" [10] and "neural edge enhancers" [11], for the reduction of quantum noise from x-ray images [12] and for the supervised semantic segmentation the organ in x-ray images [13]. By extension of them, an MTANN was developed by Suzuki et al. for distinguishing a specific structure from other structures in medical images [14], which accommodates the tasks of pattern-recognition in the computer-aided diagnosis (CAD) domain. Suzuki et al. proved that MTANN's architecture allows the model to achieve high performance through training with small amount of data [15]. In this study, we applied this advantage of the MTANN framework and further improved this characteristic by combining with the Hessian-enhancer.

The 3D MTANN is a supervised deep-learning model consisting of linear-output artificial neural network regression, which is capable of operating on voxel data directly. The 3D MTANN is trained with 3D input CT volumes and the corresponding "teaching" volumes for enhancement of a liver tumor and suppression of other structures. The input to the 3D MTANN is the voxel values in a sub-volume (i.e., a 3D image patch), V_s, extracted from 3D input CT images. In this study, in order to integrate Hessian-enhanced images into 3D MTANN, the corresponding sub-volumes, V_s' is also extracted from 3D Hessian-enhanced images. The sub-volumes of 3D input images and the sub-volumes of 3D Hessian-enhanced images are entered into the 3D MTANN as input in a pair-wise manner. The output, $O(x,y,z)$, of the 3D MTANN is a continuous scalar value, which corresponds to a single pair-wise sub-volume, represented by:

$$O(x,y,z) = NN\{I(x-i,y-j,z-k), I'(x-l,y-m,z-n) | i,j,k \in V_s, l,m,n \in V'_s\}, \quad (3)$$

where NN(·) is the output of the 3D MTANN, $I(x,y,z)$ represents a voxel value of the 3D input volume, and $I'(x,y,z)$ represents a voxel value of the 3D Hessian-enhanced volume.

The 3D MTANN employs a linear function instead of the sigmoid function as the activation function in the output layer because the performance and convergence characteristic of the model were improved significantly when the linear function is applied to continuous mapping of values in image processing [6]. To train the probability

Authorized licensed use limited to: Intel Corporation via the Virtual Library. Downloaded on February 05,2022 at 02:53:13 UTC from IEEE Xplore. Restrictions apply.

map of being a tumor, the following error function is minimized:

$$E = \frac{1}{P}\sum_c \sum_{(x,y,z)\in V_T}\{T_c(x,y,z) - O_c(x,y,z)\}^2, \quad (4)$$

where c is training case number, V_T is the training volume, $T_c(x,y,z)$ corresponds to the teaching volume, and P is the number of total training voxels in V_T. A linear-output backpropagation algorithm [6] is applied to train the 3D MTANN since a linear function was employed as the activation function in the output layer. After training, we expect the 3D MTANN to output the values close to the teaching images, which leads to a likelihood map of being a tumor as its output. The final segmentation result is obtained by a binarization step with an optimum thresholding value and morphological processing, which forms an adaptive thresholding layer.

Fig. 1 shows the training and testing phases of our 3D MTANN for semantic segmentation of liver tumors. In the training phase, original liver CT images and their Hessian-enhanced images were entered as input to the 3D MTANN in a pair-wise manner. Convolution kernels were simultaneously applied to the pair images at the same location, providing a single output image as a map of likelihood of being a tumor. When the output likelihood map was calculated through the 3D MTANN, the error between the output map and the teaching image was used to update the parameters in the 3D MTANN. Once the 3D MTANN was trained, in the testing phase, a new 3D CT volume together with its corresponding Hessian-enhanced image were entered into the 3D MTANN, obtaining a likelihood map of being a tumor. Final tumor segmentation results were obtained via an adaptive thresholding layer.

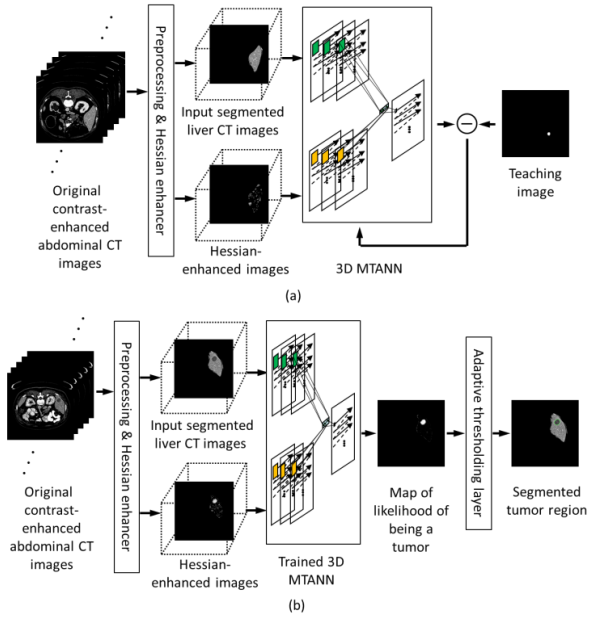

Fig. 1. Architecture of our 3D MTANN combined with a Hessian-based enhancer for tumor segmentation in a training phase (a) and a test phase (b).

4. RESULT AND DISSCUSIONS

To evaluate the advantage of our semantic segmentation scheme in training with small sample sizes, we trained 2 MTANN models with: 14 tumors from 12 patients and 28 tumors from 18 patients. The training cases were manually selected for best covering the range of size and shape of liver tumor appearances. To test the trained models independently, we excluded the whole training dataset (18 patients with 135 tumors) and used the remaining patients, namely, 24 patients with 59 tumors, for evaluation. We evaluated our models by using Dice coefficients in an independent test. Average Dice coefficients for the 2 MTANNs trained with 14 and 28 tumors were 0.703 and 0.668, respectively, with no statistically significant differences (paired t-test: P>>0.05). Notice that the meaning of no statistically significant difference in this study is that the performance of our model trained with small training data set have no significant differences with those trained with larger training data sets. This result shows our model only need small numbers of samples to reach the highest performance. The decreasing performance due the additional training cases might be caused by the bias introduced into the training cases.

The Dice coefficients for the top 10 deep-learning models that had been trained with 908 tumors in 131 patients in the MICCAI 2017 competition ranged from 0.625 to 0.702. We achieved the performance comparable to the best performance (Dice coefficient of 0.702) in the MICCAI 2017 worldwide competition with a smaller size training dataset of 12 patients than the 131 patients. Table I shows Dice coefficients for our proposed method and the top 10 deep-learning models in the MICCAI 2017 competition[5]. Fig. 2 illustrates the segmentation results of the 3D MTANN and the "gold-standard" manual segmentation with the corresponding likelihood map of being a tumor by the MTANN.

TABLE I. Comparison of Dice coefficients by our proposed method and the top 10 deep-learning models in the MICCAI 2017 competition [5].

Ranking	Researchers	Institution	Dice coefficient	# of tumors used for training	# of patients used for training
1	Tian et al.	Lenovo	**0.702**	**908**	**131**
2	Li et al.	CUHK	0.686	908	131
3	Chlebus et al.	Fraunhofer	0.676	908	131
4	Vorontsov et al.	MILA	0.661	908	131
5	Yuan et al.	MSSM	0.657	908	131
6	Ma et al.	NJUST	0.655	908	131
7	Bi et al.	Uni Sydney	0.645	908	131
8	Kaluva et al.	Predible Health	0.64	908	131
9	Han	N.A.	0.63	908	131
10	Wang et al.	KTH	0.625	908	131
	Our MTANN model 1		**0.703**	**14**	**12**
	Our MTANN model 2		0.668	28	18

Authorized licensed use limited to: Intel Corporation via the Virtual Library. Downloaded on February 05,2022 at 02:53:13 UTC from IEEE Xplore. Restrictions apply.

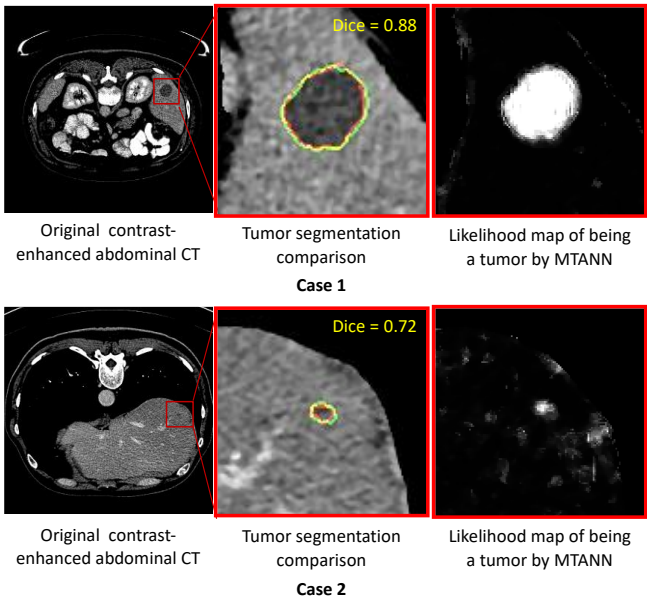

Fig. 2. Comparisons of our 3D MTANN tumor segmentation with "gold-standard" manual segmentation. Red contour: "gold-standard" manual segmentation, green: 3D MTANN, and yellow: complete agreement between the two. Corresponding likelihood maps of being a tumor by our MTANN were shown on the right.

For training a 3D MTANN model, it took 3 minutes and 22 seconds on a GPU server (CPU: Xeon E5-2698 v4, 2.20GHz with 256G RAM, GPU: Tesla V100-DGXS, Nvidia) or 18 minutes 49 seconds on a PC (CPU: Intel i7-8700K, 3.2GHz with 32G RAM, GPU: GeForce GTX 1080, Nvidia). For testing a new case by the trained 3D MTANN, it took 8.4 seconds on the server or 18.6 seconds on the PC per patient on average.

5. CONCLUSION

We developed a small-sample-size deep-learning approach to segmentation of liver tumors based on our original 3D MTANN combined with a Hessian-based enhancer. Our model required a small-size training dataset to obtain a high performance, which would be essential in deep learning applications in medical imaging where a large database is not available.

6. ACKNOLEDGEMENT

This work was supported in part by a JST-Mirai Program Grant. The authors are grateful to the members of the Suzuki Lab for their valuable discussions.

7. COMPLIANCE WITH ETHICAL STANDARDS

Ethical approval was not required because of the retrospective use of the open source data, LiTS dataset [4].

8. REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA. Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, 2018, doi: 10.3322/caac.21492.

[2] K. Suzuki *et al.*, "Quantitative radiology: automated CT liver volumetry compared with interactive volumetry and manual volumetry," *AJR. Am. J. Roentgenol.*, vol. 197, no. 4, pp. W706-12, 2011, doi: 10.2214/AJR.10.5958.

[3] K. Suzuki, R. Kohlbrenner, M. L. Epstein, A. M. Obajuluwa, J. Xu, and M. Hori, "Computer-aided measurement of liver volumes in CT by means of geodesic active contour segmentation coupled with level-set algorithms," *Med. Phys.*, vol. 37, no. 5, pp. 2159–2166, 2010, doi: 10.1118/1.3395579.

[4] A. L. Simpson *et al.*, "A large annotated medical image dataset for the development and evaluation of segmentation algorithms," 2019, [Online]. Available: http://arxiv.org/abs/1902.09063.

[5] P. Bilic *et al.*, "The Liver Tumor Segmentation Benchmark (LiTS)," no. January, 2019, [Online]. Available: http://arxiv.org/abs/1901.04056.

[6] K. Suzuki, H. Yoshida, J. Näppi, and A. H. Dachman, "Massive-training artificial neural network (MTANN) for reduction of false positives in computer-aided detection of polyps: Suppression of rectal tubes," *Med. Phys.*, vol. 33, no. 10, pp. 3814–3824, 2006, doi: 10.1118/1.2349839.

[7] K. Suzuki and K. Doi, "How Can a Massive Training Artificial Neural Network (MTANN) Be Trained With a Small Number of Cases in the Distinction Between Nodules and Vessels in Thoracic CT?," *Acad. Radiol.*, vol. 12, no. 10, pp. 1333–1341, 2005, doi: https://doi.org/10.1016/j.acra.2005.06.017.

[8] Y. Sato *et al.*, "Tissue classification based on 3D local intensity structures for volume rendering," *IEEE Trans. Vis. Comput. Graph.*, vol. 6, no. 2, pp. 160–180, 2000, doi: 10.1109/2945.856997.

[9] Z. Jin, H. Arimura, S. Kakeda, F. Yamashita, M. Sasaki, and Y. Korogi, "An ellipsoid convex enhancement filter for detection of asymptomatic intracranial aneurysm candidates in CAD frameworks," *Med. Phys.*, vol. 43, no. 2, pp. 951–960, 2016, doi: 10.1118/1.4940349.

[10] K. Suzuki, I. Horiba, and N. Sugie, "Efficient approximation of neural filters for removing quantum noise from images," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1787–1799, 2002, doi: 10.1109/TSP.2002.1011218.

[11] K. Suzuki, I. Horiba, and N. Sugie, "Neural edge enhancer for supervised edge enhancement from noisy images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1582–1596, 2003, doi: 10.1109/TPAMI.2003.1251151.

[12] K. Suzuki, I. Horiba, N. Sugie, and M. Nanki, "Neural filter with selection of input features and its application to image quality improvement of medical image sequences," *IEICE Trans. Inf. Syst.*, vol. E85-D, no. 10, pp. 1710–1718, 2002.

[13] K. Suzuki, I. Horiba, N. Sugie, and M. Nanki, "Extraction of left ventricular contours from left ventriculograms by means of a neural edge detector," *IEEE Trans. Med. Imaging*, vol. 23, no. 3, pp. 330–339, 2004, doi: 10.1109/TMI.2004.824238.

[14] K. Suzuki, S. G. Armato III, F. Li, S. Sone, and K. Doi, "Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography," *Med. Phys.*, vol. 30, no. 7, pp. 1602–1617, 2003, doi: 10.1118/1.1580485.

[15] K. Suzuki, H. Abe, H. MacMahon, and K. Doi, "Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN)," *IEEE Trans. Med. Imaging*, vol. 25, no. 4, pp. 406–416, 2006, doi: 10.1109/TMI.2006.871549.

Authorized licensed use limited to: Intel Corporation via the Virtual Library. Downloaded on February 05,2022 at 02:53:13 UTC from IEEE Xplore. Restrictions apply.