2020, 37th NATIONAL RADIO SCIENCE CONFERENCE
(NRSC 2020), September 8 - 10, 2020
German University in Cairo (GUC), Cairo, Egypt

DEEP JOINT SEGMENTATION OF LIVER AND CANCEROUS NODULES FROM CT IMAGES

Nermeen A. Elmenabawy, Ahmed Elnakib, Hossam El-Din Moustafa*

Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University, Mansoura, Egypt
eng_nermeena@yahoo.com, nakib@mans.eu.eg, and hossam_moustafa@hotmail.com

ABSTRACT

A framework is proposed for joint liver and cancerous nodule segmentation from abdomen computed tomography (CT) images. The proposed framework consists of three main units. First, a preprocessing unit is used to enhance the image contrast. Second, two different deep convolutional-deconvolutional neural networks (CDNN), namely, Alexnet and Resnet18 models, are investigated to extract the features of liver images. Finally, a pixel wise classification unit is performed to provide the final segmentation maps of the liver and tumors. Results on the challenging MICCAI'2017 liver tumor segmentation (LITS) database, using Alexnet model and 4-fold cross-validation, achieve a Dice similarity coefficient of 90.4% for liver segmentation and of 62.4% for lesion segmentation. Comparative results with related techniques for joint liver and tumor segmentations show the effectiveness of the proposed framework.

Keywords: Liver, Segmentation, Cancerous, Transfer deep learning

I. INTRODUCTION

Worldwide, the world health organization (WHO) reported an estimate of 782,000 deaths from liver cancer in 2018, occupying the fourth leading mortality causes after lung, colorectal, and stomach [1]. Locally, it is the first leading cause of cancer death among males and the second leading cause of death among females in Egypt [2]. These statistics reflects the epidemic inflation of liver cancer, globally and locally. Standard clinical procedure for liver cancer diagnostics examines visually abdomen computer tomography (CT) images to check the chance of developing a cancer. However, this procedure is subjective and time consuming. Automated accurate segmentation of the liver and/or cancerous nodules shows an ability to assist the radiologist in detection, diagnostic, and treatment decisions. The goal of this paper is to provide an automated accurate segmentation of the liver as well as cancerous nodules from abdomen CT images.

In the literature, many traditional and deep learning liver and lesions segmentation techniques have been developed. For example, Cheng and Wei [3] used a region growing algorithm to segment liver CT images, then they applied a wavelet analysis to detect tumors location. Hu et al. [4] segmented the liver by training a deep 3D convolutional neural network (CNN) to learn a liver probability map. Then, a global and local appearance descriptors are incorporated adaptively into a segmentation model to guide the liver surface evolution. Krishan and Mittal [5] detected Hepatocellular carcinoma cancer (HCC) and metastases tumors based on histogram equalization to enhance the image and supported vector machine (SVM) to perform classification. Korabelnikov et al. [6] used a CNN Alexnet model [7] for extracting deep learning features for liver tumor segmentation. Their method involves the selection of a Hounsfield range, from -2000 to 2000, as a pre-processing step and a Guassian smoothing post-processing step in order to improve the performance. Vorontsov et al. [8,9] used a pipelined framework using U-net models [10] to segment both the liver then the lesions from 2D axial CT slices. Chlebus et al. [11] used a modified U-net, composed of four resolution levels, to segment liver tumors. Gruber et al. [12] used two fully connected U-net models, sequentially, for segmenting the liver and its tumor. Bellver et al. [13] built a cascaded deep learning framework to segment the liver and its lesions from CT images, based on a base VGG16 model [14]. Bi et al. [15] used a cascaded Resnet [16] models, composed of convolutional and deconvolutional layers, to segment the liver and lesions from CT images. Han [17] used a hybrid of U-net/Resnet models, composed of 32 layers, to segment the liver tumor. Although many techniques, in the literature, have achieved considerable success in liver and/or tumour segmentations, there is still a need to investigate how to improve the accuracy.

In this work, a deep learning framework is proposed for simultaneous liver and liver cancer segmentation, based on the pretrained CNN Alexnet model. Unlike [6], the proposed framework does not apply any post-processing steps. Unlike [8], [9], [12], [13] and [15], the proposed framework uses only a one stage deep learning model, instead of cascaded designs, to perform the segmentation of the liver and tumor simultaneously. The evaluation results of the proposed framework show that incorporating efficient preprocessing steps can improve the quality of segmentation. Compared with related joint segmentation techniques of liver and cancerous nodules, the proposed

framework shows superior performance. The rest of this paper is organized as follows: Section II details the proposed framework. Section III discusses the results of the proposed system and related techniques. Finally, section IV highlights the conclusion and future works.

II. METHODOLOGY

As shown in Fig. 1, the input to the proposed system is the abdomen CT image. The proposed three-step framework applies a prepocessing step to enhance the image quality, followed by a feature extraction step using a convolutional-deconvolutional neural nework (CDNN) model, and a pixel-wise classification step to produce the final joint output segmentation map, with three labels: liver, lesion, and background (BG) (see Fig. 1). The details of each of these steps are illustrated below.

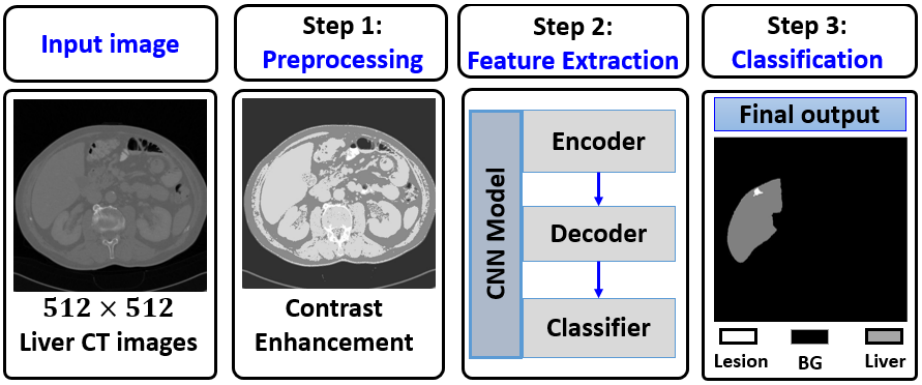

Fig. 1: The proposed framework for joint segmentation of liver and lesions, composed of three steps: pre-processing, deep learning feature extraction, and pixel-wise classification. The framework input is the original CT liver image and its output is the final joint segmentation map, with three labels: liver, lesion, and background (BG).

A. Preprocessing

In this stage, preprocessing applies two operations on the CT input image: histogram equalization and median filter smoothing. Proposed output images for each operation are exemplified in Fig. 2.

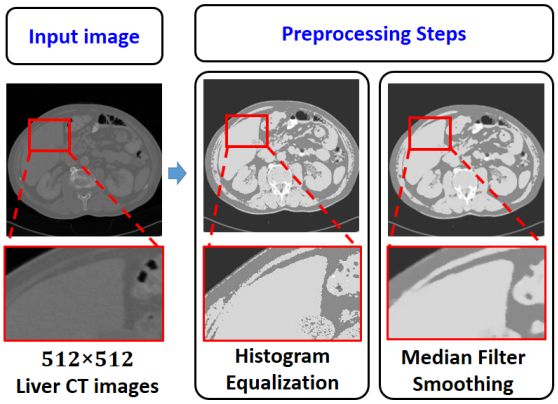

Fig. 2: The proposed preprocessing operations: histogram equalization and median filter smoothing. For each operation, the red square box is enlarged to illustrate the job of the proposed preprocessing operation.

B. Feature Extraction

To extract compact features for the joint segmentation process, two CDNN models are investigated, one is based on Alexnet model [7] and the other is based on Resnet18 model [16]. Both models have repeatedly used in many medical applications, showing good results, e.g., for liver and tumor segmentation [6],[15],[17] and for the segmentation of other different medical objects [18]. Alexnet model is composed from 62,4 million parameters and pretrained on 10 million training images from 1000 subjects [7]. The proposed CDNN model, we call it CDNN-Alex, is composed from an encoder, decoder, and a pixel classification layer (see Fig. 1). Following [19], the encoder is composed of the pretrained Alexnet layers (five convolutional layers and two out of the three fully connected layers, see Fig. 3). The encoder parameters are inherited from the pretrained Alexnet, where the last fully connected layer (FC8 in Fig. 3) is removed, following a transfer learning scheme [19]. The decoder performs a transposed

convolution (deconvolution) for up sampling, 8 times in size, such that its CDNN output is of same size as the input image. A pixel classification layer is added, such that each bit of the output image is classified into one of three labels: background, liver, or tumor. More detailed of the CDNN-Alex model can be found in [19].

On the other hand, the second investigated CNN model, to extract liver and lesion features, is the pretrianed Resnet18 model [16]. This model is pretrained on more than 1.2 million labelled high-resolution images belonging to around 200 categories. Resnet18 model is composed from a 18 weighted layers encoder, a decoder, and a pixel classification layer [16]. Applying the transfer learning scheme, the encoder parameters are inherited from the pretrained Resnet, where the last fully connected layer is removed. The decoder has the same number of layers of the encoder, performing deconvolution in order to decode the extracted features to the same dimensions of the input image. Like the CDNN-Alex model, a pixel-wise classification layer is added, such that each bit of the output image is classified into one of three labels: background, liver, or tumor.

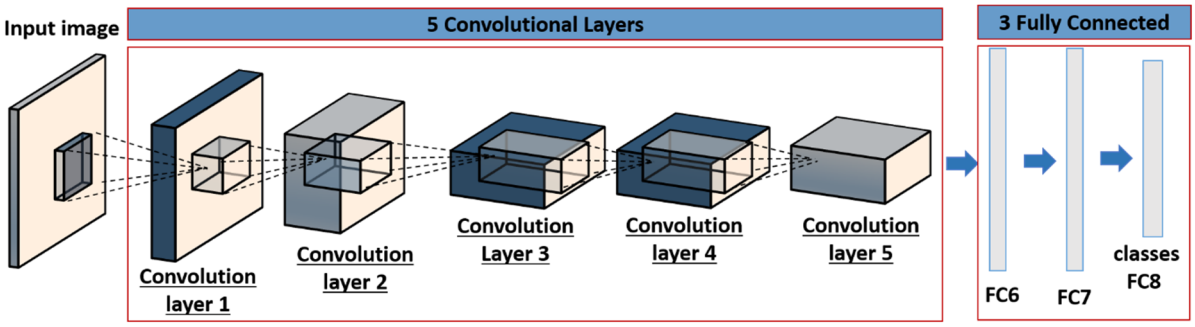

Fig. 3: Alexnet architecture [6]: five convolutional layers and three fully connected layers (FC6, FC7, and FC8)

C. Classification

To provide a joint segmentation map for both the liver and tumors pixels, a softmax layer followed by a pixel-wise classification layer are applied to the output of the CDNN-Alexnet decoder or the Resnet decoder (see Fig. 1). The pixel-wise classification layer produces the final joint output segmentation map, with three labels: liver, lesion, or background (BG).

D. Performance Metrics

To efficiently evaluate the system performance, two standard area- and distance-based metrics are used, the Dice similarity coefficient (DSC) [20] and the average symmetric surface distance ($ASSD$) [21]. Let Seg denotes the proposed system object segmentation map and G denotes the ground truth segmentation object map, provided by an expert. The DSC, an area-based metric, measures the area overlap between Seg and G images, defined as:

$$DSC\ (Seg, G) = |Seg \cap G| / 0.5(|Seg| + |G|) \times 100\ \% \quad (1)$$

where the operator $|.|$ provide the segmentation area (i.e., the number of segmented object pixels). The $ASSD$, a distance metric, describing the Euclidean distance d between two surfaces (the segmented object surface, $SegS$, and the ground truth segmentation, GS), is defined as:

$$ASSD(SegS, GS) = 1/ (|SegS| + |GS|) \times (\sum_{x \in SegS} d(x, GS) + \sum_{y \in GS} d(y, SegS)) \times 100\ \% \quad (2)$$

where x and y denote points the segmentation surface, $SegS$, and the ground truth surface, GS, rescpetiveley.

III. Experimental Results

The data specifications, experimental setup, proposed system results and discussions are detailed below.

A. LITS Database

The proposed system is tested on the Medical Image Computing and Computer-Assisted Intervention (MICCAI'2017) Liver Tumor Segmentation (LITS) challenging database [22], [23]. This database is composed of 130 abdomen CT scans, available online for training, given by seven different clinical institutions. The training CT scans are given with the reference ground truth segmentations of the liver and tumors done by trained radiologists. The number of axial slices is between 42 to 1026 per volume and a total of 16,917 images. Each cross-section is of a size of 512×512 pixels. More details of the data and its description can be found in [22] and [23].

B. Experimental Setup

In order to train the CDNN-Alexnet and Resnet18 models on the LITS database, we initialized the encoder weights by transfer all the pertained weights of the Alexnet model in [7] and Resenet18 model in [16]. For fine tuning, we trained the encoder and decoder layers on the LITS data using the Stochastic Gradient Descent (SGD) optimization, with 10^{-4} learning rate and 0.9 momentum to minimize the cross-entropy loss [24]. Matlab© 2018a is used for implementation. The maximum number of training epochs is set to 40. Training is performed by shuffling the training set for every epoch, using a mini-batch size of 1000 images.

C. Visual Results

To test the efficiency of the proposed preprocessing steps to improve the segmentation accuracy of the liver and tumors, visual qualitative results are compared for the proposed system without preprocessing and after applying the preprocessing step using the two investigated models (CDNN-Alexnet and Resnet18). As shown in Fig. 5, visual segmentation results shows that the CDNN-Alexnet shows superior performance over the Resnet18 model for all investigated cases. In addition, the visual results, in Fig. 5, support that the proposed preprocessing step can lead to improve the results for both the liver and tumor segmentations for both the two investigated models. Overall, Alexnet model associated with using the proposed preprocessing steps visually shows the best performance.

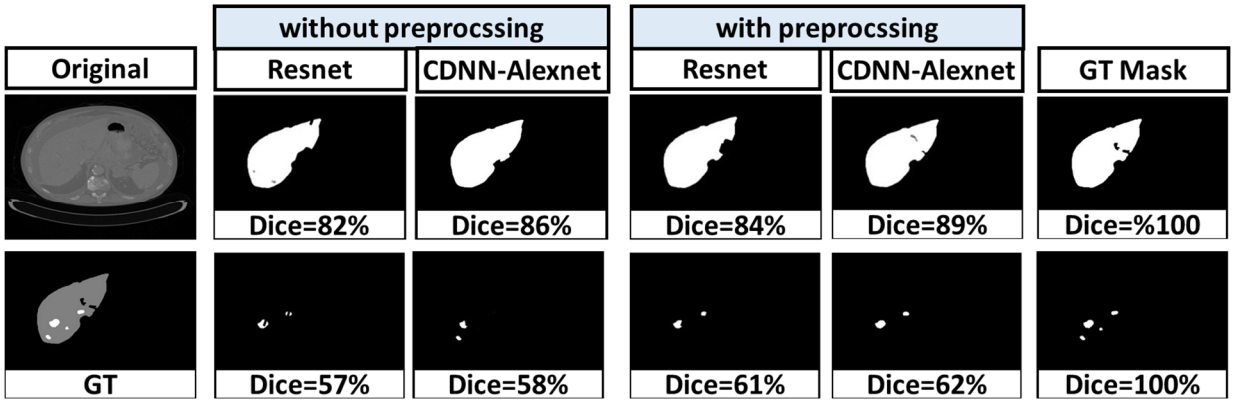

Fig. 5: Visual comparison results for a sample test CT image. First row shows the liver segmentation and second row shows the tumor segmentation. (a) The original test image and its Ground Truth (GT) segmentation. Segmentation without preprocessing using Resnet18 (b) and CDNN-Alexnet (c) models. Segmentation with preprocessing using Resnet18 (d) and CDNN Alexnet (e) models. (f) GT masks. The Dice metric for each segmentation example is given below.

D. Quantitative Results

To further investigate the potential of each CDNN model for the segmentation of the liver and tumor and the impact of the proposed preprocessing operations to improve the segmentation results, quantitative results are provided in Table 1 using 4-fold cross-validation on LITS data. Consistent with the visual results, Table 1 demonstrates that CDNN-Alexnet model shows superior performance than Resnet model. This may be due the fact that Alexnet model is of much less number of layers (7 layers) than the Resnet18 (18 layers), exhibiting much less number of training parameters, that can be trained more efficiently. In addition, Table. 1 shows that the proposed preprocessing can improve the performance of the CDNN-Alex model from a DSC of 88.54% to 90.4% for the liver, and from 59.7% to 62.4% for the tumor. These results highlight the efficiency of the proposed CDNN-Alex model with the proposed preprocessing operations for providing accurate segmentations for both liver and tumors.

E. Comparative Results

To investigate the benefits of the proposed framework, comparison results with related method on the same LITS challenging data are provided in Table 2. For precise evaluation of performance metrics of the proposed framework, the total number of 130 separate scans are divided using 4-fold cross-validation, in order to reduce the variability due to the random test data selection. The 4-fold cross-validation performs four different experiments, each using a fully different test data subjects, composed of 25% of the scans (33 test subjects' images). Note that

Table 1: Comparison results of liver and tumor segmentations using CDNN-Alexnet or Resnet models without and with the proposed preprocessing steps, in terms of DSC and ASSD metrics using 4-fold cross-validation on LITS data

CDNN Model	Preprocessing	Liver		Tumor	
		DSC	*ASSD*	*DSC*	*ASSD*
Resnet18	Without	82.3	4.21	57.7	5.76
	With	84.2	3.65	61.4	4.77
CDNN-Alexnet	Without	88.5	3.87	59.7	5.32
	With	90.4	2.65	62.4	4.98

the parameter setting of [8],[11], and [13] are kept unchanged (their results on the LITS data, shown in Table, are as published in [8], [11], and [13]). To evaluate the results of [6] on the LITS data, we applied the same 4-fold cross validation on the same folds, used to our framework. We select a default Gaussian post processing smoothing parameters of zero mean of and variance of 0.25. As shown in Table 2, the proposed framework has shown a competing performance. Whereas the tumor DSC metric of the proposed framework shows improvement, the liver DSC, even of being high (90.4%), is still below the compared methods. Nevertheless, the accurate tumor segmentation is being more clinically important for the detection, diagnosis, and treatment decisions. In the future, we will investigate how to further improve the quality of the liver segmentation and test it over other databases, e.g., sliver07 database [25].

Table 2: Comparison results between the related work and the proposed framework on the challenging LITS data, composed of 130 scans.

Paper name	Experimental results(LITS database)	Methodology	DSC	
			Liver	Tumor
Vorontsov et al. [8]	Train size=105 Test size=15	Liver lesion segmentation informed by joint liver segmentation	95.0%	52.0%
Bellver et al. [13]	Train size=104 Test size=26	U-net, with VGG16 base network	96.0%	43.0%
Chlebus et al. [11]	Train size=99 Test size=30	Modified U-net	no	58.0%
Korabelnikov et al. [6]	4-fold validation Train size=97 Test size=33	CDNN-Alexnet with post processing	88.7%	59.9%
Proposed system	4-fold validation Train size=97 Test size=33	CDNN-Alexnet with preproceeing	90.4%	62.4%

IV. CONCLUSION

In our work, we used a framework for joint liver and lesion segmentations investigating different CDNN models. The proposed framework is tested on the MICCAI challenge 2017 LITS database. Carefully designed preprocessing seps show a potential to improve the performance from a DSC of 88.54 to 90.4 for the liver, and from 59.7% to 62.4% for the tumor, using a CDNN-Alexnet model. Compared with related methods, the proposed framewrok shows competing results for liver segmentation and superior performance for tumor segmentation, which is of more clinical impact for the detection, diagnosis, and treatment decisions. In the future, we plan to investigate other deep learning architectures in order to further improve the performance. In addition, other databases will be targeted to investigate the robustness of the proposed framework

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: A Cancer Journal for Clinicians*, vol. 68(6), pp. 394-424, 2018.

[2] A.S. Ibrahim, H. M. Khaled, N. N Mikhail, H. Baraka, and H. Kamel, "Cancer incidence in Egypt: results of the national population-based cancer registry program," *Journal of Cancer Epidemiology*, vol. 2014, pp. 1-18, 2014.
[3] C. H. Cheng, and L. Y. Wei, "Rough Classifier Based on Region Growth Algorithm for Identifying Liver CT Image," *Journal of Tamkang University of Science and Technology*, vol. 19(1), pp. 65-74, 2016.
[4] P. Hu, F. Wu, J. Peng, P. Liang, D. Kong, "Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution," *Physics in Medicine and Biology*, vol. 61(24), pp. 8669-8676, 2016.
[5] A. Krishan, and D. Mittal, "Detection and classification of liver cancer using CT images," *International Journal on Recent Technologies in Mechanical and Electrical Engineering*, vol. 2, pp. 93-98, 2015.
[6] A. N. Korabelnikov, A.V. Kolsanov, P. M. Zelter, K. V. Bychenkov, and A. V. Nikonorov, "Liver tumor segmentation CT data based on Alexnet-like convolution neural nets," *Information Technology and Nanotechnology (ITNT)*, vol. 2016, pp. 348-356, 2016.
[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, pp. 1097-1105.
[8] E. Vorontsov, A. Tang, C. Pal, and S. Kadoury, "Liver lesion segmentation informed by joint liver segmentation," in *2018 IEEE 15th International Symposium on Biomedical Imaging, IEEE*, 2018. pp. 98-136.
[9] E. Vorontsov, M. Cerny, P. Régnier, L. Di Jorio, C. J.Pal, R. Lapointe, F. Vandenbroucke-Menu,S. Turcotte, S. Kadoury and A.Tang. "Deep Learning for Automated Segmentation of Liver Lesions at CT in Patients with Colorectal Cancer Liver Metastases," *Radiology: Artificial Intelligence*, vol. 1(2), pp.180014, 2019.
[10] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, October 5-9, 2015, pp. 234-241, Springer.
[11] G. Chlebus, A. Schenk, J. H. Moltz, B. Van Ginneken, H. K. Hahn, and H. Meine, "Deep learning based automatic liver tumor segmentation in CT with shape-based post-processing," in *First Conference on Medical Imaging with Deep Learning (MIDL 2018)*, Amsterdam, The Netherlands 2018, pp. 1-9.
[12] N. Gruber, S. Antholzer, W. Jaschke., C. Kremser, and M. Haltmeier, "A joint deep learning approach for automated liver and tumor segmentation," *arXiv preprint arXiv:1902.07971,* 2019.
[13] M. Bellver, K. K. Maninis, J. Pont-Tuset, X. Giró-i-Nieto, J. Torres, and, L. Van Gool, "Detection-aided liver lesion segmentation using deep learning," *arXiv preprint arXiv:1711.11069*, 2017.
[14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
[15] L. Bi, J. Kim, A. Kumar, and D. Feng, "Automatic liver lesion detection using cascaded deep residual networks," *arXiv preprint arXiv:1704.02703*, 2017.
[16] K. He, X. Zhang, S. Ren and J. Sun, "Identity mappings in deep residual networks," in *European Conference on Computer Vision. Cham*, Springer, vol. 9908, October 2016, pp. 630-645.
[17] X. Han, "Automatic liver lesion segmentation using a deep convolutional neural network method," *arXiv preprint arXiv:1704.07239,* 2017.
[18] S. Khan, and S.P. Yong, "A deep learning architecture for classifying medical images of anatomy object," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2017, pp.1661-1668, IEEE.
[19] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431-3440.
[20] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus, S. J. Haker, and R. Kikinis, "Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports," *Academic Radiology*, vol. 11(2), pp. 178-189, 2004.
[21] H. Stahl, K. Stultz, H.Stahl and K.Stultz, "Free-surface deformation measurement," in *35th Aerospace Sciences Meeting and Exhibit,*1997, p.778-785.
[22] P. Bilic, P. F. Christ, E. Vorontsov, et al., "The liver tumor segmentation benchmark (LITS)," *arXiv preprint* arXiv:1901.04056, 2019.
[23] The LITS website (2017). [Online]. Available: https://competitions.codalab.org/competitions/17094.
[24] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc.COMPSTAT'2010, Physica-Verlag HD*, 2010, pp. 177-186.
[25] H. Müller and D. Unay, "Retrieval from and understanding of large-scale multi-modal medical datasets a review," *IEEE transactions on multimedia*, vol. 19(9), pp. 2093-2104, 2017.