

# Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks

Wen Li<sup>1,2</sup>, Fucang Jia<sup>1\*</sup>, Qingmao Hu<sup>1</sup>

<sup>1</sup>Research Lab for Medical Imaging and Digital Surgery, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

<sup>2</sup>Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China  
Email: \*fc.jia@siat.ac.cn

Received October 2015

---

## Abstract

Liver tumors segmentation from computed tomography (CT) images is an essential task for diagnosis and treatments of liver cancer. However, it is difficult owing to the variability of appearances, fuzzy boundaries, heterogeneous densities, shapes and sizes of lesions. In this paper, an automatic method based on convolutional neural networks (CNNs) is presented to segment lesions from CT images. The CNNs is one of deep learning models with some convolutional filters which can learn hierarchical features from data. We compared the CNNs model to popular machine learning algorithms: AdaBoost, Random Forests (RF), and support vector machine (SVM). These classifiers were trained by handcrafted features containing mean, variance, and contextual features. Experimental evaluation was performed on 30 portal phase enhanced CT images using leave-one-out cross validation. The average Dice Similarity Coefficient (DSC), precision, and recall achieved of  $80.06\% \pm 1.63\%$ ,  $82.67\% \pm 1.43\%$ , and  $84.34\% \pm 1.61\%$ , respectively. The results show that the CNNs method has better performance than other methods and is promising in liver tumor segmentation.

## Keywords

Liver Tumor Segmentation, Convolutional Neural Networks, Deep Learning, CT Image

---

## 1. Introduction

According to the World Health Organization (WHO) reports, liver cancer has been the second major cause of death in all cancers [1]. An accurate and robust tumors segmentation method is needed for effective diagnosis and treatment.

Currently, computed tomography (CT) is one of the most widely used imaging modalities for detection and diagnosis of liver tumors, featured by its high spatial resolution and fast scanning speed. In routine clinical practices, the segmentation can be done manually by operators with good expertise and experience, but the process is

---

\*Corresponding author.

always time consuming and different operators often produce variable results. Therefore, the development of an automatic segmentation method is very challenging, mainly due to the irregularity of tumors shapes, the great diversity of tumors intensities, and the ambiguity of boundaries between tumor and surrounding normal liver tissues.

A variety of machine learning methods have been developed for automatic or semi-automatic segmentation of liver tumors at Liver Tumor Segmentation challenge 2008 [2]. Specifically, Häme *et al.* used fuzzy clustering and fitted a geometric deformable model on the clustering result to obtain tumor regions [3]; Massoptier *et al.* applied K-means clustering to extract liver lesions [4]; Shimizu *et al.* trained AdaBoost classifier to discriminate the label of voxels through a number of handcrafted features [5]; Zhou *et al.* proposed a semi-automatic segmentation method using support vector machine (SVM) classifier to extract tumor region on one slice, then projected the tumor contour to its neighboring slices and trained SVM classifier again till all slices were processed [6]. Random forest (RF) also was applied to the segmentation of brain tumor and won the BRATS 2013 [7]. All these methods require handcrafted features to train the classifiers to achieve good segmentation. Actually, the extraction and selection of features are highly problem-dependent. It is not easy to select features to represent the characteristic of liver tumor data accurately. In addition, the size of feature dimension largely influences the performance of machine learning methods.

In recent years, deep learning has become a hot research topic in the field of image processing. Many studies have shown that CNNs could obtain outstanding performance on very challenging tasks, such as visual object recognition [8], image classification [9]. The CNNs model was also applied for segmenting knee cartilage [10]. Deep CNNs introduced by LeCun *et al.* [11], is a supervised learning model formed by multi-layer neural networks. Each intermediate layer is a successively higher level abstraction. It can be able to capture highly nonlinear mappings between inputs and outputs. CNNs are fully data-driven and can retrieve hierarchical features by building high-level features from low-level ones.

In this work, we aimed to use CNNs to segment liver tumors in CT images. The overall framework for training and testing process is showed in **Figure 1**. The procedures include pre-processing, patch extraction, classifier training, classification, post-processing. To assess the performance of our model, we compared our model to four popular machine learning algorithms: AdaBoost, Random Forests (RF), and support vector machine (SVM).

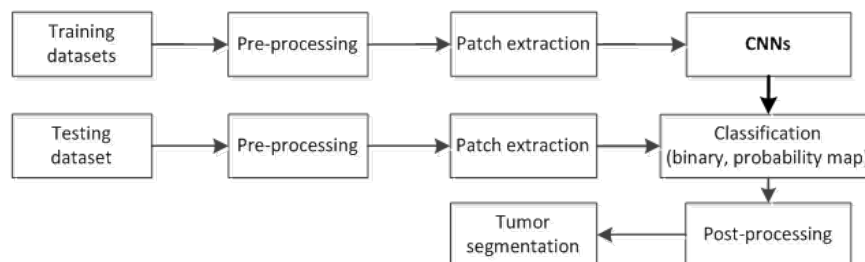
## 2. Method

### 2.1. Data Preprocessing

Simple preprocessing on CT images was performed. Firstly, Gaussian smoothing filter was used to reduce noise. Secondly, images were normalized with mean of zero and unit variance to put the data in the same scale. Thirdly, the preprocessed images were down-sampled by a factor of 2 to train the models. Testing was performed on the same preprocessing operations and down-sampling process.

### 2.2. Image Patch Sampling

We trained CNNs using image patches centered at each pixel. These patches were divided into tumor and normal liver tissue. A given patch is labeled as positive sample if it contains at least 50% or more of liver tumor pixels, otherwise it is labeled as negative sample. In consideration of the training time and the balance of training samples, we randomly sampled the same number of negative samples as much as the positive samples.



**Figure 1.** The flowchart of training and testing process for liver tumor segmentation.

### 2.3. Deep CNNs Architecture

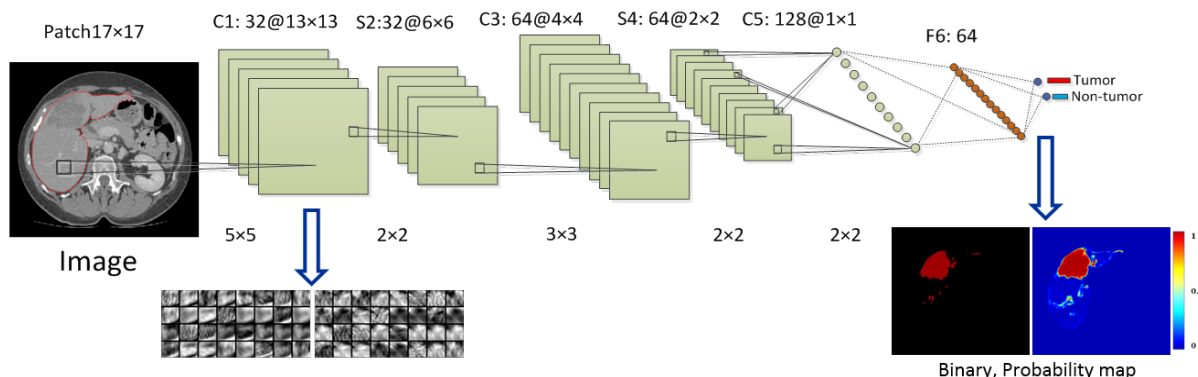
The CNNs model combines local receptive fields, shared weights, sub-sampling to ensure some degree of shift, scale and distortion invariance [11]. This model can learn lots of feature maps by convolving input image with a linear filter, adding a bias term and applying a non-linear function. Specially, in order to reduce features' dimension and avoid overfitting, max-pooling layer is introduced. The whole net optimizes parameters by minimizing cost function using stochastic gradient descent (SGD).

In this study, we designed five CNNs of different patch size to segment liver tumors. A detailed illustration of CNNs with input patch size of  $17 \times 17$  and seven layers is shown in **Figure 2**, which included three convolutional layers, two max-pooling layers, a fully-connected layer and a softmax classifier. The max-pooling operation was adopted after the convolutional layers. The first convolutional layer C1 is composed of 32 feature maps which are connected to all of image patched of  $17 \times 17$  through filters of size  $5 \times 5$  and stride of one pixel. The size of the feature maps generated in this layer was  $13 \times 13$ . The second layer S2 is max-pooling with kernel size of  $2 \times 2$  and stride of 2. Receptive fields were not overlapping. Layer C3 took the output of S2 as input with size  $6 \times 6$ . We again used the same convolutional operation to obtain 64 feature maps with the size of  $4 \times 4$ . Then we did the same operation like S2, in Layer S4 we got the feature maps with size  $2 \times 2$ . The last convolutional layer C5 consisted of 128 feature maps of size  $1 \times 1$ . The sixth fully-connected layer F6 was applied in the top of CNNs in order to discover the relationships between high-level features obtained from previous layers, where there were 64 neurons in this layer. The final layer contained two units fully connected with the layer F6, one neuron activated by softmax regression produced a value between 0 and 1 which can be interpreted as the probability of the pixel centered at the patch being tumor or not. These CNNs architectures are showed in **Table 1**. Theano [12] was used to implement the CNNs model.

**Table 1.** Details of the different CNNs architectures.

Patch size		Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6	Layer 7
13×13	Layer type	C	MP	C	MP	FC	Softmax	-
	Feature maps	64	64	128	128	64	2	-
	Filter size	3×3	2×2	3×3	2×2	1×1	-	-
	Weights	640	-	73856	-	8256	130	-
15×15	Layer type	C	MP	C	MP	C	FC	Softmax
	Feature maps	32	32	64	64	128	64	2
	Filter size	3×3	2×2	3×3	2×2	2×2	1×1	-
	Weights	320	-	18496	-	32896	8256	130
17×17	Layer type	C	MP	C	MP	C	FC	Softmax
	Feature maps	32	32	64	64	128	64	2
	Filter size	5×5	2×2	3×3	2×2	2×2	1×1	-
	Weights	832	-	18496	-	32896	8256	130
19×19	Layer type	C	MP	C	MP	C	FC	Softmax
	Feature maps	32	32	64	64	128	64	2
	Filter size	5×5	2×2	3×3	2×2	2×2	1×1	-
	Weights	832	-	18496	-	32896	8256	130

a. Layer type: C-convolutional, MP-max-pooling, FC-fully-connected.



**Figure 2.** The architecture of CNNs for liver tumor segmentation.

### 3. Experiments and Results

#### 3.1. CT Datasets and Performance Measures

The data used in this study were retrieved from Zhujiang Hospital affiliated with Southern Medical University. Twenty-six portal phase enhanced CT images were used. The images had in-plane resolution of 0.69 mm to 0.95 mm and had different spacing ranging from 0.4 mm to 2.5 mm.

For each test image, all pixels were classified as tumor or non-tumor with a class label of 1 or 0, respectively. In order to evaluate the performance, we compared the experimental results with ground truth in terms of Dice Similarity Coefficient (DSC), precision, and recall defined as follows:

$$\text{DSC}(A, B) = (2 \times |A \cap B|) / (|A| + |B|) \times 100\%, \text{ Precision} = \text{TP} / (\text{TP} + \text{FP}) \times 100\%, \text{ Recall} = \text{TP} / (\text{TP} + \text{FN}) \times 100\%$$

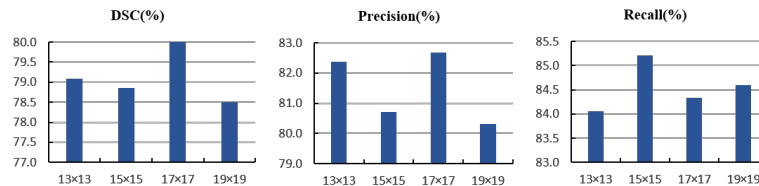
where A is the segmentation result, B is the ground truth. TP, FN, FP denotes truth positive, false negative, and false positive, respectively.

#### 3.2. Experiments and Results

In the experiments, we designed CNNs architectures with different input patch sizes to study their impact on the accuracy of liver tumor segmentation. The performance is evaluated using leave-one-out cross validation. **Figure 3** shows the comparisons of average DSC, precision, and recall values using the different CNNs. It could be seen that the patch size  $17 \times 17$  as the CNNs inputs outperformed the others with respect to DSC, and precision.

We trained our best architecture of CNNs described in Section 2.3 to segment liver tumor. The model was trained using 10 epochs. The cost function was the negative log likelihood functions and learning rate was set as 0.01. The training sets consisted of one million  $17 \times 17$  sized patches. In total, the number of trainable parameters for this architecture is 60,610. We applied our best CNNs on test datasets and displayed the segmentation result as a binary image and a probability map. AdaBoost, RF, and SVM approaches were built to evaluate the performance of our best CNNs. These automatic segmentation methods were implemented in the scikit-learn Python toolkit [13]. We extracted a 55-dimensional feature vectors associated with voxel value, mean, variance, and context features to train classifiers. The context features were obtained from the random offset vector in pixels' neighborhood. The AdBoost model was trained using 50 weak learners to produce final prediction. A RF is an ensemble model based on randomized trees. In this paper, we used RFs containing 10 trees. The SVM model used in our experiments adopted a Gaussian radius basis function (RBF). The segmentation result of these methods was refined by morphological opening and closing operations.

The experimental results of the AdaBoost, RF, SVM, and CNNs are shown in **Table 2**. We observed that CNNs performed better than the former three methods in terms of DSC and recall. These results demonstrated that automatically extracted features by CNNs are effective in segmenting liver tumors as compared to handcrafted features.



**Figure 3.** Experiment results of different architectures of CNNs used in liver tumor segmentation.

**Table 2.** Overall evaluation results for the different methods.

Methods	Performance Measures		
	DSC (%)	Precision (%)	Recall (%)
AdaBoost	75.67 ± 4.79	80.23 ± 5.59	74.26 ± 4.71
RF	79.47 ± 2.20	91.39 ± 2.13	75.11 ± 3.35
SVM	79.78 ± 3.89	83.71 ± 4.84	82.06 ± 3.26
CNNs	80.06 ± 1.63	82.67 ± 1.43	84.34 ± 1.61

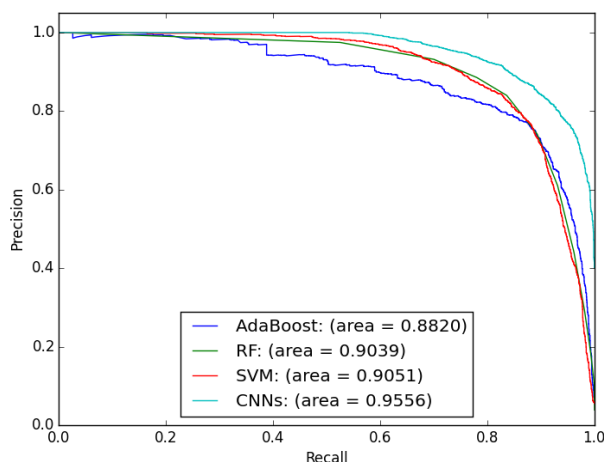
In **Figure 4**, we also show the Precision-Recall plot for each method. The areas under the Precision-Recall curve obtained by the AdaBoost, RF, SVM, and CNNs methods were 0.8820, 0.9039, 0.9051, and 0.9556, respectively. **Figure 5** shows the final segmentations of different methods on four patients' CT images. It is showed that AdaBoost, RF, and SVM model are inclined to under-segmentation, misclassify normal liver tissue into tumor, and miss small tumor. While the CNNs model obtains richer features about liver tumor boundaries and performs better than other methods.

Our method had limitation in segmenting tumors with heterogeneous intensity and fuzzy borders. When the boundary between tumor and the surrounding tissue is ambiguous, some under-segmentation may occur due to the CNNs method did not obtain efficient border information. However, CNNs can work well for inhomogeneous tumor with clear boundary.

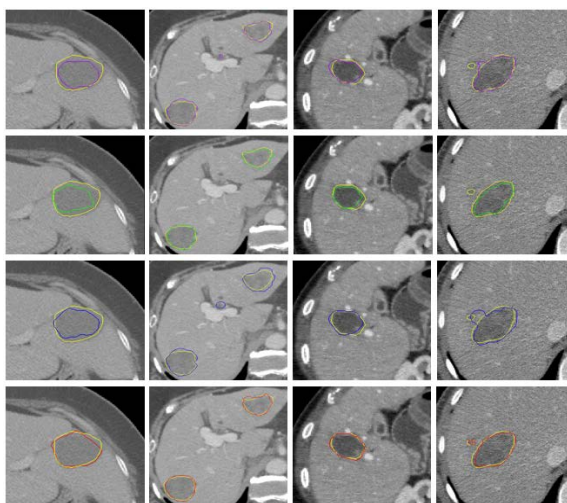
The training took about one hour on a NVIDIA Tesla K20c GPU with 2496 cores. For one test image, it needs about two minutes to segment tumors.

### 4. Discussion

In this paper a new, practical application of CNNs was presented to segment liver tumors. This method was tested on 30 CT images using leave-one-out cross validation. The experiments demonstrated that the CNNs



**Figure 4.** The performance comparison of our approach and others in the PR plane.



**Figure 5.** Liver tumor segmentation in four CT images, the methods from top to bottom are AdaBoost (purple), RF (green), SVM (blue), and CNNs (red), the ground truth (yellow).

model produced accurate and robust liver tumor segmentation. Compared to traditional machine learning methods such as AdaBoost, RF, and SVM, the CNNs method performed better. The CNNs still has limitation on segmenting tumors with inhomogeneous density and unclear boundary, especially the under-segmentation in the tumor adjacent to structures with similar densities.

## Acknowledgements

This study was supported by the following grants: NSFC-GD Union Foundation (Grant No. U1401254), Guangdong Science and Technology Grant (Grant No. 2012A080203013 and 2012A030400013), and Shenzhen Key Science and Technology Development Grant (Grant No. JCYJ20140509174140681).

## References

- [1] World Health Organization. <http://www.who.int/mediacentre/factsheets/fs297/en/>
- [2] Deng, X. and Du, G. (2008) 3D Liver Tumor Segmentation Challenge 2008. *MICCAI Workshop*.
- [3] Häme, Y. (2008) Liver Tumor Segmentation Using Implicit Surface Evolution. *Proceedings of the MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge II* 2008.
- [4] Massoptier, L. and Casciaro, S. (2008) A New Fully Automatic and Robust Algorithm for Fast Segmentation of Liver Tissue. *Proceedings of the MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge II* 2008.
- [5] Shimizu, A., Narihira, T., Furukawa, D., Kobatake, H., Nawano, S. and Shinozai, K. (2008) Ensemble Segmentation Using Adaboost with Application to Liver Lesion Extraction from A CT Volume. *Proceedings of the MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge II* 2008.
- [6] Zhou, J., Xiong, W., Tian, Q., Tian, Q., Qi, Y., Liu, J., Leow, W., Han, T., Venkatesh, S.K. and Wang, S. (2008) Semi-Automatic Segmentation of 3D Liver Tumors from CT Scans Using Voxel Classification and Propagational Learning. *Proceedings of the MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge II* 2008.
- [7] Tustison, N.J., Shrinidhi, K.L., Wintermark, M., Durst, C.R., Kandel, B.M., Gee, J.C., et al. (2015) Optimal Symmetric Multimodal Templates and Concatenated Random Forests for Supervised Brain Tumor Segmentation (Simplified) with ANTsR. *Neuroinformatics*, **13**, 209-225. <http://dx.doi.org/10.1007/s12021-014-9245-2>
- [8] Cireşan, D.C., Giusti, A., Gambardella, L.M. and Schmidhuber, J. (2013) Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks. *MICCAI* 2013, LNCS 8150, 411-418. [http://dx.doi.org/10.1007/978-3-642-40763-5\\_51](http://dx.doi.org/10.1007/978-3-642-40763-5_51)
- [9] Krizhevsky, A., Sutskever, I. and Hinton, G. (2012) Image Net Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 1097-1105.
- [10] Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E. and Nielsen, M. (2013) Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network. *MICCAI* 2013, LNCS 8150, 246-253. [http://dx.doi.org/10.1007/978-3-642-40763-5\\_31](http://dx.doi.org/10.1007/978-3-642-40763-5_31)
- [11] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. <http://dx.doi.org/10.1109/5.726791>
- [12] Theano. <http://www.deeplearning.net/software/theano/>
- [13] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011) Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*, **12**, 2825-2830.