

Automatic liver tumor segmentation in follow-up CT studies using Convolutional Neural Networks

R. Vivanti¹ MSc, A. Ephrat¹ MSc, L. Joskowicz¹ PhD, O.A. karaaslan²
,N. Lev-Cohain² MD, J. Sosna² MD

¹ The Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem, Israel.,

² Department of Radiology,
Hadassah Hebrew University Medical Center, Jerusalem, Israel,
Refael.vivanti@mail.huji.ac.il

Abstract. We present a new, fully automatic algorithm for liver tumors segmentation in follow-up CT studies. The inputs are a baseline CT scan and a delineation of the tumors in it and a follow-up scan; the outputs are the tumors delineations in the follow-up CT scan. The algorithm consists of four steps: 1) deformable registration of the baseline scan and tumors delineations to the follow-up CT scan; 2) automatic segmentation of the liver; 3) training a Convolutional Neural Network (CNN) as a voxel classifier on all baseline; 4) segmentation of the tumor in the follow-up study with the learned classifier. The main novelty of our method is the combination of follow-up based detection with CNN-based segmentation. Our experimental results on 67 tumors from 21 patients with ground-truth segmentations approved by a radiologist yield an average overlap error of 16.26% (std=10.33).

1 Introduction

Radiological follow-up of tumors is essential liver tumor therapy. The analysis of Tumor volume changes in longitudinal CT scans is required for treatment evaluation. Today, most radiologists estimate the tumor size with linear measurements methods such as RECIST [1]. It is well known that this estimate can be off by as much as 50%, especially for tumors with irregular shapes. Previous research shows that true volumetric measurements are the most accurate information for tumor monitoring [2].

Tumor delineation is the main bottleneck of tumor volume computation. Manual delineation is time-consuming, is user-dependent, and requires expert knowledge. Semi-automatic segmentation methods, e.g., live wire, region growing, and level sets also require user interaction and may lead to significant intra- and inter- observer variability. Automatic tumor segmentation poses significant challenges and is not part of the clinical workflow, with the exception of a few tumor types. Model-based methods are also limited, as they require the construction of generic tumor priors for the

segmentation. Moreover, most methods process each follow-up scan independently without taking advantage of the availability of the previous scans of the same patient.

In the past decade, researchers have developed a variety of methods for semi-automatic and automatic segmentation of liver tumors. Bourquain et al. [3] describe an interactive region-growing method for the vessels and tumors. Li et al. [4] use a machine learning technique to classify the intensity profiles of the liver tumors. The method is biased to blob-like tumors, so it is less accurate for tumors with irregular borders. Both methods require many seeds per CT slice and are thus of limited clinical use. Other methods use machine learning for liver tumors segmentation. However, the features that they use for the classification are hand-crafted and different for each method. Freiman et al. [5] use an SVM classifier to automatically produce many seeds for a graph based liver tumor segmentation. The hand-crafted features were the mean, std, minimum and maximum value in a 5x5x5 window. Zhou et al. [6] also use an SVM classifier, with the area median and the voxel value as features. Liver tumors segmentation was the subject of the 2008 MICCAI 3D Liver Tumors Segmentation Challenge Workshop. It consisted of 14 groups describing interactive, semi-automatic, and automatic liver tumors segmentation algorithms [7].

Recent works directly address the follow-up tumor segmentation task. In these works, the tumor delineation of the baseline scan serves as a patient-specific prior for the automatic tumor segmentation of the follow-up scan [8-10]. They show that the robustness and accuracy of the tumor volume and tumor volume difference measures significantly improve when the patient-specific tumor delineation from the baseline is used. Weizman et al. [8] uses a prior from the baseline MRI scan for Optic Path Gliomas segmentation. Vivanti et al. [9] use a similar method for lung tumors.

Only a few works address the follow-up of liver tumors. Cohen et al. [10] use the baseline tumor delineation after affine registration between the two scans to initialize 2D region-growing segmentation for each scan slice. Their work is limited to liver metastases. Moltz et al. [11] segment sphere-shaped liver metastases based on rigid registration between the scans. Militzer et al. [12] perform follow-up liver tumors segmentation with a pre-computed generative growth model created with Probabilistic Boosting Trees from many follow-up cases. To the best of our knowledge, [12] is the only work that does not assume spherical liver tumors in the follow-up framework.

Among the various machine learning techniques currently in use, Convolutional Neural Networks (CNN) [13] were brought to their full potential by making them both large and deep. They have proved their effectiveness in a wide variety of tasks, ranging from handwritten character recognition to neuronal structures segmentation. One of the advantages of CNN over SVM methods is that of automatically learns the features, thus obviating the need to customize hand-crafted features.

In this paper we present a new automatic algorithm for liver tumor segmentation in follow-up CT studies. The inputs are the baseline scan and the tumors delineation, and a follow-up scan; the outputs are the tumors delineations in the follow-up CT scan. The algorithm consists of four steps: 1) deformable registration of the baseline scan and tumors delineations to the follow-up CT scan; 2) automatic segmentation of the liver; 3) construction of a voxel classifier by training a Convolutional Neural Network (CNN) on all baseline scans and; 4) final segmentation of the tumor in the follow-up

study with the learned classifier. The main novelty of our method is the combination of follow-up based detection with CNN-based segmentation.

The advantages of our method are: 1) it is fully automatic; 2) it addresses a wide variety of liver tumors and metastases; 3) it performs local deformable registration to model more accurately the tumor transformation; and 4) it uses CNN to simultaneously learn features and builds a classifier based on them.

2 Method

The basic premise of our method is that the radiologist-validated tumor delineation in the baseline scan is a high-quality prior for the tumor location and size in the follow-up scan. The tumor location and size prior is automatically constructed by registration of the baseline and follow-up scans, thus allowing us to handle a large variety of tumors sizes (Fig. 1), and to obviate the need for detection method.

2.1 Registration

The first step is to automatically compute a liver mask using the stand-alone liver segmentation method in [15]. The method relies on Bayesian classification, adaptive morphological operations, and active contours for liver segmentation. We perform this segmentation for both the baseline and the follow-up scans. Although not always accurate, this mask provides an adequate coarse Region of Interest (ROI).

The next step is to define a ROI that contains the follow-up tumor with high probability. This ROI is obtained by registering the baseline scan and its tumors delineations to the follow-up scan. The registration between the baseline and the follow-up scan is performed in two steps. The first is a global deformable registration between the baseline and follow-up scans in the liver ROI automatically computed in the baseline scan using the liver mask. The liver global ROI deformable registration consists of a rigid affine registration followed by a deformable registration with B-Splines.

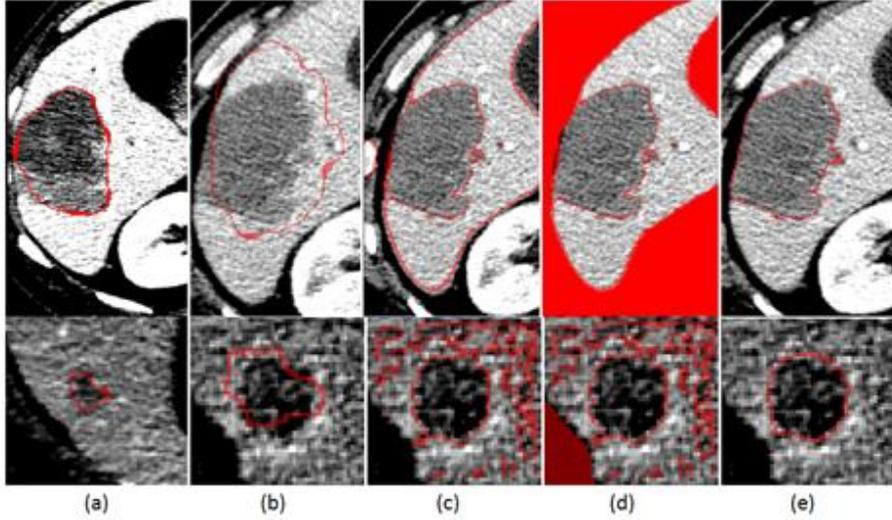
The second step is a separate local deformable registration for each baseline tumor delineation. The baseline tumor delineation is enclosed in an axis-aligned bounding box that defines the local tumor ROI. The local registration is performed for each baseline tumor in three stages: 1) pure translation registration; 2) rigid affine registration, and; 3) deformable registration by Mutual Information. The baseline delineation is transferred to the follow-up scan using the concatenation of the resulting transformations, and bounded in an axis-aligned 3D bounding box. This local registration step allows modeling more accurately the tumors changes.

Finally, the follow-up tumor ROI is doubled in each direction to account for possible tumor growth of up to eightfold in volume and to compensate for residual registration errors. The registration stage obviates the need for a separate detection step.

2.2 Deep Learning

We use a Convolutional Neural Network (CNN) to classify each voxel as being either ‘tumor’ or ‘healthy’. The classification is based on voxel intensities in an axis-

aligned square centered at the voxel. The liver mask is used to exclude the voxels



from the baseline set that are outside the liver.

We define first an Artificial Neural Network (ANN) as a directed weighted graph whose computation units are the graph vertices and whose arcs are weighted. In a feed-forward computation, for each vertex v_i , the values from each input vertex v_j is

Fig. 1. Illustration of the main steps of the segmentation process on two tumors (top and bottom row): (a) baseline tumor with delineation (red) on which the CNN is trained; (b) follow-up tumor with transformed baseline delineation superimposed on it. The deformable registration between the baseline and the follow-up scans is used to set the ROI that contains the follow-up tumor; (c) tumor voxel classification based on the CNN; (d) liver mask for the removal of false positives, and; (e) final segmentation after segmentation leaks removal.

multiplied by the weight of the connecting edge w_j . The output of v_i is a function of the sum $\sum_j w_j v_j$, e.g., the Rectified Linear Unit function $\text{ReLU}(x) = \max(0, x)$.

In the training step, the ANN is discriminatively trained by determining its edge weights with the standard back-propagation algorithm. The weights are updated by stochastic gradient descent with the equation:

$$\Delta w_{ij}(t+1) = \Delta w_{ij}(t) + \eta \frac{\partial C}{\partial w_{ij}}$$

where η is the learning rate and C is the cost function. At iteration t , a single tagged training example is used to adjust the weights by back propagation. After each epoch – a one-time pass over the entire training set – the learning rate η is reduced to allow finer weights adjustment in the following epochs.

In a CNN, one or more layer is convolutional: the nodes are grouped in kernels, and the weights of the input edges to these nodes are the values of this kernel. When feed-forwarding through each node, the computation of the sum $\sum_j w_j v_j$ is also the convolution of one kernel element and the input values. The actual values of the kernels, together with all other edge's weights are determined by back propagation. The

convolution layer replaces the correlation with manually-determined kernels in the feature-extraction step in other machine learning methods. The advantage of CNN is that it can simultaneously learn both the appropriate kernels for feature extraction and a voxel-classifier based on those features.

We use a CNN with seven hidden layers (Fig. 2). The input layer has one node for each pixel a the 35x35 patch. The first hidden layer is convolutional, with 48 kernels. Each kernel computes a convolution of the input with a 4x4 kernel followed by a ReLU function and a 2x2 pooling layer. Layers three and five are convolutional, with 48 5x5 kernels followed by a ReLU function and a 2x2 pooling layer. Layer seven is fully-connected with 200 nodes followed by a ReLU function. The output layer is the classification layer with two fully connected output nodes based on the softmax function:

$$p_j = \frac{\exp(x_j)}{\sum_k \exp(x_k)}$$

where p_j is the probability to be in a class, and x_j and x_k are the total inputs from nodes j and k in the former layer, respectively. The number of kernels, nodes, and their functions were determined experimentally after several trial runs.

To separate between training and test sets, we train our network on the baseline scans, and test it on the follow-up tumors. The training set is derived from the baseline tumor in four stages: 1) remove non-liver voxels from the baseline scans using the liver mask; 2) normalize the baseline tumor ROI to compensate for different contrast agent doses so its mean and std intensity values are equal to those of the follow-up tumor ROI; 3) remove from the training set examples from the larger class, so the ‘tumor’ and ‘healthy’ classes are of equal size, and; 4) shuffle the training set. We then train the CNN with this training set for many epochs until convergence, using dropout (randomly zeroing) of half of the edges after each epoch to avoid over-fitting.

2.3 Segmentation

Once the CNN is trained, the follow-up tumor ROI is segmented by classifying all of its voxels in four steps: 1) run the trained CNN in feed-forward to classify each patch; 2) classify non-liver voxels as healthy tissue with the liver mask of the follow-up scan; 3) remove the remaining segmentation leaks with the method described in [8], and; 4) remove small holes and “islands” with morphological operations.

3 Experimental Results

We evaluate our method on 67 tumors from pairs of CT scans from 21 patients. The scans were acquired on a 64-row CT scanner (Phillips Brilliance 64) and are of size $512 \times 512 \times 350-500$ voxels, $0.6-1.0 \times 0.6-1.0 \times 0.7-3$ mm, with contrast agent. The cases were carefully chosen from the hospital archive to represent the variety of patient ages, conditions, and pathologies. The tumors include hypodense, mixed tumors, and metastases of varying sizes and shapes with volume > 1 cc. The mean time difference between the baseline and the follow-up scans is 3.78 months (std=2.44), in a time range of 1.03-11.2 months. The tumor volumes range is 1.1-4477.86cc. The mean volumetric change is 126.48cc (std=224.05). An expert radiologist approved the tumors ground-truth delineations in both the baseline and the follow-up scans.

We quantify the follow-up tumor segmentation error with the DICE volumetric overlap error (VOE) and the average symmetric surface distance (ASSD) over the entire liver. A VOE $> 70\%$ is considered as failure and not included in the mean.

We compare our results to Freiman et al. [5] since their code is available to us, report good stand-alone segmentation results [7] and use machine learning with hand-crafted features. We use the liver tumor ground truth center of mass as a seed for their method. Note that our validation data set is significantly larger and more diverse than that of [7].

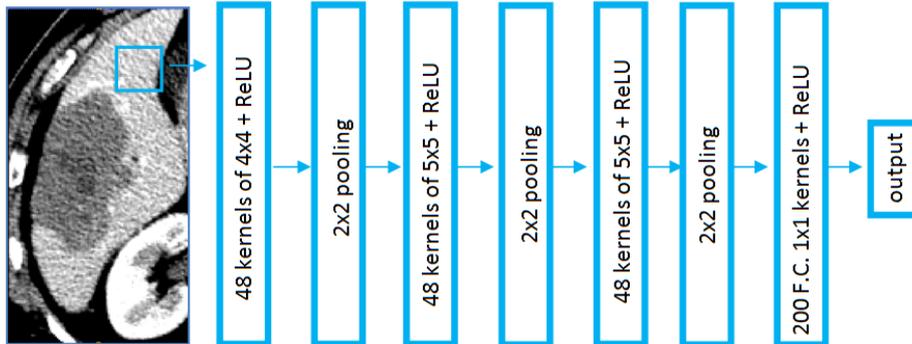


Fig. 2. Illustration of the CNN as a voxel classifier: nine layers, of which 7 are hidden layers, **Input:** 35x35 voxel-centered patch. **F.C.** – Fully Connected. **Output:** voxel classification as ‘tumor’ or ‘healthy’

We have implemented our method with the following settings. For registration, we use the Elastix registration toolbox [16] with the gradient descent optimizer with up to 200 iterations. For B-Spline registration, we set the grid spacing to 12mm. For CNN, we use the Caffe deep learning framework with GPU acceleration [17]. In each epoch, a batch of 100 examples is processed simultaneously. We stop the convergence after 6,000 epochs. All computations were performed on an Intel® Core™ i7-4930K CPU @3.40GHz, 3701 MHz, 6 Cores, 32 GB RAM running Windows 7 x64 operating system and NVIDIA GeForce GTX TITAN GPU.

Table 1 summarizes the results. Our method achieves a VOE of 16.75% (std=9.88), a significant improvement of 60.29% in comparison to the tumor stand-alone segmentation in [5]. The ASSD is 2.05 mm (std=1.68), an improvement of 81.65%. The success rate was improved by 89.98%. The minimum and maximum values were also improved. The running time of our method was less than 5 minutes for all cases.

To quantify the contribution of the segmentation step, we compute the accuracy of the patient-specific prior to the registration step. The VOE and ASSD without the segmentation step are 42.46% (std=17.1) and 3.32 mm (std=1.74) respectively. To quantify the inter-observer delineation variability, we asked a second radiologist to delineate 10 datasets. The mean VOE and ASSD between the delineations are 11.83% (std=11.12) and 1.16 mm (std=1.12) respectively.

We conclude that the follow-up framework effectively focuses the segmentation on the tumor ROI using the baseline tumor delineation and contributes to the robustness and accuracy of the follow-up segmentation. This is achieved by providing a strong prior for the follow-up tumor segmentation. Stand-alone methods such as that in [5] must detect the tumor ROI in the entire image base on weak or non-existent priors, which sometimes fails altogether and may decrease the delineation accuracy.

	VOE [%]				ASSD [mm]				Success %
	Mean	Std	Min	Max	Mean	Std	Min	Max	
Ours	16.75	9.88	4.53	36.10	2.05	1.68	0.28	5.14	90.47
[5]	42.18	19.43	8.49	65.06	11.17	7.89	3.72	28.96	42.85
Reg.	42.46	17.10	18.02	69.05	3.32	1.74	1.21	6.35	100
2nd obs.	11.83	11.12	1.29	29.32	1.16	1.12	0.07	3.52	100

Table 1. Results: **VOE** – Volume Overlap Error. **ASSD** – Average Symmetric Surface Distance. **Ours** - our results, **[5]** - results of [5] on our database, **Reg.** - using the transformed delineation from the baseline scan into the follow-up scan as a segmentation. **2nd obs.** - delineations of a second radiologist, to measure the inter-observer variability.

4 Conclusions

We have presented a new automatic liver tumor segmentation method for follow-up CT studies. The inputs to the method are baseline CT scan of the liver, the delineation of the tumor in it and the follow-up scan. The output is the delineation of the tumor in the follow-up scan. Our method uses a cascade of registration methods to define a well-fitted tumor ROI on the follow-up scan based on the baseline delineation. A Convolutional Neural Network is trained on all baseline liver masking to clas-

sify tumor and healthy voxels. The CNN is used as a voxel classifier to produce the follow-up tumor segmentation. The segmentation leaks in the resulting tumor segmentation are then removed to produce the final result.

The novelty of our work is in the use of CNN with automatic features learning – in contrast with previous work that use various hand-crafted features. Importantly, the follow-up framework obviates the need for tumor detection step, significantly increasing robustness and accuracy as compared to stand-alone segmentation methods. Our method yields an overlap error of 16.75%, an improvement of 60.29% in comparison to [5]. Our registration approach includes an additional local step that focuses on the tracked tumor and helps refine the ROI. Our results on 67 tumors pairs from 21 patients show a considerable improvement over stand-alone SVM based methods and may provide relevant clinical measurements for liver tumors. We plan to apply our method to other organs' tumors, and to additional imaging modalities, e.g. MRI.

References

1. Eisenhauer, E., Therasse, et al.: New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eu. J. cancer* 45(2) (2009) 228-247
2. Miller, A.B., Hoogstraten, B., Staquet, M., Winkler, A.: Reporting results of cancer treatment. *Cancer* 47(1) (2011) 207–214
3. Bourquain, H., Schenk, A., Link, F., Preim, B., Peitgen, O.H.: Hepavision2a software assistant for preoperative planning in living related liver transplantation and oncologic liver surgery. *Proc.16th Conf. Comp. Assisted Radiology and Surgery* (2002) 341–6
4. Li, Y., Hara, S., Shimura, K.: A machine learning approach for locating boundaries of liver tumors in ct images. *Proc. 18th Int. Conf. on Pattern Recognition* (2006) 400–3
5. Freiman, M., Cooper, O., Lischinski, D., Joskowicz, L.: Liver tumors segmentation from CTA images using voxels classification and affinity constraint propagation. *International Journal of Computer Assisted Radiology and Surgery* 6(2) (2010) 247-255
6. Zhou, J., Xiong, W., Tian, Q., Qi, Y., Liu, J., Leow, W. K., Han, T., Venkatesh, S.K., Wang, S.C.: Semi-automatic segmentation of 3D liver tumors from CT scans using voxel classification and propagational learning. *MICCAI Workshop* 41 (2008) 43
7. Deng, X., Du, G.: *Proc. of the 3D Segmentation in the Clinic: grand challenge II - Liver Tumor Segmentation*. www.grand-challenge2008.bigr.nl/proceedings/liver/articles (2008)
8. Weizman, L., Ben-Sira, L., Joskowicz, L., Precel, R., Constantini, S., Ben-Bashat, D.: Automatic segmentation and components classification of optic pathway gliomas in MRI. *Medical Image Computing and Computer-Assisted Intervention* (2010) 103-110

9. Vivanti, R., Joskowicz, L., Karaaslan, O.A., Sosna, J.: Automatic lung tumor segmentation with leaks removal in follow-up CT studies. *Int. J. Comp. Assisted Radiology and Surgery* 10.1007/s11548-015-1150-0, to appear (2015)
10. Cohen, A.B., Diamant, I., et al.: Automatic detection and segmentation of liver metastatic lesions on serial CT examinations. *Proc. SPIE Med. Imag. Conf.* (2014) 903519-27
11. Moltz, J.H., Schwier, M., Peitgen, H.O.: A general framework for automatic detection of matching lesions in follow-up CT. *Proc. IEEE Int. Symp. Biomed. Imag.* (2009) 843-6
12. Militzer, A., Tietjen, C., Hornegger, J.: Learning a prior model for automatic liver lesion segmentation in follow-up CT images. *Dept Informatik Technical Reports, CS-2013-03, ISSN 2191-5008* (2013)
13. Fukushima, K.: Neocognitron: A self-organizing neural network for a mechanism of pattern recognition unaffected by shift in position. *Bio. Cybernetics* 36(4) (2011) 193–202
14. Ji, S., Xu, W., Yang, M., Yu, K.: 3D convolutional neural networks for human action recognition. *Pattern Anal. and Machine Intel., IEEE Trans. on*, 35(1) (2013) 221-231
15. Freiman, M., Eliassaf, O., et al.: An iterative Bayesian approach for nearly automatic liver segmentation: algorithm and validation. *Int J Comp. Aided Rad. and Surg.* 3 (2008) 439-46
16. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim J.P.: Elastix: a toolbox for intensity-based medical image registration. *Trans. Med. Imaging* 29(1) (2010) 196–205
17. Jia, Y.: Caffe: An open source convolutional architecture for fast feature embedding. <http://caffe.berkeleyvision.org> (2013)