

---

# COMPUTER TOMOGRAPHY IMAGE SEGMENTATION WITH SUBMANIFOLD SPARSE CONVOLUTIONAL NETWORKS

---

**Dmitry Retinskiy and Tatiana Khanova**  
xperience.ai  
Nizhny Novgorod, Russia  
{dmitry, tatiana}@xperience.ai

## ABSTRACT

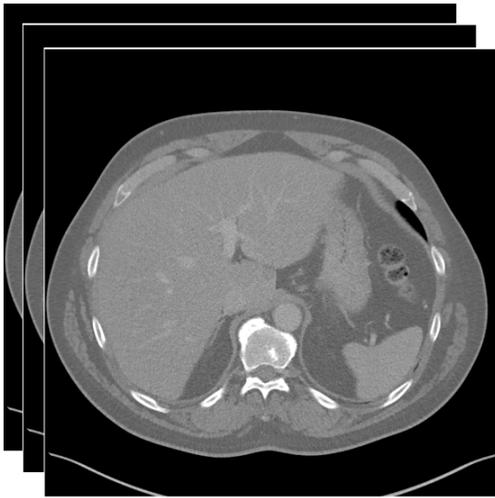
We present method for effective kidney and kidney's tumor segmentation based on the 3-dimensional model reconstructed from the computed tomography (CT) input images. Analysis of the 3D model is done using submanifold sparse convolutional network (SSCN).

**Keywords:** SSCN · 3D Semantic Segmentation

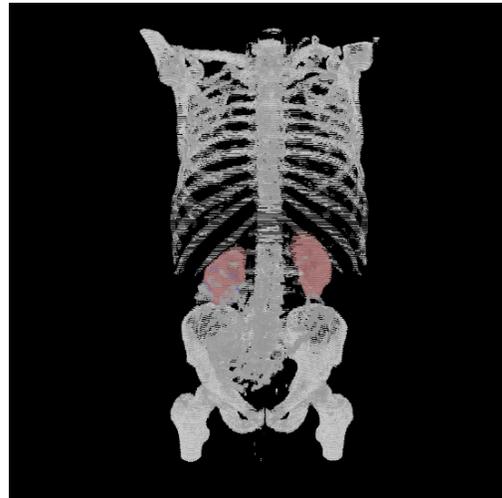
## 1 Introduction

Nowadays conventional methods of medical image analysis are applied in 2 dimensions. Methods based on 2.5D and 3D representation are just arising, see a recent survey [3]. In theory, information in 3D should provide better segmentation quality, as additional spatial dimension allows identifying and separating organs in the body with higher precision. In this paper we describe several techniques on how to approach 3D segmentation using SSCN and describe general processing pipeline from converting input CT scan to 3D point cloud, training DL network and mapping results from 3D back to the CT scan format.

High-level pipeline is designed as follows. We use ScanNet example from SSCN [1] as a baseline. First, we preprocess training data [4]: for each example we go through the image deck in the scan, take points with top 3% intensity and add them into 3D point cloud along with the corresponding labels. As it can be seen from the data, kidney almost completely covered by the top 10% intensity points. To fit training process into GPU memory we take only 3%, which, as experiments show, is quite sufficient for the task. We train 2 models: 1st model detects and extracts kidney with the tumor as one object, 2nd model takes point cloud formed at the previous step and performs fine-tuning and final segmentation into 3 classes: tissue, kidney and tumor. The last step produces CT images basing on the model output: for each image in the deck we label its points, separate them into clusters, apply convex hull and fill space inside with the corresponding label. Finally labels in each slice in the object are written back to the corresponding segmentation file in CT scan in the NIFTI format [5].



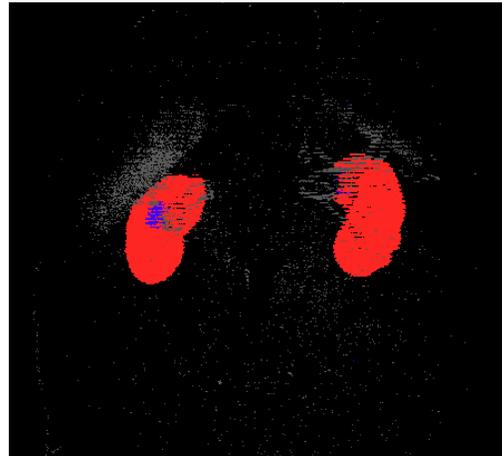
(a) Input deck of CT images



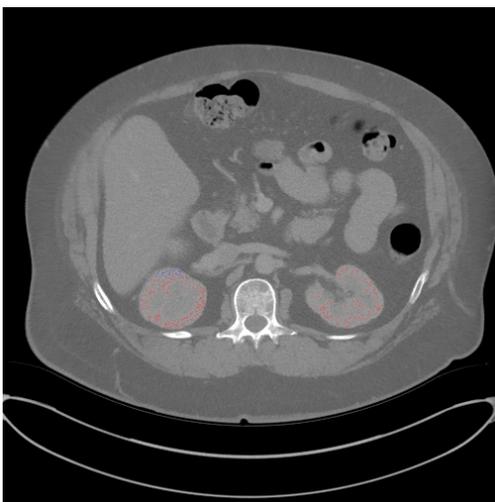
(b) 3D point cloud



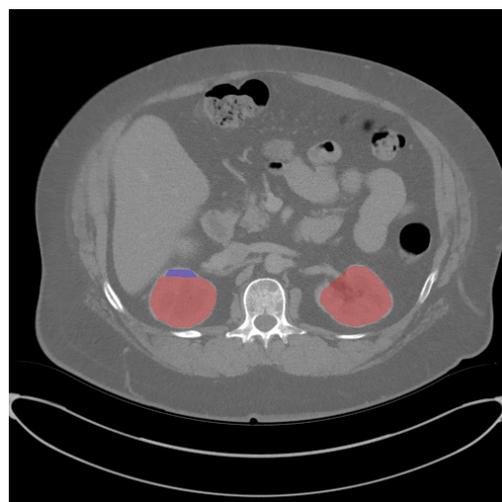
(a) Kidney cloud



(b) Segmented kidney/tumor/tissue cloud



(a) Segmented points on the CT image



(b) Applied convex hull for the points

## 2 Implementation details

### 2.1 Data pre-processing

CT scan consists of slices (e.g. from 40 to 700) and each slice is an image, where each dot is characterized by intensity value. Our goal during pre-processing is to produce point cloud containing enough information for the SSCN network to train and, at the same time, to fit training process into available GPU memory. Kidney in the input data was fully represented by the points with top 10% intensity, so the remaining 90% of the points could be safely removed. Further, during training and inference phases we divide these 10% into several parts and process sequentially to fit into resource constraints where applicable. Result of the pre-processing step is the set of point clouds for each training example, where each point is defined by 3 coordinates and ground-truth label (for the training set).

### 2.2 Training

Our model is based on the ScanNet example from SSCN repository [1]. In short, ScanNet is built on the Unet model with addition of submanifold sparse convolutions [2]. We have two training phases. During 1st phase we train the network to detect kidney and tumor points as belonging to one object (having 2 classes: kidney and tissue) and extract corresponding point cloud. At the 2nd phase this kidney cloud serves as an input to the network which is trained to perform segmentation into 3 classes (tissue, kidney, tumor). This two-step approach allows us to increase tumor segmentation quality significantly.

### 2.3 Data post-processing

Final step maps labeled points to the original CT images. To be able to perform such mapping we store original point coordinates with each point in the cloud, note that point cloud can be transformed and cut during training/inference steps. But mapping of the points is not enough to get good segmentation of the image due to significant sparsity of the points, e.g. you get only some point-wise contours of the object, while you expect to get a dense shape. So, the next step applies convex hull and fills points inside the hull with the corresponding label. Since we usually have several objects of interest on the image, before applying convex hull we perform clusterization of the labeled points using DBSCAN algorithm, which, in addition, gets rid of the possible outliers.

## 3 Experimental results

Below results correspond to the dataset with 170 training and 40 validation cases. We believe that the proposed method has a lot of room for improvement, e.g. tuning training parameters, augmentation techniques, etc. would allow reaching higher IOU for the tumor class.

Network	Tissue IOU	Kidney IOU	Tumor IOU
1st step: 2-classes network	0.979	0.819	-
2nd step: 3-classes network	0.918	0.965	0.548

## References

- [1] SSCN GitHub repository: <https://github.com/facebookresearch/SparseConvNet>
- [2] Ben Graham. Sparse 3D convolutional neural networks. In *arXiv:1505.02890v2*
- [3] Hesamian, M.H., Jia, W., He, X. et al. Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. In *Journal of Digital Imaging (2019) 32: 582*. <https://doi.org/10.1007/s10278-019-00227-x>
- [4] Nicholas Heller, Niranjana Sathianathan, Arveen Kalapara, Edward Walczak, Keenan Moore, Heather Kaluzniak, Joel Rosenberg, Paul Blake, Zachary Rengel, Makinna Oestreich, Joshua Dean, Michael Tradewell, Aneri Shah, Resha Tejpal, Zachary Edgerton, Matthew Peterson, Shaneabbas Raza, Subodh Regmi, Nikolaos Papanikolopoulos, Christopher Weight. The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes. In *arXiv:1904.00445v1 [q-bio.QM]*
- [5] NIFTI format description: <https://nifti.nimh.nih.gov>