Automatic Segmentation of The Prostate in 3D MRI

Christian Eschen, MSc student in Medicine and Technology, Technical University of Denmark
Supervisor: Professor Ole Winther, Technical University of Denmark

Abstract

Automatic segmentation of the prostate in 3D MRI is of high clinical demand since it has high diagnostic value regarding benign and malignant prostatic hypertrophy. Semantic segmentation of the prostate in 3D MRI is a challenging task since there is a large variation of prostate shapes, there are different parameters used for acquisition of T2-weighted MRI of the prostate and the amount of data is limited.
We propose a 3D dense fully convolutional neural network for segmentation of the prostate in 3D MRI.
We find that access to a broad variety of feature maps in the final layer is necessary for successfully training of deep convolutional neural networks in 3D. We show state of the art on the Promise-12 challenge with a dice coefficient equal 0.8667 using 3D dense fully convolutional neural networks.

Data

The Promise-12 challenge data consist of 50 T2-weighted 3D MRI images of variable resolution, number of voxels and field of view.
We separate the data into 45 volumes for training and 5 for validation.

Online data augmentation

Sub volumes are uniformed sampled from the background and the prostate with equal probability. The size of the sub volumes are 64x64x16. All sub volumes are z-score normalized.
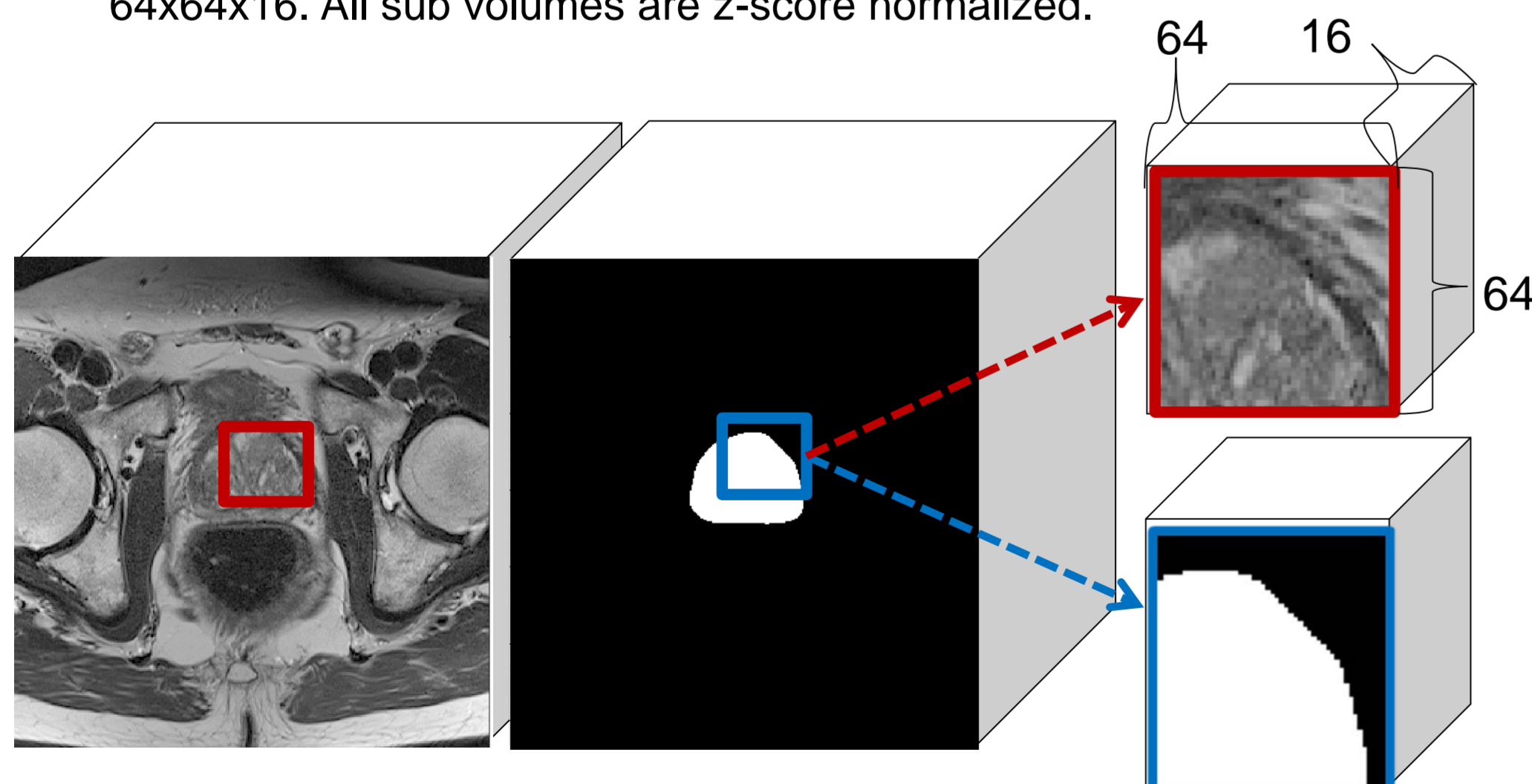

Figure 1: Online data augmentation: Uniform sampling of both classes with equal probability.

Methods

A fully 3D convolutional neural network with dense connections was utilized. The dense block[1] consists of iterative concatenation of each layer after Batch Normalization, ReLU and convolution:

$$x_l = H_1([x_0, x_1, ..., x_{l-1}])$$

We employ a 3D dense fully convolutional neural network, which is an extension of the Tiramisu[2] 2D model architecture to 3D.
The numbers of layers pr dense block are [9,10,12,15,17,20,17], respectively.
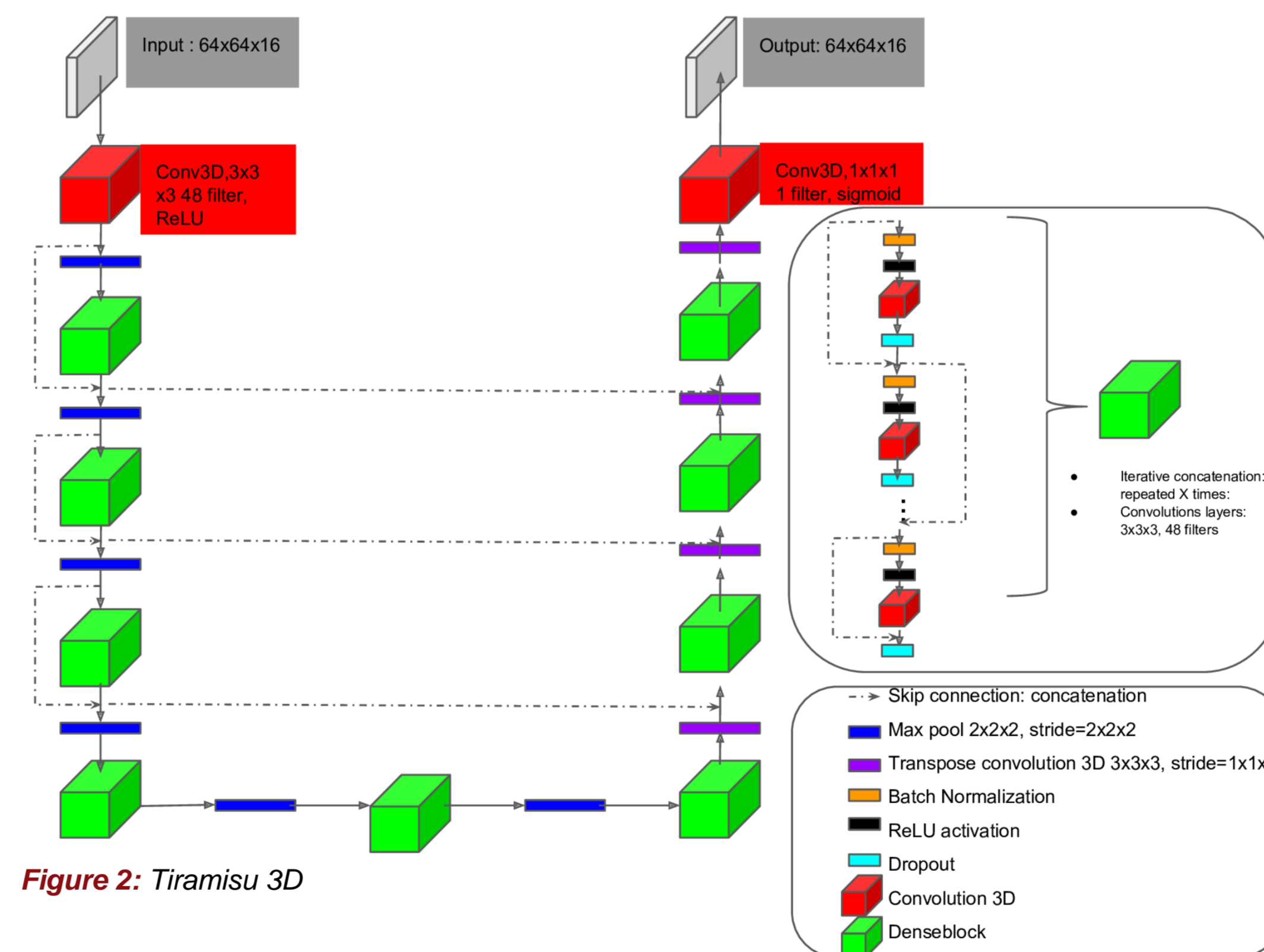

Figure 2: Tiramisu 3D

The network was trained with binary crossentropy and with the Adam optimizer for 1000 epochs. The learning rate was initially set to 0.0001 with decay 0.0001, and drop of learning rate with 0.1 each 300 epoch. The training was carried out asynchronously on 4 Titan X Pascal GPU's with batch size 2 on each GPU.

Results

The training and the validation loss converged. The large variance on the validation set is due to the low validation data size.
The inference was carried out using the overlapping sub volumes of size 64x64x16 with stride 32x32x8. The probabilities for the overlapping sub volumes was averaged. We report a validation dice coefficient on 0.8667 on whole image inference.
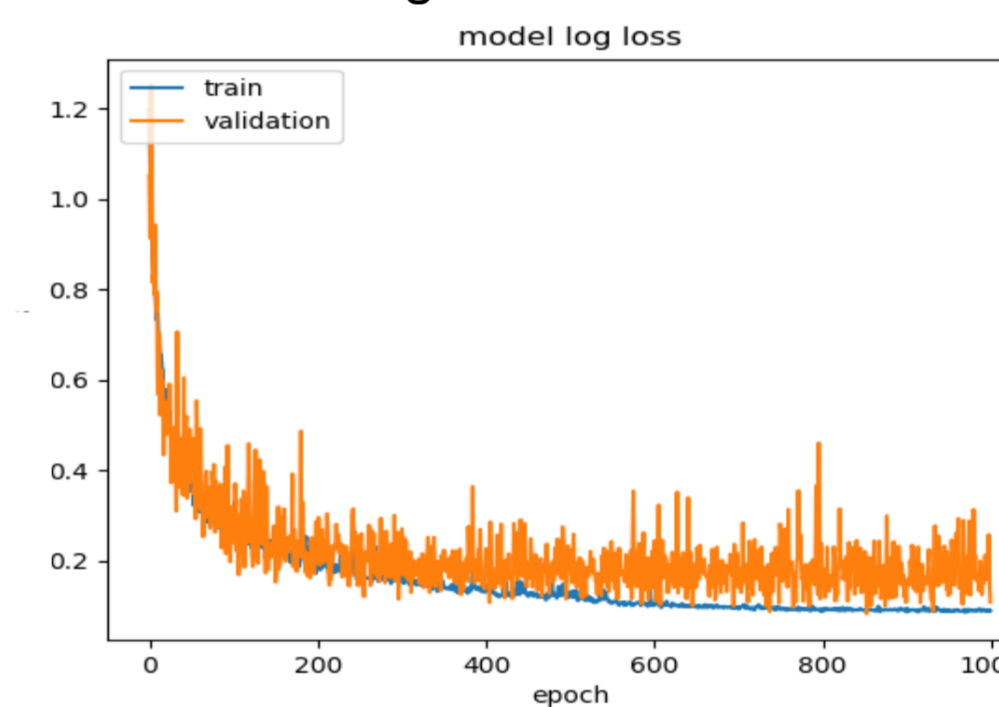

Figure 3: Training and validation loss

Rank	Name	Score
1	Philips DL-MBS	87.21
2	AutoDenseSeg	87.19
3	WHU-CS-sigma-RPI	87.04
4	CUMED	86.65

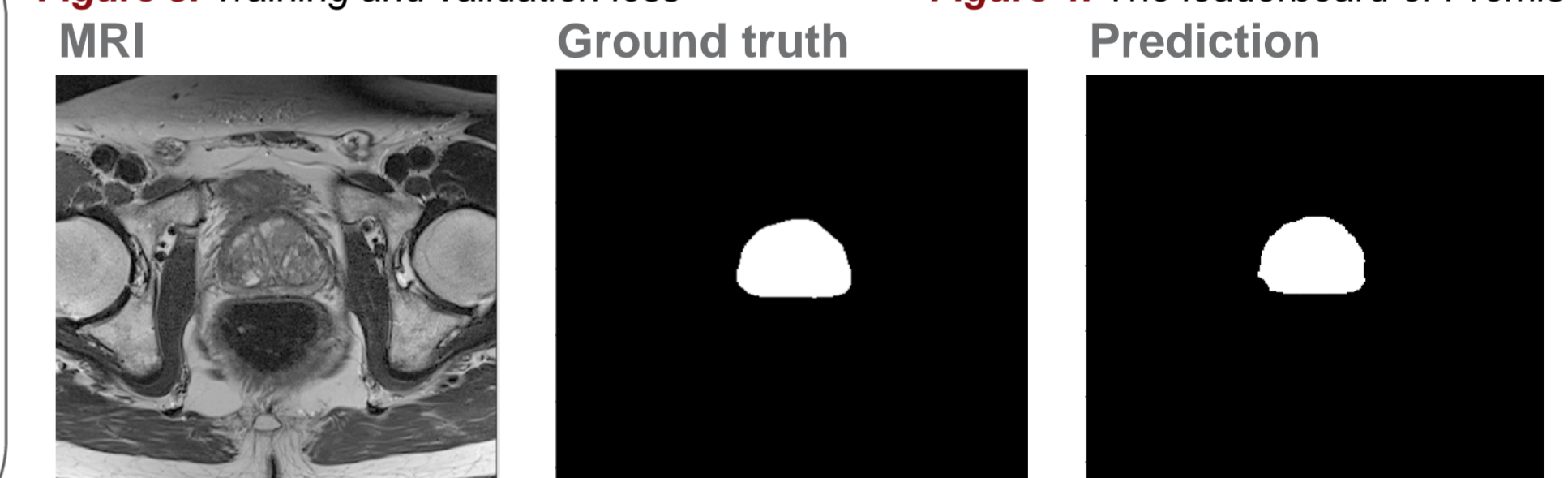*Figure 4: The leaderboard of Promise-12*

Figure 5: Inference and predictions

Conclusion

We propose the 3D Tiramisu architecture which shows state of art on the Promise-12 challenge dataset. We have observed that increased numbers of layers pr dense block through the network is important for the performance . The final layer before softmax needs access to a broad variety of feature maps, which yields deep supervision.

References

1. Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. Vol. 1. No. 2. 2017.
2. Jégou, Simon, et al. "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation." Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on. IEEE, 2017.