3D U-Net:
Learning Dense Volumetric Segmentation from Sparse Annotation

Özgün Cicek[1,2], Ahmed Abdulkadir[1,4], Soeren S. Lienkamp[2,3], Thomas Brox[1,2], and Olaf Ronneberger[1,2,5]

[1] Computer Science Department, University of Freiburg, Germany
[2] BIOSS Centre for Biological Signaling Studies, Freiburg, Germany
[3] University Hospital Freiburg, Renal Division, Faculty of Medicine, University of Freiburg, Germany
[4] Department of Psychiatry and Psychotherapy, University Medical Center Freiburg, Germany
[5] Google DeepMind, London, UK

[o] Vision
COMPUTER VISION University of Freiburg

UNI FREIBURG

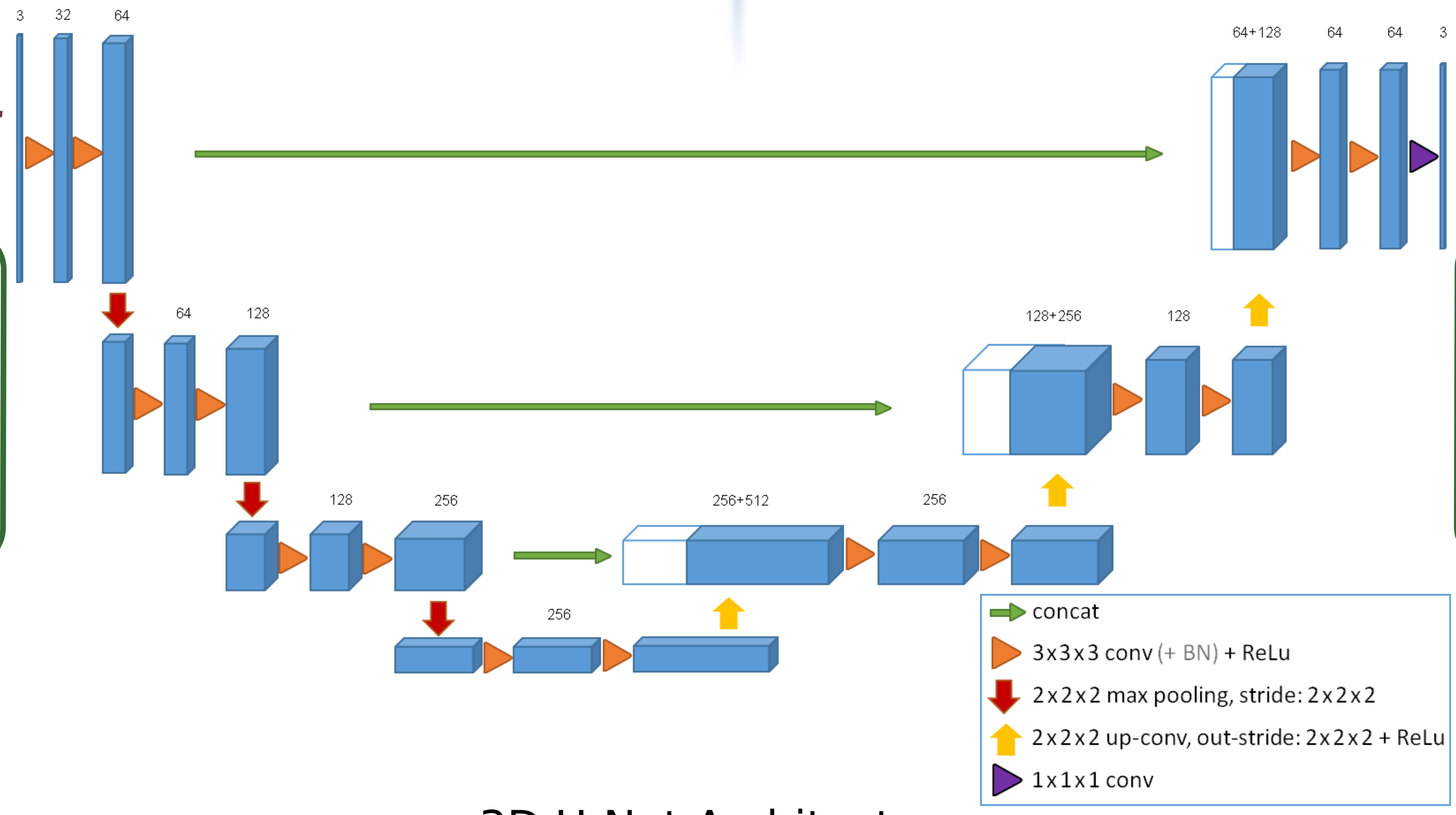
Abstract. We introduce a 3D convolutional neural network (CNN) for volumetric segmentation that learns from sparsely annotated volumetric images. We outline two attractive use cases of this method: (1) semi-automated segmentation, (2) fully-automated segmentation. The proposed network extends the previous u-net architecture from Ronneberger et al. [1] into 3D. We evaluate the proposed method on a complex 3D structure, *Xenopus* kidney, and achieve good results for both use cases.

Challenge 1. Deep learning approaches require rich data-sets annotated by experts. However, annotation of volumetric data is tedious since only 2D slice-wise annotations are possible. Annotating a volume slice-by-slice is also redundant considering the similarity between the neighboring slices.

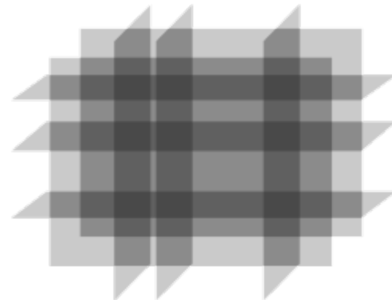
Challenge 2. Training CNNs requires millions of samples. However, acquiring millions of biomedical images is not feasible. Therefore, we adopt the data augmentation of u-net [1] which smoothly deforms the images on-the-fly to introduce infinitely many and biomedically plausible images.

On-the-Fly Data Augmentation

- cropping
- rotation
- smooth elastic deformation
- intensity transformation

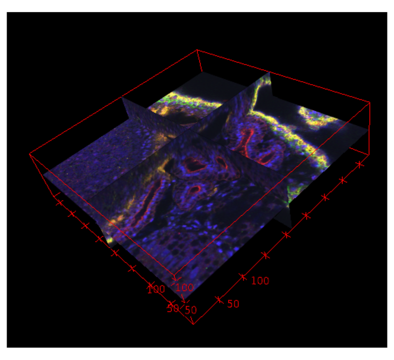


Learning from Sparse Annotation Using "Weighted Softmax Loss"

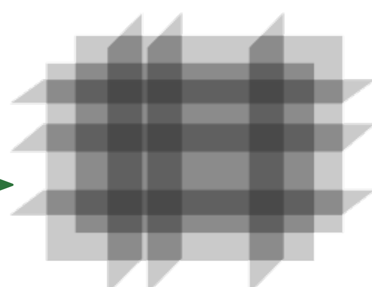


3D U-Net Architecture
without Bottleneck [2] and Optionally with Batch Normalization (BN) [3]
which Extends U-Net [1] to 3D by Replacing Each Operation by Its 3D Counterpart.

Use Case 1: Semi-Automated Segmentation

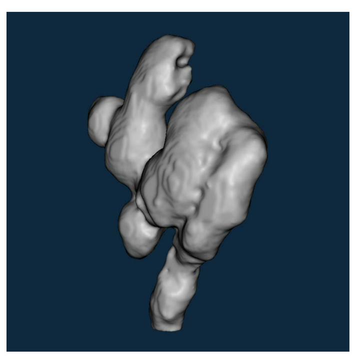


raw data



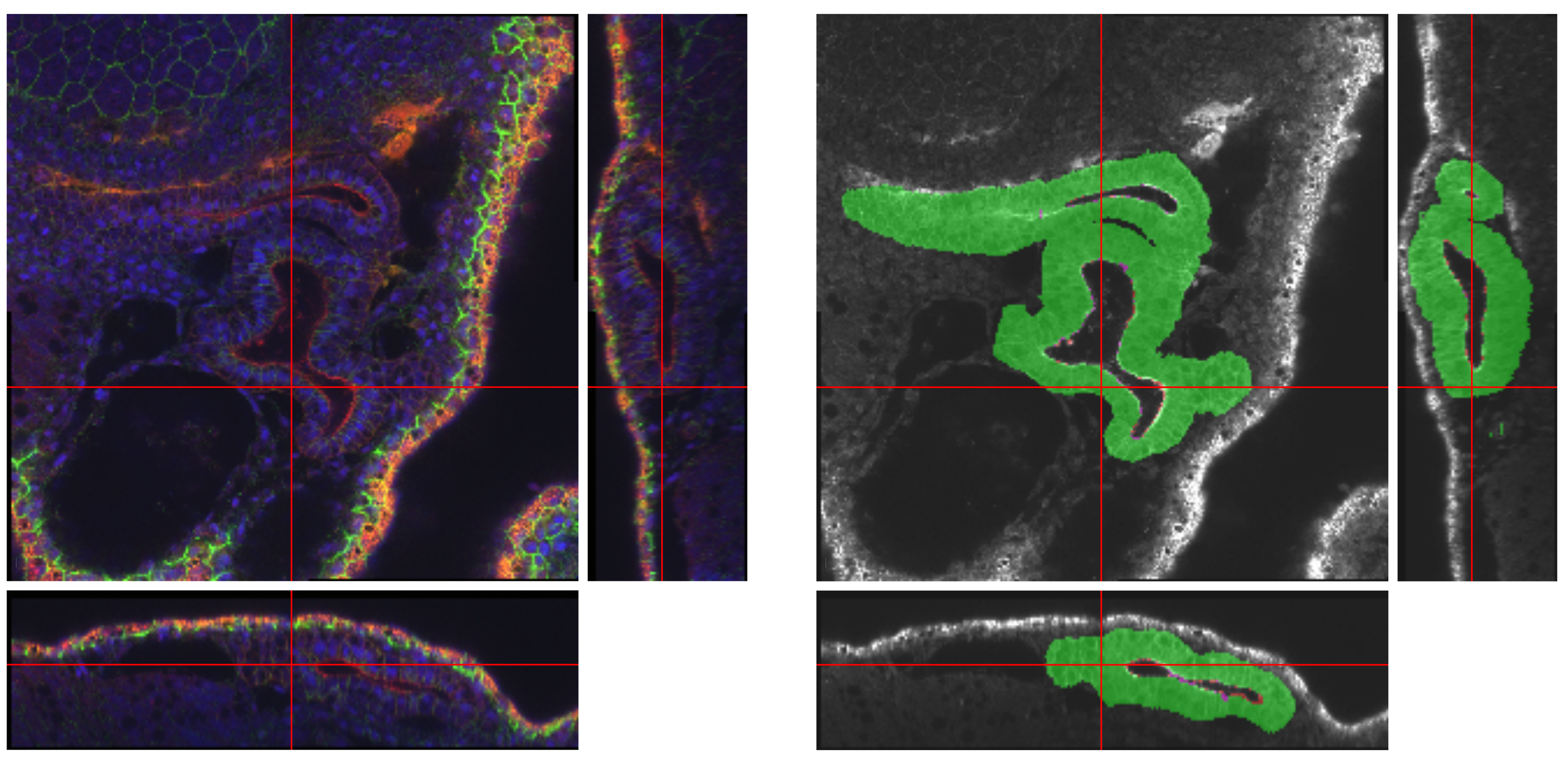
manual sparse annotation by Slicer3D [4,5]

train and apply 3D U-Net



dense segmentation

Qualitative Results: (left) raw image - (right) segmentation



channel0: Tomato-Lectin
channel1: DAPI
channel2: Beta-Catenin

true positives
false negatives
false positives

Quantitative Results

Table 1: Cross validation results for semi-automated segmentation (IoU)

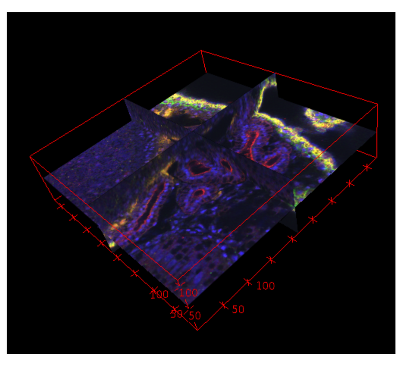
test slices	3D w/o BN	3D with BN	2D with BN
subset 1	0.822	0.855	0.785
subset 2	0.857	0.871	0.820
subset 3	0.846	0.863	0.782
average	0.842	0.863	0.796

Sparsity

Table 2: Effect of # of slices for semi-automated segmentation (IoU)

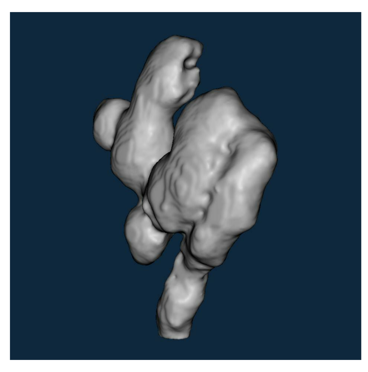
GT slices	GT voxels	IoU S1	IoU S2	IoU S3
1,1,1	2.5%	0.331	0.483	0.475
2,2,1	3.3%	0.676	0.579	0.738
3,3,2	5.7%	0.761	0.808	0.835
5,5,3	8.9%	0.856	0.849	0.872

Use Case 2: Fully-Automated Segmentation



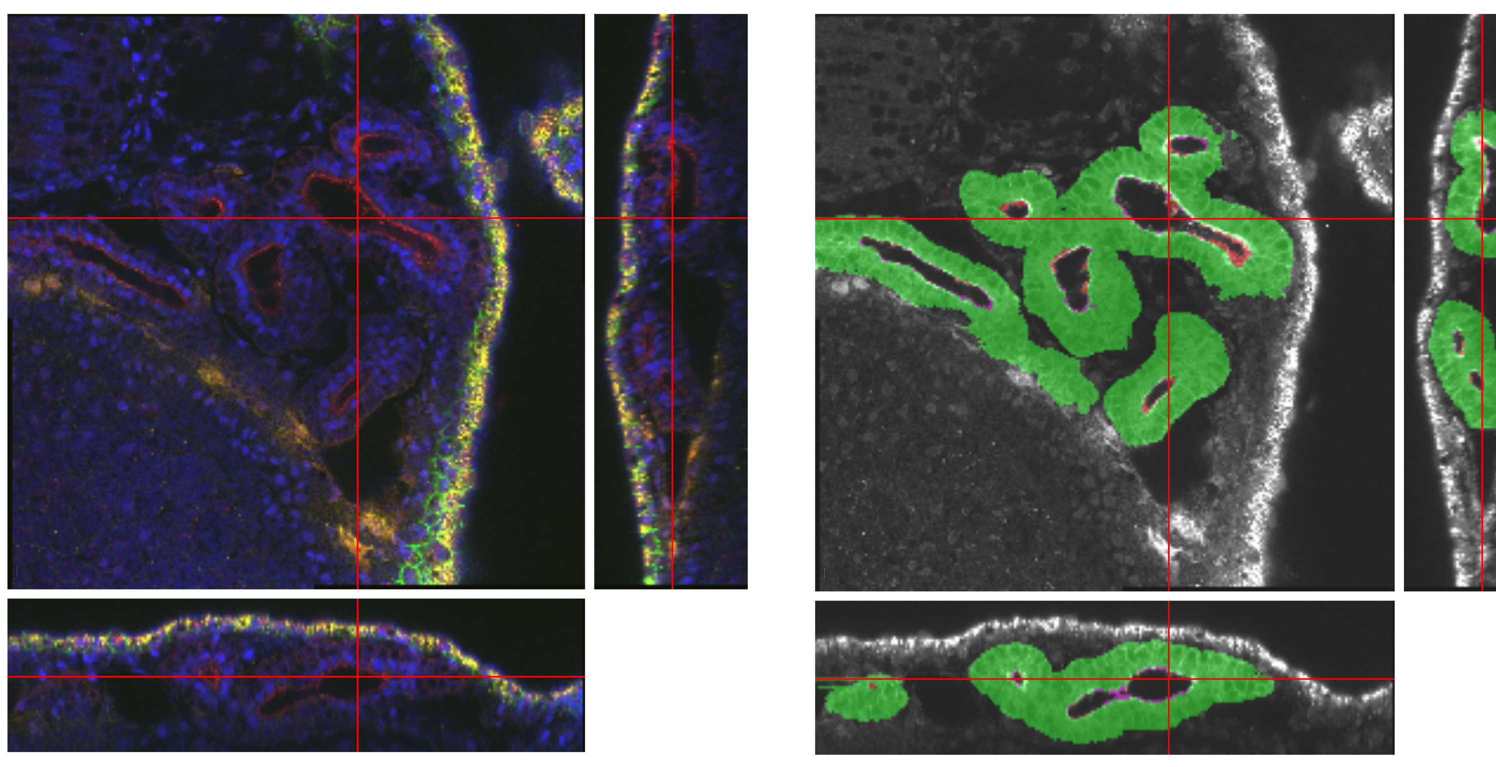
raw data

apply trained 3D U-Net



dense segmentation

Qualitative Results: (left) raw image - (right) segmentation



channel0: Tomato-Lectin
channel1: DAPI
channel2: Beta-Catenin

true positives
false negatives
false positives

Quantitative Results

Table 3: Cross validation results for fully-automated segmentation (IoU)

test volume	3D w/o BN	3D with BN	2D with BN
1	0.655	0.761	0.619
2	0.734	0.798	0.698
3	0.779	0.554	0.325
average	0.723	0.704	0.547

References

[1] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: MICCAI. LNCS, vol. 9351, pp. 234-241. Springer (2015).
[2] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. To appear in: CVPR (2016).
[3] Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: ICML, pp. 448-456 (2015).
[4] Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., et al.: 3D slicer as an image computing platform for the quantitative imaging network. J. Magn Reson Imaging 30(9), 1323-1341 (2012).
[5] https://www.slicer.org
[6] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. In: Proc. ACMMM. pp. 675-678 (2014).

Software

The implementation of the 3D u-net (based on BVLC Caffe [6]) is available as open-source at: http://lmb.informatik.uni-freiburg.de/resources/opensource/unet.en.html

Acknowledgement

We thank the DFG (EXC 294 and CRC-1140 KIDGEM Project Z02 and B07) for supporting this work. AA acknowledges funding by the grant KF3223201LW3 of the ZIM (Zentrales Innovationsprogramm Mittelstand). SSL acknowledges funding from DFG (Emmy Noether-Programm). We also thank Elitsa Goykovka for the useful annotations and Alena Sammarco for the excellent assistance in imaging.

MICCAI 2016 Athens GREECE
19TH INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING & COMPUTER ASSISTED INTERVENTION
October 17-21 2016
INTERCONTINENTAL ATHENAEUM
ATHENS / GREECE

DFG

bioss CENTRE FOR BIOLOGICAL SIGNALLING STUDIES

KIDGEM

IIF INSTITUT FÜR INFORMATIK FREIBURG