DKFZ-GE

Authors: Fabian Isensee*, Paul F Jäger*, Simon AA Kohl, Jens Petersen, Klaus H Maier-Hein

Email: f.isensee@dkfz.de

Platform: Linux (tested on Ubuntu)

Prerequisites: Python 3.6 or higher, CUDA-capable GPU (Pascal or newer)

DKFZ-GE: SUMMARY

nnU-Net [1] is our publicly available framework for automated design of segmentation methods. Given a set of training images, nnU-Net automatically configures the entire segmentation pipeline, from preprocessing, network architecture, training scheme, inference scheme all the way to post-processing. Here, we apply nnU-Net to four datasets to showcase its generalization capabilities and by extension its usefulness for analyzing datasets in the biological domain. Our submission simply used the segmentation pipelines generated by nnU-Net without further manual intervention. Since nnU-Net revolves around supervised learning, we restrict ourselves to the datasets with sufficiently densely annotated training data available: **Fluo-C3DH-A549**, **Fluo-C3DH-A549-SIM, Fluo-N2DH-SIM+**, and **Fluo-N3DH-SIM+**. We did not use external data and we did not annotate additional images.

DKFZ-GE: PREPROCESSING

Each image is normalized separately by subtracting its mean and dividing by its standard deviation. nnU-Net is designed for semantic segmentation. To overcome this limitation for **Fluo-N3DH-SIM+** and **Fluo-N2DH-SIM+** (which require instance segmentation), we treat the border and center of the cells as a separate semantic classes, thus converting the instance segmentation task into a 2-class (3 if one counts the background) semantic segmentation problem. Due to the borders separating adjacent cell labels, instance segmentations can be retrieved by connected component analysis (see below). The thickness of the border class is set to 0.5 μm for **Fluo-N3DH-SIM+** and 0.7 μm for **Fluo-N2DH-SIM+**. **Fluo-C3DH-A549** and **Fluo-C3DH-A549-SIM** contain only one cell and can be treated as a semantic segmentation problem. Due to the similarity between the simulated and real dataset, we pooled the training cases to create a single large dataset which we refer to as *Fluo-C3DH-A549(-SIM)*.

*Contributed equally

DKFZ-GE: SEGMENTATION

nnU-Net is an out-of-the box framework that fully automatically designs segmentation pipelines. All network architectures generated by nnU-Net follow a plain 3D U-Net [2, 3] template which is automatically adapted to each dataset. We apply nnU-Net independently to **Fluo-N3DH-SIM+**, **Fluo-N2DH-SIM+**, and *Fluo-C3DH-A549(-SIM)*. Only images with complete annotations are added to the training dataset. For **Fluo-N3DH-SIM+** and *Fluo-C3DH-A549(-SIM)*, nnU-Net does not utilize time information and processes each frame independently. For **Fluo-N2DH-SIM+**, we provide the four time steps prior to the frame of interest as additional input channels (as imaging modalities).

For *Fluo-C3DH-A549(-SIM)*, we manually select nnU-Net's 3D full resolution configuration and train four U-Net models. Each model was trained on three of the four available time series. The configured 3D U-Net processes input patches of shape 24×256×256 and a batch size of 2. Due to the lack of sufficiently diverse training cases, only single models were trained for **Fluo-N3DH-SIM+** and **Fluo-N2DH-SIM+** using all available training images from their respective dataset. For **Fluo-N3DH-SIM+** we again manually selected the 3D full resolution configuration. The configured 3D U-Net processes input patches of shape 32×192×384 and a batch size of 2. Being a 2D dataset, only the 2d configuration is available for **Fluo-N2DH-SIM+**. Here, nnU-Net configures a patch size of 896×768 and a batch size of 4 for its 2D U-Net.

nnU-Net trains the networks with the sum of cross-entropy and Dice loss [4, 5]. Stochastic gradient descent with large (0.99) Nesterov momentum is used for optimization. The learning rate is initialized to 0.01 and decayed with the *polylr* schedule as in [6]. Data augmentation is applied on the fly during training. Training is done for a total of 1000 epochs where we define one epoch as 250 iterations (thus 25k iterations total). The nnU-Net default would be to run five-fold cross-validation on the training cases and utilize this information to empirically determine which configuration to use (for these datasets this could have been either *2d* or *3d_fullres*) and what post-processing (based on all but largest component suppression) to use.

We use the four models from the cross validation on *Fluo-C3DH-A549(-SIM)* as an ensemble to predict the corresponding test cases. Final predictions are obtained by averaging the individual softmax outputs. The test cases of **Fluo-N3DH-SIM+** and **Fluo-N2DH-SIM+** are predicted by the corresponding single model (no ensembling).

DKFZ-GE: POST-PROCESSING

Post-processing was not applied: nnU-Net selected not to use it for *Fluo-C3DH-A549(-SIM)* and the lack of a cross-validation on **Fluo-N3DH-SIM+** and **Fluo-N2DH-SIM+** prevented the determination of post-processing altogether. However, for **Fluo-N3DH-SIM+** and **Fluo-N2DH-SIM+**, nnU-Net predicts the cell center region as well as the border region. These must be converted back to instance segmentation by identifying connected components on the cell center map. The resulting cell instances are then successively grown into the border region until no border region remains.

REFERENCES

1. Isensee F, Jäger PF, Kohl SA, Petersen J, Maier-Hein KH. Automated design of deep learning methods for biomedical image segmentation. arXiv:1904.08128, 2019.
2. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention*, 234-241 (2015).
3. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention*, 424-432 (2016).
4. Drozdzal M, Vorontsov E, Chartrand G, Kadoury S, Pal C. The importance of skip connections in biomedical image segmentation. In *Deep Learning and Data Labeling for Medical Applications*, 179-187 (2016).
5. Milletari F, Navab N, Ahmadi SA. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *Proceedings of the 4th International Conference on 3D Vision*, 565-571 (2016).
6. Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **40**, 834-848 (2017).