What is a nnU-Net?

Anton Zhitomirsky

December 31, 2023

This is a sumamry of the contents of the file [3] and example of its implementation (from the github [2]).

1 Summary of [3]

Similarly to the original U-Net architecture they don't change much, but since they are now dealing with 3D data, they consider a pool of basic U-Net architectures: 2D U-Net (1.1), a 3D U-Net (1.2) and a U-Net Cascade (1.3).

1.1 2D U-Net

Contrary to belief that using a 2D network in the context of a 3D network may be counter-intuitive, the paper found that 3D segmentation methods deteriorate in performance if the dataset is anisotropic [3] (when the property varies according to direction [1]).

1.2 3D U-Net

The architecture would typically thrive on contextual information, but we are limited by GPU memory, which allows us to train this architecture only on image patches. This may pose a problem for generating PTV areas because of the large structures, which, in size, may be compared to a Liver which has been cited as 'large images such as Liver, may impede training' [3].

1.3 U-Net Cascade

While the 2D and 3D U-Nets generate segmentations at full resolution, the cascade first generates low resolution segmentations and subsequently refines them.
This addresses the shortcoming in Section 1.2 because we first train the 3D U-Net on downsampled images (which compresses the amount of contextual information you need) (stage 1) and then upsamples to the original voxel spacing and passed as additional (one hot encoded) input channels to a second 3D U-Net, which is trained on patches at full resolution (stage 2).

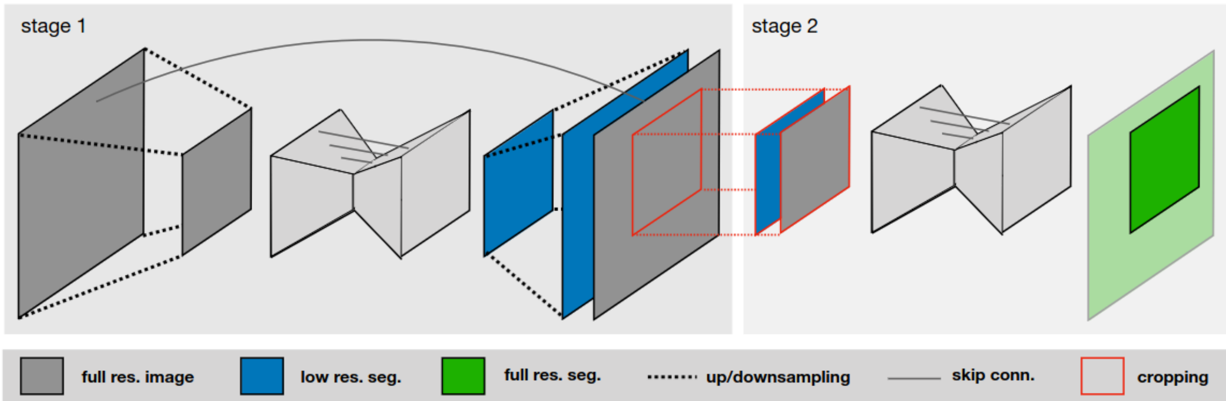

Fig. 1. U-Net Cascade (on applicable datasets only). Stage 1 (left): a 3D U-Net processes downsampled data, the resulting segmentation maps are upsampled to the original resolution. Stage 2 (right): these segmentations are concatenated as one-hot encodings to the full resolution data and refined by a second 3D U-Net.

2 Summary of [2]

This paper focuses more on the out-of-the-box aspect of training a neural network on a dataset. We have a three step recipe to follow to allow for no intervention from the user when it comes to defining custom parameters based on the structure the u-net is analysing:

2.1 Fixed Parameters

During development of nnU-Net we identified a robust configuration (that is, certain architecture and training properties) that can simply be used all the time. This includes, for example, nnU-Net's loss function, (most of the) data augmentation strategy and learning rate.

- Architectural template: closely follows U-Net, uses Leaky ReLU instead of ReLU.
- Training schedule: 1000 epochs with 250 mini-batches
- Inference: sliding window

2.2 Rule-based Parameters

Uses the dataset fingerprint to adapt certain segmentation pipeline properties by following hard-coded heuristic rules. For example, the network topology (pooling behavior and depth of the network architecture) are adapted to the patch size; the patch size, network topology and batch size are optimized jointly given some GPU memory constraint.

- intensity normalisation
- Resampling
- target spacing
- Adaptation of network topology, patch size and batch size.

- Initialization
- Architectural topology
- Adaptation to GPU memory budget.
- Batch size
- Configuration of the 3D U-Net cascade.

2.3 Empirical Parameters

trial-and-error. For example the selection of the best U-net configuration for the given dataset (2D, 3D full resolution, 3D low resolution, 3D cascade) and the optimization of the postprocessing strategy.

- Ensembling and selection of U-Net configuration(s)
- Post-processing

2.4 Data Fingerprint

- As a first processing step, nnU-Net crops the provided training cases to their non-zero region (improved computational efficiency)
- captures fingerprint like:
  - image size (before and after cropping)
  - image spacing (phyiscal size of voxels)
  - modalities (x-ray, ultrasound, ct, etc...)
  - as well as mean and std of all voxels

References

[1] URL: `https://sciencenotes.org/isotropic-vs-anisotropic-definition-and-examples/`.

[2] Fabian Isensee et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation". In: (2021). URL: `https://www.nature.com/articles/s41592-020-01008-z`.

[3] Fabian Isensee et al. "nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation". In: (2018). URL: `https://arxiv.org/pdf/1809.10486.pdf`.