Retina U-Net for Aneurysm Detection in MR Images

Michael Baumgartner, Paul F. Jaeger, Fabian Isensee, and Klaus H. Maier-Hein

Division of Medical Image Computing, German Cancer Research Center (DKFZ), Heidelberg, Germany

1 Methods

This section gives a short overview of the methods used for our submission to the ADAM 2020 challenge.

1.1 Retina U-Net

Based on the previous success of object detection models in a variety of difficult tasks in the medical [7] and natural image processing domain [4] we use Retina U-Net [2] as our base model. Instead of a Feature Pyramid Network [3] we employ a Path Aggregation Network [5] to directly enrich the detection features with the information from the semantic segmentation branch. Furthermore, additional convolutions are inserted after the element wise addition of the feature maps to improve the representational power of the network. The final architecture is depicted in Fig 1.

The network is trained for 600 epochs with 250 iterations per epoch. We use the AdamW [6] optimizer with an initial learning rate of 3×10^{-4} and a weight decay of 3×10^{-5} to update the weights of the network. The first 20 epochs are used to linearly ramp up the learning rate from 1×10^{-6} to 3×10^{-4} followed by a PolyLR [1] schedule.

The anchor sizes are determined by the Cartesian product of: $[6.0, 8.0, 10.0]$, $[5.0, 7.0, 10.0]$ and $[4.0, 6.0, 24.0]$ voxels. This results in 27 anchors per position.

1.2 Preprocessing

We use the provided pre-aligned and bias field corrected data for training and inference. To give the neural network a consistent field of view of the physical space, we resample all the data to $0.357 \times 0.357 \times 0.5$ mm and normalize each scan and modality to have zero mean and unit standard deviation. Due to limited GPU memory we use patches with $224 \times 224 \times 56$ voxels and a batch size of two during training and inference.

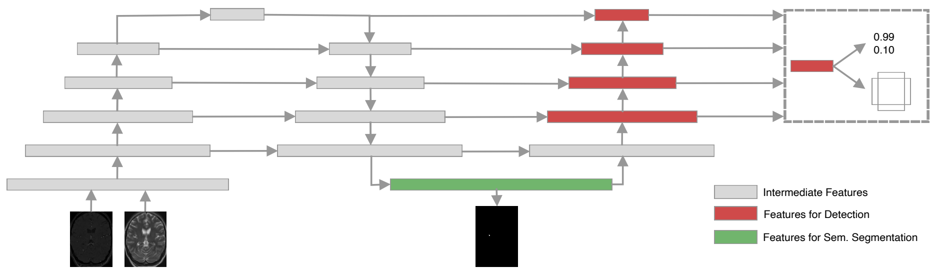

Fig. 1. Shows the U-Net like encoder with a Feature Pyramid Network. Following the Retina U-Net architecture, the decoder is extended to incorporate the additional semantic segmentation information (green). Finally, a Path Aggregation Network is used to generate the features which are used for the detection prediction (red).

1.3 Postprocessing

We use a 5 fold cross-validation during training and ensemble the final model of each fold with weighted box clustering [2]. The final predictions are obtained by computing the center point of the bounding boxes. Predictions with a probability lower than 0.5 are removed to balance the sensitivity and the number of false positive per scan.

References

1. Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. TPAMI (2016)
2. Jaeger, P.F., Kohl, S.A.A., Bickelhaupt, S., Isensee, F., Kuder, T.A., Schlemmer, H., Maier-Hein, K.H.: Retina u-net: Embarrassingly simple exploitation of segmentation supervision for medical object detection. Machine Learning for Health (ML4H) at NeurIPS 2019 (2018)
3. Lin, T., Dollár, P., Girshick, R.B., He, K., Hariharan, B., Belongie, S.J.: Feature pyramid networks for object detection (2016)
4. Lin, T., Maire, M., Belongie, S.J., Bourdev, L.D., Girshick, R.B., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft COCO: common objects in context (2014)
5. Liu, S., Qi, L., Qin, H., Shi, J., Jia, J.: Path aggregation network for instance segmentation (2018)
6. Loshchilov, I., Hutter, F.: Fixing weight decay regularization in adam (2017)
7. Setio, A.A.A., Traverso, A., de Bel, T., Berens, M.S., van den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M.E., Geurts, B., van der Gugten, R., Heng, P.A., Jansen, B., de Kaste, M.M., Kotov, V., Lin, J.Y.H., Manders, J.T., Sóñora-Mengana, A., García-Naranjo, J.C., Papavasileiou, E., Prokop, M., Saletta, M., Schaefer-Prokop, C.M., Scholten, E.T., Scholten, L., Snoeren, M.M., Torres, E.L., Vandemeulebroucke, J., Walasek, N., Zuidhof, G.C., van Ginneken, B., Jacobs, C.: Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The luna16 challenge. Medical Image Analysis **42**, 1 – 13 (2017)