Segmentation Architecture Ablation Experiment

We originally tested our method on using the U-Net architecture [1], which is a popular architecture choice in image segmentation tasks, particularly for medical imaging applications [2]. Nevertheless, our method can be extended to other architectures. Given an arbitrary segmentation network, our method requires replacing fixed resizing layers with variable resizing layers and using the hypernetwork to predict all the convolutional parameters of the network.

In this experiment we test our method on other popular alternative segmentation architectures. While the majority of segmentation architectures feature the same basic components (convolutional layers, resizing layers, skip connections), we aimed to cover other architectual components not included in the U-Net design such as pyramid spatial pooling layers and squeeze-excitation layers. The ablation includes the following architectures:

- **Residual UNet**. This is popular U-Net variant where the operations at each resolution feature a residual connection similar to the ResNet architecture [3]. Previous work has highlighted the importance of this type of connections in biomedical image segmentation [4].
- **FPN**. Feature Pyramid Networks [5] constuct a pyramid of features at several resolution levels, combining them in a fully convolutional manner and performing predictions at multiple resolutions during training.
- **PSPNet**. Pyramid Spatial Pooling [6] layers gather context information at multiple resolution levels in parallel, and several competitive natural image segmentation models include them in their architectures [7]. For our implementation, we perform the dynamic resizing operations in the encoder part and used the multi resolution PSP blocks in the decoder part.
- **SENet**. Squeeze Excitation networks [8] incorporate an attention mechanism to dynamically reweigh feature map channels during the forward pass. Squeeze-Excitation layers have been shown to be performant in segmentation tasks [9, 10, 11]. In our implementation, we include Squeeze-Excitation layers to both the encoder and the decoder stages of the network.

For each considered architecture we compare a single hypernetwork with a variable rescaling factor to a set of individual baselines with varying rescaling factors. We train baselines at 0.05 φ increments, i.e. $\varphi = 0, 0.05, 0.1, ..., 1$. For the hypernetwork model we sample from $p(\varphi) = \mathcal{U}(0,1)$. We evaluate on the OASIS2d semantic segmentation task introduced in the paper, which features 24 brain structure labels. We evaluate using Dice score and average over the labels. Results in Figure 1 show segmentation quality as the rescaling factor φ varies. We find similar results to those in the paper (Figure 3) with our single hypernetwork model matching the set of individually trained networks.

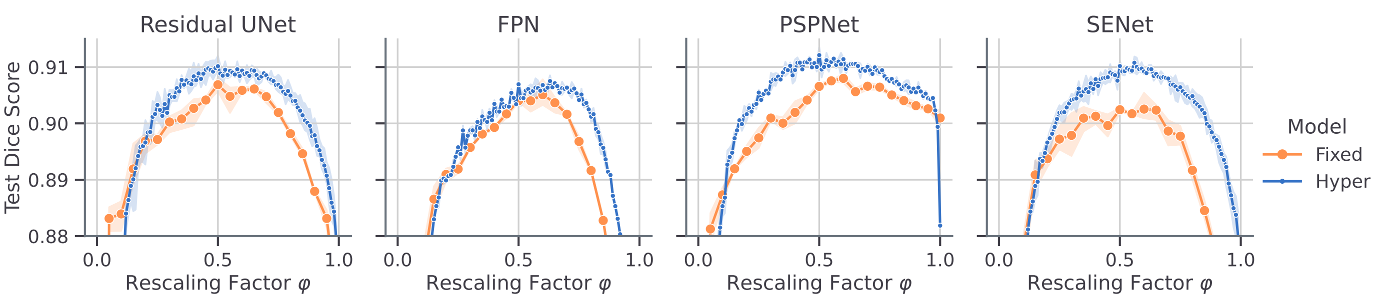

Figure 1: **Segmentation Architecture Ablation** Test Dice Score on the OASIS segmentation task for various choices of primary network architecture. Each plot features a family of networks trained with fixed amounts of feature rescaling (Fixed) and a single instance of the proposed hypernetwork method (Hyper) trained on the entire range of rescaling factors. Results are averaged across three random initializations, and the shaded regions indicate the standard deviation across them.

References

[1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

[2] Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2):203–211, 2021.

[3] Nahian Siddique, Paheding Sidike, Colin Elkin, and Vijay Devabhaktuni. U-net and its variants for medical image segmentation: theory and applications. *arXiv preprint arXiv:2011.01118*, 2020.

[4] Michal Drozdzal, Eugene Vorontsov, Gabriel Chartrand, Samuel Kadoury, and Chris Pal. The importance of skip connections in biomedical image segmentation. In *International Workshop on Deep Learning in Medical Image Analysis, International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*.

[5] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017.

[6] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2881–2890, 2017.

[7] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 801–818, 2018.

[8] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[9] Zilong Zhong, Zhong Qiu Lin, Rene Bidart, Xiaodan Hu, Ibrahim Ben Daya, Zhifeng Li, Wei-Shi Zheng, Jonathan Li, and Alexander Wong. Squeeze-and-attention networks for semantic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 13065–13074, 2020.

[10] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018.

[11] Leonardo Rundo, Changhee Han, Yudai Nagano, Jin Zhang, Ryuichiro Hataya, Carmelo Militello, Andrea Tangherloni, Marco S Nobile, Claudio Ferretti, Daniela Besozzi, et al. Use-net: Incorporating squeeze-and-excitation blocks into u-net for prostate zonal segmentation of multi-institutional mri datasets. *Neurocomputing*, 365:31–43, 2019.