Transfer learning for Domain Adaptation

John Huang, Alex Kumar, Junyao (Bill) Zheng

Duke PRATT SCHOOL of ENGINEERING

Introduction

Applying machine learning models to the medicinal domain introduces a new set of challenges, particularly the relatively small dataset size. In order to train a model from the ground-up the amount of data should be magnitudes larger than the number of trainable parameters. This is not the case for the domain of medical imagery where only hundreds of images are available for models with millions of parameters. A common solution to this issue has been transfer learning, the process of using a model trained on a large dataset as the starting point for a model trained on a smaller domain specific dataset.

The goal of this project is to perform semantic segmentation, or pixel-wise image classification, using the *MMSegmentation* library which contains pre-trained segmentation models. We tested out various pretrained models of different backbones and methods on the diabetic macular edema (DME) dataset. After exploring the dataset and the models, we attempted to fine-tune the candidate models in order to achieve optimal performance on the DME task. After the fine-tuning step, our team explored possible improvements in the experimental pipeline, including modifying the data preprocessing pipeline and loss function.

Contributions (all worked together on all parts, but with a focus on):

- All: Developing transfer learning pipeline in MMSegmentation library
- John: Evaluating base model performance on DME dataset
- Alex: Hyperparameter tuning on candidate models
- Bill: Investigating training procedures to boost performance with limited data medical applications

Motivation

Semantic segmentation is done by fully convolutional network (FCN) model architectures. Unlike simpler image classification tasks, good performance is hard to achieve without enormous amounts of training data. With limited data taks, this becomes very challenging and introduces the need for transfer learning. Transfer learning is the process of re-utilizing weights from a model trained on a much larger and general image dataset to initialize an application-specific model for faster convergence. This is needed in fields such as medical imaging since the datasets are usually very small. There are many different architecture variations on convolutional networks, like U-net, Mask R-CNN, Semantic FPN, and more. There are also the usual data transformations, such as horizontal flipping and adversarial patches, to augment the dataset and combat overfitting.

Method

Transfer Learning with MMSegmentation and Diabetic Macular Edema Dataset

MMSegmentation is an open source semantic segmentation library with extensive resources on pre-trained models with modular designs and high efficiency. The initial step involved selecting pre-trained models based on the following criteria: dataset used to pre-train the model, and performance on the corresponding datasets.

UNet is the only backbone architecture among MMSegmentation's library of pre-trained models that was trained on medical imaging datasets. UNet was trained on four different medical imaging datasets (CHASE, STARE, HRF, and DRIVE) and had 3 different hybrid variations: Fully Convolutional Networks (FCN), Pyramid Scene Parsing Network (PSPNet), and DeepLabV3. This resulted in 12 different baseline models to work with.

Domain adaptation is then performed on the variety of UNet models to evaluate their performance on a smaller dataset comprising optical coherence tomography (OCT) images of diabetic macular edema and the corresponding segmentation of fluids. The dataset is obtained from a group of ten patients, containing 110 images of dimension 768 x 496.

Next, strategies to enhance the performance of the models are explored. These include data augmentation, fine-tuning hyperparameters, and experimenting with different loss functions.

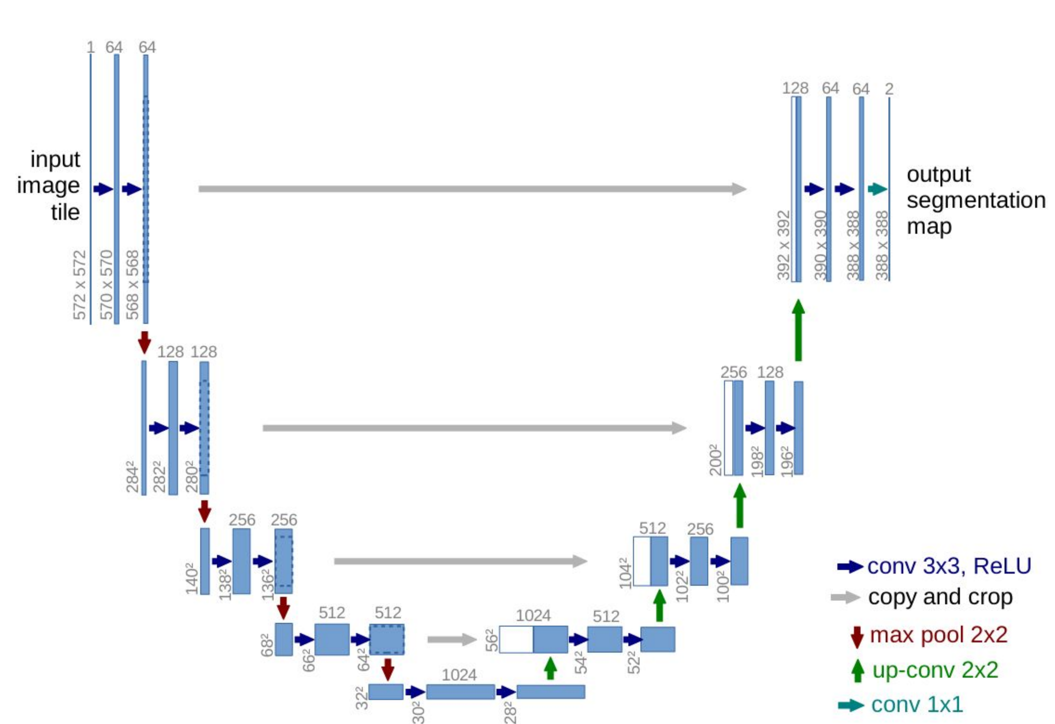

Fig.1: UNet Architecture

Results

Effect of Batch Size on Performance

Our investigation into the impact of batch size on model performance revealed a linear relationship. However, this trend ceased to improve beyond a batch size of 8 due to hardware constraints. Possible explanations include: noise reduction, where larger batches likely contributed to noise reduction in gradient estimates, fostering smoother convergence during training, and greater generalizability, where the linear improvement suggests that larger batches enhanced the model's generalizability, providing a more accurate estimation of the gradient.

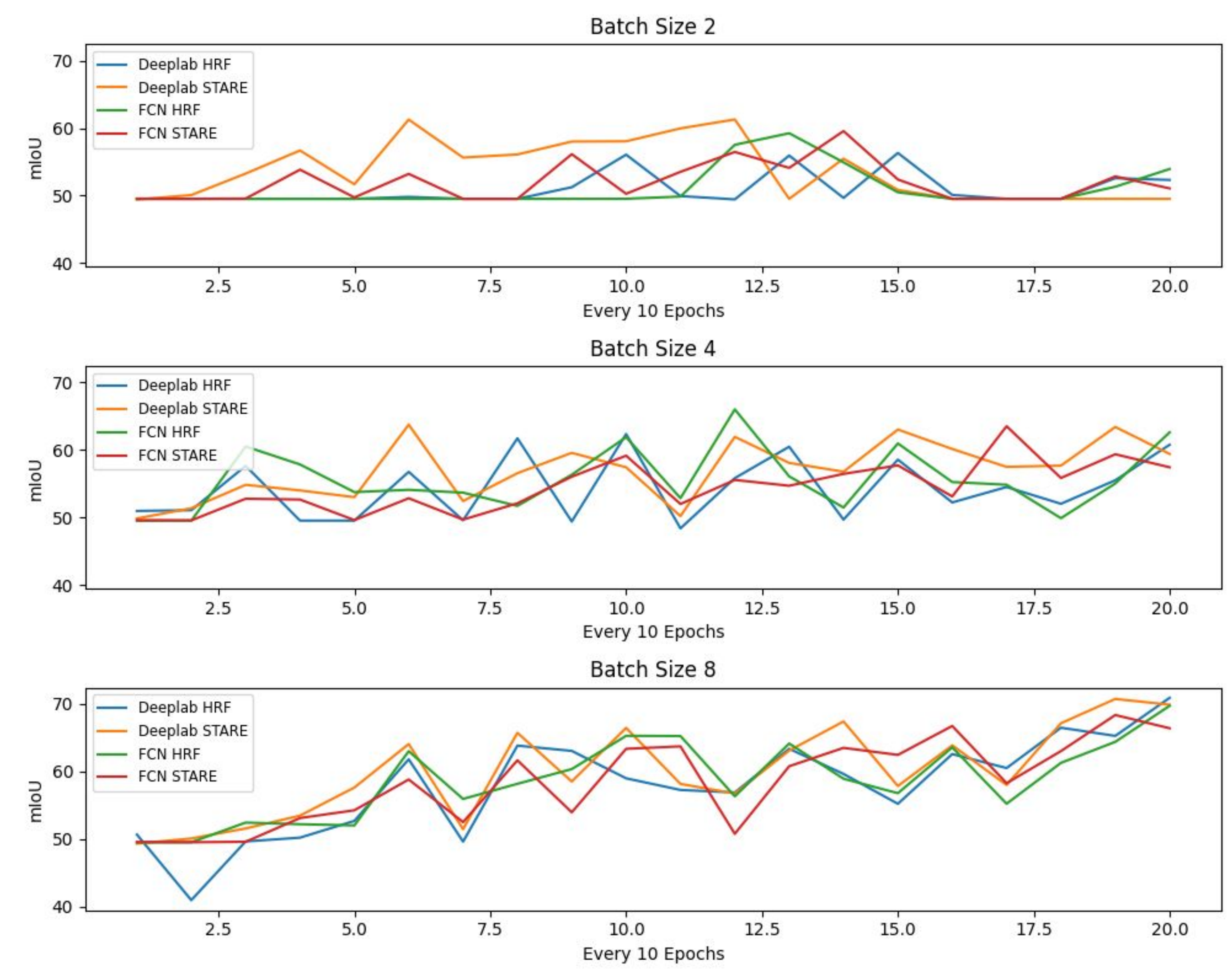

Fig. 2: mIoU of Models Trained with Different Batch Size

Impact of Data Augmentation

Our experiments with data augmentation demonstrated a notable increase in performance. We explain that it helped address overfitting. Given the limited size of the DME dataset, data augmentation played a crucial role in combating overfitting by generating more diverse training samples.

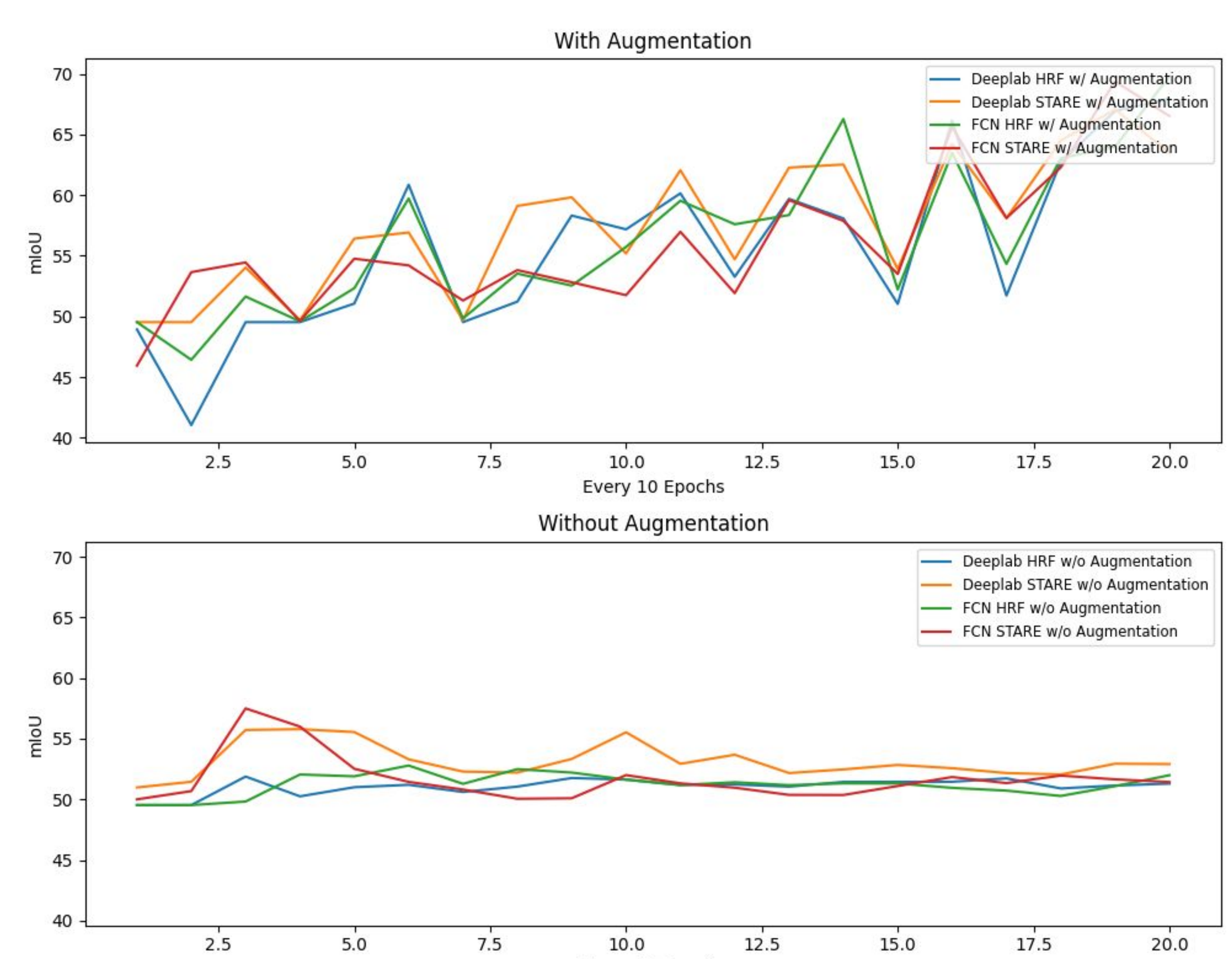

Fig.3: mIoU of Models Trained With and Without Data Augmentation

Final Models

Using the insights gained from the conducted experiments, we trained various pre-trained models with data augmentation, larger batch-sizes, and differing losses. However, we observed a notable random variance in the performance outcomes of the models, even when trained under identical parameters. Overall, the models that performed well also performed similarly and is shown by the resulting predicted masks in Fig.3 and Table 1 results.

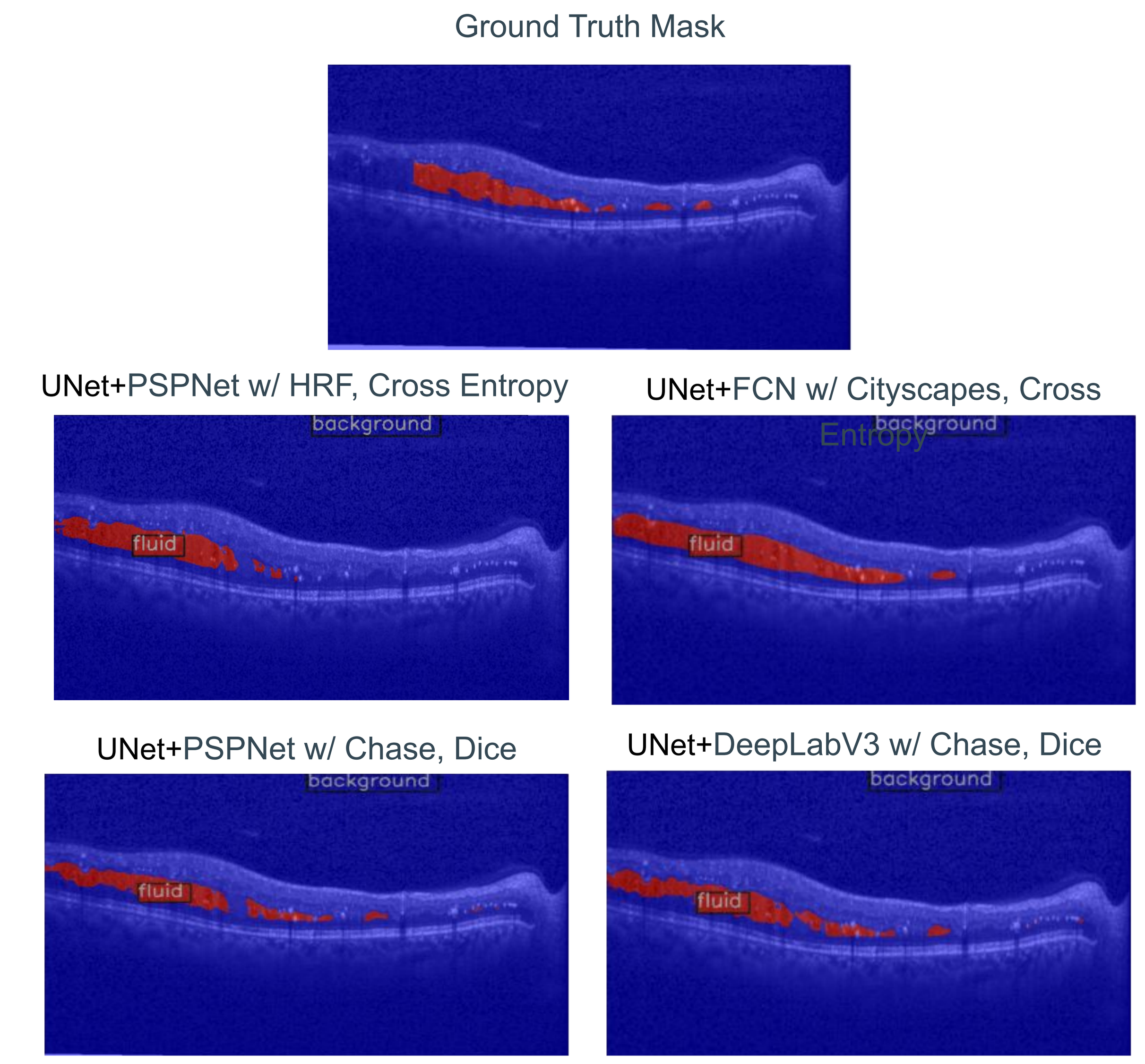

Fig.4: Samples of Ground Truth and Predictions Made by Various Models

Models	IoU	Accuracy	Dice
UNet+PSPNet w/ HRF, Cross Entropy Loss	44.07	64.99	61.18
UNet+FCN w/ Cityscapes, Cross Entropy	43.67	70.9	60.79
UNet+DeepLabV3 w/ Chase, Dice	43.97	62.89	61.08
UNet+PSPNet w/ Chase, Dice	45.29	65.25	62.34

Table 1: Evaluation of Models with Different Metrics

Conclusion

Through this project, we have successfully demonstrated the capability of adapting a image segmentation model pre-trained on larger datasets towards a more specific task through transfer learning. Moreover, we further explored the potential of the models by fine-tuning the hyperparameters and tested different loss functions, achieving an IoU of 45.29% and a Dice score of 62.34% on the DME dataset. These metrics highlights the model's enhanced performance in segmenting medical images, demonstrating the effectiveness of our methodology in optimizing model performance for specialized tasks such as medical imaging.

Future Works

As demonstrated by our results, many pretrained base models are good candidates for finetuning on the DME task using transfer learning. Surprisingly, models pre-trained on Cityscapes performed very well. This marks an area of future exploration, emphasizing the importance of pretraining on a massive dataset. We propose exploring ensemble methods, theoretically capturing valuable pretrained features across the various candidate base models. We are also identifying a need to analyze which type and level of features are maintained after the finetuning process. This would provide greater insight into the transfer learning process, recognizing which features are valuable for different downstream tasks.

References

Contributors, Mmsegmentation. *MMSegmentation: OpenMMLab Semantic Segmentation Toolbox and Benchmark*. 2020, https://github.com/open-mmlab/mmsegmentation.

Ronneberger, Olaf, et al. 'U-Net: Convolutional Networks for Biomedical Image Segmentation'. *arXiv [Cs.CV]*, 2015, http://arxiv.org/abs/1505.04597. arXiv.

Zhao, Hengshuang, et al. 'Pyramid Scene Parsing Network'. *arXiv [Cs.CV]*, 2017, http://arxiv.org/abs/1612.01105. arXiv.

Chen, Liang-Chieh, et al. 'Rethinking Atrous Convolution for Semantic Image Segmentation'. *arXiv [Cs.CV]*, 2017, http://arxiv.org/abs/1706.05587. arXiv.

Long, Jonathan, et al. 'Fully Convolutional Networks for Semantic Segmentation'. *arXiv [Cs.CV]*, 2015, http://arxiv.org/abs/1411.4038. arXiv.

Chiu, Stephanie J., et al. 'Kernel Regression Based Segmentation of Optical Coherence Tomography Images with Diabetic Macular Edema'. *Biomed. Opt. Express*, vol. 6, no. 4, Optica Publishing Group, Apr. 2015, pp. 1172–1194, https://doi.org10.1364/BOE.6.001172.