Novel Convolutional Embeddings for Domain Adaptation in Medical Segmentation

SURF
Summer Undergraduate Research Fellowship
SURF Project Code: 2022053

Supervisor: Erick Purwanto
Authors: Mingzirui Wu, Sitang Gong, Dongheng Lin, Qiaochu Zhao, Suqi Zhang, Doyeon Kim

Abstract

Medical Artificial Intelligence recently focuses on **annotating organs** on cell image dataset to perform clinical diagnoses and to suggest suitable treatments. While there are some progress on **single-organ dataset** [1], demands from Medical Community for a more accurate AI model capable of segmenting **multi-organ dataset** arise. Our team attempted new methods of data processing and novel models to segment Functional Tissue Units (FTU). We introduced the **Convolutional Embedding** algorithm and obtained an accuracy of 0.88 and similarity coefficient of 0.81 in **HuBMAP + HPA - Hacking the Human Body 2022 Kaggle Competition** [2]. We have **seized a GOLD prize**, and achieved **the 5th rank out of 920 competitors**.

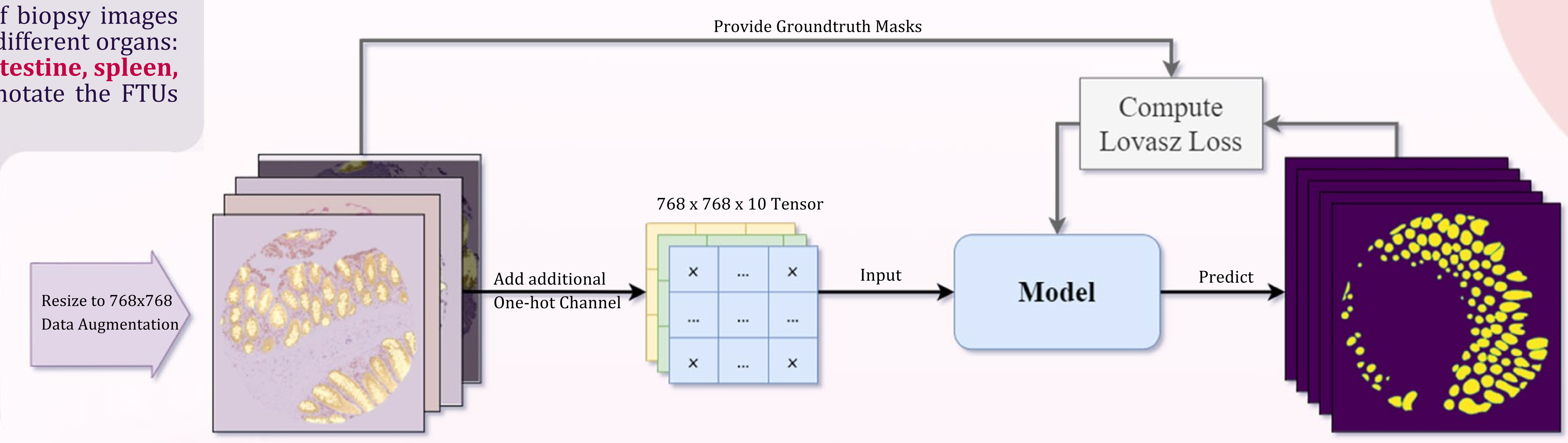
Background & Method

Introduction

Given a dataset consists of biopsy images and their masks from five different organs: **kidney, prostate, large intestine, spleen, and lung**, we want to annotate the FTUs within them.

General Process

Original Images and Masks

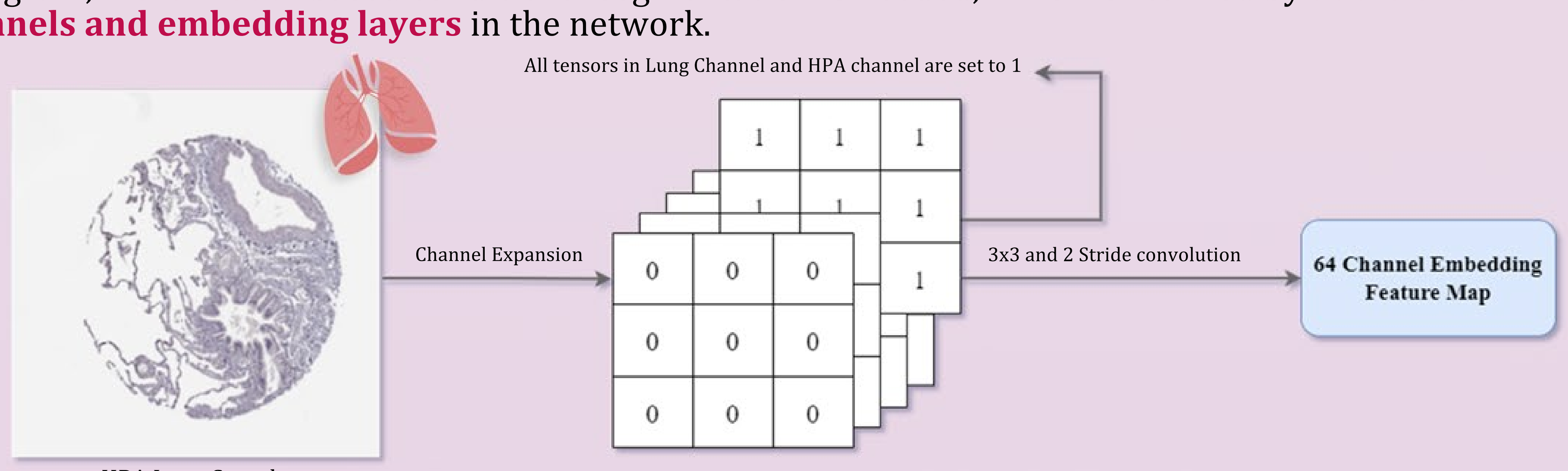


In order to cope with dataset diversity and organ type inconsistency:

- Firstly, we have applied complicated **Data Augmentation and Color Stain Normalization** combinations to cover all the features that may appear in the diverse testing set.
- Secondly, we implemented **Multimodal Mechanism** by adding one-hot organ channels to each sample which will then be converted to **Organ Embeddings** by the model.

Convolutional Embedding Representation

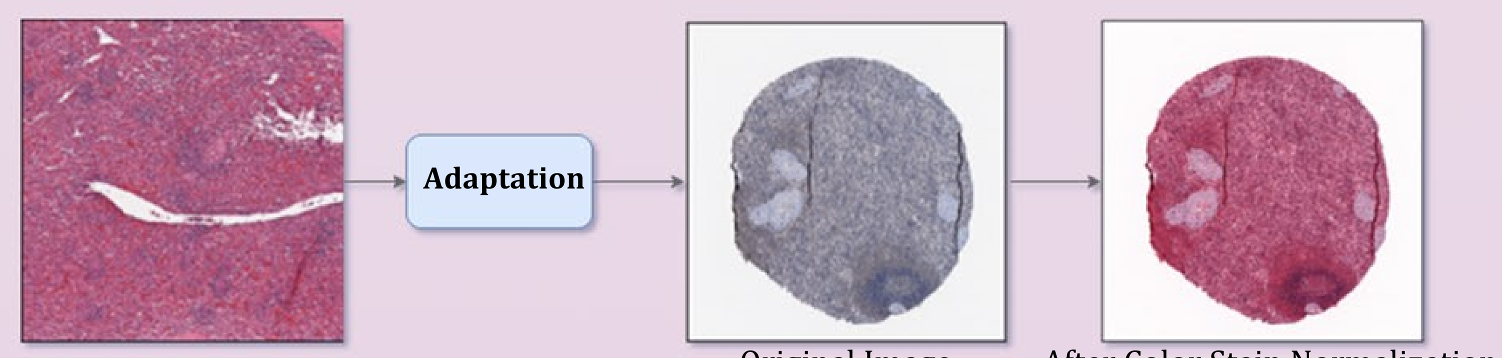
Given the samples are of different types, which shows great **inter-class diversity** as illustrated in the histogram, **multimodal** based conditional segmentation is needed, which is realized by **additional channels and embedding layers** in the network.



HPA Lung Sample

Color Stain Normalization Preprocessing

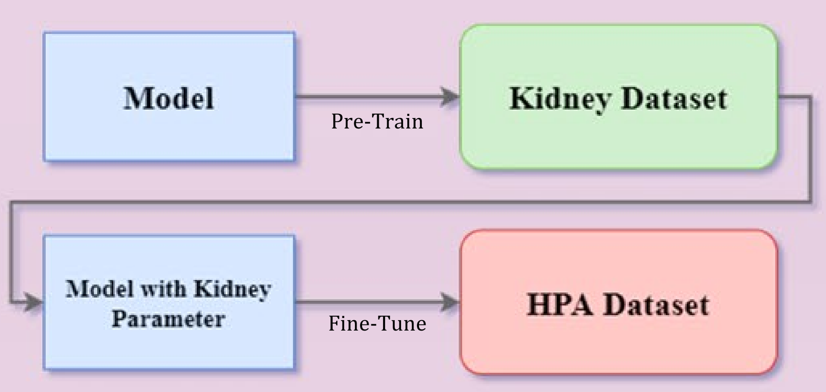
There are some differences between training images and testing images in terms of **color stain**. Thus, **Macenko** algorithm is used to produce **color normalized** training images and then combine them with the original training set.

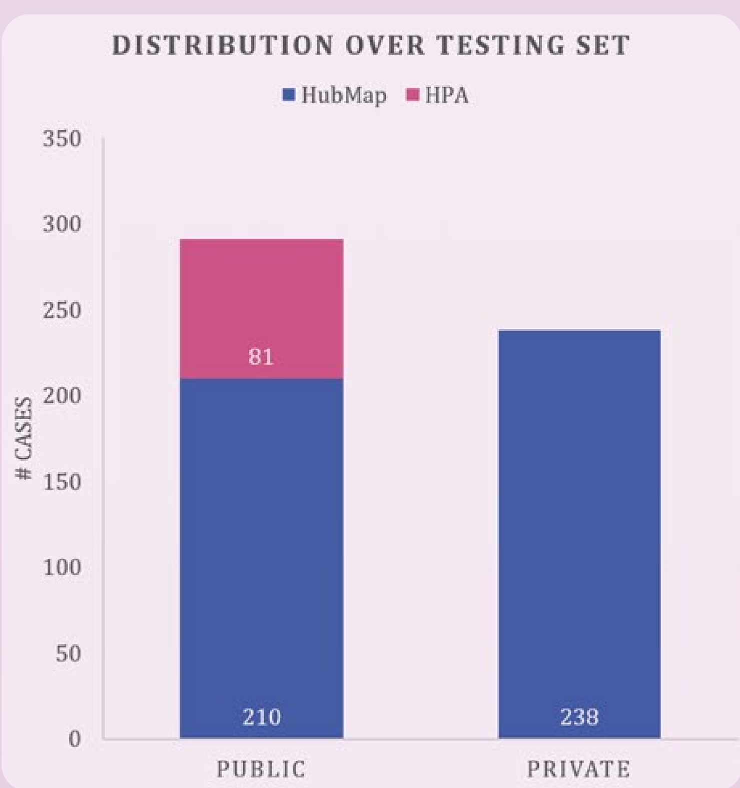


Test Image   Original Image   After Color Stain Normalization

HubMap Transfer Learning

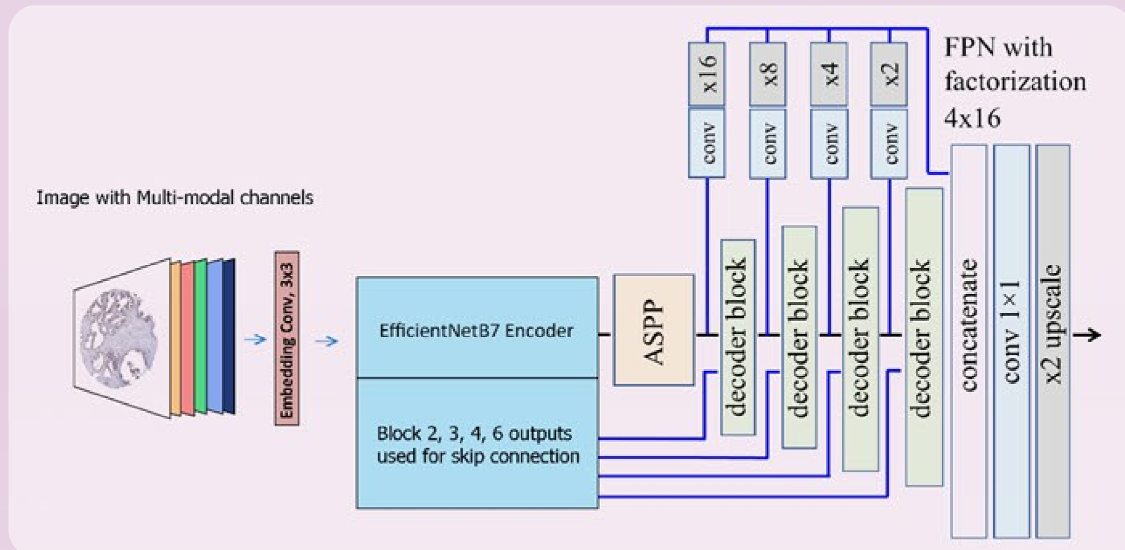
With consideration to the difference between our training set, **HPA dataset**, and the testing dataset contains many **HubMap samples**. The research team **pretrained** the model on a previous competition, and then **fine-tune** the model on current competition.





Model Architecture

The novel network is based on the **EffUnet** [3], in which the **EfficientNet encoder** creates a representation of features at different levels [4], while the **decoder** combines the features and generates a prediction as a segmentation mask.



Dice Coefficient

This competition is evaluated on the **mean Dice coefficient**, which can be used to compare the pixel-wise agreement between a predicted segmentation and its corresponding ground truth.

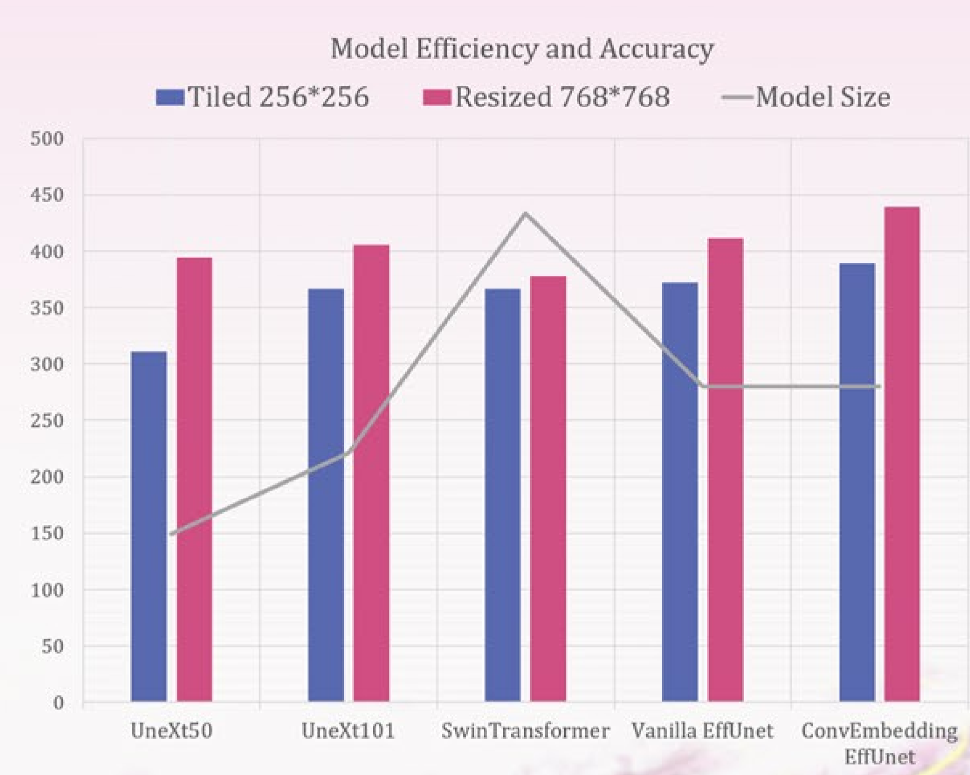
$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

Lovász Loss

Lovász loss is a **differentiable surrogate** of Dice coefficient and can be used for optimizer [5]. The loss formula is given by:

$$\text{loss}(\boldsymbol{f}) = \frac{1}{|\mathcal{C}|} \sum_{c \in \mathcal{C}} \overline{\Delta_{J_c}}(\boldsymbol{m}(c))$$

Results



Discussion

Applying Transformers to Computer Vision is a prevailing research area. Many researchers claims that **Vision Transformer** is always a superior to **Convolutional Networks** in terms of **Domain Adaptation**, it lacks embedding information capabilities to overcome data diversity.

However, our team has tested several networks in the competition, and as a result, our novel **Convolutional Embedding** method is not only outperforming in terms of **accuracy**, but also more **efficient** than those **Transformer counterparts.** The model achieves a Dice score of 0.88 on local validation and 0.81 in testing set which is **11% higher** than SwinTransformer's score, while the size of the model is **as small as 60%** of the SwinTransformer [7] solution.

Conclusion

Our novel **Convolutional Embedding strategy** successfully enhanced the final accuracy result to **81%**, showing that embeddings of **organ metadata** helps in **improving performance** of a simpler convolutional model in Medical Segmentation.

Achievement:

Gold Prize

5 SURF2022_SC440 +2 0.81

We have reached **5th rank out of 920 teams** with a score very close to the top placeholder's score **0.82**.

Reference

[1] Shubham, S. , Jain, N. , Gupta, V. , Mohan, S. , Ariffin, M. M. , & Ahmadian, A. . (2021). Identify glomeruli in human kidney tissue images using a deep learning approach. Soft Computing, 1-12.
[2] HuBMAP + HPA - Hacking the Human Body, Segment multi-organ functional tissue units, https://www.kaggle.com/competitions/hubmap-organ-segmentation
[3] Tan, M. , & Le, Q. V. . (2019). Efficientnet: rethinking model scaling for convolutional neural networks.
[4] Baheti, B. , Innani, S. , Gajre, S. , & Talbar, S. . (2020). Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE
[5] Berman, M. , Triki, A. R. , & Blaschko, M. B. . (2018). The Lovasz-Softmax Loss: A Tractable Surrogate for the Optimization of the Intersection-Over-Union Measure in Neural Networks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE.
[6] Liu, Z. , Lin, Y. , Cao, Y. , Hu, H. , Wei, Y. , & Zhang, Z. , et al. (2021). Swin transformer: hierarchical vision transformer using shifted windows.

With great thanks to: Juntuo Wang, Botao Jiang, Jin Bei, Yufeng Gao, Anchi Zhang.