Histopathology-Specific Foundation Model for Tissue Segmentation and Clustering

Si-Jin Kim, Kazi Rakib Hasan, Yu Ando, Seokhwan Ko, Junghwan Cho*, Nora Jee-Young Park and Hyungsoo Han

Department of Biomedical Science, Kyungpook National University, Daegu 41994, Korea

***Contact e-mail: cho.junghwan@gmail.com**

Introduction

The rise of generative AI models has garnered significant attention, offering flexible interpretation of data across various domains. In particular, large language models(LLMs), such as chatGPT, have made notable contributions to information retrieval in diverse fields. However, the development and management of these models present significant challenges due to the vast volume and diversity of data, necessitating substantial data resources and computational power.

In healthcare, obtaining access to large and diverse medical datasets remains a complex endeavor. Additionally, the utilization of diverse domain information can introduce homogenization effects in downstream tasks. Although homogenization has its advantages, it can also perpetuate underlying model flaws and hinder adaptation to specific tasks. To mitigate the issue, we present building a histopathology foundation model(hPFM) that extensively utilizes domain-specific data for solving domain-specific problems.

Methods

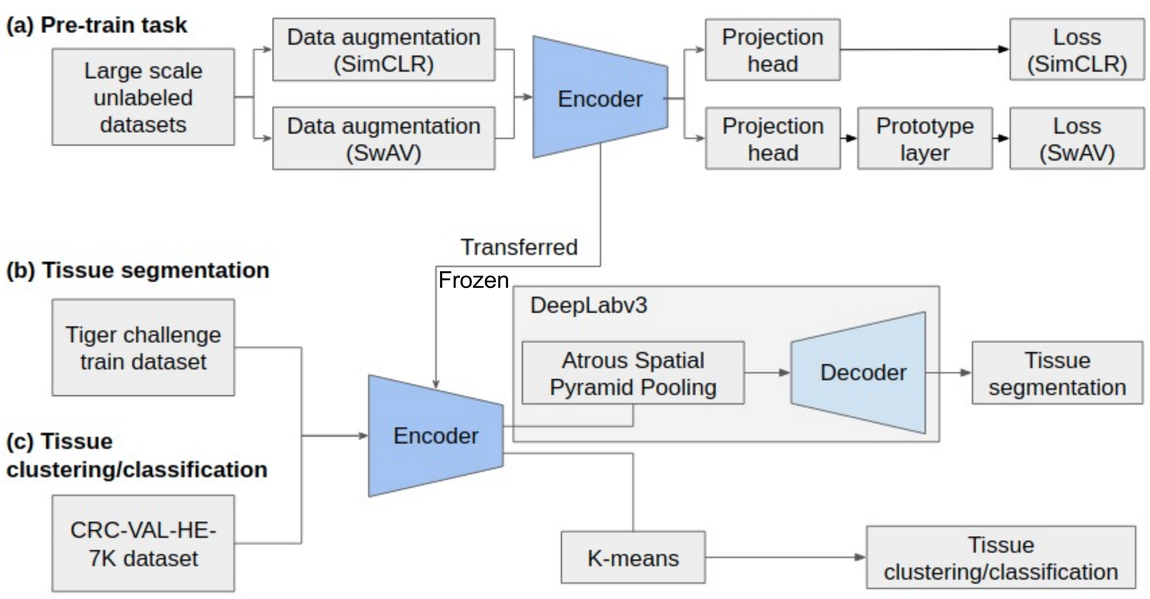

Figure 1. Overall pipeline for building histopathology-specific foundation models and downstream tasks

(a) Pre-train task : This task is to train self-supervised learning models such as SwAV and SimCLR to learn common histopathological representations/features on unlabeled datasets. the weight of the pre-trained self-supervised learning model is transferred to the encoder of the tissue segmentation and clustering model before learning is performed.

(b) Tissue segmentation : DeepLabv3 as a tissue segmentation model uses the training dataset of Tiger Challenges organized for predicting Tumor-Infiltrating Lymphocytes(TILs) within breast cancer tissues. The dataset consists of three different tissue area classes (tumor, stroma, and others).

(c) Tissue clustering/classification : we utilized a K-means method to cluster the nine tissue types (such as adipose, background, debris, lymphocytes, mucus, smooth muscle etc.) from CRC-VAL-HE-7K dataset.

Table 1. A large scale of unlabeled histopathology datasets (patch type) for pre-training

Name of pretrained Datasets	Organs	Image size
andrewjanowczyk epi	Breast	1000 x 1000
andrewjanowczyk nuclei	Breast	2000 x 2000
andrewjanowczyk tubule	Colon	775 x 522
andrewjanowczyk mitosis	Breast	2000 x 2000
CoNSeP	Colon	1000 x 1000
MiMM	Bone	2560 x 1920
monuseg_2018_train_data	Various	1000 x 1000
BACH	Breast	2048 x 1536
NCT-CRC-HE-100K	Colon	224 x 224
Gleason19	Prostate	5120 x 5120
Lymph	Various	1388 x 1040
BreakHis	Breast	700 x 460
LC25000	Lung	768 x 768
breastpathQ	Breast	512 x 512
Kather_texture_2016_image_tiles_5000	Colon	150 x 150

- **Type of organ used for pretrain task** : Breast, Colon, Bone, Prostate, Lung
- **Total of image samples** : about 150,000
- **Total of cropped image samples** : about 600,000

Table 2. Experimental setup for pre-train task

	SwAV	SimCLR
Architecture	ResNet(18, 34, 50)	ResNet(18, 34)
Optimizer	Adam	Adam
Epoch	300	300
Batch size	2048	1024
Cropped image size	256 x 256	256 x 256
Number of cropped samples	about 600,000	about 600,000
Train samples	80% of total dataset	80% of total dataset
Validation samples	20% of total dataset	20% of total dataset
Number of prototypes	28	None
Applying augmentation	- Random Resized Crop - Random Horizontal Flip - Random Grayscale - Gaussian Blur - Color Jitter - Multi crop	- Random Resized Crop - Random Horizontal Flip - Random Grayscale - Gaussian Blur - Color Jitter

- **Random Horizontal Flip** : (p = 0.5)
- **Random Grayscale** : (p = 0.2)
- **Color Jitter** : (p = 0.8), saturation(0.4), contrast(0.4), brightness(0.4), hue(0.2)
- **Multi crop** : number of crops : [2, 4], Min scale crop : [0.33, 0.10], Max scale crop : [1, 0.33]

Results

Comparing the performance of a pre-trained model specific to histopathology with that of models pre-trained on ImageNet and randomly initialized models

- **ImageNet pre-trained model** : Model pretrained with ImageNet data provided by Torchvision
- **Random initialized model** : Model with random weight initialization

Table 3. Downstrem results on the task of the tissue segmentation and clustering/classification

Task	Metric	ImageNet pretrained ResNet			Random Initialized ResNet			Self-supervised Learning SimCLR ResNet		SwAV ResNet		
		18	34	50	18	34	50	18	34	18	34	50
Segmentation	Dice	0.80	0.79	0.81	0.65	0.60	0.62	0.81	0.80	0.83	0.83	**0.85**
Clustering	NMI	0.70	0.76	0.78	0.45	0.31	0.41	**0.88**	0.82	0.75	0.69	0.74
	ARI	0.62	0.71	0.73	0.39	0.22	0.28	**0.82**	0.76	0.68	0.58	0.63
	ACC	0.74	0.79	0.81	0.57	0.45	0.52	**0.87**	0.82	0.78	0.68	0.76

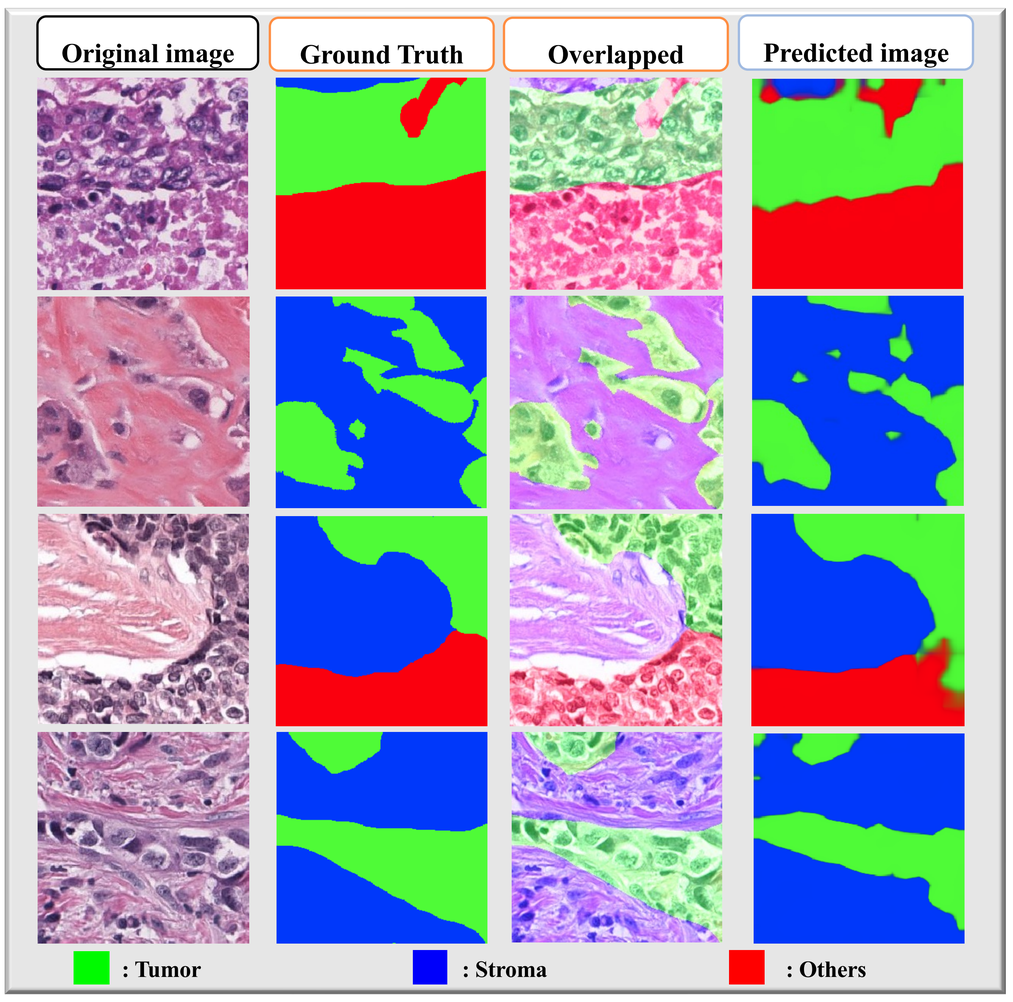

Figure 2. Comparison of predicted segmentation map to ground truth using a SwAV pre-trained model (ResNet-50)

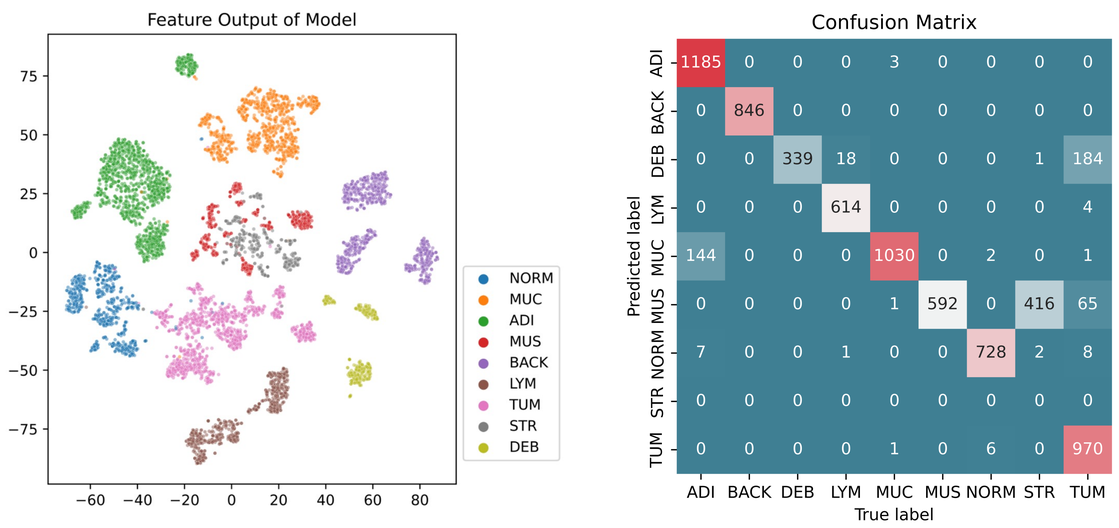

Figure 3. Tissue clustering results obtained from the SimCLR pre-trained model (ResNet-18): on the left is a t-SNE plot, and on the right is a confusion matrix

- **Tissue segmentation :** The SwAV pre-trained model shows higher dice scores (0.85) on all ResNet architectures than the other models.
- **Tissue clustering :** The SimCLR pre-trained model(ResNet-18) shows better clustering performance results (NMI : 0.88, ARI : 0.82, ACC : 0.87) on all ResNet architectures than the other models.

Conclusions

- In this work, a histopathology-specific foundation model was constructed using SSL learning schemes with a large scale of unlabeled image.
- Utilizing the pre-trained foundation model, we conducted two downstream tasks: tissue segmentation and clustering.
- Our experimental results demonstrate superior performance on both tasks compared to using an ImageNet pre-trained model.
- Our findings highlight the effectiveness of our pre-trained model in histopathology applications, showcasing its potential as a valuable tool in this domain.

Acknowledgement

This research was supported by a grant from the Brain Pool Program through the National Research Foundation of Korea funded by the Ministry of Science and ICT (No.2022H1D3A2A01096490) and the Ministry of Education (2021R1I1A3056903).