Path-CT Registration with Self-Supervised Vision Transformer

Anshuk Gottipati, UNC Charlotte
M.S. Yaying Shi, Ph.D. Yonghong Yan, College of Computing and Informatics

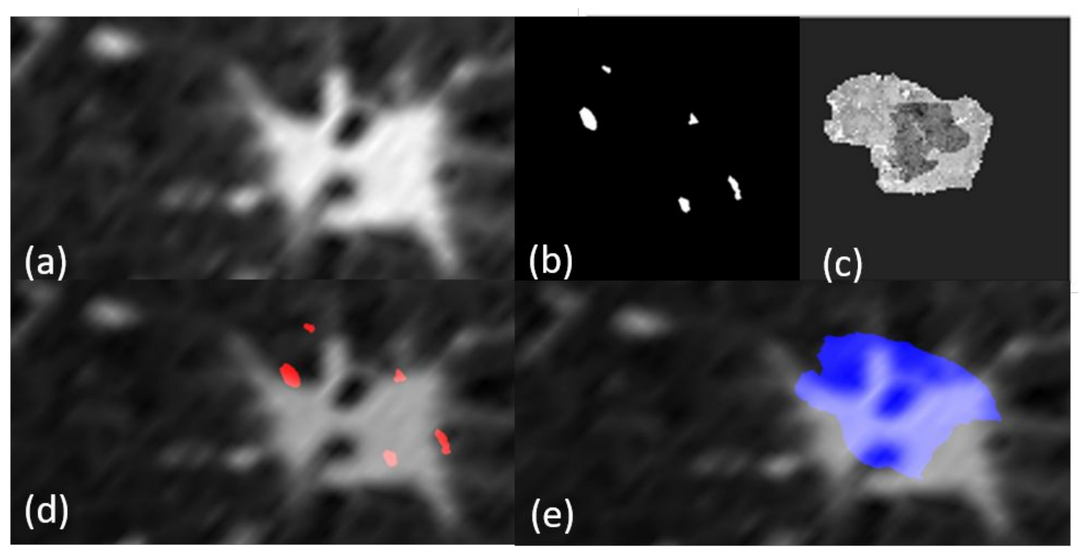
UNIVERSITY OF NORTH CAROLINA CHARLOTTE

Introduction

Medical Image Processing is used to:

- Combine different imaging modalities
- Diagnose Diseases
- Provide detailed images

We train two self-supervised learning feature extractors for CT and Pathology images.

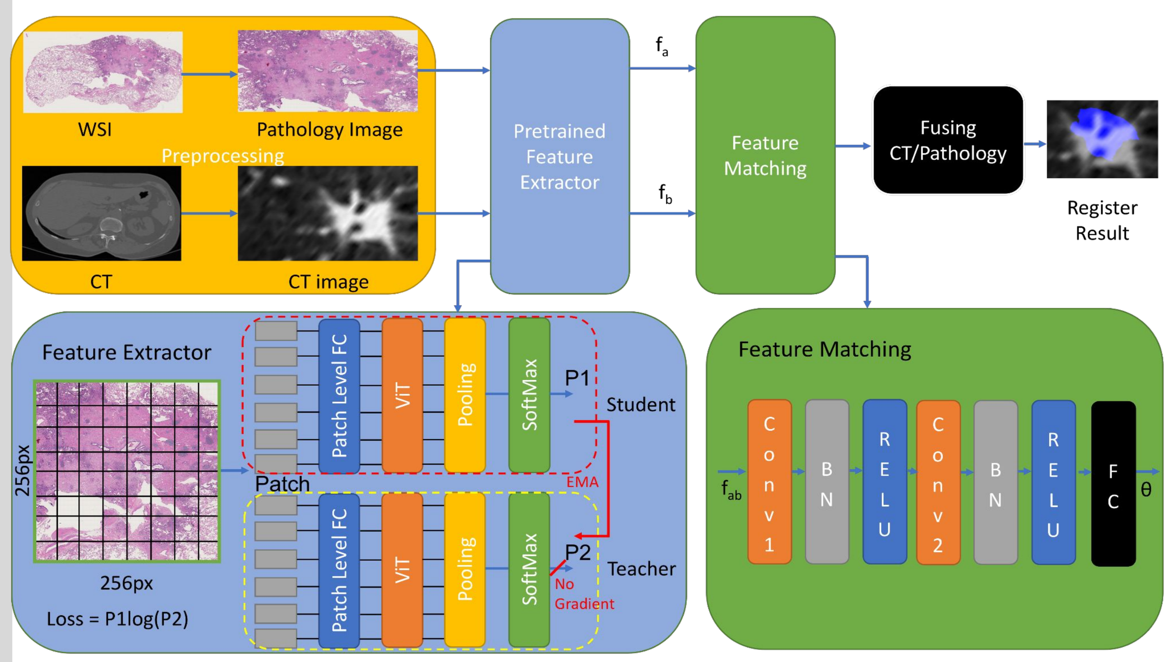


Objectives

- Registering CT images to Pathology images should produce more information.
- Manual registration methods rely on use capability and is time consuming.
- Its is important to seek an automated registration method.
- Conventional methods like PI-RADS, determine the transformation field through intense computation, which is less efficient.
- Hope to have a new direction in medical image registration using self-supervised vision transformers

Method

- Pre-processing involved extending pathology images to a 3D feature representation and matching them with the 3D CT slides.
- Use A self-supervised feature extractor, for both CT images and for whole-slide pathology images
- A feature matching sub-network that aligns and maps the distinctive features extracted from both image types.
- Post-processing for fusing pathology and CT patches based on the correlation maps produced from the previous step.



Results

Quantitative

- Evaluated using Euclidean distance and Dice Score.
- Compared with the original work the Dice score was improved from 65.9% to 72.6%

Case ID	Euclidean Distance (mm)↓		Dice ↑	
	Original	Our	Original	Our
LungFCP-01-0001	1.75	1.782	0.731	0.692
LungFCP-01-0002	2.15	1.764	0.624	0.739
LungFCP-01-0003	2.02	1.811	0.595	0.687
LungFCP-01-0004	1.81	1.447	0.689	0.764
LungFCP-01-0005	1.42	1.671	0.692	0.745
LungFCP-01-0006	2.67	1.915	0.624	0.728

Quantitative result of registering pathology images with CT image on all the six cases

Ablation Study

- If the feature extractor matches the same results as the ground truth we consider it correct.
- Categorized classes into background, lesion and blood vessel and the accuracy of the feature extractor is in the below table.

Dataset	Epochs	Accuracy
TCIA-PROSTATE	100	0.88
TCIA-PROSTATE	800	0.89
TCGA-PRAD	100	0.95
TCGA-various	100	0.96

Accuracy of pre-trained feature extractors on different dataset

Pretraining using DINO framework as the feature extractor.

Conclusions

The performance of our framework is notable achieving 72.9% Dice Score and 1.73 Euclidean distance therefore improving on the Matlab implementation.

Additionally, out framework offers a novel approach to fuse high-resolution images ot low-resolution counterparts, but I am knowledgeable about this.

References

Yaying Shi, Christian Micklisch, Erum Mushtaq, Salman Avestimehr, Yonhhong Yan, and Xiaodong Zhang, "An ensemble approach to automatic brain tumor segmentation," in *International MICCAI Brainle-sion Workshop*. Springer, 2022, pp. 138-148