CVPR JUNE 17-21, 2024 SEATTLE, WA

Correlation-aware Coarse-to-fine MLPs for Deformable Medical Image Registration

Mingyuan Meng[1, 2], Dagan Feng[1], Lei Bi[2], and Jinman Kim[1]

[1] School of Computer Science, The University of Sydney, Sydney, Australia

[2] Institute of Translational Medicine, Shanghai Jiao Tong University, Shanghai, China

THE UNIVERSITY OF SYDNEY

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

Introduction

Background

Deformable image registration is a fundamental step for medical image analysis. Recently, transformers have been used for registration and outperformed Convolutional Neural Networks (CNNs). Transformers can capture long-range dependence among image features, which have been shown beneficial.

Existing limitation

Due to the high computation/memory loads of self-attention operations, transformers are typically used at downsampled feature resolutions and cannot capture fine-grained long-range dependence at the full image resolution. This limits deformable registration as it necessitates precise dense correspondence between each image pixel. Multi-layer Perceptrons (MLPs) are efficient in computation/memory usage, enabling the feasibility of capturing fine-grained long-range dependence at full resolution. Nevertheless, MLPs have not been extensively explored for image registration and are lacking the consideration of inductive bias crucial for medical registration tasks.

Contribution

We propose the first correlation-aware MLP-based registration network (CorrMLP), which introduces a correlation-aware multi-window MLP block in a novel coarse-to-fine registration architecture.

Method

CorrMLP

- The CNN-based encoder extracts two four-level hierarchical feature pyramids from the fixed and moving images separately.
- The MLP-based decoder leverages the extracted feature pyramids to perform four steps of coarse-to-fine registration using the proposed CMW-MLP blocks.
- The correlation information between images (image-level) and between registration steps (step-level) is leveraged as supplementary information to guide registration.

Correlation-aware Multi-window MLP (CWM-MLP) block

- Each CWN-MLP block takes two sets of feature maps as input and then explores the potential correspondence between them, which is purposely optimized to capture correlation-aware multi-range dependence for deformable medical image registration.
- A 3D correlation layer to calculate the local correlations, followed by a multi-window MLP module to capture correlation-aware multi-range dependence.
- Multi-range dependence are captured via multiple local window-based MLP branches, enabling the block to handle both large and small local deformations.

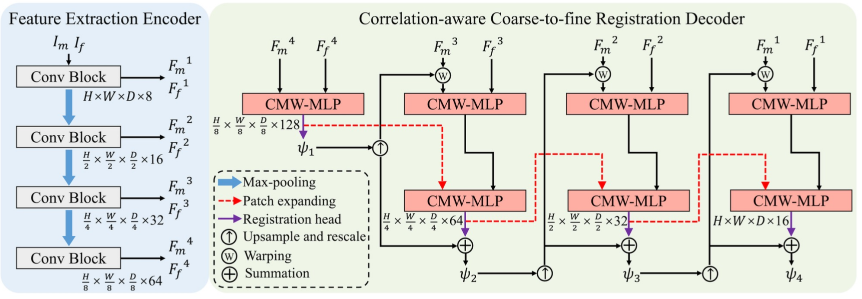

Figure 1: The overall architecture of our CorrMLP. It consists of a CNN-based hierarchical feature extraction encoder and a correlation-aware coarse-to-fine registration decoder based on CMW-MLP blocks.

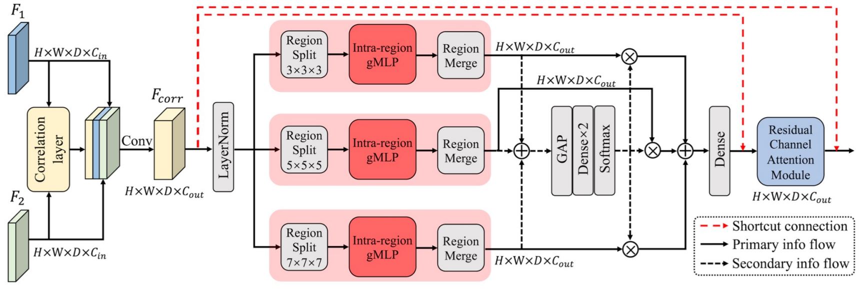

Figure 2: The architecture of our CMW-MLP block. It contains a 3D correlation layer to calculate the local correlations, a multi-window MLP module to capture multi-range dependence, and a residual channel attention module to highlight crucial feature channels.

Results

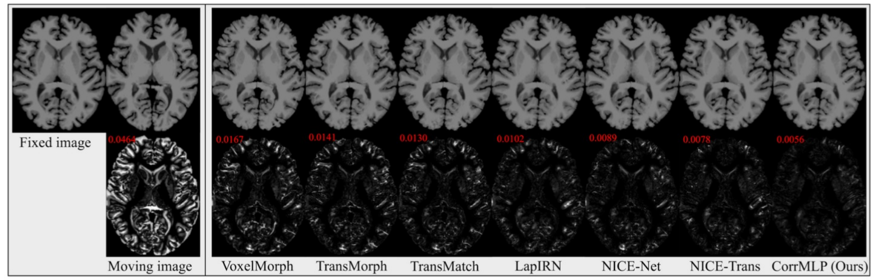

Figure 4: Qualitative comparison for brain image registration. Below each image is an error map that shows the intensity differences from the fixed image, with the mean absolute error placed in the upper left corner. A cleaner error map indicates a better registration result.

Method		Mindboggle dataset		Buckner dataset		Runtime	
		DSC ↑	NJD (%) ↓	DSC ↑	NJD (%) ↓	CPU (s)	GPU (s)
Before registration		0.347*	/	0.406*	/	/	/
SyN [17]	Traditional	0.534*	1.956	0.567*	1.874	3427	/
NiftyReg [18]	Traditional	0.569*	2.364	0.611*	2.175	159	/
VoxelMorph [7]	CNN, direct	0.552*	2.532	0.589*	2.220	2.84	0.23
Swin-VoxelMorph [13]	Transformer, direct	0.566*	2.254	0.605*	2.016	5.67	0.52
TransMorph [12]	Transformer, direct	0.571*	2.400	0.608*	2.183	3.68	0.35
TransMatch [15]	Transformer, direct	0.578*	2.036	0.622*	1.995	3.06	0.28
LapIRN [9]	CNN, coarse-to-fine	0.605*	2.164	0.632*	2.112	4.97	0.46
ULAE-net [35]	CNN, coarse-to-fine	0.610*	2.000	0.640*	1.940	5.37	0.51
Dual-PRNet++ [32]	CNN, coarse-to-fine	0.608*	2.424	0.636*	2.195	4.61	0.44
SDHNet [36]	CNN, coarse-to-fine	0.598*	1.872	0.634*	1.843	3.24	0.26
NICE-Net [11]	CNN, coarse-to-fine	0.618*	2.043	0.643*	1.963	3.55	0.32
NICE-Trans [22]	Transformer, coarse-to-fine	0.625*	2.324	0.649*	2.277	4.02	0.37
CorrMLP (Ours)	MLP, coarse-to-fine	**0.642**	**1.821**	**0.661**	**1.788**	5.48	0.49

Table 1: Quantitative comparison for brain image registration. The best results in each dataset are in bold. ↑: the higher is better. ↓: the lower is better. *: P<0.05, in comparison to CorrMLP.

- Figure 4 and Table 1 present the qualitative and quantitative comparison between the CorrMLP and existing registration methods for brain image registration, where our CorrMLP achieved significantly higher DSCs than all the comparison methods without sacrificing transformation smoothness.
- Table 3 presents the DSC results of the ablation study on architecture designs. The baseline MLPMorph has already surpassed VoxelMorph and TransMorph by a large margin, demonstrates the superiority of MLPs over transformers and CNNs on deformable image registration: MLP blocks can capture fine-grained long-range dependence at high resolutions. By employing MLP blocks in a correlation-aware coarse-to-fine registration architecture, our CorrMLP also outperformed MLPMorph by a large margin.
- Table 4 presents the DSC results of the ablation study on MLP blocks. Replacing our CMW-MLP blocks with five different existing MLP blocks all resulted in lower DSCs, indicating the effectiveness of our CMW-MLP block. Removing any MLP branch degraded the registration performance.

Method	Mindboggle	Buckner	ACDC
VoxelMorph [7]	0.552	0.589	0.754
TransMorph [12]	0.571	0.608	0.768
MLPMorph (Ours)	0.604	0.632	0.780
No correlation	0.628	0.650	0.800
Only image-level correlation	0.637	0.657	0.806
Only step-level correlation	0.634	0.655	0.805
CorrMLP (Ours)	**0.642**	**0.661**	**0.810**

Table 3: DSC results of the ablation study on architecture designs. The best results are in bold.

MLP block	Mindboggle	Buckner	ACDC
S^2-MLP [26]	0.621	0.644	0.794
sMLP [27]	0.622	0.645	0.794
Hire-MLP [28]	0.620	0.643	0.793
Swin-MLP [20]	0.624	0.646	0.797
Multi-axis gated MLP [29]	0.625	0.647	0.798
MW-MLP (Ours)	0.628	0.650	0.800
No 3×3×3 MLP branch	0.639	0.657	0.808
No 5×5×5 MLP branch	0.635	0.654	0.805
No 7×7×7 MLP branch	0.637	0.655	0.806
CMW-MLP (Ours)	**0.642**	**0.661**	**0.810**

Table 4: DSC results of the ablation study on MLP blocks. The best results are in bold.

Conclusion

We have shown the effectiveness of MLPs for deformable medical image registration by developing the first MLP-based coarse-to-fine registration network (CorrMLP). In the CorrMLP, we introduce a correlation- aware multi-window MLP (CMW-MLP) block and use it in a novel coarse-to-fine registration architecture that takes into account both image-level and step-level correlations. Extensive experiments on both brain and cardiac image registration show that, with the CMW-MLP block and the correlation-aware coarse-to-fine registration architecture, our CorrMLP can outperform state-of-the-art registration methods. Furthermore, we suggest that our CMW-MLP block can serve as a general MLP block applying to various network architectures for image registration tasks, and our CorrMLP also can apply to multi-modal registration tasks such as multi-parametric brain MRI registration.

Acknowledgement: This work was supported by Australian Research Council (ARC) under Grant DP200103748.

Code: https://github.com/MungoMeng/Registration-CorrMLP