Medical Image Registration Meets Vision Foundation Model: Prototype Learning and Contour Awareness

Hao Xu[1], Tengfei Xue[1], Jianan Fan[1], Dongnan Liu[1], Yuqian Chen[2,4], Fan Zhang[3], Carl-Fredrik Westin[2,4], Ron Kikinis[2,4], Lauren J. O'Donnell[2,4], Weidong Cai[1]

[1] University of Sydney, Australia [2] Harvard Medical School, USA [3] UESTC, China [4] BWH, USA

IPMI 2025

Introduction

- Medical image registration methods, which rely solely on intensity-based similarity, often struggle in cases with complex anatomy or ambiguous boundaries.
- Segment Anything Model (SAM) provides high-quality segmentation masks, enabling direct incorporation of structural priors into the registration process.
- We propose a novel SAM-assisted registration framework that incorporates prototype learning and contour awareness to enhance medical image registration.

Methodology

1. SAM Mask Generation

- Segmentation masks are generated for both the fixed and moving images using SAM with text prompts. These masks serve as anatomical priors to guide the registration process.

2. Prototype Contrast and Alignment

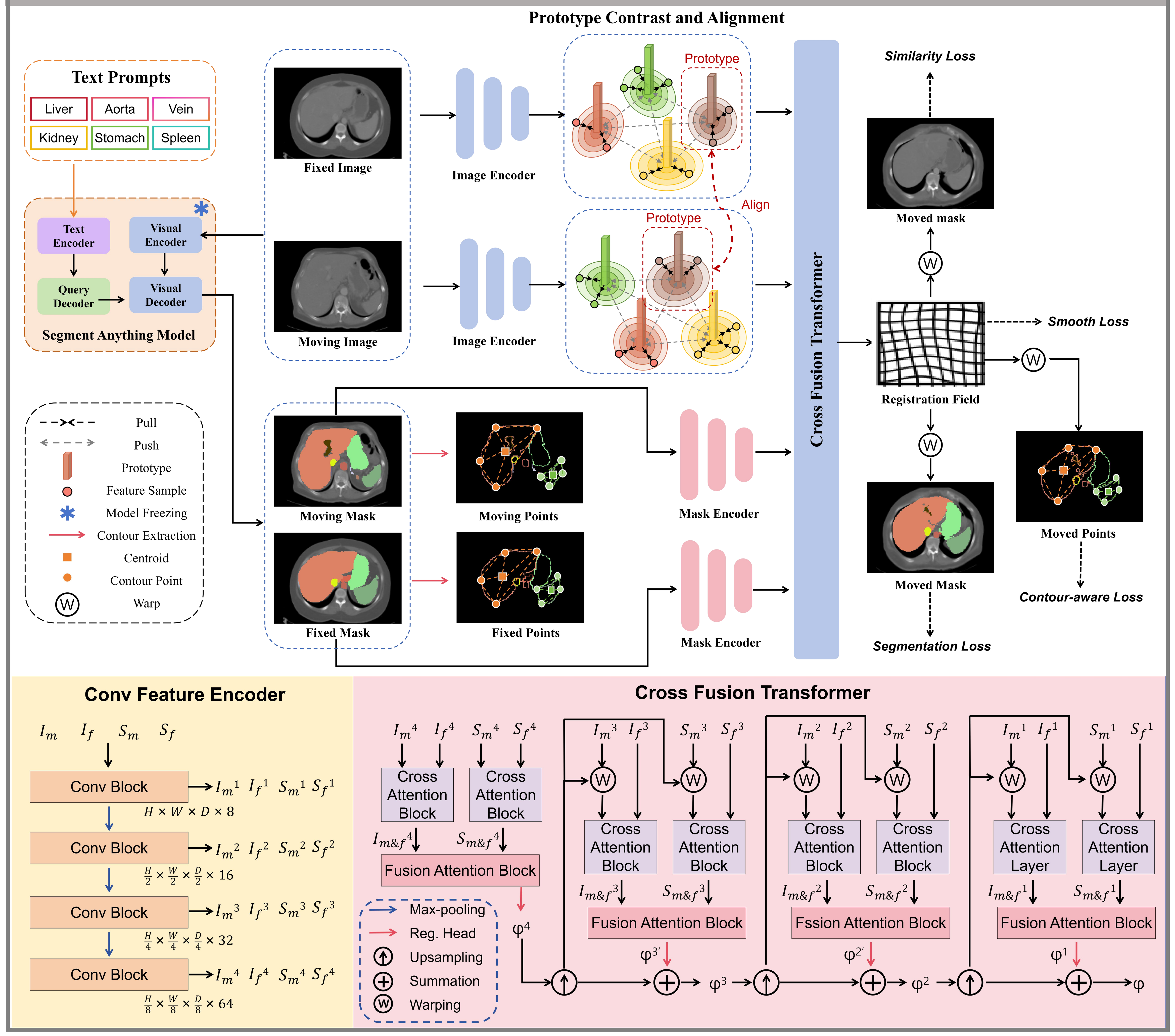
- Extract feature prototypes of different anatomical regions of the image according to the SAM mask.
- Introduce contrast loss to aggregate features of similar regions and align prototypes of the same region in different images.

3. Contour-Aware Loss

- Extract the contour point sets of the fixed and moving mask.
- Introduce the Chamfer Loss to minimize the Euclidean distance between the fixed and moved point sets.

$$\mathcal{L}_{\text{contour}} = \frac{1}{|C_m|}\sum_{i \in C_m} \min_{j \in C_f} \|i-j\|^2 + \frac{1}{|C_f|}\sum_{j \in C_f} \min_{i \in C_m} \|j-i\|^2$$

Framework



Evaluation on the Abdomen CT dataset

Methods	DSC (%) ↑													SDlogJ ↓
	Spl	Kid R	Kid L	Eso	Liv	Sto	Aor	IVC	Vei	Pan	Adr R	Adr L	AVG	
Initial	30.5	23.0	27.2	10.9	50.2	19.2	21.9	20.1	1.9	7.1	3.7	3.8	17.1	-
VoxelMorph [1]	61.1	55.3	55.6	30.7	70.1	31.2	44.4	44.0	16.7	19.3	18.3	14.4	38.4 ± 16.6	0.143
TransMorph [4]	60.3	54.3	54.0	30.1	70.7	33.5	45.4	46.1	19.3	18.5	18.2	16.3	39.0 ± 16.2	0.254
TransMatch [5]	65.3	58.4	56.3	33.4	72.3	36.4	49.3	54.5	21.3	20.6	19.9	18.7	42.2 ± 14.7	0.101
CorrMLP [21]	68.2	60.1	63.7	38.4	73.4	45.6	51.2	56.9	24.7	29.1	24.6	24.0	46.7 ± 13.2	0.099
SAM Masks [33]	92.8	88.1	89.5	72.2	95.5	89.4	88.4	84.5	68.5	79.7	64.2	62.6	81.2	-
VoxelMorph [1] + SAM	64.2	61.1	59.6	40.7	70.3	44.4	52.1	53.5	19.2	34.1	24.8	25.9	45.8 ± 12.7	0.065
TransMorph [4] + SAM	71.0	68.7	68.6	49.2	73.1	47.4	66.3	61.1	21.6	38.2	33.9	31.7	52.6 ± 12.2	0.088
TransMatch [5] + SAM	74.0	70.4	69.6	**70.2**	74.3	54.3	67.1	67.9	32.8	52.9	**46.3**	47.1	60.5 ± 11.4	0.095
CorrMLP [21] + SAM	77.1	72.8	75.7	55.1	77.7	57.2	**76.1**	73.1	35.1	50.0	44.6	41.8	61.3 ± 11.3	0.103
Ours	**83.9**	**79.3**	**76.0**	55.2	**83.8**	**82.5**	74.5	**73.2**	**36.9**	**56.4**	44.8	**48.5**	**66.3 ± 10.6**	0.091

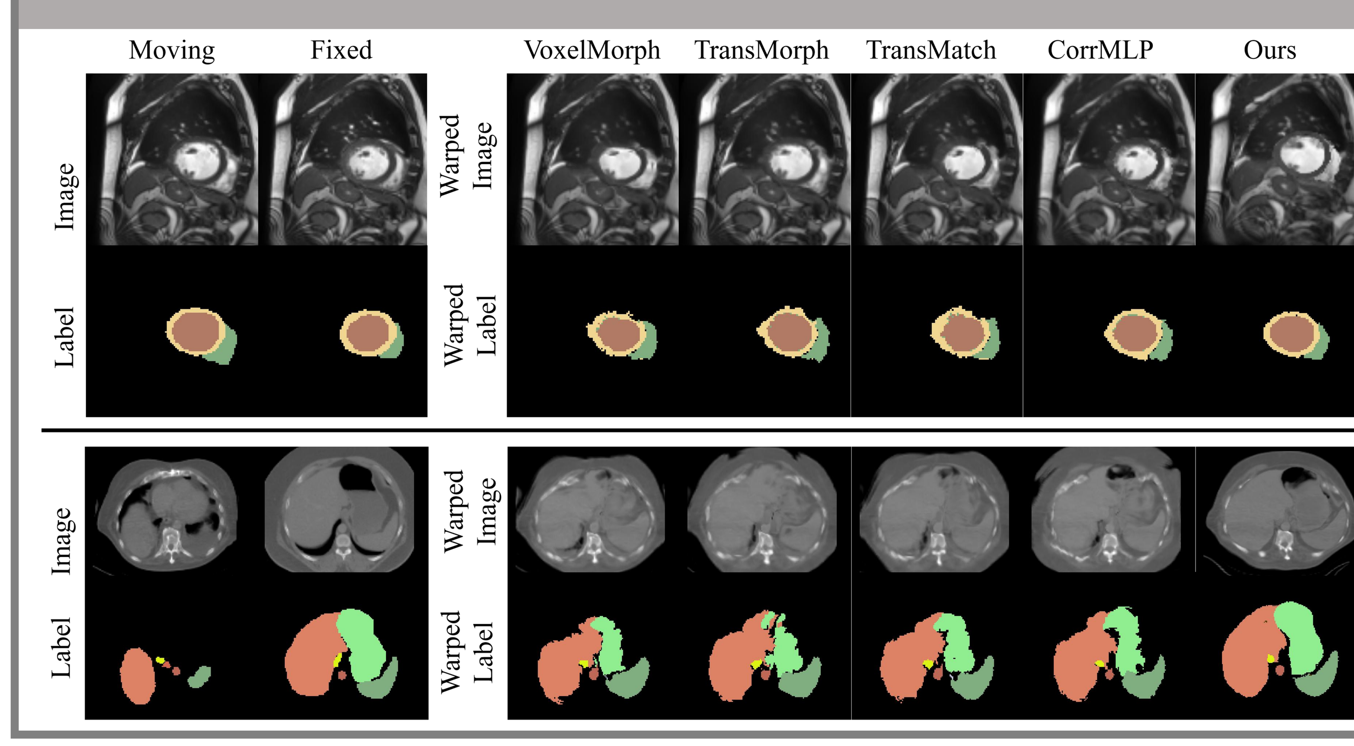
Evaluation on the ACDC MRI dataset

Methods	DSC (%) ↑				SDlogJ ↓
	LV	Myo	RV	AVG	
Initial	58.1	35.8	74.5	56.1	-
VoxelMorph [1]	83.2	60.1	80.5	74.6 ± 7.1	0.041
TransMorph [4]	82.5	58.8	80.4	73.9 ± 7.4	0.031
TransMatch [5]	82.3	58.6	82.1	74.4 ± 6.9	0.037
CorrMLP [21]	82.8	72.9	83.2	79.7 ± 4.6	0.054
SAM Masks [33]	89.2	79.2	76.2	81.6	-
VoxelMorph [1] + SAM	86.2	57.9	82.0	75.4 ± 6.3	0.077
TransMorph [4] + SAM	86.0	60.7	81.3	76.0 ± 5.5	0.026
TransMatch [5] + SAM	87.6	61.8	82.8	77.5 ± 5.1	0.073
CorrMLP [21] + SAM	83.4	74.2	**84.2**	80.6 ± 4.2	0.047
Ours	**91.9**	**77.6**	83.9	**84.6 ± 3.7**	0.049

Ablation Study

Dataset	$L_{prototype}$	$L_{contour}$	DSC (%) ↑	SDlogJ ↓
Abdomen	✗	✗	62.8	0.089
	✓	✗	63.3	0.086
	✓	✓	66.3	0.091
ACDC	✗	✗	83.0	0.053
	✓	✗	83.5	0.045
	✓	✓	84.6	0.049

Visualization



Scan the QR code for the full paper