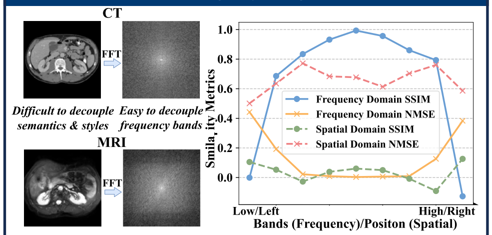
FAMNet: Frequency-aware Matching Network for Cross-domain Few-shot Medical Image Segmentation

Yuntian Bo, Yazhou Zhu, Lunbo Li, Haofeng Zhang

NJUST

AAAI-25 / IAAI-25 / EAAI-25
FEBRUARY 25 – MARCH 4, 2025 | PHILADELPHIA, USA

Motivation

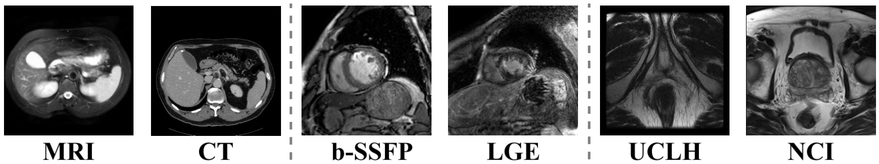


- Domain shift during inference is ignored by existing methods.
- Even within the same organ or region, spatial domain similarity varies significantly across domains, while frequency domain distinctions are more subtle, with variations primarily in high and low-frequency bands.

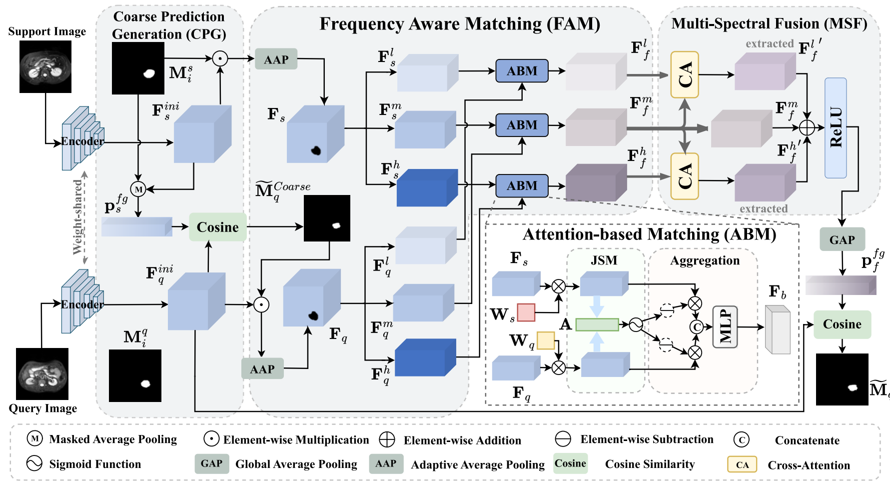
Contribution

- We extend few-shot medical image segmentation to a new task termed cross-domain few-shot medical image segmentation.
- We propose a Frequency-aware Matching Network (FAMNet), which includes two key components: a novel Frequency-aware Matching (FAM) module, to concurrently address both intra-domain and inter-domain variances, along with a Multi-spectral Fusion (MSF) module to further suppress domain-specific information.
- Our method achieves state-of-the-art performance on three cross-domain medical image datasets.

Problem: Domain Shift in Medical Images



Method: FAMNet



Quantitative Results

Method	Ref.	Abdominal CT → MRI					Abdominal MRI → CT				
		Liver	LK	RK	Spleen	Mean	Liver	LK	RK	Spleen	Mean
PANet	ICCV'19	39.24	26.47	37.35	26.79	32.46	40.29	30.61	26.66	30.21	31.94
SSL-ALP	TMI'22	70.74	55.49	67.43	**58.39**	63.01	71.38	34.48	32.32	51.67	47.46
ADNet	MIA'22	50.33	39.36	37.88	39.37	41.73	64.25	37.39	25.62	42.94	42.55
QNet	IntelliSys'23	58.82	42.69	51.67	44.58	49.44	70.98	38.64	30.17	43.28	45.77
CATNet	MICCAI'23	44.58	43.67	50.27	46.34	46.21	54.52	41.73	40.24	45.84	45.60
RPT	MICCAI'23	49.22	42.45	47.14	48.84	46.91	65.87	40.07	35.97	51.22	48.28
PATNet	ECCV'22	57.01	50.23	53.01	51.63	52.97	**75.94**	46.62	42.68	63.94	57.29
IFA	CVPR'24	48.81	45.79	51.46	51.42	49.37	50.05	36.45	32.69	43.08	40.57
Ours	—	**73.01**	**57.28**	**74.68**	58.21	**65.79**	73.57	**57.79**	**61.89**	**65.78**	**64.75**

Method	Ref.	Cardiac LGE → b-SSFP				Cardiac b-SSFP → LGE			
		LV-BP	LV-MYO	RV	Mean	LV-BP	LV-MYO	RV	Mean
PANet	ICCV'19	51.43	25.75	25.75	36.66	36.24	26.37	23.47	28.69
SSL-ALP	TMI'22	83.47	22.73	66.21	57.47	65.81	25.64	51.24	47.56
ADNet	MIA'22	58.75	36.94	51.37	49.02	40.36	37.22	43.66	40.41
QNet	IntelliSys'23	50.64	37.88	45.24	44.58	31.08	34.03	39.45	34.85
CATNet	MICCAI'23	64.63	42.41	56.13	54.39	45.77	43.51	46.02	45.10
RPT	MICCAI'23	60.84	42.28	57.30	53.47	50.39	40.13	50.50	47.00
PATNet	ECCV'22	65.35	50.63	68.34	61.44	66.82	**53.64**	59.74	60.06
IFA	CVPR'24	64.04	43.22	74.58	62.28	68.07	36.07	**60.42**	54.85
Ours	—	**86.64**	**51.84**	**76.26**	**71.58**	**77.37**	52.05	54.75	**61.39**

Method	Ref.	Prostate UCLH → NCI				Prostate NCI → UCLH			
		Bladder	CG	Rectum	Mean	Bladder	CG	Rectum	Mean
PANet	ICCV'19	51.85	38.89	36.28	42.34	49.90	36.27	43.58	43.25
SSL-ALP	TMI'22	**54.86**	42.59	41.65	46.37	66.6	43.12	56.94	55.55
ADNet	MIA'22	53.92	46.11	42.26	47.43	62.84	51.34	60.45	58.21
QNet	IntelliSys'23	39.66	35.71	34.25	36.54	41.22	43.21	39.19	41.21
CATNet	MICCAI'23	47.61	45.29	42.51	45.14	51.98	48.81	55.01	51.93
RPT	MICCAI'23	52.17	41.22	46.93	46.77	62.35	51.66	64.63	59.55
PATNet	ECCV'22	50.04	**52.71**	44.57	49.10	69.91	**53.02**	62.79	61.91
IFA	CVPR'24	41.02	39.15	36.45	38.87	52.54	39.85	34.27	42.22
Ours	—	53.06	48.48	**52.57**	**51.37**	**77.64**	48.73	**68.91**	**65.09**

- Our method achieves state-of-the-art performance, outperforming existing few-shot medical image segmentation models and cross-domain few-shot semantic segmentation models by 7.46%, 2.78%, 10.14%, 1.33%, 2.27%, and 3.18% in Dice score across six cross-domain directions.

Frequency-aware Matching

$$\phi_s = \mathrm{SC}(\mathcal{F}(\rho(\mathbf{F}_s))),$$

$$\mathbf{F}_s(i) = \rho^{-1}(\mathcal{F}^{-1}(\mathcal{B}_p(\phi_s, I(i)))) = \begin{cases} \mathbf{F}_s^l, & i=1 \\ \mathbf{F}_s^m, & i=2 \\ \mathbf{F}_s^h, & i=3 \end{cases},$$

- 1) The FAM module uses forward/inverse Fourier transform to decompose the foreground features of support and query into different frequency components (taking support feature as example).

$$\begin{cases} \mathbf{F}'_s = \mathbf{F}_s\mathbf{W}_s \\ \mathbf{F}'_q = \mathbf{F}_q\mathbf{W}_q \end{cases},$$

$$\mathbf{A}(\mathbf{F}'_s, \mathbf{F}'_q) = \sigma\left(\frac{\mathbf{F}'_s \cdot \mathbf{F}'_q}{\|\mathbf{F}'_s\|\|\mathbf{F}'_q\|}\right),$$

$$\begin{cases} \mathbf{F}''_{s,0} = \mathbf{A}' \odot \mathbf{F}'_s \\ \mathbf{F}''_{q,0} = \mathbf{A}' \odot \mathbf{F}'_q \end{cases},$$

$$\begin{cases} \mathbf{F}''_{s,1} = (1-\mathbf{A}') \odot \mathbf{F}'_s \\ \mathbf{F}''_{q,1} = (1-\mathbf{A}') \odot \mathbf{F}'_q \end{cases},$$

- 2) During typical training, models rely on domain-specific frequency bands (DSFBs) to adapt, but this can cause overfitting to certain prominent features, like high-frequency edges in CT images, leading to performance degradation in unseen domains like MRI, where edges are less distinct.

  This stage learns attention weights and applies forward/inverse attention to different frequency bands, helping the model focus on domain-agnostic frequency bands (DAFBs) and reduce reliance on DSFBs.

$$\mathbf{F}_b = \mathrm{MLP}\left(\mathrm{Cat}\left(\mathbf{F}''_{s,fb}, \mathbf{F}''_{q,fb}\right), \varphi\right), \quad fb \in \{0,1\}$$

- 3) The features belonging to the support and query are fused, producing a new debiased feature that serves as the final representation for a specific frequency band.

Resources

Paper

Code

Video

Feel free to contact me: yuntian.bo@njust.edu.cn