FAReg: Unsupervised Deformable MRI Registration Using Fourier Analysis

Yuhao Liu*, Boyue Song*, Jinping Tang†, Ge Zhu†
School of Computer and Big Data (Cybersecurity School), Heilongjiang University
liuyh0202@foxmail.com, songboyue96@gmail.com, tangjinping@hlju.edu.cn, zhuge@hlju.edu.cn

***Abstract*—Deformable brain MRI registration remains challenging because of complex anatomy and prevalent imaging artifacts. Most existing methods estimate dense displacement fields only in the spatial domain and lack explicit frequency-domain inductive bias, which limits modeling of periodic structures. Many boundary-aware designs also rely on fixed weight maps and are easily misled by pseudo-boundaries caused by partial volume effects and Gibbs ringing, reducing robustness under spatially heterogeneous artifacts. To address these issues, we propose FAReg, an unsupervised registration framework that integrates explicit frequency-domain modeling with adaptive boundary optimization. FAReg introduces a Frequency-domain Feature Fusion Module (FFFM), which explicitly decomposes features with learnable sine–cosine bases to capture quasi-periodic anatomy, and an Edge-Gradient Dynamic Reweighting (EGDR) mechanism, which uses a spatially adaptive gradient-based map to suppress artifact-induced gradients while reinforcing anatomically consistent boundaries. Extensive experiments on two public brain MRI datasets demonstrate that FAReg consistently outperforms existing registration methods in both structural alignment accuracy and boundary fidelity.**

***Index Terms*—Frequency-domain Feature Fusion Module, Edge-Gradient Dynamic Reweighting, Periodic Structure Modeling, Unsupervised MRI Registration**

I. Introduction

MRI registration aims to align anatomical structures between a moving and a fixed image by estimating a spatial transformation, and it underpins numerous downstream tasks such as deformable population template construction [1], [2] and multi-atlas segmentation [3]–[5]. Although deep learning approaches—including CNN-based [6] and Transformer-based [7] frameworks—have substantially improved registration efficiency and accuracy, achieving anatomically consistent alignment remains challenging under complex tissue patterns, noise, and acquisition artifacts.

Brain MRI exhibits repetitive and quasi-periodic anatomical patterns, such as cortical folding and white-matter fiber textures, which appear in the frequency domain as stable directional energy concentrations forming distinctive "spectral signatures" [8]. These characteristics provide geometric and phase-consistency cues for registration. However, most CNN-based and Transformer-based methods operate mainly in the spatial domain without explicit frequency regularization, making phase coherence hard to preserve when aligning highly repetitive structures and often causing local mismatches or deformation discontinuities. Motivated by Fourier analysis, we introduce a Frequency-domain Feature Fusion strategy that augments spatial representations with frequency-aware features. Our design explicitly partitions early features into a Fourier branch and a spatial branch: the former extracts periodic patterns via learnable sine–cosine projections, and the latter preserves local intensity variation and boundary detail. The two branches are concatenated and fused by a lightweight convolution to form a unified representation for deformation estimation.

Another major source of registration error arises near anatomical boundaries, including the cortical surface and ventricular interfaces, due to partial volume effects (PVE) and Gibbs ringing artifacts. PVE blurs tissue transitions, while Gibbs ringing introduces oscillations near sharp edges [9]; together they create pseudo-boundaries that mislead gradient-driven optimization. Although existing gradient- or edge-aware formulations improve boundary sensitivity [10]–[12], they are often coupled with fixed image-derived weighting that may overemphasize pseudo-edges and suppress low-contrast but anatomically valid boundaries. To address this limitation, we propose an Edge-Gradient Dynamic Reweighting (EGDR) strategy that learns a spatially adaptive edge-reliability map and reweights boundary-related terms to suppress artifact-induced gradients and emphasize anatomically consistent edges.

Based on the above analysis, we develop an unsupervised MRI registration framework FAReg that jointly models frequency-domain structural regularities and improves boundary alignment under artifact contamination. FAReg differs from prior approaches at two levels. For feature modeling, unlike methods that rely only on spatial convolutions or attention to implicitly learn periodicity, or methods that use Fourier representations only at deformation decoding, FAReg introduces the Frequency-domain Feature Fusion Module (FFFM). FFFM incorporates learnable sine–cosine bases at the earliest network stage, providing an explicit frequency-domain inductive bias for periodic anatomical structures. For optimization, unlike static gradient weighting, FAReg's Edge-Gradient Dynamic Reweighting (EGDR) learns a spatially adaptive reliability map that suppresses artifact-induced gradients and reinforces true anatomical boundaries. Through the synergy of FFFM and EGDR, FAReg achieves end-to-end frequency-aware and boundary-reliable registration. The main contributions are summarized as follows:

1) We propose a Frequency-domain Feature Fusion Module

*Co-first authors. †Corresponding authors.

(FFFM) that introduces decoupled sine–cosine projections to explicitly model the inherent periodic patterns in brain MRI. By fusing frequency-domain representations with spatial features at the initial stage, FFFM provides a frequency-aware inductive bias that enhances the encoding of both global periodic regularity and local anatomical details for deformation estimation.

2) We propose an Edge-Gradient Dynamic Reweighting (EGDR) strategy to enhance registration robustness and alleviate boundary degradation caused by PVE and Gibbs artifacts with an adaptive edge-weighting map derived from the fixed image's gradient magnitude.
3) We develop the FAReg framework by integrating the above modules with a Fourier Resampling Module, enabling full-resolution displacement reconstruction while preserving periodic structure consistency.
4) We propose the novel Peak Frequency Energy Concentration (PFEC) metric to quantitatively assess the concentration of periodic features in the frequency domain.

II. Related Works

A. Learning-based Medical Image Registration

Unsupervised medical image registration has evolved from classical optimization to deep learning paradigms. Diffeomorphic methods, including SyN [13] and LDDMM [14], offer strong theoretical guarantees but are computationally expensive; later ANTs evaluations further verified robust similarity metrics in neuroimaging benchmarks [15]. Learning-based models such as VoxelMorph [6] and TransMorph [7], with follow-ups including DeepFlash [16], HyperMorph [17], and CycleMorph [18], improve efficiency, hyperparameter adaptation, and cycle-consistent regularization; newer designs such as LessNet [19] and TCDE-Net [20] further strengthen lightweight multi-scale representation. However, these methods remain predominantly spatial-domain and lack explicit periodic or frequency-domain modeling.

B. Frequency-domain Representation and Periodic Modeling

Frequency-domain modeling has been explored in registration and related vision tasks to capture periodic structures. Fourier-Net [21] and Fourier-Net+ [22] represent deformation in band-limited Fourier space and reconstruct dense fields by inverse Fourier transform, improving 3D efficiency. Yet they mainly apply Fourier modeling at decoding rather than explicitly encoding periodic patterns during feature extraction. Sinusoidal representations, including Fourier Neural Networks [23] and Snake [24], improve periodic signal modeling but remain limited for complex anatomy. FAN [25] further introduces learnable Fourier bases to disentangle periodic and nonlinear components, motivating frequency-aware feature learning for medical image registration.

C. Boundary-aware Registration under Imaging Artifacts

Accurate boundary alignment remains challenging because partial volume effects (PVE) and Gibbs ringing degrade gradient reliability near tissue interfaces [9]. Classical studies show that reducing Gibbs artifacts [26] and modeling PVE [27] improve boundary fidelity. Accordingly, gradient-aware and edge-aware measures, including GMSD [10], NTG [11], and gradient-guided learning [12], enhance boundary sensitivity. However, these methods usually use fixed image-derived weighting and are less robust to spatially heterogeneous artifacts. Recent evidence that mitigating PVE and Gibbs artifacts improves structural analysis [28], [29] further motivates adaptive boundary-aware mechanisms.

III. Method

Accurately estimating dense displacement fields is challenging due to periodic anatomy and artifact-corrupted boundaries. As shown in Fig. 1, FAReg first applies a Frequency-domain Feature Fusion Module (FFFM) to the fixed image f and moving image m to extract fused frequency–spatial features, then uses an encoder to predict a compact band-limited deformation representation, which is deterministically upsampled to full resolution by the Fourier Resampling Module (FRM). The network is trained end-to-end with an Edge-Gradient Dynamic Reweighting (EGDR) loss to improve boundary alignment while enforcing global smoothness.

A. Frequency-domain Feature Fusion Module

The Frequency-domain Feature Fusion Module (FFFM) is designed to overcome the inherent limitation of conventional spatial-domain encoders in explicitly modeling periodic or repetitive anatomical structures (such as cortical folding patterns). From a theoretical perspective, Fourier analysis [30], [31] provides a rigorous framework for decomposing signals into elementary frequency components, thereby exposing latent periodic structures embedded within complex data. Motivated by this property, FFFM introduces frequency-domain inductive bias at the early stage of feature extraction, enabling the network to jointly encode global oscillatory patterns and local anatomical details.

Our design builds upon the Fourier Analysis Network (FAN) [25], which embeds Fourier series directly into the network architecture to impose periodicity as an intrinsic inductive bias and improve generalization in periodic modeling. Specifically, the FAN layer $\theta(x)$ is defined as:

$$\theta(x) \triangleq [\cos(W_f x) \| \sin(W_f x) \| \sigma(B_s + W_s x)], \qquad (1)$$

where $W_f \in \mathbb{R}^{d_x \times d_f}$, $W_s \in \mathbb{R}^{d_x \times d_s}$, and $B_s \in \mathbb{R}^{d_s}$ are learnable parameters; d_f and d_s denote the channel dimensions of the Fourier-domain and spatial-domain branches, respectively, and $p = d_f : d_s$ denotes the ratio of their channel capacities.

We adapt and extend FAN for 3D medical image registration through the following three specific stages.

a) Feature mapping and channel partitioning: As shown in Fig. 1, the input image pair is first processed by a $3 \times 3 \times 3$ convolution to obtain the initial features $\mathbf{X} = Conv_{3\times3\times3}(Concat(f, m)) \in \mathbb{R}^{B\times C_{in}\times H_x\times W_x\times D_x}$ with B representing the batch size, C_{in} the number of input channels, and $H_x \times W_x \times D_x$ the spatial size of the feature map.

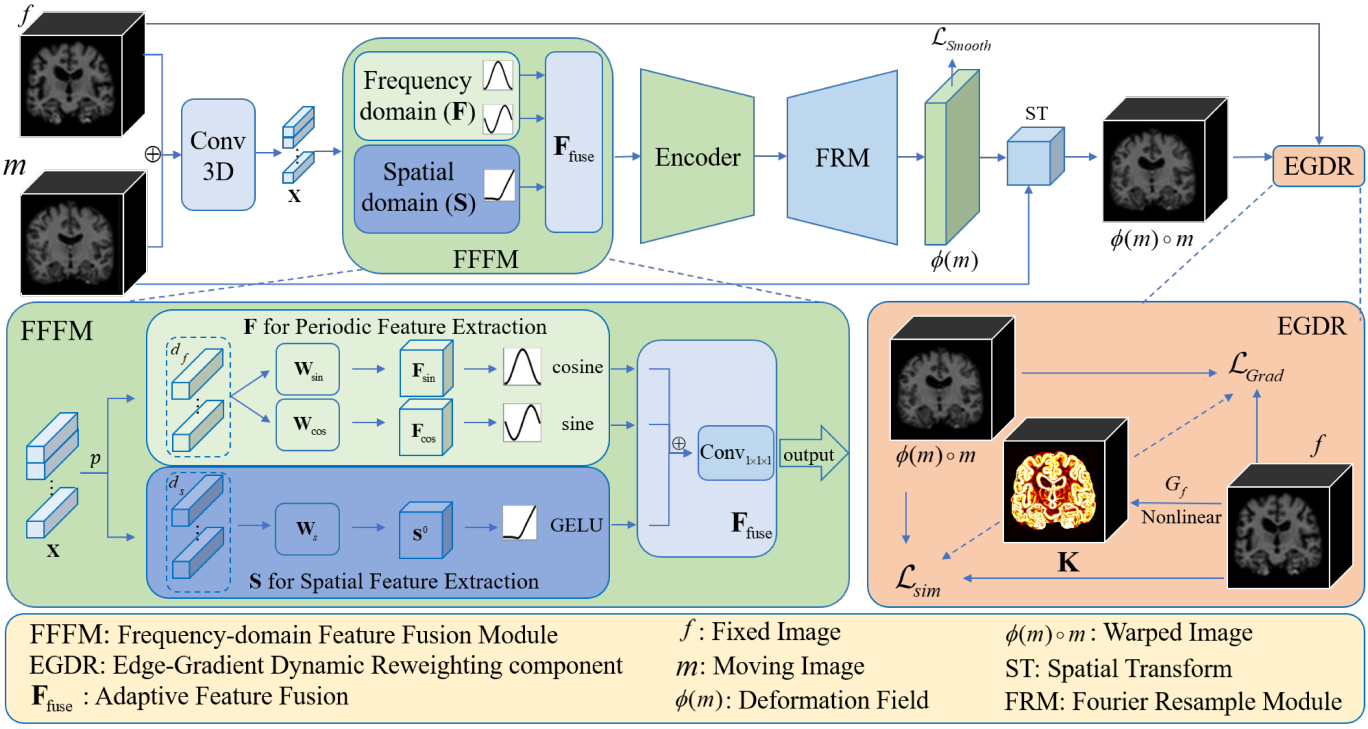

Fig. 1: Overview of FAReg. The concatenated image pair is fed to FFFM to fuse Fourier and spatial features, then encoded into a low-dimensional displacement representation that is upsampled by the Fourier Resampling Module (FRM) to the full-resolution field ϕ for warping. In FFFM, $p = d_f : d_s$ sets the channel allocation ratio between the frequency-domain (F) and spatial-domain (S) branches, where d_f channels are processed by $\mathbf{W}_{\sin}$ and $\mathbf{W}_{\cos}$, and d_s channels by $\mathbf{W}_S$. Training is regularized by EGDR, which derives an edge-reliability map $\mathbf{K}$ from the fixed-image gradient (Eq. (7)) to suppress artifact-induced gradients and emphasize reliable anatomical boundaries.

b) Dual-branch feature extraction with independent bases: In the original FAN, sine and cosine activations share a single projection matrix $\mathbf{W}_f$, constraining both bases to the same feature subspace and limiting their ability to model distinct periodic directions or frequencies. Inspired by [32], we decouple the shared projection in standard FAN by employing two independent matrices, $\mathbf{W}_{\sin}$ and $\mathbf{W}_{\cos}$. This allows the sine and cosine bases to model distinct periodic directions and frequencies:

$$\mathbf{F}_{\sin} = \mathbf{W}_{\sin} \cdot \mathbf{X}_f + \mathbf{b}_{\sin}, \mathbf{F}_{\cos} = \mathbf{W}_{\cos} \cdot \mathbf{X}_f + \mathbf{b}_{\cos}, \quad (2)$$

The frequency-domain feature $\mathbf{F}$ is formed by concatenating these outputs:

$$\mathbf{F} = \sin(\mathbf{F}_{\sin}) \oplus \cos(\mathbf{F}_{\cos}), \quad (3)$$

where $\oplus$ denotes concatenation along the channel dimension.

In parallel with the frequency-domain branch, the spatial component is generated via a separate mapping matrix $\mathbf{W}_S$:

$$\mathbf{S}^0 = \mathbf{W}_S \cdot \mathbf{X}_s + \mathbf{b}_S, \quad (4)$$

serving to preserve local non-periodic structures, intensity variations, and boundary information in the input features.

To further enhance the discriminability of spatial structures, a smooth nonlinear activation function, GELU, is applied to $\mathbf{S}^0$:

$$\mathbf{S} = \mathrm{GELU}(\mathbf{S}^0). \quad (5)$$

GELU strengthens local intensity differences while maintaining output smoothness, thereby facilitating the network in distinguishing subtle yet important boundary variations among different anatomical regions.

c) Adaptive feature fusion: The features $\mathbf{F}$ and $\mathbf{S}$ are concatenated and fused via a $1 \times 1 \times 1$ convolution. This operation performs a learnable, channel-wise reweighting, adaptively blending periodic and spatial information based on local context to form a unified feature representation $\mathbf{F}_{\text{fuse}}$ for the encoder:

$$\mathbf{F}_{\text{fuse}} = \mathrm{Conv}_{1\times1\times1}(\mathbf{F} \oplus \mathbf{S}). \quad (6)$$

B. Encoder

After the FFFM, the extracted features jointly encode periodic structures and spatial textures, with the Fourier branch capturing global repetitive patterns and the spatial branch preserving local intensity and boundary details. We therefore introduce an encoder to extract multi-scale contextual information and enhance spatial consistency.

The encoder adopts a hierarchical convolutional architecture as in Fourier-Net [21], consisting of stacked 3D convolutional blocks with progressive downsampling to expand the receptive field and capture long-range spatial correlations. The encoder module outputs a spatial-domain representation of a band-limited displacement field, denoted as $\mathbb{D}_\phi$, which avoids direct regression in the Fourier domain while preserving the benefits of band-limited modeling. As illustrated in [21], displacement fields reconstructed from a limited set of low-frequency Fourier coefficients can achieve performance comparable to full-resolution registration.

C. Fourier Resampling Module

To reconstruct the full-resolution dense displacement field ϕ from the band-limited representation $\mathbb{D}_\phi$ produced by the encoder, we adopt a deterministic Fourier Resampling Module (FRM) following [21]. Specifically, $\mathbb{D}_\phi$ is first transformed into the frequency domain using a discrete Fourier transform (DFT), where its spectrum is zero-padded to the target resolution, and then mapped back to the spatial domain via an inverse DFT to obtain ϕ. This analytic reconstruction scheme provides a principled decoder consistent with Fourier-Net for displacement field recovery.

D. Edge-Gradient Dynamic Reweighting Component

While periodic feature modeling effectively captures repetitive anatomical structures and improves global registration consistency, Gibbs artifacts [26] and partial volume effects [27] still lead to unreliable and error-prone alignment near anatomical boundaries. Due to the sensitivity of gradients to imaging artifacts and spatially varying reliability, we introduce an Edge-Gradient Dynamic Reweighting strategy to explicitly suppress artifact-induced gradients while reinforcing true anatomical boundaries during optimization, thereby enhancing the robustness and accuracy of registration in complex boundary regions.

To dynamically suppress edge-related artifacts, a naive strategy is to use the gradient magnitude as a voxel-wise weight map to reweight the boundary-consistency term. Specifically, for a fixed image f, we compute the gradient magnitude $G_f = |\nabla f|_2$, where ∇ is implemented using a 3D Sobel operator followed by robust normalization. However, due to the heavy-tailed gradient distribution in medical images [33], directly weighting the loss by G_f would allow a small number of high-contrast edges to dominate the optimization. We therefore construct an adaptive edge-weight factor K for each voxel, which is a nonlinear function of G_f, to suppress weak gradients originating from noise, homogeneous regions, or minor artifacts. The weight factor K is defined as follows:

$$K = \sigma\left(\kappa \cdot (G_f - \tau)\right), \quad K \in [0,1], \tag{7}$$

where κ controls the sharpness of the response, and τ suppresses weak gradients arising from noise, artifacts, or homogeneous regions. The threshold τ is set to 0.4 based on an empirical analysis of the gradient magnitude distribution across the datasets. In particular, the threshold τ specifies the minimum gradient-magnitude level to be considered reliable; gradients with $G_f < \tau$ (often arising from noise, PVE-induced blurring, or mild ringing) are largely suppressed, while gradients with $G_f > \tau$ are progressively emphasized. σ denotes the sigmoid function, which serves as a nonlinear mapping to squash the response into $[0,1]$.

As illustrated in Fig. 2, K increases monotonically with G_f. Compared with directly using G_f as the weight map, the proposed tunable nonlinear mapping produces a more desirable weighting profile $K(G_f)$. Voxels affected by pseudo-boundaries typically exhibit weaker and less reliable gradients than true anatomical interfaces, and are therefore assigned smaller weights after mapping. Consequently, deformation optimization is guided toward anatomically consistent boundaries in f, effectively suppressing misleading gradient cues and improving boundary robustness. Applying Eq. (7) voxel-wise over the entire image domain yields a spatially varying weight map, denoted as $\mathbf{K}$, as shown in Fig. 1.

Based on K defined over the image domain Ω, to emphasize reliable anatomical boundaries, we propose a modified weighted similarity loss $\mathcal{L}_{sim}$ built upon the standard NCC measure [13], which is defined as Eq. (8):

$$\mathcal{L}_{sim} = 1 - \frac{\sum\limits_{x\in\Omega} K\left(m' - \mu_{m'}\right)\left(f - \mu_f\right)}{\sqrt{\sum\limits_{x\in\Omega} K\left(m' - \mu_{m'}\right)^2}\sqrt{\sum\limits_{x\in\Omega} K\left(f - \mu_f\right)^2 + \epsilon}}, \tag{8}$$

where $\mu_{m'}$ and μ_f denote the corresponding K-weighted mean intensities of m' and f, respectively:

$$\mu_{m'} = \frac{\sum_{x\in\Omega} K \cdot m'}{\sum_{x\in\Omega} K + \epsilon}, \quad \mu_f = \frac{\sum_{x\in\Omega} K \cdot f}{\sum_{x\in\Omega} K + \epsilon}, \tag{9}$$

The constant ϵ is added for numerical stability.

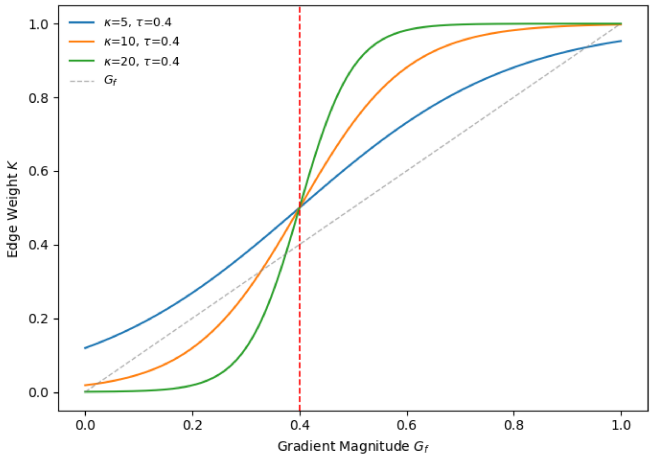

Fig. 2: Schematic Diagram of Edge Dynamic Weighting.

In addition, PVE may attenuate true boundary responses by smoothing intensity transitions, leading to low-gradient yet anatomically valid boundaries that are difficult to capture during optimization. To ensure these boundaries remain well

captured and aligned, we introduce a weighted gradient-consistency loss that is conceptually inspired by the Normalized Total Gradient (NTG) measure [11], denoted as $\mathcal{L}_{Grad}$ in Eq. (10):

$$\mathcal{L}_{Grad} = \frac{1}{\mid \Omega \mid} \sum_{x \in \Omega} K \cdot \parallel \nabla m' - \nabla f \parallel_2^2, \tag{10}$$

where ∇ is the spatial gradient operator. This term penalizes discrepancies in the gradient domain between the warped and fixed images. By reusing the weight factor K at each voxel, the penalty is concentrated on voxels identified as reliable boundaries, thereby reinforcing structural alignment while remaining robust to artifact-induced noise.

To further maintain spatial regularity, we incorporate a Laplacian-based smoothness constraint:

$$\mathcal{L}_{Smooth} = \frac{1}{\mid \Omega \mid} \sum_{x \in \Omega} \|\nabla \phi(m)\|_2^2. \tag{11}$$

The final loss function is formulated as a weighted combination of all components:

$$\mathcal{L}_{total} = \mathcal{L}_{sim} + \lambda_s \mathcal{L}_{Smooth} + \lambda_g \mathcal{L}_{Grad}, \tag{12}$$

where λ_s and λ_g control the contributions of the smoothness regularizer and the gradient-consistency term, respectively.

IV. Experiments and Results

A. Datasets

OASIS-1 Dataset [34]: OASIS-1 contains T1-weighted brain MRI scans from 416 subjects. We use the preprocessed version provided by [17], which includes 414 subjects and is tailored for inter-subject registration. Preprocessing involves bias-field correction, skull stripping, rigid alignment, and cropping to $160 \times 192 \times 224$. Thirty-five anatomical labels generated by FreeSurfer are available for quantitative evaluation. The dataset is split into 201, 201, and 12 volumes for training, testing, and validation, respectively.

IXI Dataset: The IXI dataset released by [7] for atlas-to-subject registration consists of 576 brain MRI scans, each cropped to $160 \times 192 \times 224$. The anatomical atlas follows [18]. The dataset is divided into 403 training, 115 testing, and 58 validation volumes.

B. Implementation Details

The proposed framework is implemented in PyTorch and optimized using Adam with a learning rate of 1×10^{-4} and a batch size of 1. All models are trained for 1000 epochs. For OASIS-1, we set $p = 1$, $\lambda_s = 1$, $\lambda_g = 0.2$, $\kappa = 10$, and $\tau = 0.4$; for IXI, we use $p = 0.5$, $\lambda_s = 5$, $\lambda_g = 0.5$, $\kappa = 15$, and $\tau = 0.4$. All experiments are conducted on a single NVIDIA RTX 3090 GPU.

Following FFFM, the encoder comprises five convolutional blocks. The first three blocks perform downsampling, each containing two $3 \times 3 \times 3$ convolutions, where the second layer uses stride 2 and doubles the channel dimension. The final two blocks perform symmetric upsampling, each consisting of one fractional convolution with stride 2 followed by two convolutional layers that halve the number of channels.

C. Evaluation Metrics

Registration accuracy is evaluated using the Dice similarity coefficient (Dice) and the 95% Hausdorff distance (HD95), which measure volumetric overlap and boundary alignment, respectively.

The Dice coefficient [35] is defined as:

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|}, \tag{13}$$

where A and B are binary segmentation masks. Dice ranges from 0 to 1, with higher values indicating better volumetric agreement between the warped image labels and the fixed-image ground truth.

HD95 evaluates boundary-based geometric consistency. Let S_A and S_B denote the surface point sets of A and B, respectively. To reduce sensitivity to outliers, HD95 is computed as:

$$\begin{aligned} \mathrm{HD95}(S_A, S_B) = \max \Big(& \mathrm{perc}_{95}\{\min_{b \in S_B} \|a - b\|_2 : a \in S_A\}, \\ & \mathrm{perc}_{95}\{\min_{a \in S_A} \|b - a\|_2 : b \in S_B\} \Big). \end{aligned} \tag{14}$$

HD95 is reported in millimeters, with lower values indicating more accurate boundary alignment.

D. Quantitative Analysis

To evaluate the proposed FAReg, we compare it with representative learning-based registration methods, including VoxelMorph [6], DeepFlash [16], TransMorph [7], Fourier-Net variants [21], [22], Diff-LessNet [19], and TCDE-Net [20]. Among them, Fourier-Net [21] serves as the baseline, as we follow its encoder–decoder architecture and adopt its model-driven Fourier decoding scheme as the Fourier Resampling Module (FRM), while introducing additional feature-level and optimization-level enhancements. As shown in Table I, the proposed FAReg achieves the best overall performance on both public brain MRI datasets.

TransMorph [7] captures global deformations but shows inferior boundary alignment. TCDE-Net [20] improves multi-scale representation through its dual-encoder design, yet lacks explicit frequency-domain modeling for periodic anatomical structures. The Fourier-Net family [21], [22] improves efficiency via Fourier-based displacement decoding, but limiting frequency modeling to the reconstruction stage can reduce structural consistency in complex regions.

The proposed FAReg incurs slightly higher inference cost due to the FFFM and EGDR modules compared with lightweight models such as Fourier-Net(+) [21], [22], LessNet [19], and VoxelMorph [6]. Compared with the baseline, this additional cost leads to consistent improvements in Dice and HD95, indicating the benefit of explicitly modeling frequency-domain features and incorporating adaptive edge-aware constraints for improved topological preservation and boundary accuracy.

TABLE I: Performance comparison of different methods on the IXI and OASIS datasets. The best Dice and HD95 values are indicated in boldface.

Method	IXI			OASIS			
	Dice↑	Parameters(M)	Mult-Adds(G)	Dice↑	HD95↓	Parameters(M)	Mult-Adds(G)
Initial	0.386±0.195	-	-	0.572±0.053	3.831	-	-
VoxelMorph [6](2018)	0.732±0.123	0.301	398.81	0.805±0.022	1.546	0.301	91.16
DeepFlash [16](2020)	0.743±0.132	-	-	0.825±0.016	1.537	-	-
TransMorph [7](2022)	0.754±0.124	46.771	657.64	0.858±0.014	1.494	46.771	150.39
Diff-Fourier-Net [21](2023)	0.761±0.131	4.198	169.07	0.843±0.013	1.491	3.891	31.00
Diff-Fourier-Net+ [22](2023)	0.750±0.130	0.880	19.30	0.860±0.013	1.487	0.211	0.52
Diff-LessNet$_{12}$ [19](2025)	0.768±0.127	1.064	481.43	0.868±0.016	1.472	0.963	113.42
TCDE-Net [20](2025)	0.779±0.118	52.732	117.82	0.876±0.011	1.443	50.934	102.59
FAReg(ours)	**0.787±0.127**	5.645	174.39	**0.883±0.013**	1.397	5.986	45.53

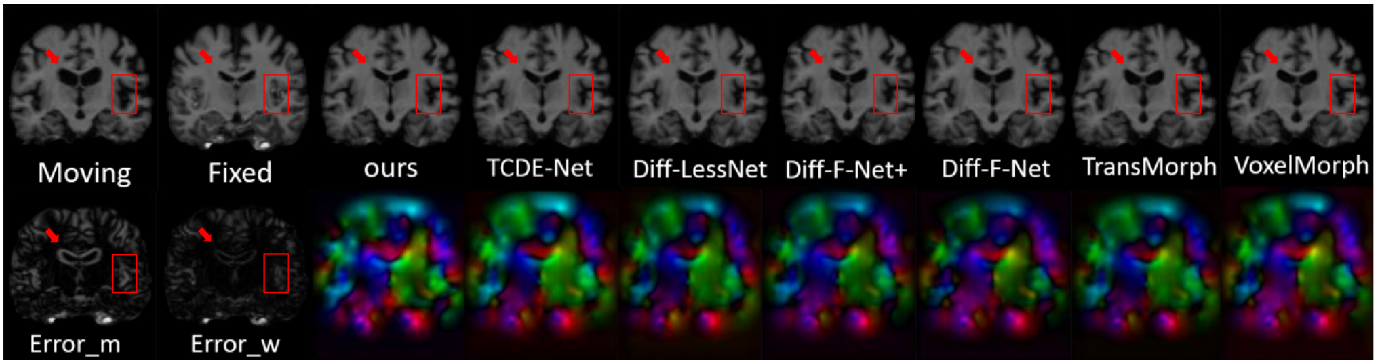

Fig. 3: Visualization of registration results on the OASIS dataset. Except for the first two columns, the top row shows the warped images produced by the network, and the bottom row presents the corresponding displacement fields. The first two columns of the top row display the moving and fixed images, while the first two columns of the bottom row show the difference maps (between the fixed and moving images, and between the fixed and warped images), where higher intensity indicates larger differences. The boxed regions highlight areas notably affected by partial volume effects (PVE), and the arrows indicate locations with the most pronounced boundary variations.

E. Qualitative Analysis

Fig. 3 compares warped images, deformation fields, and difference maps produced by the proposed FAReg and other learning-based methods. The proposed FAReg achieves the highest visual consistency with the fixed image and markedly reduced errors, particularly near high-contrast anatomical boundaries and regions with complex structural transitions. Although periodic structures are difficult to assess visually, the improved boundary alignment and error suppression can be attributed to the proposed EGDR mechanism. The effectiveness of FFFM in modeling periodic anatomical patterns is further validated through dedicated ablation experiments presented in the following section.

F. Ablation Studies

a) Ablation study on the Effectiveness of FFFM and EGDR: Table II reports ablations of the FFFM and EGDR, evaluated by Dice (overlap) and HD95 (boundary error). Removing FFFM substantially degrades Dice, confirming that early Fourier feature decomposition is critical for modeling periodic structures and maintaining globally coherent deformations. Removing EGDR mainly increases HD95, indicating that adaptive edge reweighting effectively suppresses artifact-induced gradients and improves boundary alignment. Further isolating EGDR shows that the K-weighted similarity term strengthens reliable-edge discrimination, while the gradient-consistency term ($\mathcal{L}_{\mathrm{Grad}}$) stabilizes optimization in PVE-affected regions; using both yields the best overall performance.

TABLE II: Ablation study on the proposed FFFM and the edge-gradient weighting mechanism using the OASIS dataset. When "FFFM" is indicated as "✗", the convolution layers of the original network are used in place of FFFM; when "$\mathcal{L}_{sim}$ with K" is indicated as "✗", only NCC is employed as the similarity loss.

FFFM	EGDR		Dice↑	HD95↓
	$\mathcal{L}_{sim}$ with K	$\mathcal{L}_{Grad}$		
✗	✗	✗	0.843±0.013	1.491
✗	✓	✗	0.858±0.015	1.472
✗	✗	✓	0.852±0.013	1.479
✗	✓	✓	0.862±0.016	1.470
✓	✗	✗	0.872±0.012	1.464
✓	✓	✗	0.880±0.017	1.451
✓	✗	✓	0.873±0.015	1.445
✓	✓	✓	**0.883±0.013**	**1.397**

b) Ablation study on Periodic Feature Modeling Capability of FFFM: To evaluate whether FFFM enhances periodic structure modeling, we conduct frequency-domain analyses on feature maps extracted after the first convolutional layer from

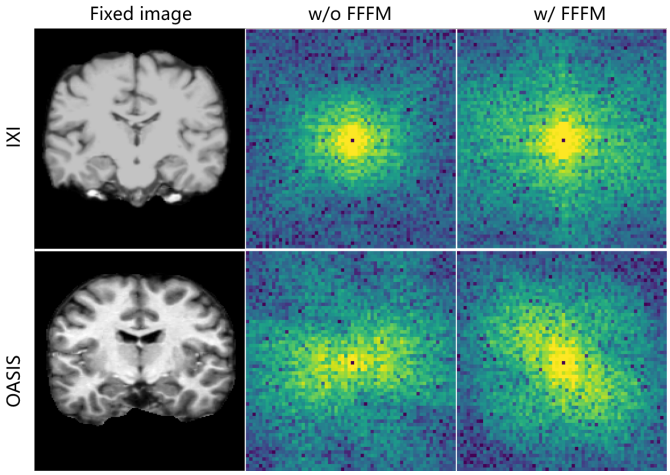

Fig. 4: A schematic illustration of the frequency-domain analysis on both datasets. The first column shows the fixed image, and the remaining columns present log-PSD heatmaps of first-layer feature maps before and after introducing FFFM.

models trained for 300 epochs. We compute PSD and log-PSD and visualize them with log-PSD heatmaps and radially averaged spectra. As shown in Fig. 4, conventional CNN features exhibit diffuse spectra without dominant frequencies, indicating weak frequency selectivity, whereas FFFM produces structured stripe-like patterns with stronger mid-frequency energy concentration, suggesting improved representation of repetitive anatomical structures.

TABLE III: Quantitative evaluation of periodic feature representation using Spectral Entropy (SE) and Peak Frequency Energy Concentration (PFEC) on IXI and OASIS for models with/without FFFM.

Model	IXI		OASIS	
	SE↓	PFEC↑	SE↓	PFEC↑
w/o FFFM	2.745±0.053	0.520±0.003	2.318±0.060	0.538±0.002
w/ FFFM	1.237±0.008	0.797±0.006	1.171±0.007	0.803±0.004

Moreover, we introduce two complementary metrics to quantify the learned features in the frequency domain. One is spectral entropy (SE) [36], which measures the global dispersion of the normalized spectrum, as defined in Eq. (15):

$$\mathrm{SE} = -\sum_{r} \frac{P(r)}{\sum_r P(r)} \log_2 \frac{P(r)}{\sum_r P(r)}, \tag{15}$$

where $P(r) = \frac{1}{N(r)} \sum_{(u,v)\in R(r)} \mathrm{PSD}(u,v,w)$, and PSD is computed as $\mathrm{PSD} = |\mathrm{FFT}(\mathbf{H})|^2$, where $\mathbf{H}$ denotes the feature map output from the first layer; r is the frequency radius, $R(r)$ denotes the frequency ring with radius r, and $N(r)$ is the number of frequency points within the ring. Higher SE indicates a more uniform frequency distribution.

However, SE alone cannot adequately characterize a dominant periodic component, so we propose Peak Frequency Energy Concentration (PFEC) to quantify spectral energy around the peak frequency. Unlike conventional PSD-based measures, PFEC emphasizes concentration around the dominant frequency and thus better reflects periodic anatomical structures. This property makes it useful not only for feature analysis in registration tasks, but also for other applications involving periodic textures or quality assessment of frequency-domain representations. PFEC is defined as:

$$\mathrm{PFEC} = \frac{\sum_{r_{\text{peak}}-\delta}^{r_{\text{peak}}+\delta} P(r)}{\sum_r P(r)} \tag{16}$$

where the maximum of the radial power spectrum is located at $r_{\text{peak}} = \arg\max_r P(r)$. The numerator computes the spectral energy within a local window $[r_{\text{peak}} - \delta, r_{\text{peak}} + \delta]$ centered around the peak frequency, while the denominator accounts for the total energy over all frequencies. This metric directly reflects the model's ability to focus on the dominant periodic frequency components.

As reported in Table. III, FFFM achieves lower SE and significantly higher PFEC than the baseline, indicating more concentrated spectral distributions and stronger emphasis on periodic components. All metrics yield p-values < 0.001, demonstrating statistical significance. Together with the log-PSD visualizations, these results validate the effectiveness of FFFM in modeling periodic anatomical structures.

c) Ablation Study on the Placement of FFFM within the Network: We also investigate the placement of FFFM within the network. As shown in Table. IV, this experiment focuses on Dice as the primary metric, since it is designed to evaluate how FFFM placement affects overall structural overlap rather than boundary-specific alignment. Inserting FFFM at the input layer yields the best performance, whereas its effectiveness decreases at deeper layers, indicating that periodic and high-frequency information is progressively attenuated by successive convolutions and should be modeled at the earliest stage.

TABLE IV: Dice scores with FFFM placed at different network layers.

Loc(FFFM)	1	2	3	4	5	6
Dice↑	**0.883**	0.878	0.874	0.873	0.873	0.872

V. Conclusion

We propose FAReg, an unsupervised deformable MRI registration framework that combines explicit frequency-domain modeling (FFFM) with adaptive boundary-aware optimization (EGDR) to improve dense displacement field estimation. FAReg captures quasi-periodic anatomical structures through learnable Fourier feature fusion and suppresses artifact-induced misalignment via spatially adaptive gradient reweighting. Extensive evaluation on two public brain MRI datasets, OASIS and IXI, shows that FAReg consistently achieves higher Dice overlap and lower boundary error (HD95) than state-of-the-art registration methods while maintaining competitive computational efficiency. FAReg still has limitations: the effectiveness of FFFM decreases in deeper layers due to attenuation of high-frequency signals, and EGDR requires careful tuning of threshold and sharpness across datasets and

imaging protocols. We hope this work encourages further integration of frequency-domain inductive bias and adaptive boundary modeling in learning-based medical image analysis.

ACKNOWLEDGMENT

This work was supported in part by the Research Funds for Universities of Heilongjiang Province under grant NO. 2022-KYYWF-1047, NO. 2025-KYYWF-ZR0387 and NO. 2025-KYYWF-ZR0425.

REFERENCES

[1] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE transactions on image processing*, vol. 5, no. 10, pp. 1435–1447, 1996.

[2] K. A. Ganser, H. Dickhaus, R. Metzner, and C. R. Wirtz, "A deformable digital brain atlas system according to talairach and tournoux," *Medical Image Analysis*, vol. 8, no. 1, pp. 3–22, 2004.

[3] V. K. Reed, W. A. Woodward, L. Zhang, E. A. Strom, G. H. Perkins, W. Tereffe, J. L. Oh, T. K. Yu, I. Bedrosian, G. J. Whitman *et al.*, "Automatic segmentation of whole breast using atlas approach and deformable image registration," *International Journal of Radiation Oncology* Biology* Physics*, vol. 73, no. 5, pp. 1493–1500, 2009.

[4] M. Cabezas, A. Oliver, X. Lladó, J. Freixenet, and M. B. Cuadra, "A review of atlas-based segmentation for magnetic resonance brain images," *Computer methods and programs in biomedicine*, vol. 104, no. 3, pp. e158–e177, 2011.

[5] P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy," *Neuroimage*, vol. 46, no. 3, pp. 726–738, 2009.

[6] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "An unsupervised learning model for deformable medical image registration," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9252–9260.

[7] J. Chen, E. C. Frey, Y. He, W. P. Segars, Y. Li, and Y. Du, "Transmorph: Transformer for unsupervised medical image registration," *Medical image analysis*, vol. 82, p. 102615, 2022.

[8] S. Sharma and Y. Zhang, "Fourier transform power spectrum is a potential measure of tissue alignment in standard mri: A multiple sclerosis study," *PLoS One*, vol. 12, no. 4, p. e0175979, 2017.

[9] J. Tohka, "Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review," *World journal of radiology*, vol. 6, no. 11, p. 855, 2014.

[10] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE transactions on image processing*, vol. 23, no. 2, pp. 684–695, 2013.

[11] S.-J. Chen, H.-L. Shen, C. Li, and J. H. Xin, "Normalized total gradient: A new measure for multispectral image registration," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1297–1310, 2017.

[12] H. Qiu, K. Hammernik, C. Qin, C. Chen, and D. Rueckert, "Embedding gradient-based optimization in image registration networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 56–65.

[13] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain," *Medical image analysis*, vol. 12, no. 1, pp. 26–41, 2008.

[14] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *International journal of computer vision*, vol. 61, no. 2, pp. 139–157, 2005.

[15] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ants similarity metric performance in brain image registration," *Neuroimage*, vol. 54, no. 3, pp. 2033–2044, 2011.

[16] J. Wang and M. Zhang, "Deepflash: An efficient network for learning-based medical image registration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 4444–4452.

[17] A. Hoopes, M. Hoffmann, B. Fischl, J. Guttag, and A. V. Dalca, "Hypermorph: Amortized hyperparameter learning for image registration," in *International conference on information processing in medical imaging*. Springer, 2021, pp. 3–17.

[18] B. Kim, D. H. Kim, S. H. Park, J. Kim, J.-G. Lee, and J. C. Ye, "Cyclemorph: cycle consistent unsupervised deformable image registration," *Medical image analysis*, vol. 71, p. 102036, 2021.

[19] X. Jia, W. Lu, X. Cheng, and J. Duan, "Decoder-only image registration," *IEEE Transactions on Medical Imaging*, 2025.

[20] X. Yang, D. Li, L. Deng, S. Huang, and J. Wang, "Tcde-net: An unsupervised dual-encoder network for 3d brain medical image registration," *Computerized Medical Imaging and Graphics*, vol. 123, p. 102527, 2025.

[21] X. Jia, J. Bartlett, W. Chen, S. Song, T. Zhang, X. Cheng, W. Lu, Z. Qiu, and J. Duan, "Fourier-net: Fast image registration with band-limited deformation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 1, 2023, pp. 1015–1023.

[22] X. Jia, A. Thorley, A. Gomez, W. Lu, D. Kotecha, and J. Duan, "Fourier-net+: Leveraging band-limited representation for efficient 3d medical image registration," *arXiv preprint arXiv:2307.02997*, 2023.

[23] M. Uteuliyeva, A. Zhumekenov, R. Takhanov, Z. Assylbekov, A. J. Castro, and O. Kabdolov, "Fourier neural networks: A comparative study," *Intelligent Data Analysis*, vol. 24, no. 5, pp. 1107–1120, 2020.

[24] L. Ziyin, T. Hartwig, and M. Ueda, "Neural networks fail to learn periodic functions and how to fix it," in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, ser. NIPS '20. Red Hook, NY, USA: Curran Associates Inc., 2020.

[25] Y. Dong, G. Li, Y. Tao, X. Jiang, K. Zhang, J. Li, J. Deng, J. Su, J. Zhang, and J. Xu, "Fan: Fourier analysis networks," *arXiv preprint arXiv:2410.02675*, 2024.

[26] R. Archibald and A. Gelb, "A method to reduce the gibbs ringing artifact in mri scans while keeping tissue boundary integrity," *IEEE Transactions on Medical Imaging*, vol. 21, no. 4, pp. 305–319, 2002.

[27] M. A. G. Ballester, A. P. Zisserman, and M. Brady, "Estimation of the partial volume effect in mri," *Medical image analysis*, vol. 6, no. 4, pp. 389–405, 2002.

[28] M. Fenski, J. Gröschel, P. Gatehouse, C. Kolbitsch, and J. Schulz-Menger, "Artifacts in cardiac t1 and t2 mapping techniques–influence on reliable quantification," *Journal of Cardiovascular Magnetic Resonance*, p. 101934, 2025.

[29] W. Mazur-Rosmus and A. T. Krzyżak, "The effect of elimination of gibbs ringing, noise and systematic errors on the dti metrics and tractography in a rat brain," *Scientific Reports*, vol. 14, no. 1, p. 15010, 2024.

[30] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*. Princeton university press, 1971, vol. 1.

[31] J. Duoandikoetxea, *Fourier analysis*. American Mathematical Society, 2024, vol. 29.

[32] S. Jeong and H. Y. Kim, "Convolutional fourier analysis network (cfan): A unified time-frequency approach for ecg classification," *arXiv preprint arXiv:2502.00497*, 2025.

[33] Y. Jia and T. Darrell, "Heavy-tailed distances for gradient based image descriptors," *Advances in neural information processing systems*, vol. 24, 2011.

[34] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults," *Journal of cognitive neuroscience*, vol. 19, no. 9, pp. 1498–1507, 2007.

[35] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[36] Y. Tian, H. Zhang, W. Xu, H. Zhang, L. Yang, S. Zheng, and Y. Shi, "Spectral entropy can predict changes of working memory performance reduced by short-time training in the delayed-match-to-sample task," *Frontiers in human neuroscience*, vol. 11, p. 437, 2017.