SBPR-Net: Synchronous Bidirectional Pyramid Registration Network for Medical Image Registration

1st Jialin Zheng
College of Electronics and Information Engineering
Shenzhen University
Shenzhen, China
2310433073@email.szu.edu.cn

2nd Wenming Cao
Guangdong Key Laboratory of Intelligent Information Processing
Shenzhen University
Shenzhen, China
wmcao@szu.edu.cn

3rd Deliang Lian
College of Electronics and Information Engineering
Shenzhen University
Shenzhen, China
liandl@szu.edu.cn

4th Zhiyue Yan*
Guangdong Key Laboratory of Intelligent Information Processing
Shenzhen University
Shenzhen, China
2150432014@email.szu.edu.cn

***Abstract*—In recent years, significant advancements have been achieved in deformable image registration. Unlike conventional approaches that directly estimate a flow field at the network output, this paper introduces a novel Synchronous Bidirectional Pyramid Registration Network (SBPR-Net) for unsupervised non-rigid registration. The proposed network implements a bridging registration strategy by simultaneously processing the moving and fixed images, generating two intermediate flow fields that map each image to a common virtual image. The final moving-to-fixed flow field is then derived indirectly through a parameter-free mathematical relationship between these intermediate flow fields. This design enables flexible and efficient deformation estimation of registration. To improve robustness, a novel constraint is introduced to handle virtual image uncertainty, while a pyramid architecture facilitates coarse-to-fine flow field prediction. Extensive experiments on OASIS and IXI brain MRI datasets demonstrate that SBPR-Net outperforms state-of-the-art methods, confirming the effectiveness of the bridging registration strategy.**

***Index Terms*—deformable image registration, bidirectional registration, deep convolutional neural network, unsupervised learning.**

I. Introduction

Deformable image registration is fundamental in medical image analysis, with applications in diagnostics, disease monitoring, and surgical navigation [1]–[3]. It aligns anatomical structures across images from different patients, time points, or scanners to aid clinical decisions, involving dense nonlinear correspondence estimation based on shared features. Unlike one-way correspondence methods, Avants *et al.* [4] introduced a symmetric registration framework that performs correspondence matching simultaneously from both images. Their method optimizes deformations from both the moving and fixed images toward a virtual image to accelerate registration. However, such traditional methods are often computationally intensive, limiting their practicality in supervised settings. Recently, unsupervised learning-based registration techniques [5]–[8] have gained attention due to their independence from manually annotated anatomical landmarks, which are costly to obtain. Methods such as VoxelMorph [5], VTN [9], and TransMorph [10] employ a U-Net architecture that combines encoder-decoder structures with skip connections to extract multi-scale features and directly predict the flow field at the network output. Nevertheless, these methods rely on direct regression of the flow field, which becomes challenging when handling complex deformations, particularly those involving large displacements.

Inspired by optical flow research on progressive large displacement refinement, Ranjan *et al.* [11] and Hui *et al.* [12] warp moving images/features at each pyramid level, while Miao *et al.* [8] use sequential warping for coarse-to-fine alignment. Such pyramid strategies predict sub-fields at each resolution to progressively align moving and fixed features via multi-stage registration. Integrating this with symmetric design [4], we hypothesize that driving both features toward an intermediate virtual image could enhance efficiency and stability. While some learning-based studies [13]–[15] adopt bidirectional concepts, they estimate direct flow fields between images (both directions) rather than dual-path registration toward a virtual image, leaving them inefficient for complex/large deformations.

In this paper, we propose a Synchronous Bidirectional Pyramid Registration Network (SBPR-Net) for unsupervised medical image registration. The network indirectly estimates the final flow field by predicting flows from both the moving and fixed images toward a common virtual image. Since the

Author marked with * is corresponding author. We acknowledge the support of this research by The Fundamental Research Foundation of Shenzhen under Grant No. JCYJ20230808105705012.

model outputs two intermediate flow fields—from moving-to-virtual and fixed-to-virtual—we establish a mathematical relationship to derive the desired moving-to-final field. The SBPR-Net is designed to perform bidirectional pyramid registration in a synchronous manner. To improve the efficiency and stability of the bridging registration strategy, we introduce a bridging constraint and a cosine constraint. The pyramid registration framework enables a progressive, coarse-to-fine estimation of flow fields. We evaluate the proposed method on two public brain MRI datasets, and extensive experiments demonstrate that SBPR-Net outperforms current state-of-the-art deformable registration methods, confirming its feasibility and effectiveness. The main contributions of our work are summarized as follows:

- We propose a bridging registration strategy where the network performs registration from both the moving and fixed images toward a virtual image and estimates the final flow field indirectly. A mathematical model is established to formulate the relationship among the flow fields involved in the bridging process.
- We introduce a bridging constraint and a cosine constraint to handle the uncertainty of the virtual image in the bridging registration model, which significantly improves the robustness and efficiency of the strategy.
- We validate the effectiveness of our approach on two widely-adopted MRI datasets. Experimental results demonstrate the superiority of our method over existing state-of-the-art approaches and verify the feasibility of the proposed framework.

II. Approach

In this chapter, we first investigate the feasibility of the bridging registration strategy by establishing a corresponding mathematical model to derive the relationship among u_{MV}, u_{FV} and u_{MF} — focusing particularly on whether the expression of u_{MF} can be derived from u_{MV} and u_{FV}. Next, we introduce the network architecture design of our bridging pyramid registration network. Finally, we analyze potential issues that may arise during the bridging registration process and present corresponding solutions in the loss function summary. Table I lists the symbols used in this section.

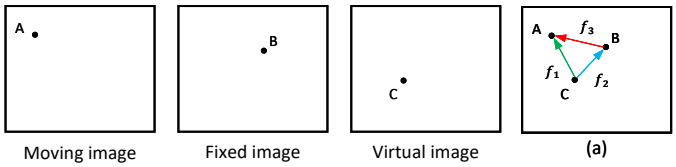

Fig. 1. Dots (A, B, C) in moving, fixed and virtual image represent the same structure but locate in three different position respectively. Figure (a) shows the spatial relationship of specific vectors in flow fields in the same volumetric domain.

A. Relationship among Flow Fields

As shown in Fig. 1, there are images respectively representing moving, virtual and fixed image. The dots in the three images represent the same structure but are located at three different locations (A, B and C) in these images. This means that the structure corresponding to the dots in any two of the three images should be spatially aligned after performing registration. For convenience of description, we put these three dots at different locations into a picture as shown in Fig. 1(a).

Considering the registration at Id_C, where we carry out the registration process from both the moving and fixed images to the virtual image. According to Fig. 1(a), the vectors f_1 and f_2 in their respective flow flied are satisfied as follows:

$$\begin{aligned} u_{MV}(Id_C) &= f_1 \\ u_{FV}(Id_C) &= f_2 \\ Id_A &= Id_C + f_1 \\ Id_B &= Id_C + f_2 \\ f_3 &= f_1 - f_2 \end{aligned} \tag{1}$$

Then, considering the registration at Id_B, where we carry out the registration from moving image register to fixed image, the vector f_3 in flow field u_{MF} is satisfied as Eq. (2). And we can obtain the substituting relationship as bellow:

$$\begin{aligned} u_{MF}(Id_B) &= f_3 \\ u_{MF}(Id_C + f_2) &= f_1 - f_2 \end{aligned} \tag{2}$$

$$u_{MF}(Id_C + f_2) = u_{MV}(Id_C) - u_{FV}(Id_C) \tag{3}$$

In Eq. (3), u_{MF} and $u_{MV} - u_{FV}$ can be regarded as the moving image and the warped moving image in the registration process, respectively. So, according to Spatial Transformer Network (STN) [16], we can derive the following expression:

$$u_{MV} - u_{FV} = STN(u_{MF}, u_{FV}). \tag{4}$$

Since our objective is to solve for u_{MF}, what we actually need is u_{VF} rather than u_{FV}. According to inverse registration [17], we can obtain u_{VF} by the inverse network as shown in Eq. (5). So we can obtain bridging registration function (BRE), which explains the relationship of u_{MV}, u_{FV} and u_{MF} during bridging registration strategy as shown in Eq. (6):

$$u_{VF} = STN(u_{FV}, -u_{FV}) \tag{5}$$

$$\begin{aligned} u_{MF} &= STN(u_{MV} - u_{FV}, u_{VF}) \\ &= STN(u_{MV} - u_{FV}, STN(u_{FV}, -u_{FV})) \end{aligned} \tag{6}$$

B. Network based on Bridging Registration Strategy

1) Encoder of SBPR-Net: According to the pyramid registration framework, the encoder stage performs feature extraction on the moving image I_{M_0} and fixed image I_{F_0} through separate processing streams. This yields multi-resolution feature maps $F_{M_{1-4}}$ and $F_{F_{1-4}}$ via successive encoder blocks. Since the encoder blocks serve identical purpose in this dual-stream architecture, we implement weight sharing across them. The detailed structure of the encoder block is illustrated in Fig. 3. Furthermore, we generate two sets of pooled images I_{M_i} and I_{F_i} ($i = 1$ to 4) by applying successive average pooling operations to the input moving and fixed images. The spatial dimensions of these pooled images are reduced to $\frac{1}{2^i}$ of the original image size.

TABLE I
INTERPRETATION OF SYMBOLS INVOLVED IN APPROACH.

Symbol	Description
A	A point in the moving image. Similarly, points B and C denote points in the fixed image and virtual image respectively. It is assumed that these three points (A, B, C) depict corresponding anatomical structures within their respective images.
Id_A	The coordinate of point A in the image. Similarly, Id_B and Id_C represent the coordinates of points B and C.
u_{MF}	The flow field from the moving image to the fixed image. Similarly, u_{MV} and u_{FV} are those from the moving image to the virtual image and from the fixed image to the virtual image, respectively.
f_i	The vector at a point in the flow field.
STN	The Spatial Transformer Network performs image resampling according to a prescribed flow field.
I_{M_i}	When i = 0, it denotes the original input moving image I_M. When i = 1, 2, 3, 4, it represents the images obtained by progressively downsampling the original moving image through average pooling operations. Desciption of symbol I_{F_i} is analogous to that of symbol I_{M_i}.
F_{M_i}	The features obtained by progressively downsampling the input moving image through encoder blocks. The subscript i denotes the downsampling stage. Description of symbol F_{F_i} is analogous to that of symbol F_{M_i}.
u_{MV_i}	The flow field from the moving image to the virtual image estimated at each level i of the pyramid registration framework. Description of symbol u_{FV_i} is analogous to that of symbol u_{MV}.
I_{WM_i}	I_{WM_i} denotes the warped image, which is the result of warping the corresponding pooled fixed image I_{M_i} with the flow field u_MV_i. Description of symbol I_{WF_i} is analogous to that of symbol I_{WM_i}.

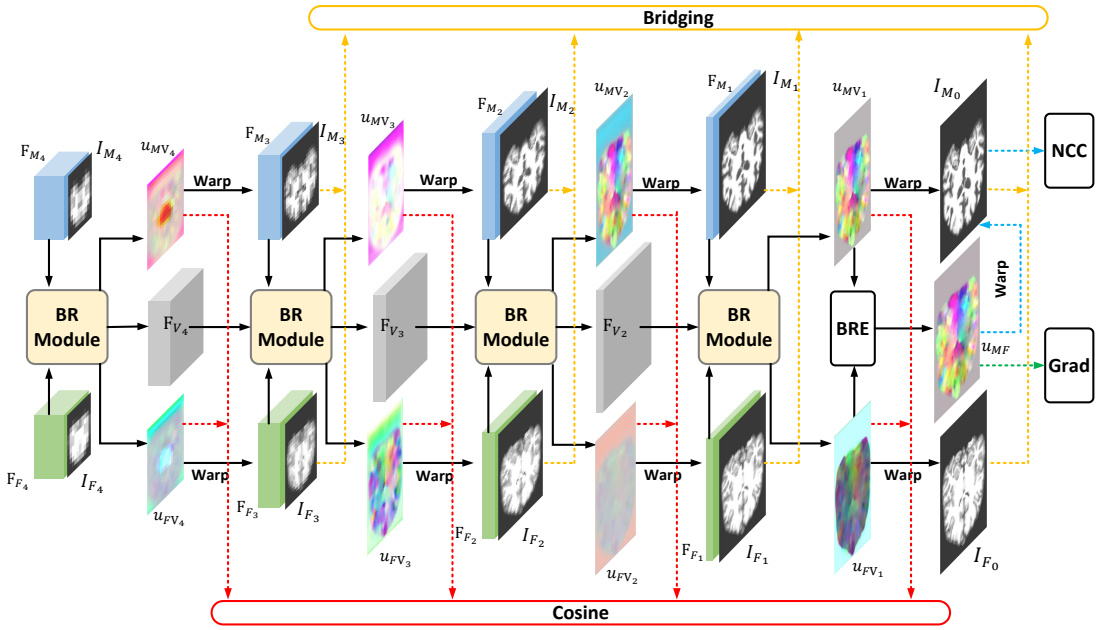

Fig. 2. Detail of the decoder of our proposed SBPR-Net. The blue blocks and their adjacent images in the diagram correspond to the moving features and pooled moving images from the encoder, while the green blocks and their associated images represent the fixed features and pooled fixed images. BRM represents the Bridging Registration Module. BRE stands for Bridging Registration Equation, referring to Eq. (6). The yellow, red, blue, and green dashed lines represent the data flow forming the Briding constraint, Cosine constraint, Similarity constraint, and Smoothness constraint in the loss function, respectively.

2) Decoder of SBPR-Net: This section details the decoder of our proposed SBPR-Net. The blue blocks in the diagram represent the moving image features F_{M_i} ($i = 1$ to 4) extracted from the encoder. Each F_{M_i} is concatenated with the pooled image I_{M_i} at the corresponding scale. Similarly, the green blocks represent the fixed image features F_{F_i} ($i = 1$ to 4), each of which is concatenated with its respective pooled image I_{F_i} at the same resolution level. The network comprises a 4-stage ($i = 4, 3, 2, 1$) pyramid structure, with each stage containing a Bridging Registration Module (BRM). As shown in Fig. 2, each BRM fuses the following input features: the moving image feature F_{M_i}, the fixed image feature F_{F_i}, and (when available) the virtual feature $F_{V_{i+1}}$ from the previous coarse stage. The module produces three outputs: the flow fields u_{MV_i} and u_{FV_i}, and a context-preserving virtual feature F_{V_i} (where $i = 2$ to 4). During forward propagation, the resulting flow fields u_{MV_i} and u_{FV_i} are applied to warp the moving feature F_{M_i} and fixed feature F_{F_i}, respectively. The warped features are then separately split out the corresponding warped pooled images I_{WM_i} and I_{WF_i} (where $i = 0$ to 3). These resulting images contribute to the Bridging Constraint term in the loss function, as indicated by the yellow dashed lines. Meanwhile, the flow fields u_{MV_i} and u_{FV_i} generated by each BRM are used to compute the Cosine Constraint loss, illustrated by the red dashed lines. Final flow field u_{MF} is generated by u_{MV_1} and u_{FV_1} through the BRE which represents Eq. (6). And

it also warps the moving image I_{M_0} to get warped moving image I_W for similarity calculation with I_{F_0} as shown by the blue dashed lines. In addition, the flow field u_{MF} will sent for smoothness constraint.

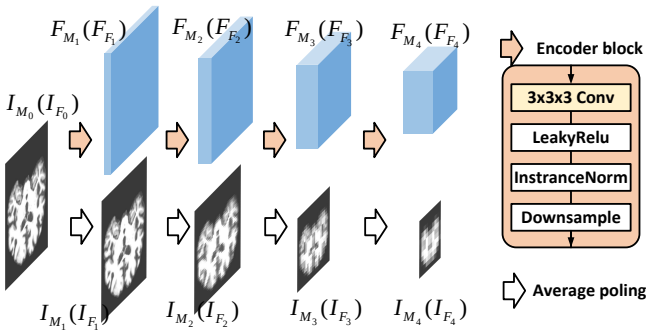

Fig. 3. Detail of the encoder of our proposed SBPR-Net, which is a pure CNN-based feature extractor, which contains the convolution (3x3x3 Conv), Instance Normalization (InstanceNorm), LeakyRelu and Average Pooling (Downsample) with a stride of 2. I_{M_i} and I_{F_i} ($i = 1$ to 3)are the pooled images obtained by successively applying average pooling to the original moving and fixed images.

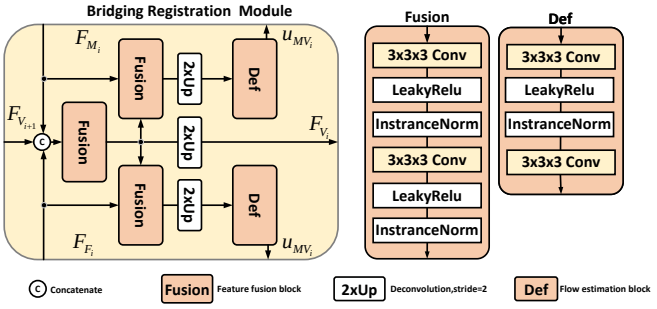

Fig. 4. Detail of the BRM, which consists feature fusion block (Fusion) and flow estimation block (Def). BRM accepts the feature maps F_{M_i}, F_{F_i} and $F_{V_{i+1}}$ as input and predicts the flow field u_{MV_i} and u_{FV_i} and virtual feature information F_{V_i} ($i = 4, 3, 2$).

3) Bridging Registration Module: Fig. 4 show the detail of Bridging Registration Module (BRM), which is responsible for fusing and generating the virtual feature F_{V_i}, as well as the flow fields u_{MV_4}, u_{FV_4} used for bidirectional pyramid structure registration. Features F_{M_4}, F_{F_4} and $F_{V_{i+1}}$ (if available) are concatenated and then fed into feature fusion block to predict virtual feature of current stage F_{F_i}. And then, F_{M_i} and F_{V_i} are used to predict the flow field u_{MV_i} to register pool moving image I_{M_i} to current virtual image I_{V_i}. Likewise, F_{F_i} and F_{V_i} are used to predict the flow field u_{FV_i} to register pool fixed image I_{F_i} to I_{V_i}.

C. Loss Functions

1) Bridging Constraint: Eq. (6) depends on an important condition, where the virtual images converted from fixed image and moving image should be the same. It means that they I_{WM_i}, I_{WF_i} should be able to achieve pixel-to-pixel equality, so we introduce the mean square error for each warped image pair from all decoder stages as follows:

$$\mathcal{L}_{bridging}(I_{WM_{0-3}}, I_{WF_{0-3}}) = \frac{1}{4\Omega}\sum_{i=0}^{3}\sum_{p\in\Omega}(I_{WM_i}(p) - I_{WF_i}(p))^2 \tag{7}$$

2) Cosine Constraint: The bridging registration strategy carries out two separate registration processes on the moving image and the fixed image simultaneously. And the bridging constraint will guarantee such processes to converge at a uncertain virtual image. On one hand, such uncertainty provides the model with a broader solution space, allowing for more flexible estimation of the target flow field. On the other hand, excessive uncertainty may compromise the model's registration performance, notably in terms of registration efficiency and the diffeomorphic property of the flow field, among other aspects. As shown in Fig. 5 (a), the red line is the most efficient way to register moving image to fixed image. Therefore, we prefer point C to be located on the red line in the figure rather than at positions like $C1$, $C2$, or $C3$. This is because such a configuration allows the bidirectional registration tasks starting from both the moving image and fixed image to reach the convergence point (virtual image) earlier, thereby completing the registration process more efficiently.

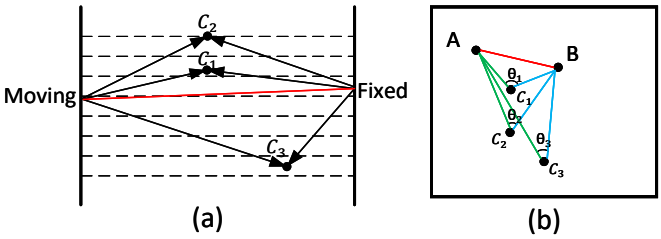

Fig. 5. Figure (a) shows that bridging registration strategy just likes the process of bridge-building, where the registration process is carried out simultaneously from both sides. Eventually the two registration tasks will converge at distinct middle states as C_1, C_2, and C_3, that correspond to the different virtual images The red line in figure (a) shows the most efficient way to register moving image to fixed image. Figure (b) illustrates the general impact of the virtual image in different states (C_1, C_2, C_3) on the vector relationships within the flow fields (u_{MV}, u_{FV}, u_{MF}). When the state of the virtual image becomes less plausible, the angle θ between corresponding vectors at the same position in flow fields(u_{MV}, u_{FV} should decrease accordingly.

Referring to the mathematical model depicted in Fig. 1(a) above, the geometric manifestation of virtual image uncertainty is illustrated in Fig 5(b). The angle θ in the figure represents the angle between vectors at corresponding spatial locations in flow field u_{MV_i} and flow field u_{MV_i}. A larger θ angle indicates that the corresponding virtual image is closer to the ideal state, where registration efficiency is highest. Therefore, we can incorporate the cosine theorem into the loss function to partially mitigate this uncertainty and the proposed

inverse-consistent constraint can be defined as follows:

$$\mathcal{L}_{cos}(u_{MV_{1-4}}, u_{MF_{1-4}}) = \frac{1}{4\Omega}\sum_{i=1}^{4}\sum_{p\in\Omega}\frac{\mu_{MV_i}(p)\mu_{FV_i}(p)}{|\mu_{MV_i}(p)||\mu_{FV_i}(p)|} \tag{8}$$

3) Similarity Constraint: For the $\mathcal{L}_{ncc}$, we adopt negative local normalized cross-correlation (ncc) [18] as the similarity metric to quantify the registration result for each inputting image pair in training. Specifically, the $\mathcal{L}_{ncc}$ between I_F and I_W is defined as follows:

$$\mathcal{L}_{sim}(I_F, I_W) = \\ -\sum_{p\in\Omega}\frac{(\sum_{p_i}(I_F(p_i)-\bar{I}_F(p))(I_W(p_i)-\bar{I}_W(p)))^2}{(\sum_{p_i}(I_F(p_i)-\bar{I}_F(p))^2)(\sum_{p_i}(I_W(p_i)-\bar{I}_W(p))^2)} \tag{9}$$

where the $\bar{I}_p$ denote the local mean intensity of image I in the location p, the p_i iterates over a n^3 neighboring region around p, with n = 9 in our setting.

4) Smoothness Constraint: To encourage the smoothness of the final flow field u_{MF}, we employ a diffusion regularizer which is defined as follows:

$$\mathcal{L}_{grad}(u_{MF}) = \sum_{p\in\Omega}||\nabla u_{MF}(p)||^2 \tag{10}$$

where $\nabla u_{MF}(p)$ is the gradient of the flow u_{MF} at the voxel p.

5) Objective Function: We trained the proposed bridging pyramid registration network by solving optimization problem as follows:

$$\min_{u_{MV_{1-4}}, u_{FV_{1-4}}} \mathcal{L}\begin{pmatrix} I_{WM_{1-3}}, I_{WF_{0-3}}, I_F, I_W, \\ u_{MF}, u_{MV_{1-4}}, u_{FV_{1-4}} \end{pmatrix} \tag{11}$$

$$\begin{aligned}&\mathcal{L}(I_{WM_{0-3}}, I_{WF_{0-3}}, I_F, I_W, u_{MF}, u_{MV_{1-4}}, u_{FV_{1-4}}) \\ &= \mathcal{L}_{sim}(I_F, I_W) \\ &+ \alpha\mathcal{L}_{grad}(u_{MF}) \\ &+ \beta\mathcal{L}_{bridging}(I_{WM_{0-3}}, I_{WF_{0-3}}) \\ &+ \gamma\mathcal{L}_{cos}(u_{MV_{1-4}}, u_{FV_{1-4}})\end{aligned} \tag{12}$$

where $\mathcal{L}_{ncc}$, $\mathcal{L}_{Grad}$, $\mathcal{L}_{bridging}$ and $\mathcal{L}_{cos}$ are similarity loss, smoothness loss, mean bridging loss and cosine loss respectively, and α, β and γ are hyper-parameter to control their contribution.

D. Experiments

E. Datasets

To evaluate SBPR-Net, we compared it with state-of-the-art methods on OASIS and IXI datasets, described below. **OASIS Dataset** [19]: Includes 405 T1-weighted MRI scans from cognitively normal subjects (18–96 years) with manual segmentations of 28 neuroanatomical regions. Scans were preprocessed via FreeSurfer (skull stripping, spatial normalization, affine registration) and cropped to 160×192×224 voxels, split into 255 training scans (64,770 pairs) and 150 test scans (725 randomly sampled pairs). **IXI Dataset** [20]: Contains 518 high-resolution MRI scans with annotations for 26 brain structures. With preprocessing identical to OASIS, volumes were resampled to 160×192×224 voxels, including 403 training scans (162,006 pairs) and 115 test scans (550 pairs).

F. Evaluation Metrics

We evaluate registration performance across three key aspects: similarity, diffeomorphic properties, and computational efficiency. For similarity, we use Dice Similarity Coefficient (DSC) to quantify spatial overlap between warped moving and fixed image segmentations (higher values indicate better alignment). 95th Percentile Hausdorff Distance (HD95) complements DSC by measuring boundary alignment robustness with reduced outlier sensitivity (lower values denote better performance). Diffeomorphic properties are evaluated via Negative Jacobian Determinant (NJD) following VoxelMorph-diff [21]. For efficiency, we measure inference time per image pair: traditional methods on a single-threaded CPU; deep learning methods via GPU inference (batch size=1). Time measurements exclude data loading/preprocessing to focus on computational efficiency.

G. Baseline Methods and Implementation Details

To validate the effectiveness of our method, we adopted the following popular methods as our comparison: SYN [4], VoxelMorph [22], Dual-PRNet [8], LapIRN [23], TransMorph [20], TransMatch [24], CycleMorph [25]. We implemented these representative methods and our BRP-Net with PyTorch on a NVIDIA GeForce RTX 3090 GPU card with 24 GB memory. For both the baseline methods and our proposed model, we consistently set the base feature channel dimension to 16 across all experiments to ensure fair comparison. We utilize the Adam optimizer for all methods with a learning rate of 0.0001 and a batch size of 1. We run SBPR-Net with parameters α as 0.5, β as 25, γ as 1 on $\mathcal{L}$. We set the epoch of experiment in OASIS and IXI to 20 and 20 respectively and obtained the model with the best performance.

H. Quantitative Registration Results and Analysis

We evaluate SBPR-Net against representative methods for inter-subject brain MRI registration on OASIS and IXI datasets: conventional SyN, U-shaped CNN VoxelMorph, pyramid-based Dual-PRNet/LapIRN, transformer-based TransMorph/TransMatch, and bidirectional CycleMorph. Table II shows SBPR-Net achieves the highest Dice scores and lowest HD95 values, validating the proposed bridging registration strategy and network architecture. Comparing with conventional SyN, SBPR-Net offers better accuracy with fast inference on both datasets. Compared to pyramid-based Dual-PRNet, it improves Dice from 78.34% to 80.42% on OASIS dataset and from 76.21% to 80.13% on IXI dataset. Furthermore, it reduces HD95 by 0.35 and 0.75 respectively. Comparing with transformer-based TransMatch, SBPR-Net is slightly inferior in diffeomorphic properties but surpasses it in accuracy and efficiency: Dice up 0.87% (OASIS) and 1.77% (IXI), inference time reduced from 0.54s to 0.36s. Comparing with bidirectional CycleMorph, SBPR-Net gains

TABLE II
QUANTITATIVE COMPARISON RESULTS OF DIFFERENT REGISTRATION METHODS ON THE OASI AND IXI DATASETS. THE BEST-PERFORMING OUTCOMES ARE HIGHLIGHTED IN **BOLD** FONT FOR EMPHASIS..

Methods	**OASIS**			**IXI**			
	DSC(%)↑	**HD95**↓	**NJD**(%)↓	**DSC**(%)↑	**HD95**↓	**NJD**(%)↓	**Times**↓
Affine	61.1	4.03	-	64.54	5.23	-	-
(2008-MIA)SyN	76.13	3.19	**0**	73.14	3.81	**0**	46.32
(2018-CVPR)VoxelMorph	77.06	2.99	1.98	74.98	3.66	1.80	**0.18**
(2019-ICCV)LapIRN	77.93	2.76	1.42	75.75	3.31	1.83	0.38
(2021-TMI)Dual-PRNet	78.34	2.71	1.80	76.21	2.96	1.90	0.23
(2022-MIA)TransMorph	78.75	2.62	0.56	76.15	2.73	0.82	0.37
(2024-TMI)TransMatch	79.55	2.44	0.02	78.36	2.22	0.03	0.54
(2021-MIA)CycleMorph	77.60	2.66	0.54	75.39	3.28	0.64	0.67
SBPR-Net	**80.42**	**2.36**	0.34	**80.13**	**2.21**	0.33	0.36

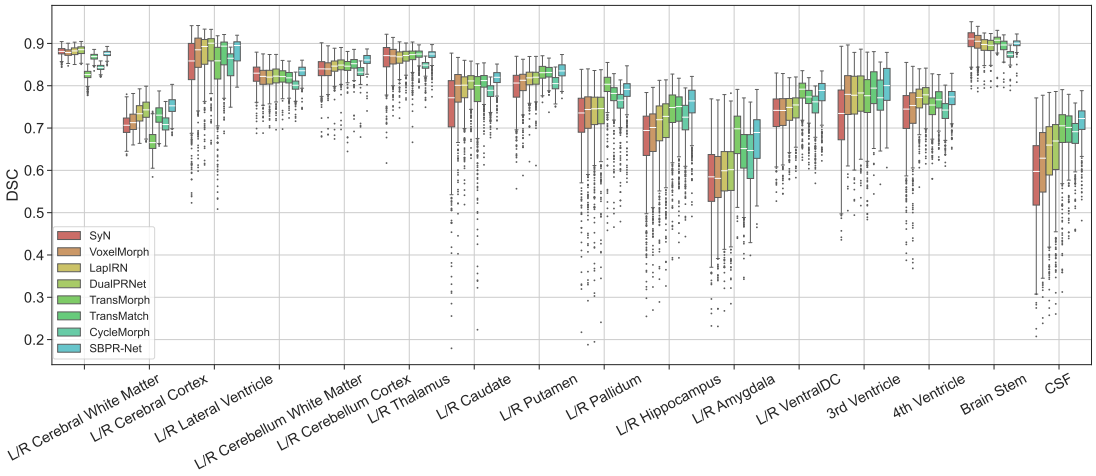

Fig. 6. Boxplots about the distributions of DSC over all anatomical structures on OASIS dataset produced by different registration methods. For clarity, some ROIs with symmetric structures on left and right sides are merged to be considered, resulting in the presented results for 16 distinct ROIs

2.82% (OASIS) and 4.74% (IXI) in Dice, and reduces HD95 from 2.36 to 2.66 (OASIS) and 2.21 to 3.28 (IXI), further confirming the bridging strategy's efficacy. For robustness assessment, Fig. 6 presents Dice boxplots across OASIS anatomical structures. SBPR-Net outperforms on both large regions (e.g., Cerebral White Matter, Cerebral Cortex) and small/complex ones (e.g., Cerebellum Cortex, Thalamus). Its consistent performance across scales highlights capability in handling global/local alignment, confirming robustness in deformable registration.

I. Qualitative Registration Results and Analysis

To qualitatively evaluate registration performance against baseline methods, we visualize intermediate and final alignment results from representative image pairs in the IXI dataset. Fig. 7 shows the warped moving images, corresponding segmentation labels, estimated flow fields, and difference maps between the original and warped moving images. The results demonstrate that our approach achieves superior alignment accuracy and produces smoother flow fields, with improved preservation of anatomical details and more realistic spatial transformations compared to other methods.

As shown in the first row of Fig. 7, the zoomed-in regions of the warped moving images generated by our method retain finer structural details and exhibit higher similarity to the fixed image. This observation is further supported by the difference maps (fourth row), where our results display a darker background and fewer red regions, indicating lower residual errors and more accurate alignment. These qualitative comparisons confirm the ability of our model to effectively capture local anatomical features during registration, leading to improved overall performance.

J. Ablation Study

1) Analysis of Cosine and Bridging Constraints: To validate the impact of the proposed cosine and bridging constraints that support our bridging registration strategy, we conducted four ablation studies on two datasets. These experiments investigate how different weighting coefficients for the constraint terms affect registration performance, with the similarity weight α fixed at 0.5. Results are summarized in Table III. Experimental evidence confirms the critical importance of the bridging constraint, which demonstrates the most significant impact on performance. When both constraints are ablated, registration accuracy experiences the most

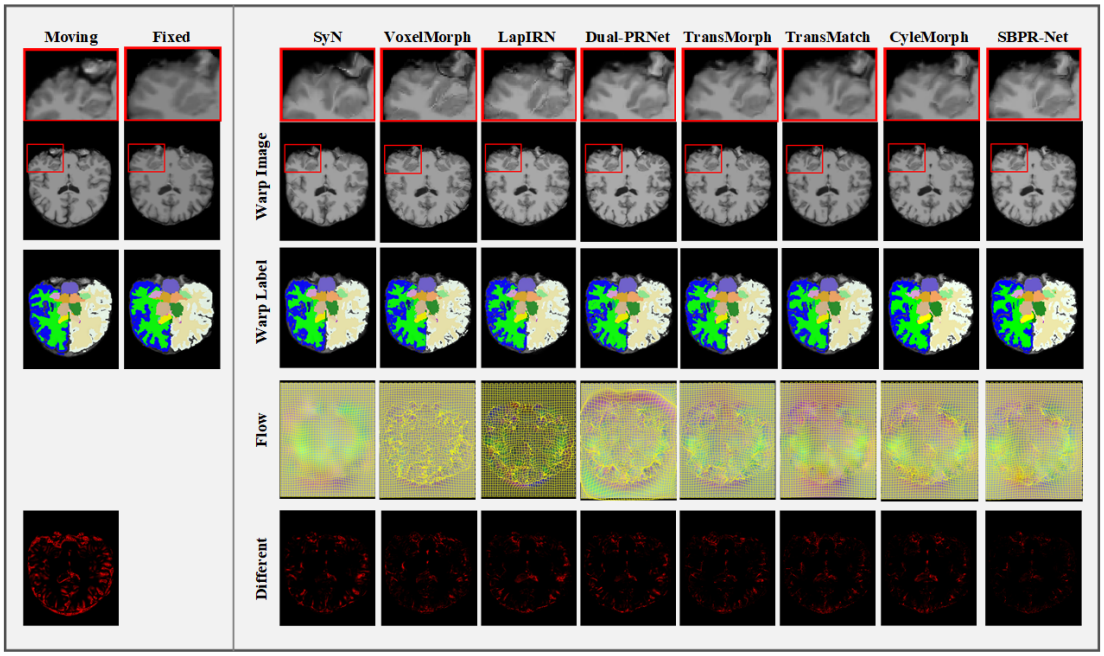

Fig. 7. Qualitative comparison of registration accuracy and corresponding flow fields from different methods on the IXI dataset. The left two columns show the original moving and fixed image pair. From top to bottom, the rows present: a zoomed region of the warped moving image, the full warped moving image, the warped moving segmentation label, the estimated flow field, and the difference map between the warped and fixed images. In the difference maps, darker backgrounds and fewer red regions indicate higher similarity and lower residual error.

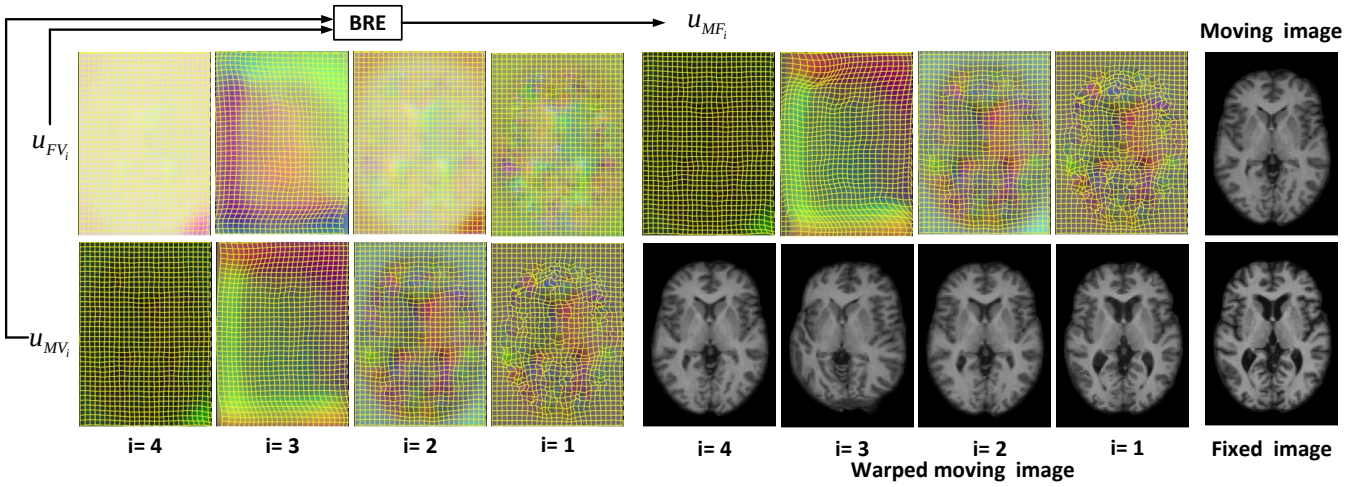

Fig. 8. Visualization of the flow fields u_{MV_i}, u_{FV_i}, and u_{MF_i} estimated via Eq. 6 at different decoder stages. BRE represent Eq. 6. Due to the varying resolutions of u_{MF_i} across stages, all u_{MF_i} fields are upsampled to the original moving image resolution. The bottom-right corner of the figure shows the corresponding warped moving images obtained by applying each upsampled u_{MF_i} to the moving image.

substantial degradation on the IXI dataset, with Dice score decreasing by 7.57%. This underscores its fundamental role as the prerequisite for the Bridging Registration Equation (BRE) to hold, justifying the relatively large weight assigned to this constraint during training. Although the cosine constraint shows a less pronounced effect than the bridging constraint, its absence still reduces Dice scores by 2.59% and 1.99% on the two datasets, respectively, compared to the full constraint configuration. This demonstrates its importance in enabling stable and efficient flow field estimation.

2) Visual Analysis for Validating the Bridging Registration Equation: To validate the effectiveness of the Bridging Registration Equation (BRE), we conducted experiments using moving and fixed image pairs from the OASIS dataset. The flow fields u_{MV_i} and u_{FV_i} were extracted from each decoder stage, and the corresponding u_{MF_i} was computed using the BRE formulation. Leveraging the bidirectional pyramid strategy in our SBPR-Net, which maintains pyramidal properties for coarse-to-fine registration, we upsampled all u_{MF_i} fields to the original image resolution and generated the corresponding

TABLE III

QUANTITATIVE EVALUATION RESULTS OF THE CONTRIBUTION OF COSINE CONSTRAINT AND BRIDGING CONSTRAINT. α, β AND γ ARE THE WEIGHTING COEFFICIENT IN LOSS FUNCTION, WHILE α WAS FIXED TO 0.5. DICE SCORE IS COMPUTED FOR ALL TEST SCANS ON THE OASIS AND IXI DATASETS. THE BEST-PERFORMING OUTCOMES ARE HIGHLIGHTED IN **BOLD** FONT FOR EMPHASIS.

β	γ	**DSC**(%)↑	
		OASIS	**IXI**
0	0	73.93	72.56
0	1	73.79	73.51
25	0	77.83	78.14
25	1	**80.42**	**80.13**

warped moving images. As shown in Fig. 8, both u_{MV_i} and u_{FV_i} contribute to u_{MF_i} estimation, with the network progressively refining the flow field u_{MF_i} through successive stages to achieve accurate image alignment.

III. CONCLUSION

This paper presents SBPR-Net, a novel unsupervised framework for deformable image registration that implements a bridging registration strategy. The core innovation lies in simultaneously processing the moving and fixed images through a dual-path architecture, where both are aligned to an intermediate virtual image, enabling indirect estimation of the final flow field via a parameter-free mathematical formulation. Extensive experiments on two public brain MRI datasets (OASIS and IXI) demonstrate that SBPR-Net achieves superior registration accuracy compared to state-of-the-art methods, while maintaining computational efficiency. The proposed approach effectively addresses challenges in handling large deformations and preserves anatomical details through its coarse-to-fine pyramid structure.

ACKNOWLEDGMENT

REFERENCES

[1] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE transactions on medical imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.

[2] G. Haskins, U. Kruger, and P. Yan, "Deep learning in medical image registration: a survey," *Machine Vision and Applications*, vol. 31, no. 1, p. 8, 2020.

[3] E. Ferrante and N. Paragios, "Slice-to-volume medical image registration: A survey," *Medical image analysis*, vol. 39, pp. 101–123, 2017.

[4] 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.

[5] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "Voxelmorph: a learning framework for deformable medical image registration," *IEEE transactions on medical imaging*, vol. 38, no. 8, pp. 1788–1800, 2019.

[6] J. Chen, Y. He, E. C. Frey, Y. Li, and Y. Du, "Vit-v-net: Vision transformer for unsupervised volumetric medical image registration," *arXiv preprint arXiv:2104.06468*, 2021.

[7] S. Zhao, Y. Dong, E. I. Chang, Y. Xu *et al.*, "Recursive cascaded networks for unsupervised medical image registration," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 10 600–10 610.

[8] M. Kang, X. Hu, W. Huang, M. R. Scott, and M. Reyes, "Dual-stream pyramid registration network," *Medical image analysis*, vol. 78, p. 102379, 2022.

[9] A. Dosovitskiy, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[10] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 10 012–10 022.

[11] A. Ranjan and M. J. Black, "Optical flow estimation using a spatial pyramid network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4161–4170.

[12] T.-W. Hui, X. Tang, and C. C. Loy, "Liteflownet: A lightweight convolutional neural network for optical flow estimation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8981–8989.

[13] L. Kong, T. Yang, L. Xie, D. Xu, and K. He, "Cascade connection-based channel attention network for bidirectional medical image registration," *The Visual Computer*, vol. 39, no. 11, pp. 5527–5545, 2023.

[14] T. C. Mok and A. Chung, "Fast symmetric diffeomorphic image registration with convolutional neural networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 4644–4653.

[15] 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.

[16] M. Jaderberg, K. Simonyan, A. Zisserman *et al.*, "Spatial transformer networks," *Advances in neural information processing systems*, vol. 28, 2015.

[17] J. Zhang, "Inverse-consistent deep networks for unsupervised deformable image registration," *arXiv preprint arXiv:1809.03443*, 2018.

[18] Y. R. Rao, N. Prathapani, and E. Nagabhooshanam, "Application of normalized cross correlation to image registration," *International Journal of Research in Engineering and Technology*, vol. 3, no. 5, pp. 12–16, 2014.

[19] 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.

[20] 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.

[21] A. V. Dalca, G. Balakrishnan, J. Guttag, and M. R. Sabuncu, "Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces," *Medical image analysis*, vol. 57, pp. 226–236, 2019.

[22] 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.

[23] T. C. Mok and A. C. Chung, "Large deformation diffeomorphic image registration with laplacian pyramid networks," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part III 23*. Springer, 2020, pp. 211–221.

[24] Z. Chen, Y. Zheng, and J. C. Gee, "Transmatch: a transformer-based multilevel dual-stream feature matching network for unsupervised deformable image registration," *IEEE transactions on medical imaging*, vol. 43, no. 1, pp. 15–27, 2023.

[25] 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.