Article

Partial Differential Equation-Constrained Diffeomorphic Registration from Sum of Squared Differences to Normalized Cross-Correlation, Normalized Gradient Fields, and Mutual Information: A Unifying Framework

Monica Hernandez [1,2,]*[iD] **, Ubaldo Ramon-Julvez** [2] **and Daniel Sierra-Tome** [2]

1   Aragon Institute of Engineering Research (I3A), 50018 Zaragoza, Spain
2   Department of Computer Sciences, University of Zaragoza (UZ), 50018 Zaragoza, Spain;
     ubi1616@gmail.com (U.R.-J.); sierradaniel.ind@gmail.com (D.S.-T.)
*   Correspondence: mhg@unizar.es

Abstract: This work proposes a unifying framework for extending PDE-constrained Large Deformation Diffeomorphic Metric Mapping (PDE-LDDMM) with the sum of squared differences (SSD) to PDE-LDDMM with different image similarity metrics. We focused on the two best-performing variants of PDE-LDDMM with the spatial and band-limited parameterizations of diffeomorphisms. We derived the equations for gradient-descent and Gauss–Newton–Krylov (GNK) optimization with Normalized Cross-Correlation (NCC), its local version (lNCC), Normalized Gradient Fields (NGFs), and Mutual Information (MI). PDE-LDDMM with GNK was successfully implemented for NCC and lNCC, substantially improving the registration results of SSD. For these metrics, GNK optimization outperformed gradient-descent. However, for NGFs, GNK optimization was not able to overpass the performance of gradient-descent. For MI, GNK optimization involved the product of huge dense matrices, requesting an unaffordable memory load. The extensive evaluation reported the band-limited version of PDE-LDDMM based on the deformation state equation with NCC and lNCC image similarities among the best performing PDE-LDDMM methods. In comparison with benchmark deep learning-based methods, our proposal reached or surpassed the accuracy of the best-performing models. In NIREP16, several configurations of PDE-LDDMM outperformed ANTS-lNCC, the best benchmark method. Although NGFs and MI usually underperformed the other metrics in our evaluation, these metrics showed potentially competitive results in a multimodal deformable experiment. We believe that our proposed image similarity extension over PDE-LDDMM will promote the use of physically meaningful diffeomorphisms in a wide variety of clinical applications depending on deformable image registration.

Keywords: PDE-LDDMM; Gauss–Newton–Krylov; optimal control optimization; band-limited vector fields; normalized cross-correlation; normalized gradient fields; mutual information

1. Introduction

In the past two decades, diffeomorphic registration has become a fundamental problem in medical image analysis [1]. The diffeomorphic transformations estimated from the solution of the image registration problem constitute the inception point in Computational Anatomy studies for modeling and understanding population trends and longitudinal variations, and for establishing relationships between imaging phenotypes and genotypes in Imaging Genetics [2–8]. Moreover, diffeomorphic registration can be as useful as any other deformable image registration framework in the fusion of multi-modal information from different sensors, the capture of correlations between structure and function, the guidance of computerized interventions, and many other applications [9–11].

A relevant issue in deformable image registration is the quest for the most sensible transformation model for each clinical domain. On the one hand, there are domains where

the underlying biophysical model of the transformation is known. The incompressible motion of the healthy heart is a relevant example [12]. On the other hand, there are also important clinical contexts where the deformation model is not known, although there is active research on finding the most plausible transformation among those explained by a physical model [13–15]. The most relevant examples are the deformation between healthy and diseased brains or the longitudinal evolution of the brain changes in healthy and diseased individuals.

Although the differentiability and invertibility of the diffeomorphisms constitute fundamental features for Computational Anatomy, the diffeomorphic constraint does not necessarily guarantee that a transformation computed with a given method is physically meaningful for the clinical domain of interest. In order to obtain physically meaningful diffeomorphisms, the diffeomorphic registration methods should be able to impose a plausible physical model to the computed transformations.

PDE-constrained LDDMM (PDE-LDDMM) registration methods have arisen as an appealing paradigm for computing diffeomorphisms under plausible physical models [13–22]. In the cases where the model is known, the PDE-LDDMM formulation allows the introduction of the priors of the particular model [23–26]. The PDE-constrained formulation is also helpful in the quest of plausible transformations for a clinical application with an unknown deformation model [15]. In addition, PDE-LDDMM is a well-suited approach for the estimation of registration uncertainty [27,28].

The different PDE-LDDMM methods differ on the variational problem formulation, diffeomorphism parameterization, regularizers, image similarity metrics, optimization methods, and additional PDE constraints. From them, the use of Gauss–Newton–Krylov optimization [13,14,22], the addition of nearly incompressible terms in the variational formulation [14,19], the use of variants involving the deformation state equation [18,22], and the introduction of the band-limited parameterization and HPC or GPU implementations [19,20,22,29], constitute the most successful contributions to the realistic and efficient computation of physically meaningful diffeomorphisms so far. Some attention has been given to the image similarity metric, where the sum of squared differences (SSD) between the final state variable and the target image has been mostly used [13,16,22]. Only two variations of PDE-LDDMM with normalized gradient fields (NGFs) and mutual information (MI) have been proposed in [18,30] (*ArXiv* paper). These two methods use gradient-based techniques for optimization.

SSD is based on image subtraction, so it is only well suited in uni-modal registration for images where the intensity of the reciprocal structures do not vary much. Indeed, SSD is not robust to noise, intensity inhomogeneity, and partial volume effects. Moreover, SSD is not suitable for multi-modal registration, even for transformation models with a few degrees of freedom such as those in rigid or affine image registration [31,32]. Therefore, there is a need for PDE-LDDMM methods, preferably with Gauss–Newton–Krylov optimization, which, apart from SSD, can support alternative image similarity metrics better behaved than SSD. The solution to this problem will promote the use of PDE-LDDMM in a wide variety of clinical applications depending on deformable image registration where images are acquired from different sensors.

This work proposes a unifying framework for introducing different image similarity metrics in the two best-performing variants of PDE-LDDMM [22,33]. From the Lagrangian variational problems, we have identified that a change in the image similarity metric involves changing the initial adjoint and the initial incremental adjoint variables. We have derived the equations of these variables needed for gradient-descent and Gauss–Newton–Krylov optimization with Normalized Cross Correlation (NCC), its local version (lNCC), Normalized Gradient Fields (NGFs), and Mutual Information (MI).

NCC, lNCC, and MI have accompanied SSD in different deformable image registration methods since their inception [34]. NGFs is an interesting metric for a wide variety of deformable registration problems [35]. These metrics are available in relevant deformable image registration packages such as the Insight Toolkit (www.itk.org, accessed on 1

January 2022), NiftyReg (https://sourceforge.net/projects/niftyreg, accessed on 1 January 2022), or Fair [32], among others. In the framework of diffeomorphic image registration, ANTS registration was implemented for SSD, lNCC, and MI (www.nitrc.org/projects/ants, accessed on 1 January 2022) and the performance of the different image similarity metrics was evaluated in [36]. We have selected these metrics as a starting point, although our framework is extensible to other image similarity metrics proposed in the literature, provided that the first and second-order variations of the image similarity metric can be written in the expected form [37,38].

Our experiments focused on the spatial (SP) and band-limited (BL) stationary parameterization of diffeomorphisms, although the proposed methods can be straightforwardly extended to the non-stationary parameterization [39]. We have obtained successful Gauss–Newton–Krylov methods for SSD, NCC, and lNCC with evaluation results greatly overpassing gradient-descent and competing with the respective version of ANTS diffeomorphic registration [40]. For NGFs, the second-order method did not provide satisfactory results in comparison with gradient-descent. For MI, the memory load of the second-order method hindered a proper evaluation in 3D datasets. We extensively studied the performance of our methods in NIREP16 and Klein et al.'s evaluation frameworks [41,42], obtaining an interesting insight into the impact of the different image similarity metrics in the PDE-LDDMM framework.

Since the advances that made it possible to learn the optical flow using convolutional neural networks (FlowNet [43]), dozens of deep-learning data-based methods were proposed to approach the problem of deformable image registration in different clinical applications [44]. Some of them are specifically devised for diffeomorphic registration where the different LDDMM ingredients are used as a backbone for diffeomorphism parameterization and the definition of the loss functions [45–54]. These methods use SSD and NCC metrics in the image similarity loss function, and the proposed models are usually limited to a single modality where the appearance of the image pairs needs to be similar to the training data. From them, only SynthMorph proposed a model valid for multi-modal registration through the extensive generation of simulated data for training, which yields a fast inference for diffeomorphism computation once the difficulties with training have been overcome [53].

Although the authors of SynthMorph provided an extensive study on the generalization capability of their models in different multi-modal experiments, the loss function is restricted to the Dice Similarity Coefficient (DSC) on image segmentations and, therefore, interesting questions such as the actual ability of deep-learning models to deal with multimodality or influence of the image similarity loss function on the registration accuracy have not been answered. In addition, results from the Learn2Reg challenge question the superiority of deep learning approaches and open new research directions into hybrid methods for which contributions to traditional optimization-based methods like ours may be of interest [55].

In the following, Section 2 reviews the foundations of PDE-LDDMM and BL PDE-LDDMM, from the original variant proposed in [13,16] to the variants used in this work. Section 4.5 analyzes the change of image similarity metrics in PDE-LDDMM and derives the equations needed for gradient-descent and Gauss–Newton–Krylov optimization for the considered metrics. Section 4 gathers the experimental setup for the evaluation of the methods, the numerical and implementation details of the proposed methods and the benchmarks. Section 5 shows the evaluation results. Finally, Section 7 gathers the most remarkable conclusions of our work.

2. PDE-Constrained LDDMM

2.1. LDDMM

LDDMM was proposed by Beg et al. in [56]. In this section, we recall the most relevant aspects of this interesting method, which underpin PDE-LDDMM. Let $\Omega \in \mathbb{R}^d$, $d = 2, 3$ be the image domain. The LDDMM registration problem is formulated between the source

and the target images, $I_0 : \Omega \to \mathbb{R}$ and $I_1 : \Omega \to \mathbb{R}$. These images are square-integrable functions on Ω. $Diff(\Omega)$ represents the Riemannian manifold of smooth diffeomorphisms on Ω. The tangent space at the identity diffeomorphism is denoted with V. The Riemannian metric of $Diff(\Omega)$ is defined from a scalar product in V

$$\langle v, w \rangle_V = \langle Lv, w \rangle_{L^2} = \int_\Omega \langle Lv(x), w(x) \rangle d\Omega, \tag{1}$$

through the invertible self-adjoint operator $L = (Id - \alpha \Delta)^s, \alpha > 0, s \in \mathbb{N}$ with inverse K.

LDDMM aims at finding a smooth map $\varphi : \Omega \to \mathbb{R}^d$ with smooth inverse such that the warped initial image $I_0 \circ \varphi$ is non-rigidly aligned with I_1. The diffeomorphism φ is parameterized in the tangent space V from a time-varying vector field flow $v : \Omega \times [0,1] \to \mathbb{R}^d$ and a path in $Diff(\Omega)$ $\phi : \Omega \times [0,1] \to \mathbb{R}^d$ such that v and ϕ satisfy the PDE

$$\partial_t \phi(t) + D\phi(t) \cdot v_t = 0 \tag{2}$$

with initial condition $\phi(0) = id$. It holds that $\varphi = \phi(1)$.

The solution to the registration problem is obtained from the minimization of a variational problem

$$E(v) = \frac{1}{2} E_{\text{reg}}(v) + \frac{1}{\sigma^2} E_{\text{img}}(I_0, I_1, \varphi), \tag{3}$$

where $E_{\text{reg}}(v)$ is the regularization term, $E_{\text{img}}(I_0, I_1, \varphi)$ is the image similarity metric that quantifies the differences between $I_0 \circ \varphi$ and I_1 after registration, and σ weights the contribution of both terms to the total energy.

2.2. Original PDE-Constrained LDDMM

PDE-constrained LDDMM (PDE-LDDMM) was originally formulated as a constrained variational problem from the minimization of

$$E(v) = \frac{1}{2} \int_0^1 \langle v_t, v_t \rangle_V dt + \frac{1}{\sigma^2} \|m(1) - I_1\|_{L^2}^2, \tag{4}$$

subject to

$$\partial_t m(t) + \nabla m(t) \cdot v_t = 0, \tag{5}$$

with initial condition $m(0) = I_0$ [14,16]. The solution of Equation (5), $m(1)$, is the warped initial image. The image similarity metric is the sum of squared differences (SSD) between the intensities of $m(1)$ and I_1.

Although in the great majority of LDDMM methods the optimization is approached with gradient-descent [16–18,40,56–58], Gauss–Newton–Krylov optimization has emerged as the method of choice for PDE-LDDMM due to the excellent numerical accuracy and the extraordinarily fast convergence rate [13,14,19]. The first and second-order differentials of the PDE-constrained variational problem are computed using the method of Lagrange multipliers, as follows.

Let us define the Lagrange multipliers $\lambda : \Omega \times [0,1] \to \mathbb{R}$ and $\eta : \Omega \to \mathbb{R}$ associated with Equation (5) and its initial condition. The Lagrangian functional corresponds to the expression

$$E_{\text{Lag}}(v) = E(v) + \int_0^1 \langle \lambda(t), \partial_t m(t) + Dm(t) \cdot v_t \rangle_{L^2} dt + \langle \eta, m(0) - I_0 \rangle_{L^2}. \tag{6}$$

The first-order variation of the Lagrangian yields the expression of the gradient

$$\nabla_v E_{\text{Lag}}(v) = Lv + \lambda \cdot \nabla m, \tag{7}$$

where

$$\partial_t m(t) + \nabla m(t) \cdot v_t = 0 \tag{8}$$

$$-\partial_t \lambda(t) - \nabla \cdot (\lambda(t) \cdot v_t) = 0, \tag{9}$$

subject to the initial and final conditions $m(0) = I_0$ and $\lambda(1) = -\frac{2}{\sigma^2}(m(1) - I_1)$. Equation (4) subject to Equation (5) is referred to as an optimal control problem, where v is the control variable, Equation (8) is the state equation and Equation (9) is the adjoint equation.

The second-order variation of the Lagrangian functional yields the expression of the Hessian-vector product, written in Gauss–Newton positive-definite approximated form,

$$H_v E_{\text{Lag}}(v)\delta v = L\delta v + \delta\lambda \cdot \nabla m, \tag{10}$$

where

$$-\partial_t \delta\lambda(t) - \nabla \cdot (\delta\lambda(t) \cdot v_t) = 0, \tag{11}$$

with final condition $\delta\lambda(1) = -\frac{2}{\sigma^2}\delta m(1)$. The variation of m, δm, satisfies the PDE with initial condition $\delta m(0) = 0$

$$\partial_t \delta m(t) + \nabla\delta m \cdot v_t + \nabla m \cdot \delta v = 0. \tag{12}$$

Optimization using gradient-descent in V is driven by the update equation

$$v^{n+1} = v^n - \epsilon K \nabla_v E_{\text{Lag}}(v), \tag{13}$$

while Gauss–Newton–Krylov optimization yields the update equation

$$v^{n+1} = v^n + \epsilon\delta v^n, \tag{14}$$

where δv^n is computed from preconditioned conjugate gradient (PCG) on the system

$$H_v E_{\text{Lag}}(v^n)\delta v^n = -\nabla_v E_{\text{Lag}}(v^n), \tag{15}$$

where $H_v E_{\text{Lag}}(v^n)$ is the positive-definite Hessian approximation. The preconditioner used in this work is K [13].

2.3. Variants of PDE-LDDMM

The original PDE-LDDMM has been recently completed with two alternative variants [22]. They can be considered theoretically but not numerically equivalent formulations of the original one. These variants have been shown to improve the original formulation in terms of registration accuracy and efficiency, in combination with the use of the band-limited parameterization and GPU implementation [20,22,33].

2.3.1. Variant I

The first variant departs from the original variational formulation (Equation (4)) by replacing the solution of the state equation (Equations (5) and (8)) with the identity

$$m(t) = I_0 \circ \phi(t), \tag{16}$$

where $\phi(t)$ is computed from the solution of the deformation state equation

$$\partial_t \phi(t) + D\phi(t) \cdot v_t = 0 \tag{17}$$

with initial condition $\phi(0) = id$. Analogously, the solution of the adjoint equation (Equation (9)) is replaced with the identity

$$\lambda(t) = |D\phi^{-1}(t)|\lambda(1) \circ \phi^{-1}(t), \tag{18}$$

and the solution of the incremental adjoint equation (Equation (11)) is replaced with the identity

$$\delta\lambda(t) = |D\phi^{-1}(t)|\delta\lambda(1) \circ \phi^{-1}(t) \cdot \delta\phi^{-1}(t). \tag{19}$$

The inverse diffeomorphism $\psi(t) = \phi^{-1}(t)$ and the corresponding Jacobian determinant $J(t) = |D\phi^{-1}(t)|$ are computed, respectively, from the inverse deformation state equation and the inverse Jacobian equation

$$-\partial_t \psi(t) - D\psi(t) \cdot v_t = 0 \tag{20}$$
$$-\partial_t J(t) - v_t \cdot \nabla J(t) = -J(t)\nabla \cdot v_t \tag{21}$$

with final conditions $\psi(1) = id$ and $J(1) = 1$. These identities were proposed in [16,22,56] and effectively used in [22].

2.3.2. Variant II

The second variant consists in replacing the state equation (Equation (5)) by the deformation state equation (Equation (17)) in the original variational formulation (Equation (4)) [18,22]. In this case, the Lagrangian corresponds to

$$E_{\text{Lag}}(v) = E(v) + \int_0^1 \langle \rho(t), \partial_t \phi(t) +$$

$$D\phi(t) \cdot v_t \rangle_{L^2} dt + \langle \mu, \phi(0) - id \rangle_{L^2}, \quad (22)$$

where the Lagrange multipliers are $\rho : \Omega \times [0,1] \to \mathbb{R}^d$, associated with the deformation state equation, and $\mu : \Omega \to \mathbb{R}^d$, associated with its initial condition.

For Variant II, the first-order variation of the Lagrangian yields the expression of the gradient

$$\nabla_v E_{\text{Lag}}(v) = Lv + D\phi \cdot \rho, \tag{23}$$

where

$$\partial_t \phi(t) + D\phi(t) \cdot v_t = 0 \tag{24}$$
$$-\partial_t \rho(t) - \nabla \cdot (\rho(t) \cdot v_t) = 0 \tag{25}$$

subject to the initial and final conditions $\phi(0) = id$, and $\rho(1) = \lambda(1) \cdot \nabla m(1)$.

The second-order variation of the Lagrangian functional yields the expression of the Hessian-vector product (in Gauss–Newton approximated form)

$$H_v E_{\text{Lag}}(v)\delta v = L\delta v + D\delta\phi \cdot \rho, \tag{26}$$

where

$$\partial_t \delta\phi(t) + D\delta\phi(t) \cdot v_t + D\phi(t) \cdot \delta v(t) = 0 \tag{27}$$
$$-\partial_t \delta\rho(t) - \nabla \cdot (\delta\rho(t) \cdot v_t) = 0 \tag{28}$$

subject to $\delta\phi(1) = 0, \delta\rho(1) = \delta\lambda(1) \cdot \nabla m(1)$.

2.4. Band-Limited PDE-LDDMM

The band-limited (BL) parameterization of diffeomorphisms was proposed by Zhang et al. [58–60]. In this section, we recall the most relevant aspects of this parameterization and then describe BL PDE-LDDMM. Let $\widetilde{\Omega}$ be the discrete Fourier domain truncated with frequency bounds $K_1, \ldots, K_d$. We denote with $\widetilde{V}$ the space of discretized band-limited vector fields on Ω with these frequency bounds. The elements in $\widetilde{V}$ are represented in the Fourier domain as $\tilde{v} : \widetilde{\Omega} \to \mathbb{C}^d$, $\tilde{v}(k_1, \ldots, k_d)$, and in the spatial domain as $\iota(\tilde{v}) : \Omega \to \mathbb{R}^d$, $\iota(\tilde{v})(x_1, \ldots, x_d)$. The application $\iota : \widetilde{V} \to V$ denotes the natural inclusion mapping of $\widetilde{V}$ in V. The application $\pi : V \to \widetilde{V}$ denotes the projection of V onto $\widetilde{V}$.

We denote with $Diff(\widetilde{\Omega})$ the finite-dimensional Riemannian manifold of diffeomorphisms on $\widetilde{\Omega}$ with corresponding Lie algebra $\widetilde{V}$. The Riemannian metric in $Diff(\widetilde{\Omega})$ is defined from the scalar product

$$\langle \tilde{v}, \tilde{w} \rangle_{\widetilde{V}} = \langle \check{L}\tilde{v}, \tilde{w} \rangle_{l^2}, \tag{29}$$

where $\check{L}$ is the projection of operator L in the truncated Fourier domain. Similarly, we will denote with $\tilde{Y}$ the projection of any linear operator Y in the truncated Fourier domain. In addition, we will denote with $\star$ the truncated convolution.

The BL PDE-LDDMM SSD variational problem is given by the minimization of

$$E(\tilde{v}) = \frac{1}{2} \int_0^1 \langle \check{L}\tilde{v}_t, \tilde{v}_t \rangle_{l^2} dt + \frac{1}{\sigma^2} \|m(1) - I_1\|_{L^2}^2 \tag{30}$$

subject to the state equation

$$\partial_t m(t) + \nabla m(t) \cdot \iota(\tilde{v}_t) = 0 \tag{31}$$

with initial condition $m(0) = I_0$.

The expression of the gradient is computed in the space of band-limited vectors yielding

$$\widetilde{\nabla_{\tilde{v}} E_{\text{Lag}}}(\tilde{v}) = \check{L}\tilde{v} + \pi(\lambda \cdot \nabla m). \tag{32}$$

The expression of the Hessian-vector product is computed analogously, yielding

$$\widetilde{H_{\tilde{v}} E_{\text{Lag}}}(\tilde{v})\delta\tilde{v} = \check{L}\delta\tilde{v} + \pi(\delta\lambda \cdot \nabla m). \tag{33}$$

Optimization using gradient-descent is driven by the update equation

$$\tilde{v}^{n+1} = \tilde{v}^n - \epsilon \widetilde{K} \widetilde{\nabla_{\tilde{v}} E_{\text{Lag}}}(\tilde{v}), \tag{34}$$

while Gauss–Newton–Krylov optimization yields the update equation

$$\tilde{v}^{n+1} = \tilde{v}^n + \epsilon \delta\tilde{v}^n, \tag{35}$$

where $\delta\tilde{v}^n$ is computed from preconditioned conjugate gradient (PCG) on the Hessian-gradient system defined in the BL domain.

2.4.1. BL Variant I

The BL version of Variant I is obtained with the identity $m(t) = I_0 \circ \phi(t)$, where $\phi(t) = id - \iota(\tilde{u}(t))$ and $\tilde{u}(t)$ is the solution of the BL deformation state equation

$$\partial_t \tilde{u}(t) + \widetilde{D}\tilde{u}(t) \star \tilde{v}_t = \tilde{v}_t \tag{36}$$

with initial condition $\tilde{u}(0) = \pi(0_V)$. Analogously, $\lambda(t)$ and $\delta\lambda(t)$ are obtained from Equations (18) and (19), where the inverse diffeomorphism and its Jacobian are computed from the inverse deformation state equation and the inverse Jacobian equation defined in the space of BL vector fields. The details can be found in [22].

2.4.2. BL Variant II

The BL version of Variant II is given by the minimization of Equation (30) subject to the BL deformation state equation (Equation (36)). The gradient and the Hessian of the Lagrangian are given by

$$\widetilde{\nabla_{\tilde{v}} E_{\text{Lag}}}(\tilde{v}) = \check{L}\tilde{v} + \tilde{\rho} - \widetilde{D}\tilde{u} \star \tilde{\rho} \tag{37}$$

$$\widetilde{H_{\tilde{v}} E_{\text{Lag}}}(\tilde{v})\delta\tilde{v} = \check{L}\delta\tilde{v} + \delta\tilde{\rho} - \widetilde{D}\delta\tilde{u} \star \tilde{\rho}. \tag{38}$$

The PDE equations involved in the computation of the gradient and Hessian are the convenient definitions of the spatial PDE equations in the band-limited domain. The specific details can be found in [22]. From them, the most relevant ones to recall are $\tilde{\rho}(1) = \pi(\lambda(1) \cdot \nabla m(1))$ and $\delta\tilde{\rho}(1) = \pi(\delta\lambda(1) \cdot \nabla m(1))$.

3. Extending PDE-LDDMM from SSD to NCC, lNCC, NGFs, and MI Image Similarity Metrics

3.1. Changing the Image Similarity Metric in PDE-LDDMM

From the analysis of the equations involved in the computation of Variants I and II with SSD, a change in the image similarity term for PDE-LDDMM supposes to recompute the expressions of $\lambda(1)$ and $\delta\lambda(1)$ for the given metric. This is valid also for Variant II since the adjoint variable $\rho(1)$ depends on $\lambda(1)$. The BL versions of Variants I and II do also depend on $\lambda(1)$ and $\delta\lambda(1)$.

Following the ideas in [32], our image similarity terms of interest can be written in the shape

$$E_{\text{img}}(v) = \Psi(r(m(1))), \tag{39}$$

where the dependence of the right-hand-side on v is obtained through the state equation. The first-order differential of E_{img} corresponds to

$$\delta E_{\text{img}}(v) = \langle \frac{\partial\Psi}{\partial r}\frac{\partial r}{\partial m(1)}, dm(1) \rangle. \tag{40}$$

Departing from the expression of Equation (A2) in [22], the first-order differential of the Lagrangian functional for a generic image similarity term is given by

$$\delta E_{\text{Lag}} = \int_0^1 < Lv, dv > dt + \langle \frac{\partial\Psi}{\partial r}\frac{\partial r}{\partial m(1)}, dm(1) \rangle +$$

$$+ \int_0^1 \langle d\lambda, \partial_t m + Dm \cdot v \rangle + \langle \lambda, \partial_t dm + Ddm \cdot v + Dm \cdot dv \rangle +$$

$$+ \langle d\eta, m(0) - I_0 \rangle + \langle \eta, dm(0) \rangle. \tag{41}$$

Then, integration by parts combined with the Green formula yield

$$\delta E_{\text{Lag}}(v) = \langle \frac{\partial\Psi}{\partial r}\frac{\partial r}{\partial m(1)} + \lambda(1), dm(1) \rangle + \text{other terms}. \tag{42}$$

The full expression can be found in Equation (A3) in [22]. Since, in particular, δE_{Lag} needs to vanish for all dm we have

$$\lambda(1) = -\frac{\partial\Psi}{\partial r}\frac{\partial r}{\partial m(1)}. \tag{43}$$

Accordingly,

$$\delta^2 E_{\text{img}}(v) = \langle \langle \langle \frac{\partial r}{\partial m(1)}, dm(1) \rangle, \frac{\partial^2\Psi}{\partial r^2}\frac{\partial r}{\partial m(1)} \rangle + \langle dm(1), \frac{\partial\Psi}{\partial r}\frac{\partial^2 r}{\partial m(1)^2} \rangle, dm(1) \rangle. \tag{44}$$

and

$$\delta^2 E_{\text{img}}(v) = \langle \langle \langle \frac{\partial r}{\partial m(1)}, dm(1) \rangle, \frac{\partial^2\Psi}{\partial r^2}\frac{\partial r}{\partial m(1)} \rangle$$

$$+ \langle dm(1), \frac{\partial\Psi}{\partial r}\frac{\partial^2 r}{\partial m(1)^2} \rangle + \delta\lambda(1), dm(1) \rangle + \text{other terms}. \tag{45}$$

Since $\delta^2 E_{\text{Lag}}$ also needs to vanish for all dm, we have

$$\delta\lambda(1) = -\langle \frac{\partial r}{\partial m(1)} dm(1), \frac{\partial^2 \Psi}{\partial r^2} \frac{\partial r}{\partial m(1)} \rangle - \langle dm(1), \frac{\partial \Psi}{\partial r} \frac{\partial r^2}{\partial m(1)^2} \rangle. \tag{46}$$

For Gauss–Newton–Krylov optimization, $\delta\lambda(1)$ is approximated by the positive definite expression

$$\delta\lambda(1) \approx -\langle \frac{\partial r}{\partial m(1)} dm(1), \frac{\partial^2 \Psi}{\partial r^2} \frac{\partial r}{\partial m(1)} \rangle, \tag{47}$$

that neglects the higher-order derivatives of the inner function r [32].

In practise, we compute $\delta E_{\text{img}}(v)$ and $\delta^2 E_{\text{img}}(v)$ in the form of Equations (40) and (44) and identify $\lambda(1)$ and $\delta\lambda(1)$ within the scalar products

$$\delta E_{\text{img}}(v) = \langle \nabla E_{\text{img}}(v), dm(1) \rangle \tag{48}$$

$$\delta^2 E_{\text{img}}(v) = \langle \langle dm(1), H E_{\text{img}}(v) \rangle, dm(1) \rangle. \tag{49}$$

Alternatively, for some metrics, the expression of $\delta\lambda(1)$ can be computed more straightforwardly from the differential of $\lambda(1)$. The obtained expressions are corroborated by the equations of $\lambda(1)$ and $\delta\lambda(1)$ for the SSD PDE-LDDMM problem with $\psi = \langle \cdot, \cdot \rangle_{L^2}$ and $r = m(1) - I_1$. In the following, we derive the expressions of $\lambda(1)$ and $\delta\lambda(1)$ for the image similarity metrics considered in this work.

3.2. Normalized Cross-Correlation (NCC)

In PDE-LDMMM, the NCC image similarity metric is defined as

$$E_{\text{img}}(v) = \frac{1}{\sigma^2} \int_\Omega 1 - \left(\frac{\langle \bar{m}(1), \bar{I}_1 \rangle^2}{\|\bar{m}(1)\|^2 \cdot \|\bar{I}_1\|^2} \right)(x) d\Omega, \tag{50}$$

where $\bar{I} = I - \text{mean}(I)$ for a generic image I. Let us define $A(x) = \langle \bar{m}(1)(x), \bar{I}_1(x) \rangle$, $B(x) = \|\bar{m}(1)(x)\|^2$, and $C(x) = \|\bar{I}_1(x)\|^2$. This allows us to define the NCC image similarity metric in the form of

$$E_{\text{img}}(v) = \frac{1}{\sigma^2} \int_\Omega 1 - r(x) d\Omega, \tag{51}$$

where $r = \frac{A^2}{B \cdot C}$.

Using the expression of the differential dr, $dA = \langle \bar{I}_1, d\bar{m}(1) \rangle$, and $dB = 2\langle \bar{m}(1), d\bar{m}(1) \rangle$, the expression of $\delta E_{\text{img}}(v)$ in terms of A, B and C is given by

$$\delta E_{\text{img}}(v) = -\frac{1}{\sigma^2} \frac{2A}{BC} \langle \bar{I}_1 - \frac{A}{B} \bar{m}(1), d\bar{m}(1) \rangle, \tag{52}$$

yielding

$$\lambda(1) = \frac{1}{\sigma^2} \frac{2}{C} \left(\frac{A}{B} \bar{I}_1 - \left(\frac{A}{B} \right)^2 \bar{m}(1) \right). \tag{53}$$

By the differentiation of $\lambda(1)$, the expression of $\delta\lambda(1)$ is given by

$$\delta\lambda(1) = \frac{1}{\sigma^2} \frac{2}{C} \left(\Theta \bar{I}_1 - 2\frac{A}{B} \Theta \bar{m}(1) - \left(\frac{A}{B} \right)^2 \right) d\bar{m}(1), \tag{54}$$

where Θ represents the differential of $\frac{A}{B}$, $\Theta = \frac{\bar{I}_1 - 2\frac{A}{B}\bar{m}(1)}{B}$.

3.3. Local Normalized Cross-Correlation (lNCC)

The lNCC image similarity metric departs from the NCC metric by computing the scalar products and the average of images in a neighborhood of size ν

$$E_{\text{img}}(v) = \frac{1}{\sigma^2} \int 1 - \left(\frac{\langle \bar{m}(1), \bar{I}_1 \rangle_\nu^2}{\|\bar{m}(1)\|_\nu^2 \cdot \|\bar{I}_1\|_\nu^2} \right)(x)d\Omega. \tag{55}$$

The expressions of $\lambda(1)$ and $\delta\lambda(1)$ are the restrictions of Equations (53) and (54) to the ν-neighborhood.

3.4. Normalized Gradient Fields (NGFs)

The NGFs image similarity metric is defined as

$$E_{\text{img}}(v) = \frac{1}{\sigma^2} \int_\Omega 1 - \left(\frac{\langle \nabla m(1), \nabla I_1 \rangle_{\nu\mu}}{\|\nabla m(1)\|_\nu \|\nabla I_1\|_\mu} \right)^2 (x)d\Omega, \tag{56}$$

where $\langle \cdot, \cdot \rangle_\epsilon = \langle \cdot, \cdot \rangle + \epsilon^2$ and $\| \cdot \|_\epsilon = \sqrt{\langle \cdot, \cdot \rangle + \epsilon^2}$. The image similarity metric can be written in the form of

$$E_{\text{img}}(v) = \frac{1}{\sigma^2} \int_\Omega 1 - r(x)^2 d\Omega, \tag{57}$$

where r is the quotient of the ϵ-scalar product and the ϵ-norms. Let us define $A(x) = \langle \nabla m(1)(x), \nabla I_1(x) \rangle_{\nu\mu}$, $B(x) = \|\nabla m(1)(x)\|_\nu$, and $C(x) = \|\nabla I_1(x)\|_\mu$. We use the same variable naming convention as in the NCC case for highlighting the analogies between both metrics. In this case, the residual is given by $r = \frac{A}{B \cdot C}$.

The expression of $\delta E_{\text{img}}(v)$ is given in terms of r by

$$\delta E_{\text{img}}(v) = -\langle 2rdr, dm(1) \rangle.$$

Using the expression of the differential dr,

$$dA = \langle \nabla I_1, \nabla dm(1) \rangle,$$

and

$$dB = (1/B)\langle \nabla m(1), \nabla dm(1) \rangle,$$

the expression of $\delta E_{\text{img}}(v)$ in terms of A, B and C is given by

$$\delta E_{\text{img}}(v) = -\frac{1}{\sigma^2} \frac{2A}{(BC)^2} \langle \nabla I_1 - \frac{A}{B^2} \nabla m(1), \nabla dm(1) \rangle, \tag{58}$$

which can be written in the form of Equation (40) using the identity $\langle u, \nabla v \rangle = \langle -\nabla \cdot u, v \rangle$. Thus,

$$\lambda(1) = -\frac{2}{\sigma^2} \nabla \cdot \left(\frac{A}{(BC)^2} \nabla I_1 - \left(\frac{A}{B^2 C} \right)^2 \nabla m(1) \right). \tag{59}$$

Since $\frac{\partial^2 \Psi}{\partial r^2} = 2$ for the NGF metric

$$\delta\lambda(1) = \langle \langle \frac{\partial r}{\partial m(1)}, dm(1) \rangle, 2\frac{\partial r}{\partial m(1)} \rangle. \tag{60}$$

The details of the numerical implementation are given in Appendix B.

3.5. Mutual Information (MI)

In PDE-LDDMM, the MI image similarity metric is defined as

$$E_{\text{img}}(v) = -\frac{1}{\sigma^2}\mathcal{M}(m(1), I_1) = -\frac{1}{\sigma^2}(\mathcal{H}(m(1)) + \mathcal{H}(I_1) - \mathcal{H}(m(1), I_1)) \tag{61}$$

where $\mathcal{H}$ is the entropy function. The entropy for a generic image I, and a generic pair of images (I, J), is defined, respectively, as

$$\mathcal{H}(I) = -\int p_I(r)\log p_I(r)dr \tag{62}$$

and

$$\mathcal{H}(I, J) = -\iint p_{I,J}(r, s)\log p_{I,J}(r, s)drds, \tag{63}$$

where p represents the estimated marginal and joint intensity distributions of the images, $p_I(r) = P(I = r)$ and $p_{I,J}(r, s) = P(I = r \cap J = s)$. In the following, we discretize the image variables and use the discrete expressions of the integrals.

The expression of $\delta E_{\text{img}}(v)$ is given by [61]

$$\delta E_{\text{img}}(v) = -\frac{1}{\sigma^2}\sum_{r,s}\delta p_{m(1),I_1}(r, s)\left(1 + \log\frac{p_{m(1),I_1}(r, s)}{p_{m(1)}(r)}\right). \tag{64}$$

In order to compute the differential $\delta p_{m(1),I_1}(r, s)$, we use the expression of p in analytical form

$$p_{m(1),I_1}(r, s) = \frac{1}{N_x}\sum_x \xi(r - \bar{m}(1)(x))\xi(s - \bar{I}_1(x)), \tag{65}$$

yielding

$$\delta p_{m(1),I_1}(r, s) = \frac{1}{N_x}\sum_x \delta\xi(r - \bar{m}(1)(x))\xi(s - \bar{I}_1(x)). \tag{66}$$

We use b-spline functions for the expression of ξ as proposed in [32]. The differential of ξ, $\delta\xi$, can be computed using the chain rule as

$$\delta\xi(r - \bar{m}(1)(x)) = -\langle\frac{\partial\xi}{\partial r}, d\bar{m}(1)\rangle. \tag{67}$$

Gathering the above expressions yields the expression of the initial adjoint variable

$$\lambda(1) = -\frac{1}{\sigma^2}\sum_{r,s}\frac{1}{N_x}\sum_x\frac{\partial\xi}{\partial r}\xi(s - \bar{I}_1(x))$$

$$\left(1 + \log\frac{p_{m(1),I_1}(r, s)}{p_{m(1)}(r)}\right). \tag{68}$$

In the computation of $\delta\lambda(1)$, a huge dense-matrix product requesting more than 5000 GBs of RAM memory arises. Therefore, we restrict this work to the gradient-descent version of PDE-LDDMM for MI. The derivation of $\delta\lambda(1)$ needed in Gauss–Newton–Krylov optimization is left outside the scope of the present work.

4. Experimental Setup

4.1. Datasets

We used five different databases in our evaluation:

NIREP16, was proposed in [41] for the evaluation of non-rigid registration. NIREP16 consists of 16 T1 Magnetic Resonance Imaging (MRI) images. NIREP16 images were acquired at the Human Neuroanatomy and Neuroimaging Laboratory, University of Iowa.

They were selected for the NIREP project from a database of 240 normal volunteers. Datasets correspond to 8 males and 8 females with a mean age of 32.5 ± 8.4 and 29.8 ± 5.8 years, respectively. The images are skull stripped and aligned according to the anterior and posterior commissures. Image dimension is $256 \times 300 \times 256$ with a voxel size of $0.7 \times 0.7 \times 0.7$ mm. Images are distributed with the segmentation of 32 gray matter regions at frontal, parietal, temporal, and occipital lobes. The most remarkable feature of this dataset is its excellent image quality. The geometry of the segmentations provides a specially challenging framework for deformable registration evaluation. In our previous works, a subsampled version of this dataset has been extensively used for the evaluation of different LDDMM methods. The mages of this dataset have been subsampled by reducing image dimension to $180 \times 210 \times 180$ with a voxel size of $1.0 \times 1.0 \times 1.0$ mm. Subsampling is needed to be able to run interesting but memory-demanding benchmark methods and to maintain the continuity of the evaluation results shown in previous works.

Klein datasets were proposed in [42] in the first extensive evaluation study of non-rigid registration methods. The datasets contain the T1 MRI images and segmentations from the LPBA40, IBSR18, CUMC12, and MGH10 databases. The four databases provide images with different levels of quality, providing varying difficulties for deformable registration [62]. Image dimension is $182 \times 218 \times 182$ with a voxel size of $1.0 \times 1.0 \times 1.0$ mm.

LPBA40 contains 40 skull-stripped brain images without the cerebellum and the brain stem. LPBA40 provides the segmentation of 50 gray matter structures together with left and right caudate, putamen, and hippocampus. LPBA40 protocols can be found at https://loni.usc.edu/research/atlases (accessed on 1 January 2022).

IBSR18 contains 18 brain images with the segmentation of 96 cerebral structures. This dataset provides the segmentation of brain structures of interest for the evaluation of image registration methods. The image quality is low. For example, most of the images show motion artifacts. The variability of the ventricle sizes is high.

CUMC12 contains 12 full brain images with the segmentation of 130 cerebral structures. Overall, the image quality is acceptable, although some of the images are noisy. The variability of the ventricle sizes is high.

MGH10 contains 10 full brain images with the segmentation of 106 cerebral structures. Overall, the image quality is acceptable, although some of the images are noisy. Ventricle sizes are usually all big.

In addition, we studied the performance of our methods in a multi-modal experiment, where the images were obtained from:

Oasis. The open-access series of imaging studies (https://www.oasis-brains.org/, accessed on 1 January 2022) is a project aimed at making neuroimaging data sets of the brain freely available to the scientific community. OASIS-3 compiles images from more than 1000 participants ranging from cognitively normal to various stages of cognitive decline. For each participant, the study includes different MRI sessions including T1, T2, FLAIR, and others. Our multimodal experiment selected a T2 image from an Alzheimer's disease participant as the source, and a T1 image from a cognitive normal participant as the target image.

4.2. Image Registration Pipeline

The evaluation in NIREP16 was performed consistently with our previous works on PDE-LDDMM diffeomorphic registration. The registrations were carried out from the first subject to every other subject in the database, yielding a total of 15 registrations per variant, optimization method, and image similarity metric. The subsampled NIREP16 database was obtained from the resampling of the original images into volumes of size $180 \times 210 \times 180$ with a voxel size of $1.0 \times 1.0 \times 1.0$ mm after the alignment to a common coordinate system using affine transformations. The images were scaled between 0 and 1 for SSD and NCC metrics, and between 0 and 255 for lNCC, NGFs, and MI. The affine alignment and subsampling were performed using the Insight Toolkit (ITK).

The LPBA40, IBSR18, CUMC12, and MGH10 images were preprocessed similarly to [42]. The input images were selected from the Synapse repository (https://www.synapse.org/#%21Synapse:syn3217707, accessed on 1 January 2022) in the folder hosting FLIRT affine registered images. In the first place, histogram matching was applied to all the images. The images were then scaled between 0 and 1 for SSD and NCC metrics, and between 0 and 255 for lNCC, NGFs, and MI. To perform these preprocessing steps we used the algorithms available in ITK.

Oasis images were finely aligned to the MNI152 atlas with NiftyReg and then skull stripped using the robust brain extraction software RobEx (https://www.nitrc.org/projects/robex, accessed on 1 January 2022).

4.3. Numerical Details, Parameter Configuration, and Implementation Details

The experiments were run on a cluster of two machines equipped with four NVidia GeForce GTX 1080 ti with 11 GBS of video memory and an Intel Core i7 with 64 GBS of DDR3 RAM, and two NVidia Titan RTX with 24 GBS of video memory and an Intel Core i7 with 64 GBS of DDR3 RAM, respectively. The codes were developed in the GPU with Matlab 2017a and Cuda 8.0.

Regularization parameters were selected from a search of the optimal parameters in the registration experiments performed in our previous work [22]. We selected the parameters $\sigma^2 = 1.0$, $\alpha = 0.0025$, and $s = 2$ and a unit-domain discretization of the image domain Ω [56]. We also tested the parameters $\sigma^2 = 0.03$, $\alpha = 3.0$, and $s = 3$ and a spatial-domain discretization of Ω, selected as optimal in [58]. For gradient-descent optimization, we obtained excellent evaluation results; however, the obtained maximum Jacobians were much higher than recommended. On the other hand, Gauss–Newton–Krylov showed convergence problems during PCG, with negative curvature values found at early inner iterations. This suggests that the specific selection of parameters in [58] might achieve fairly high structural overlaps with the cost of very aggressive underlying deformations which are glimpsed in the malfunctioning of Gauss–Newton.

The BL experiments were performed with band sizes of $32 \times 32 \times 32$ for BL Variants I and II. This selection was found as optimal for each method in our previous work [20,22,33].

Gradient-descent was implemented with an efficient method for the update of the step size based on offline backtracking line-search combined with a check on Armijo's condition. We used the stopping conditions in [13,32]. Otherwise, the optimization was stopped after 50 iterations.

Gauss–Newton–Krylov was also implemented with an offline backtracking line-search combined with a check on Armijo's condition. The number of PCG iterations was set to five. The PCG tolerance was selected from

$$\tau = \min\left(0.5, \sqrt{\frac{\|\nabla_v E(v^n)\|_2}{\|\nabla_v E(v^0)\|_2}}\right).$$

We used the stopping conditions in [13,32]. Otherwise, the optimization was stopped after 10 iterations. These parameters were selected as optimal in our previous work since the methods achieved state-of-the-art accuracy in a reasonable amount of time [22].

PDE-LDDMM was embedded into a multi-resolution scheme. The images were subsampled, and the velocity fields were resampled similarly to [63,64]. The PDE-LDDMM registration methods were executed on each resolution level. For the multi-resolution experiments, the pyramid was built with three levels with the same number of outer and inner iterations, as for the single-resolution.

To integrate the PDEs, we used the semi-Lagrangian Runge–Kutta schemes proposed in [33] for the SSD versions of Variants I and II. The solutions were computed at the Chebyshev–Gauss–Lobatto discretization of the temporal domain $[0, 1]$. The number of time steps was selected as five. Since Matlab lacks a 3D GPU cubic interpolator, we

implemented in a Cuda MEX file the GPU cubic interpolator with prefiltering proposed in [65].

The computation of differentials was approached using Fourier spectral methods as a machine-precision accurate alternative to commonly used finite difference approximations [66]. Spectral methods allow solving of ODEs and PDEs for high accuracy in simple domains for problems involving smooth data. To this end, the images were smoothed with a Gaussian filter as a preprocessing step. However, for the Gauss–Newton–Krylov version of NGFs, we used the matrix version of the differential operators, and then the computation of differentials must be approached with finite difference approximations. To be consistent with the input data, the images were also smoothed as a preprocessing step.

For lNCC, the size of the neighborhood was selected as four. For NGFs, the value of ϵ^2 for the ϵ-norms was equal to 1000. For MI, the number of histogram bins was selected equal to 16. The computation of the adjoint variable for MI required the use of sparse matrices and was implemented in the CPU since Matlab does not yet have GPU support for these data structures.

4.4. Benchmarks

For benchmarking, we run single- and multi-resolution versions of ANTS registration with SSD, lNCC and MI image similarities [40]. We also extended Stationary LDDMM (St-LDDMM), proposed in [67] as an efficient stationary variant of Beg et al.'s LDDMM [56], with NCC, lNCC, NGFs, and MI metrics. The details of the method extension can be found in Appendix A. In addition, we studied the accuracy obtained with QuickSilver, a supervised deep-learning based method with SSD in the loss function [46], and Voxel-Morph, an unsupervised deep learning-based model with SSD and NCC metrics in the loss function [50].

St-LDDMM was run with the same parameters than PDE-LDDMM in the common steps of the algorithms. ANTS was run with the following parameters for the single-resolution experiments

```
$synconvergence = ''[50,1 × 10⁻⁶,10]'',
$synshrinkfactors = ''1'',
and $synsmoothingsigmas = ''3vox''.
```

For the multi-resolution experiments the parameters were set to

```
$synconvergence = ''[50×50×50,1 × 10⁻⁶,10]'',
$synshrinkfactors = ''4×2×1'',
and $synsmoothingsigmas = ''3×2×1vox''.
```

The selection of the number of iterations was in agreement with the number of iterations used in gradient-descent and the number of outer × inner iterations used in Gauss–Newton–Krylov optimization for PDE-LDDMM.

In the multi-modal experiment, we compare our proposed methods with NiftyReg, a software for efficient registration developed at the Centre for Medical Image Computing at University College London, UK (https://sourceforge.net/projects/niftyreg/, accessed on 1 January 2022). NiftyReg is usually selected as a benchmark for non-rigid multimodal image registration. We also include in the comparison ANTS and SynthMorph, a VoxelMorph adaptation for building deep-learning models capable of dealing with multimodality [53].

4.5. Metrics and Statistical Analysis for Registration Evaluation

The evaluation in NIREP16 and Klein datasets is based on the accuracy of the registration results for template-based segmentation. The Dice Similarity Coefficient (DSC) is selected as the evaluation metric. Given two segmentations S and T, the DSC is defined as

$$DSC(S,T) = \frac{2\text{Vol}(S \cap T)}{\text{Vol}(S) + \text{Vol}(T)}. \tag{69}$$

This metric provides the value of one if *S* and *T* exactly overlap and gradually decreases towards zero depending on the overlap of the two volumes. The statistical distribution of the DSC results across the segmented structures are shown in the shape of box and whisker plots following the evaluation methods in [42]. The DSC distribution is taken over the DSC values over the different segmentation labels. This way of computing the DSC distribution reflects the recommendation given in [68] that the evaluation of non-rigid registration with the segmentation of sufficiently locally labeled regions of interest is strongly recommended for obtaining reliable measurements of the performance of the registration.

The evaluation in NIREP16 was completed with two different statistical analysis on the DSC values. In the first place, the analysis of variance (ANOVA) was conducted in order to assess whether the means of the DSC distributions are different for the image similarity metrics when observations are grouped by type of method (baseline vs. PDE-LDDMM) or variants (I vs. II). Baseline methods include ANTS, Stationary LDDMM, VoxelMorph, and QuickSilver. In the second place, pairwise right-tailed Wilcoxon rank-sum tests were conducted for the assessment of the statistical significance of the difference of medians for the distribution of the DSC values. The alternative hypothesis is that the median of the first distribution is higher than the median of the second one.

Finally, we include for NIREP16 the quantitative assessment provided by the mean and standard deviation of the relative image similarity error after registration,

$$MSE_{rel} = \frac{\|m(1) - I_1\|_{L^2}^2}{\|I_0 - I_1\|_{L^2}^2},$$

the relative gradient magnitude,

$$\|g\|_{\infty, rel} = \frac{\|\nabla_v E(v^n)\|_\infty}{\|\nabla_v E(v^0)\|_\infty},$$

and the extrema of the Jacobian determinant.

5. Results

In this section, we show the experiments conducted to evaluate the performance of the two PDE-LDDMM variants for the different image similarity metrics. First, we provide an extensive evaluation of our proposed methods in the NIREP16 database, where we have extensively evaluated previous LDDMM and PDE-LDDMM registration methods [20,22,33,67]. Next, we evaluate our proposed methods in Klein et al. databases. Finally, we compare the behavior of the different metrics in a challenging multimodal experiment.

5.1. Results on NIREP16

5.1.1. Evaluation

Figure 1 shows, in the shape of box and whisker plots, the statistical distribution of the DSC values that were obtained after the registration across the 32 segmented structures. In addition, Figure 2 gathers the results obtained with Gauss–Newton–Krylov optimization grouped by variant for a better assessment of the best-performing combination of variant and metrics. Our first observation is that the DSC coefficients for the multi-resolution experiments outperform the single-resolution experiments. The improvement is substantial for lNCC, NGFs, and MI metrics.

Single-resolution. Regarding the single-resolution experiments, it is striking the low performance of ANTS for all metrics. Both St-LDDMM and PDE-LDDMM perform more reasonably than ANTS. In general, PDE-LDDMM methods tend to outperform St-LDDMM.

For PDE-LDDMM and SSD, the differences between Gauss–Newton and gradient-descent are small. Gauss–Newton optimization significantly outperforms gradient-descent for NCC and lNCC. On the contrary, for NGFs, gradient-descent optimization outperforms Gauss–Newton in all cases. The spatial version of Variant II exhibits an especially lower performance. For St-LDDMM, the trends observed in PDE-LDDMM are also observed for

SSD and NCC metrics. However, gradient-descent performs similarly to Gauss–Newton for lNCC and Gauss–Newton outperforms gradient descent for the NGFs metric.

The results obtained with the spatial versions of the PDE-LDDMM variants are similar to the corresponding BL versions. Comparing the accuracy of both variants, Variant II provides in general better performance than Variant I.

For all variants, the overall best performing metric is NCC. For almost all variants, the gradient-descent version of NGFs performs similarly to the Gauss–Newton version of lNCC. For the BL methods, the Gauss–Newton version of NGFs also performs similarly to the gradient-descent version of SSD. Moreover, the gradient-descent version of MI performs similarly to the gradient-descent version of SSD.

Multi-resolution. Regarding the multi-resolution experiments, the performance of ANTS reached the level of accuracy of St-LDDMM and the PDE-LDDMM methods. ANTS with the lNCC metric greatly outperformed the other ANTS variants using SSD and MI, ranking among the best-performing methods. In general, PDE-LDDMM methods tend to outperform St-LDDMM, with the exception of the NGFs metric.

As happened with the single-resolution experiments for PDE-LDDMM, the differences between Gauss–Newton optimization and gradient-descent are small for the SSD metric. Gauss–Newton also outperforms gradient-descent for NCC and lNCC. For NGFs, gradient-descent optimization greatly outperforms Gauss–Newton in the case of Variant II. However, for Variant I, the differences between both optimization methods are small, especially for the BL version of the methods. The performance of the NGFs metric is further explored in Appendix B for a better understanding of these observations. For St-LDDMM, the trends observed with the single-resolution experiments are mostly observed.

For Variant I, the spatial version tends to outperform the BL version of the same variant slightly. However, the performance of the BL version of Variant II is similar or even improves the spatial version for almost all metrics. As happened with the single-resolution experiments, Variant II provides better performance than Variant I. The best performing metric for Variant I is lNCC, while for Variant II, the best-performing metric is still NCC, closely followed by lNCC. The resemblance of performance between MI and SSD metrics in the single-resolution experiments remains for the multi-resolution experiments. However, the excellent performance of NGFs metric with gradient-descent optimization is remarkable, ranking close to the best-performing metrics for Variant II.

Comparing ANTS with PDE-LDDMM methods, Variant I with SSD and gradient-descent performs similarly to ANTS-SSD. In the case of MI, PDE-LDDMM methods outperform ANTS-MI. Some PDE-LDDMM methods achieve results competing with ANTS-lNCC for the NCC and lNCC metrics.

Deep-learning methods. Because of the increasing relevance of deep-learning methods in the field, we added to our evaluation the performance of VoxelMorph [50] and QuickSilver [46]. VoxelMorph with SSD and QuickSilver with the correction step performed similarly to Variant II of PDE-LDDMM with the SSD metric. Diffeomorphic VoxelMorph with SSD ranked among the best-performing methods, with a box-plot distribution similar to Variant I with lNCC and BL Variant II with NCC and lNCC. Despite all LDDMM methods agreeing in the much better performance of NCC and lNCC metrics over SSD, Diffeomorphic VoxelMorph trained with a loss function based on SSD greatly outperformed the method trained with NCC. Lastly, it is a remarkable fact that, although VoxelMorph is informed during training of the performance through the DSC, our best-performing PDE-LDDMMs were able to achieve competitive results without the use of this information.

Statistical analysis. Table 1 shows the results of the analysis of variance (ANOVA) for the effects of method and image similarity metric selection on the distribution of the DSC values obtained in the multi-resolution experiments with Gauss–Newton optimization (with the exception of the methods combined with the MI metric). The methods on the first factor were grouped by type of method (baseline vs. PDE-LDDMM) and variants. The tests only showed no statistical significance for the differences between the spatial versions of
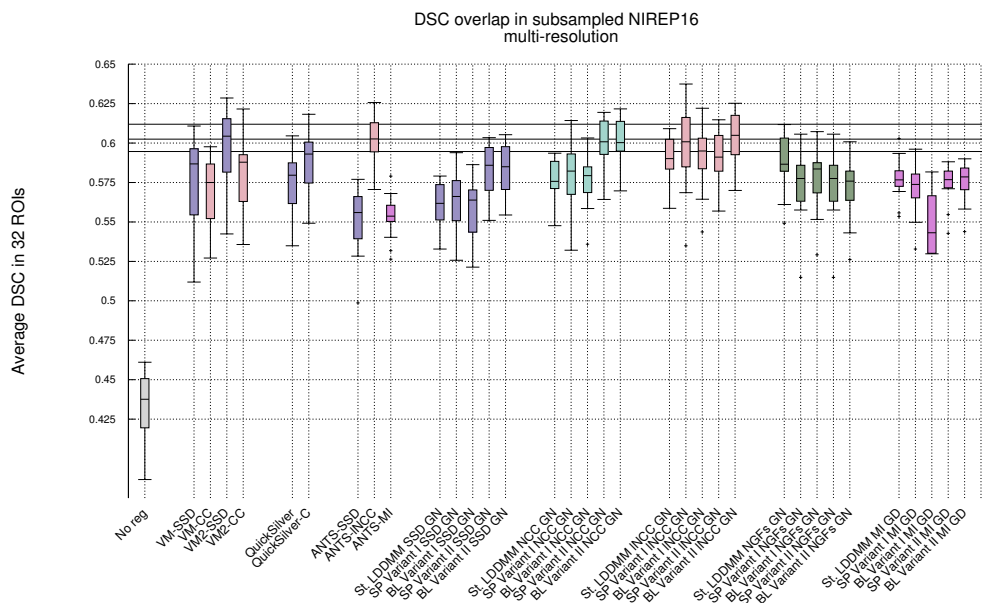
Variant I and Variant II. The selection of the E_{img} metric resulted in statistical significance for all cases.

Figure 1. NIREP16. Volume overlap obtained by the registration methods measured in terms of the DSC between the warped and the corresponding manual target segmentations. Box and whisker plots show the distribution of the DSC values averaged over the 32 NIREP manual segmentations. The whiskers indicate the minimum and maximum of the DSC values. The upper plot shows the evaluation results in the single-resolution experiments. The lower plot shows the results in the multi-resolution experiments. The horizontal lines in the plot indicate the first, second, and third quartiles of one of the best-performing baseline methods (multiresolution ANTS-lNCC), facilitating the comparison.

Figure 2. NIREP16. Same legend to Figure 1. Results obtained with PDE-LDDMM and Gauss–Newton–Krylov optimization. Boxplots grouped by variant. St-LDDMM is grouped with PDE-LDDMM results for facilitating comparison.

Table 1. NIREP16. Multi-resolution experiments and Gauss–Newton–Krylov optimization. Results of ANOVA tests for the effects of method and image similarity metric selection.

Factor I	Factor II	*p*-Value I	*p*-Value II
Baseline vs. SP Variant I	E_{img}	1.07×10^{-5}	1.09×10^{-11}
Baseline vs. BL Variant I	E_{img}	3.09×10^{-7}	1.45×10^{-19}
Baseline vs. SP Variant II	E_{img}	1.07×10^{-5}	1.09×10^{-11}
Baseline vs. BL Variant II	E_{img}	1.07×10^{-5}	1.09×10^{-11}
PDE-LDDMM methods	E_{img}	3.89×10^{-7}	1.77×10^{-21}
SP Variant I vs. SP Variant II	E_{img}	0.339	0
BL Variant I vs. BL Variant II	E_{img}	1.88×10^{-8}	3.46×10^{-13}
SP Variant I vs. BL Variant I	E_{img}	0.0284	0
SP Variant II vs. BL Variant II	E_{img}	0.0068	0

Figure 3 shows the *p*-values of pairwise right-tailed Wilcoxon rank-sum tests for the distribution of the DSC values obtained in the multi-resolution experiments with Gauss–Newton optimization (with the exception of the methods combined with the MI metric). The figure shows statistical significance for the better performance of the NCC and lNCC metrics over SSD and MI. For NGFs, obtaining statistical significance depends on the method. Among the best-performing methods, no statistical significance was found for the difference of medians.

5.1.2. Quantitative Assessment

Table 2 shows, averaged by the number of experiments, the mean and standard deviation of the MSE_{rel}, $\|g\|_{\infty,rel}$, and the extrema of the Jacobian determinant obtained with PDE-LDDMM in the NIREP16 dataset. We restrict the results to the methods with Gauss–Newton–Krylov optimization with the exception of the methods with the MI metric. For NGFs, the results with gradient descent and different Gauss–Newton approximations are analyzed in depth in Appendix B. Table 3 shows the average MSE_{rel} values and the extrema of the Jacobian determinant for VoxelMorph and QuickSilver.

For the single-resolution experiments, the best MSE_{rel} values were obtained by the NCC metric, followed by SSD. Although the MSE_{rel} values for the lNCC and MI metrics ranged higher than 20%, their performance in the evaluation reported a similar distribution. For lNCC, NGFs, and MI, the correlation between the lowest MSE_{rel} values and the highest DSC results that are usually seen for SSD in previous works does not hold anymore [22,33].

The spatial methods slightly outperformed the BL methods in terms of the MSE_{rel} values, as expected. Variant II performed better than Variant I. The relative gradient was reduced to average values ranging from 0.01 to 0.18, except for the lNCC and NGFs metrics. This means that the optimization was stopped in acceptable energy values in all these cases. Although the relative gradient obtained with lNCC was higher than recommended, the corresponding DSC distributions indicate that the lNCC methods can reach a local minimum providing good registration results. All the Jacobians remained above zero.
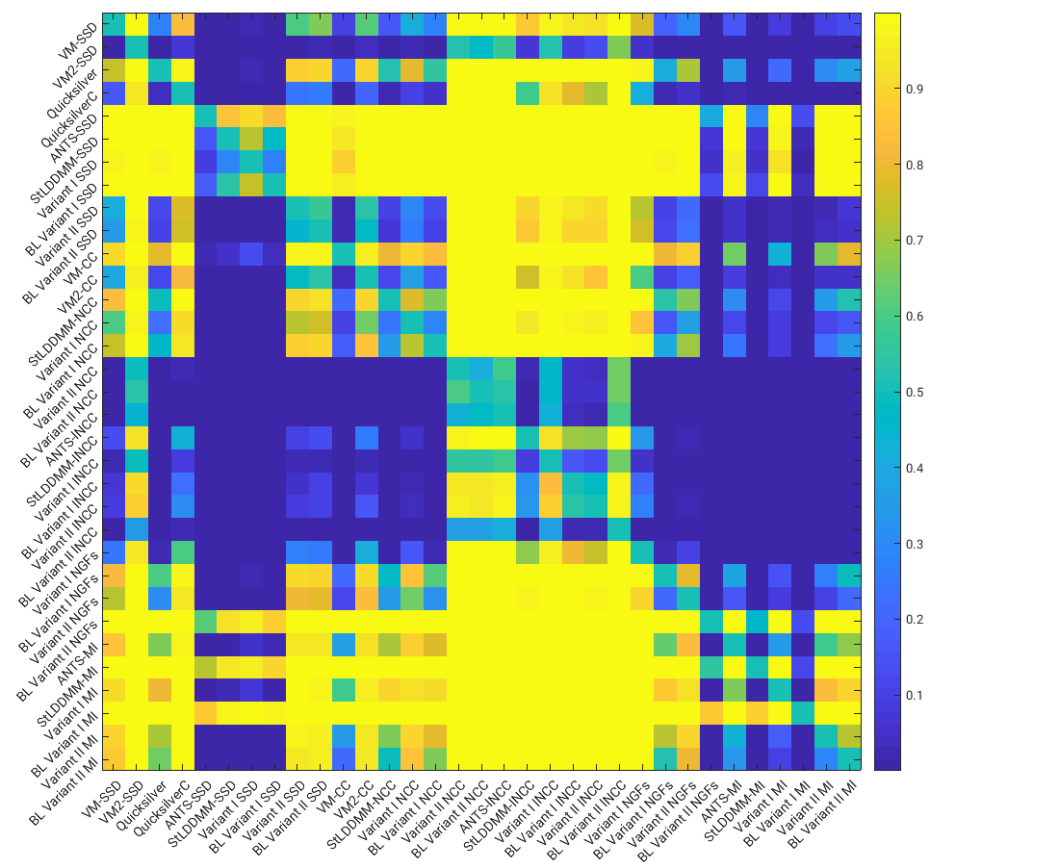

Figure 3. NIREP16. Multi-resolution experiments and Gauss–Newton–Krylov optimization. Results of the pairwise right-tailed Wilcoxon rank-sum tests.

For the multi-resolution experiments, the results regarding the MSE_{rel} values and the Jacobians were consistent with the single-resolution experiments. However, the high values of the relative gradient indicate a stagnation of the convergence in the finer resolution level that may be due to the method already starting close to the convergence point at the beginning of this resolution level.

Both VoxelMorph and QuickSilver usually obtained MSE_{rel} values greater than PDE-LDDMM with the corresponding image similarity metric. It is striking the magnitude of the Jacobian extrema obtained by VoxelMorph and its diffeomorphic version, indicating that the accuracy of the registration results shown in Figure 2 are obtained through large folds in a considerable number of locations.

Figure 4 shows the evolution of the convergence curves for the image similarity metrics E_{img} in the single-resolution experiments. For all the metrics, the trend of the E_{img} values is decreasing. The most unexpected behavior is for the curves of the lNCC metrics, where the

standard deviation remains stable and large in comparison with the energy reduction. The curves of the NGFs metrics show the stagnation of the energy values for the BL variants. This is the cause of the low DSC distributions already shown in Figure 1.

Spatial methods show slightly better E_{img} values than BL methods. Comparing the variants, Variant II provides slightly lower E_{img} values than Variant I. These results are consistent with the evaluation and the quantitative assessment shown in Figure 1 and Table 2.

Table 2. NIREP16. Convergence results. Mean and standard deviation of the relative image similarity error expressed in % (MSE_{rel}), the relative gradient magnitude ($\|g\|_{\infty,rel}$), and the Jacobian determinant extrema associated with the transformation $(\phi_1^v)^{-1}$. The best MSE_{rel} results for each variant are highlighted in bold face.

Metric	$MSE_{rel}(\%)$	$\|g\|_{\infty,rel}$	$\min(J(\phi_1^v)^{-1})$	$\max(J(\phi_1^v)^{-1})$
Single-resolution				
SP Variant I, SSD, GN	18.29 ± 2.83	0.07 ± 0.05	0.16 ± 0.05	3.70 ± 0.51
SP Variant I, NCC, GN	$\mathbf{16.51 \pm 2.13}$	0.17 ± 0.01	0.01 ± 0.02	3.92 ± 0.68
SP Variant I, lNCC, GN	20.56 ± 2.31	0.49 ± 0.32	0.04 ± 0.03	3.79 ± 0.40
SP Variant I, NGFs, GN	39.14 ± 2.91	0.71 ± 0.15	0.14 ± 0.03	4.67 ± 0.35
SP Variant I, MI, GD	24.34 ± 2.58	0.13 ± 0.06	0.18 ± 0.06	2.17 ± 0.15
SP Variant II, SSD, GN	17.10 ± 1.50	0.12 ± 0.05	0.14 ± 0.05	5.02 ± 1.01
SP Variant II, NCC, GN	$\mathbf{15.34 \pm 1.95}$	0.15 ± 0.06	0.10 ± 0.05	6.31 ± 1.65
SP Variant II, lNCC, GN	24.34 ± 2.90	0.30 ± 0.08	0.20 ± 0.03	3.71 ± 0.64
SP Variant II, NGFs, GN	46.78 ± 2.50	0.67 ± 0.13	0.34 ± 0.03	3.65 ± 0.46
SP Variant II, MI, GD	22.26 ± 2.42	0.12 ± 0.02	0.26 ± 0.05	2.61 ± 0.32
BL Variant I, SSD, GN	19.89 ± 1.76	0.01 ± 0.00	0.29 ± 0.03	3.45 ± 0.29
BL Variant I, NCC, GN	$\mathbf{17.06 \pm 1.71}$	0.03 ± 0.01	0.09 ± 0.04	4.56 ± 0.59
BL Variant I, lNCC, GN	24.30 ± 2.68	0.14 ± 0.06	0.17 ± 0.03	3.69 ± 0.49
BL Variant I, NGFs, GN	71.93 ± 1.78	0.78 ± 0.08	0.62 ± 0.02	1.56 ± 0.05
BL Variant I, MI, GD	26.21 ± 2.66	0.02 ± 0.00	0.28 ± 0.04	2.19 ± 0.14
BL Variant II, SSD, GN	17.77 ± 1.66	0.04 ± 0.01	0.13 ± 0.04	4.81 ± 0.72
BL Variant II, NCC, GN	$\mathbf{15.57 \pm 1.71}$	0.05 ± 0.02	0.09 ± 0.04	6.29 ± 1.15
BL Variant II, lNCC, GN	22.74 ± 2.77	0.16 ± 0.04	0.17 ± 0.04	4.41 ± 0.76
BL Variant II, NGFs, GN	71.77 ± 1.75	0.91 ± 0.08	0.60 ± 0.02	1.56 ± 0.05
BL Variant II, MI, GD	23.30 ± 2.61	0.03 ± 0.01	0.28 ± 0.04	2.46 ± 0.18
Multi-resolution				
SP Variant I, SSD, GN	17.98 ± 2.71	2.50 ± 3.49	0.02 ± 0.08	3.67 ± 0.82
SP Variant I, NCC, GN	14.89 ± 2.24	1.33 ± 0.89	0.04 ± 0.12	9.77 ± 20.45
SP Variant I, lNCC, GN	$\mathbf{12.94 \pm 3.11}$	0.72 ± 0.37	0.01 ± 0.02	5.17 ± 1.18
SP Variant I, NGFs, GN	29.08 ± 2.55	0.66 ± 0.26	0.05 ± 0.04	6.99 ± 1.74
SP Variant I, MI, GD	18.35 ± 2.19	0.99 ± 1.39	0.06 ± 0.06	3.26 ± 0.31
SP Variant II, SSD, GN	15.93 ± 1.51	1.13 ± 0.67	0.12 ± 0.03	9.89 ± 6.86
SP Variant II, NCC, GN	$\mathbf{13.12 \pm 1.69}$	1.02 ± 0.27	0.08 ± 0.03	14.32 ± 11.37
SP Variant II, lNCC, GN	15.23 ± 2.20	1.10 ± 0.39	0.10 ± 0.03	5.52 ± 0.81
SP Variant II, NGFs, GN	24.20 ± 2.41	0.84 ± 0.15	0.23 ± 0.03	4.76 ± 1.08
SP Variant II, MI, GD	19.15 ± 2.41	0.46 ± 0.08	0.17 ± 0.04	4.14 ± 1.30
BL Variant I, SSD, GN	19.22 ± 1.72	0.11 ± 0.05	0.14 ± 0.06	3.74 ± 0.38
BL Variant I, NCC, GN	15.81 ± 1.62	0.21 ± 0.07	0.09 ± 0.05	5.04 ± 0.76
BL Variant I, lNCC, GN	$\mathbf{14.41 \pm 2.45}$	0.82 ± 0.25	0.08 ± 0.04	6.32 ± 1.64
BL Variant I, NGFs, GN	24.10 ± 2.96	0.30 ± 0.15	0.09 ± 0.02	9.62 ± 1.90
BL Variant I, MI, GD	18.35 ± 2.19	0.31 ± 0.21	0.06 ± 0.06	3.26 ± 0.31
BL Variant II, SSD, GN	16.30 ± 1.60	0.25 ± 0.08	0.11 ± 0.04	6.30 ± 1.64
BL Variant II, NCC, GN	$\mathbf{13.65 \pm 1.72}$	0.47 ± 0.19	0.07 ± 0.03	9.46 ± 2.78
BL Variant II, lNCC, GN	13.67 ± 2.26	1.00 ± 0.48	0.08 ± 0.03	7.92 ± 1.23
BL Variant II, NGFs, GN	23.04 ± 2.25	0.51 ± 0.20	0.07 ± 0.02	7.72 ± 2.24
BL Variant II, MI, GD	19.35 ± 2.50	0.13 ± 0.15	0.17 ± 0.05	3.80 ± 0.63

Table 3. NIREP16. Quantitative results of the deep-learning methods. Mean and standard deviation of the relative image similarity error expressed in % (MSE_{rel}), and the Jacobian determinant extrema associated with the transformation ϕ^{-1}. The best MSE_{rel} results for each method are highlighted in bold face.

Metric	$MSE_{rel}(\%)$	$\min(J(\phi^{-1}))$	$\max(J(\phi^{-1}))$
VM-SSD	17.82 ± 0.98	-13.04 ± 5.72	86.12 ± 97.05
VM-CC	20.25 ± 0.93	-14.89 ± 3.87	59.69 ± 11.59
VM2-SSD	16.35 ± 0.67	-9.26 ± 5.26	53.13 ± 26.16
VM2-CC	$\mathbf{14.34 \pm 0.95}$	-7.65 ± 1.56	46.91 ± 7.83
QuickSilver	19.03 ± 1.37	0.31 ± 0.03	9.54 ± 2.44
QuickSilver-C	$\mathbf{15.75 \pm 1.27}$	0.32 ± 0.05	9.59 ± 2.66
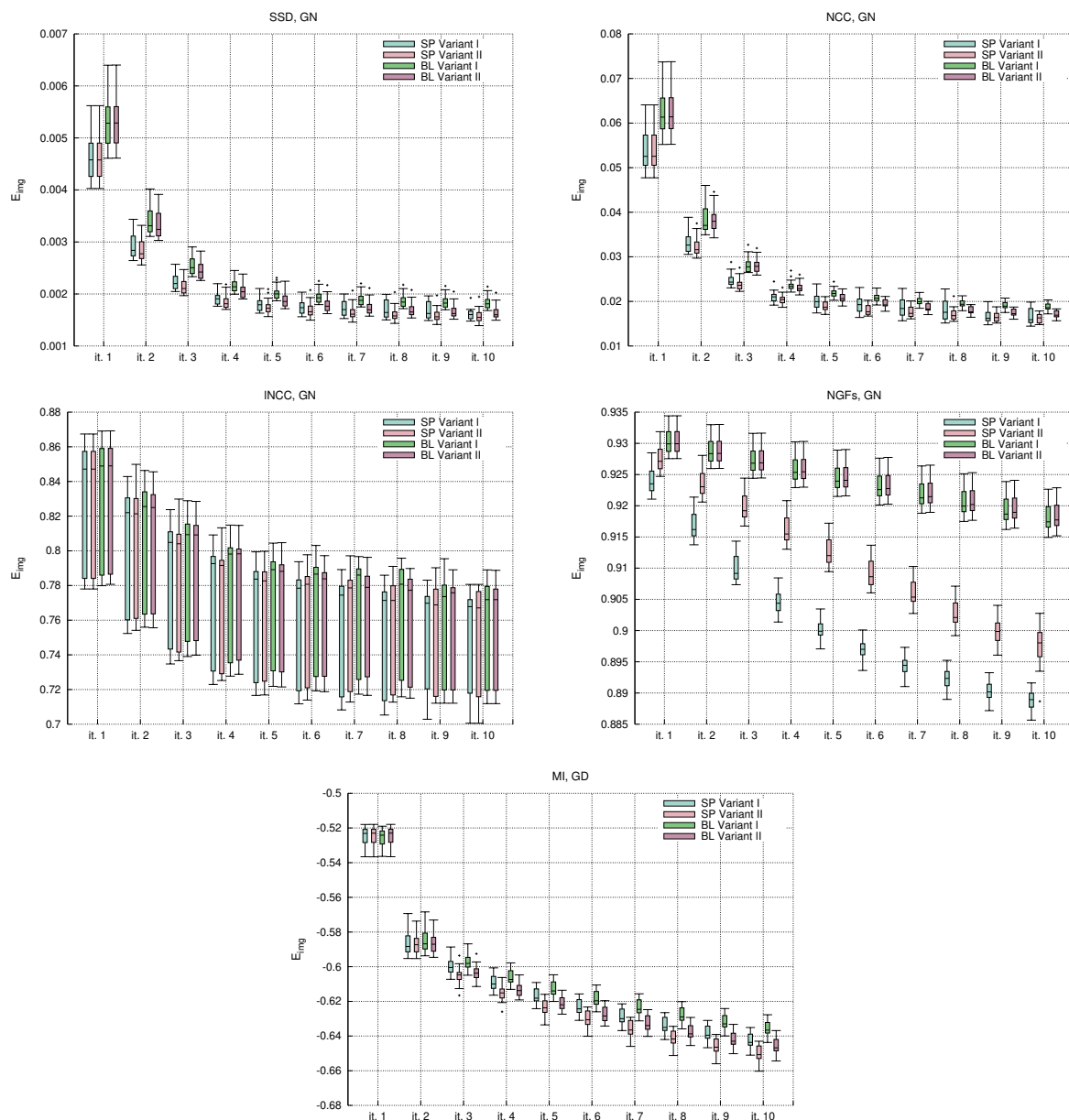
Figure 4. NIREP16. Single-resolution experiments. Box and whisker plots of the E_{img} convergence values in the single resolution experiments. For gradient-descent, the energy values are shown every five iterations.

5.1.3. Qualitative Assessment

For a qualitative assessment of the proposed registration methods, we show the registration results obtained by the different metrics for the BL version of Variant II in the multi-resolution experiments. Figure 5 shows the warped images, the difference between the warped and the target images after registration, the velocity fields, and the logarithm of the Jacobian determinant. The resemblance of the differences between the warped and the target images was high for all the metrics except for NGFs. Focusing on the registration results at the ventricle, SDD and NCC were able to achieve the best compression of the structure, while NGFs obtained the worst registration results at this location. Figure 6 shows the warped images, the difference between the warped and the target images, the displacement fields, and the logarithm of the Jacobian determinant for VoxelMorph. The resemblance of the differences between the warped and the target images was higher for SSD than NCC. The displacement fields were visually less smooth than the velocity fields obtained with PDE-LDDMM. The Jacobian determinant had negative regions all over the image. In particular, the registration results at the ventricle were achieved through large expansions and foldings in its upper boundary.
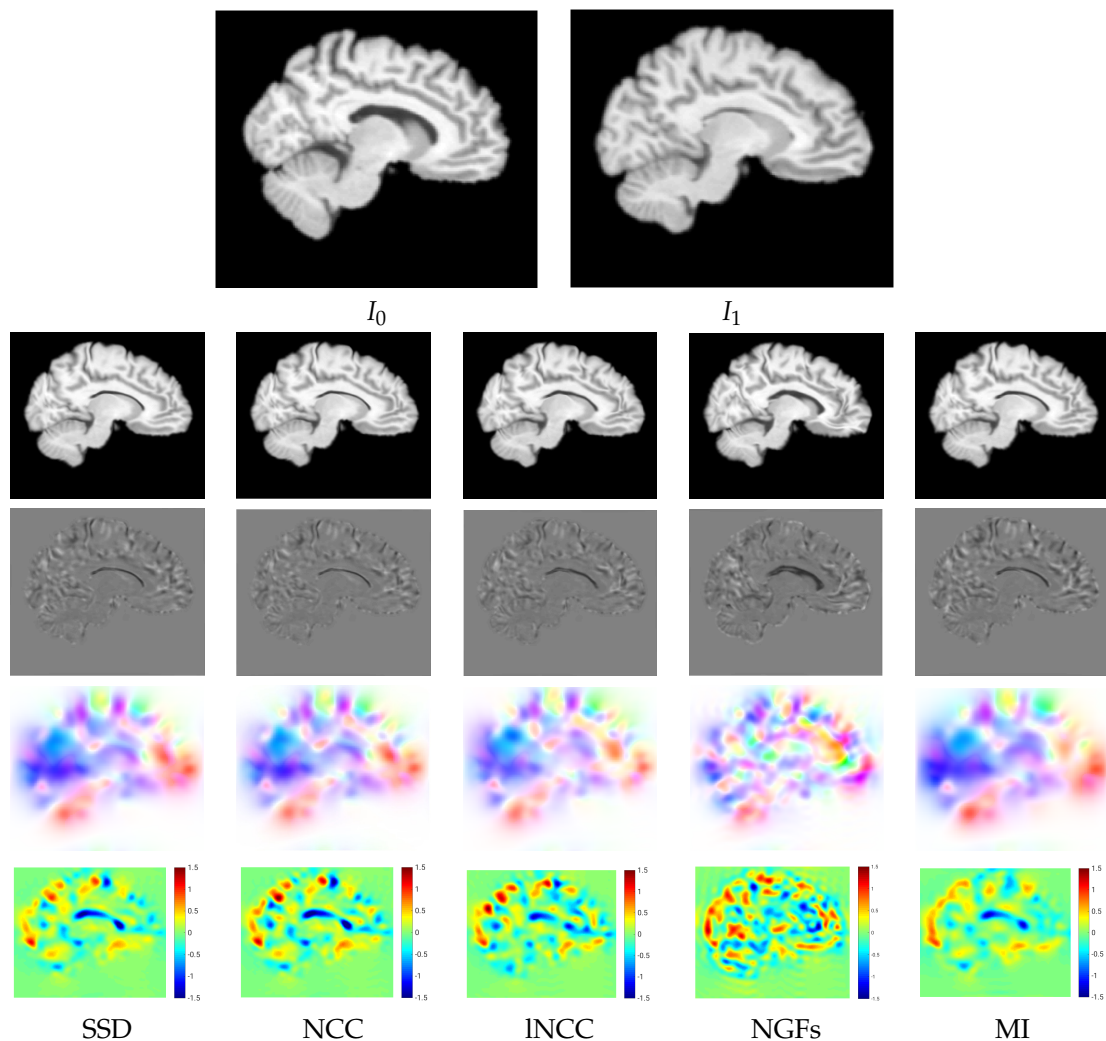
Figure 5. NIREP16. BL Variant II of PDE-LDDMM. Multi-resolution experiments. Sagittal view of the warped sources, the intensity differences after registration, the velocity fields, and the logarithm of the Jacobian determinants after registration for the different image similarity metrics. The results for SSD, NCC, lNCC, and NGFs are obtained with Gauss–Newton optimization, while for MI gradient-descent is used.
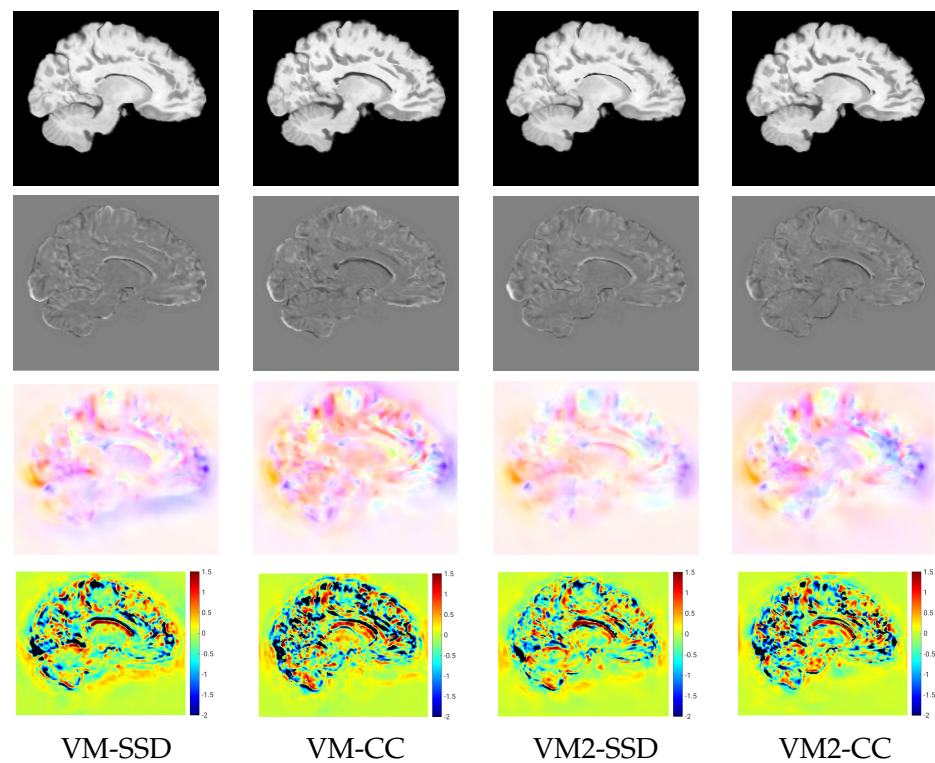
VM-SSD   VM-CC   VM2-SSD   VM2-CC

Figure 6. NIREP16. VoxelMorph experiments. Sagittal view of the warped sources, the intensity differences after registration, the displacement fields, and the logarithm of the Jacobian determinants after registration for the different image similarity metrics. For the negative Jacobian values, the logarithm is replaced by -2 and displayed in black.

5.1.4. Computational Complexity

Table 4 shows the average and standard deviation of the total computation time and the VRAM peak memory reached through the computations in the NIREP16 database for the single-resolution experiments. The BL methods achieved a substantial time and memory reduction over the spatial methods, as already demonstrated in [22,33,39]. From the Gauss–Newton methods, the methods with SSD and NCC image similarity metrics were the most efficient ones, as expected. On the other side, the methods with MI were the most time-consuming ones. Regarding memory usage, the methods using SSD and NCC were more efficient than lNCC. The memory efficiency shown by NGFs and MI metrics was due to the combination with gradient-descent and the need to perform operations involving sparse matrices on the CPU.

Table 4. NIREP16. Single-resolution experiments. Mean and standard deviation of the total GPU time, and maximum VRAM memory usage achieved by the methods through the registrations.

Method	$time_{GPU}$ (s)	Peak VRAM (MBs)
SP Variant I, SSD, GN	144.51 ± 5.87	5823
SP Variant I, NCC, GN	145.25 ± 6.93	5875
SP Variant I, lNCC, GN	169.02 ± 0.92	6159
SP Variant I, NGFs, GD	149.18 ± 1.04	5023
SP Variant I, MI, GD	347.35 ± 0.41	4997
BL Variant I, SSD, GN	77.44 ± 0.28	3263
BL Variant I, NCC, GN	79.92 ± 0.51	3279
BL Variant I, lNCC, GN	109.69 ± 1.45	3705
BL Variant I, NGFs, GD	132.41 ± 1.71	2945
BL Variant I, MI, GD	328.27 ± 0.47	2977

Table 4. *Cont.*

Method	$time_{GPU}$ (s)	Peak VRAM (MBs)
SP Variant II, SSD, GN	215.61 ± 1.66	5769
SP Variant II, NCC, GN	216.29 ± 3.25	5899
SP Variant II, lNCC, GN	238.49 ± 3.25	6065
SP Variant II, NGFs, GD	159.21 ± 1.58	4555
SP Variant II, MI, GD	354.55 ± 2.11	4555
BL Variant II, SSD, GN	100.07 ± 0.30	2271
BL Variant II, NCC, GN	100.21 ± 0.88	2557
BL Variant II, lNCC, GN	130.53 ± 1.01	2921
BL Variant II, NGFs, GD	138.28 ± 1.87	2389
BL Variant II, MI, GD	331.28 ± 1.87	2357

5.2. LPBA40, IBSR18, CUMC12, and MGH10 Evaluation Results

Figure 7 shows the statistical distribution of the DSC values obtained with PDE-LDDMM for Klein databases [42]. As a benchmark, we include the results reported in [42] for affine registration (FLIRT) and three diffeomorphic registration methods: Diffeomorphic Demons, SyN, and Dartel. We also include the results of QuickSilver and VoxelMorph.

For LPBA40, Variant I with lNCC and Variant II with metrics from SSD to NGFs reached a performance similar to SyN with many outliers significantly reduced. For each metric, Variant II outperformed the corresponding Variant I. The worst performing results were consistently achieved by NGFs and Gauss–Newton–Krylov optimization. QuickSilver performed slightly better under SDD and NCC versions of Variant I. VoxelMorph was the worst-performing method for all metrics.

For IBSR18, SP and BL Variant I with lNCC, SP Variant II with NCC and BL Variant II with NCC and lNCC metrics were the best performing PDE-LDDMM methods. Their performance was slightly over the one exhibited by QuickSilver and greatly over the one obtained with VoxelMorph. However, in all cases, these methods underperformed SyN and Dartel.

For CUMC12, the best performing PDE-LDDMM methods were SP and BL Variant I with lNCC and Variant II with NCC and lNCC. As happened with IBSR18, these methods slightly outperformed QuickSilver while greatly outperformed VoxelMorph. It is remarkable the low performance of BL variants with NGF and Gauss–Newton–Krylov optimization. All the methods underperformed SyN and Dartel methods.

Finally, for MGH10, the best performance was achieved by variants I and II with lNCC similarity metric. It is remarkable the low performance of Variant I with NGF with gradient descent underperforming Gauss–Newton–Krylov optimizers. In this case, the methods underperformed SyN, while the best-performing methods showed a DSC distribution similar to Dartel. QuickSilver and VoxelMorph achieved performance similar to the SSD version of Variant I.

These results corroborate the better performance of Variant II over Variant I obtained in the evaluation with NIREP16 for the majority of metrics. The lNCC metric is positioned as the best-performing one for the majority of methods and databases. The NGFs metric has shown better performance for gradient descent optimization in the great majority of experiments. The best PDE-LDDMM combination of variants and metrics overpassed deep-learning based methods in all the datasets.

Regarding the consistent outperformance of SyN and Dartel over all the considered methods, we found out that SyN used a probabilistic image similarity metric while Dartel used tissue probability maps as inputs. The images in IBSR18, CUMC12, and MGH10 have low contrast, and, therefore, the algorithmic choices performed by SyN and Dartel overpass the use of challenging inputs. We have also seen that performing histogram equalization for contrast enhancement as in QuickSilver original paper [46] improved the evaluation results reaching SyN and Dartel performance. However, this preprocessing reduces the influence of the used metrics in the obtained DSCs and provides less informative results.
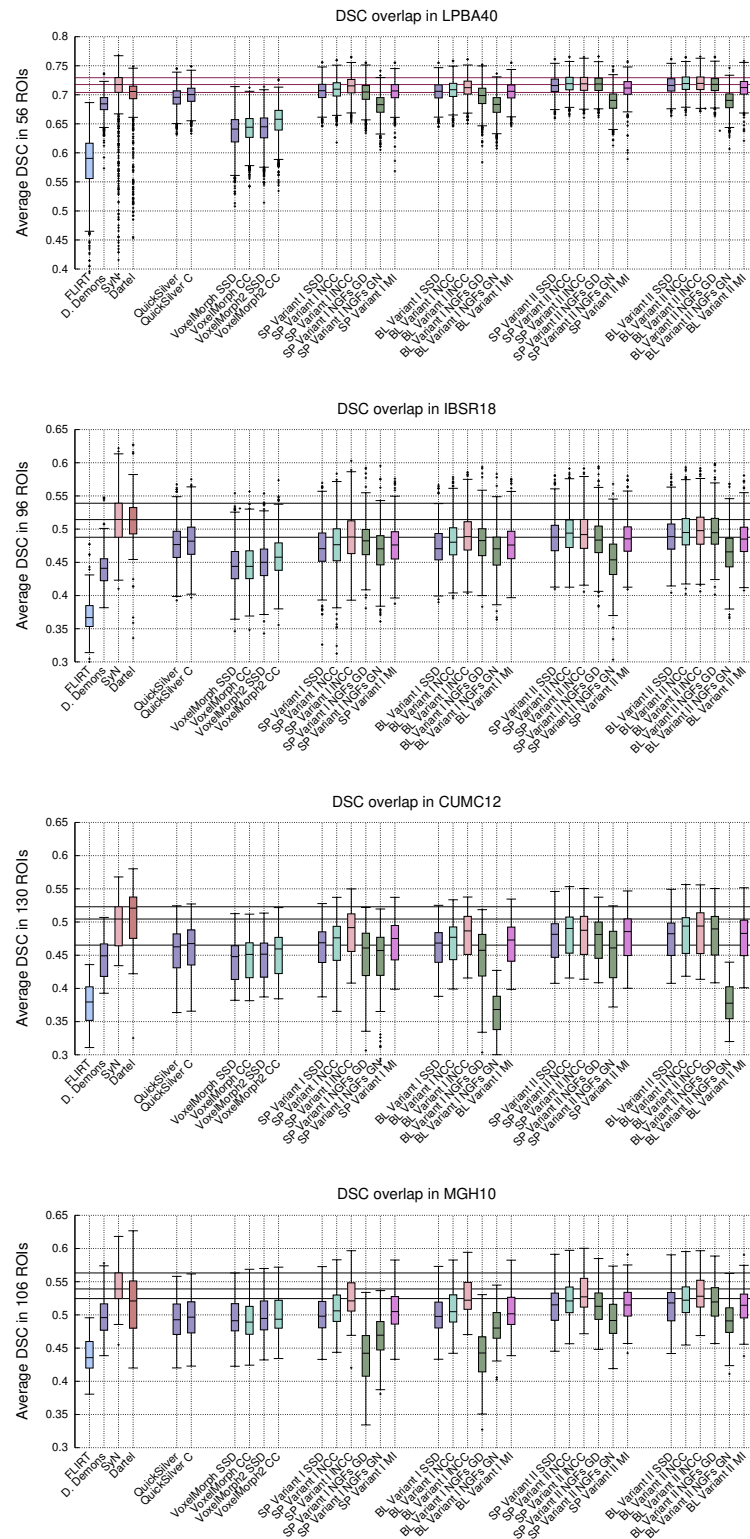
Figure 7. LPBA40, IBSR18, CUMC12, and MGH10. Distribution of the DSC values averaged over the manual segmentations in the registration experiments. The whiskers indicate the minimum and maximum of the DSC values. The horizontal lines indicate the first, second, and third quartiles of the SyN benchmark method, facilitating the comparison.

6. Multimodal Experiment

Figure 8 shows an axial view of the registration results obtained by NiftyReg, ANTS, SynthMorph, and BL PDE-LDDMM Variant II in the Oasis multimodal experiment. The worst-performing metrics are SSD (as expected) and NCC. All methods with lNCC, NGFs, and MI metrics provide acceptable registration results with subtle differences between the warps located at the ventricle front horns and the atrium and the distribution of gyri and sulci in the cerebral cortex. The most visually accurate methods are NiftyReg, ANTS-MI, SynthMorph, and PDE-LDDMM from lNCC and MI. It should be noticed that the registration results of NiftyReg are obtained at the cost of folding the transformations. SynthMorph provides excellent registration results but the used similarity metric is DSC over the segmented images, overpassing the direct use of images from different modalities.
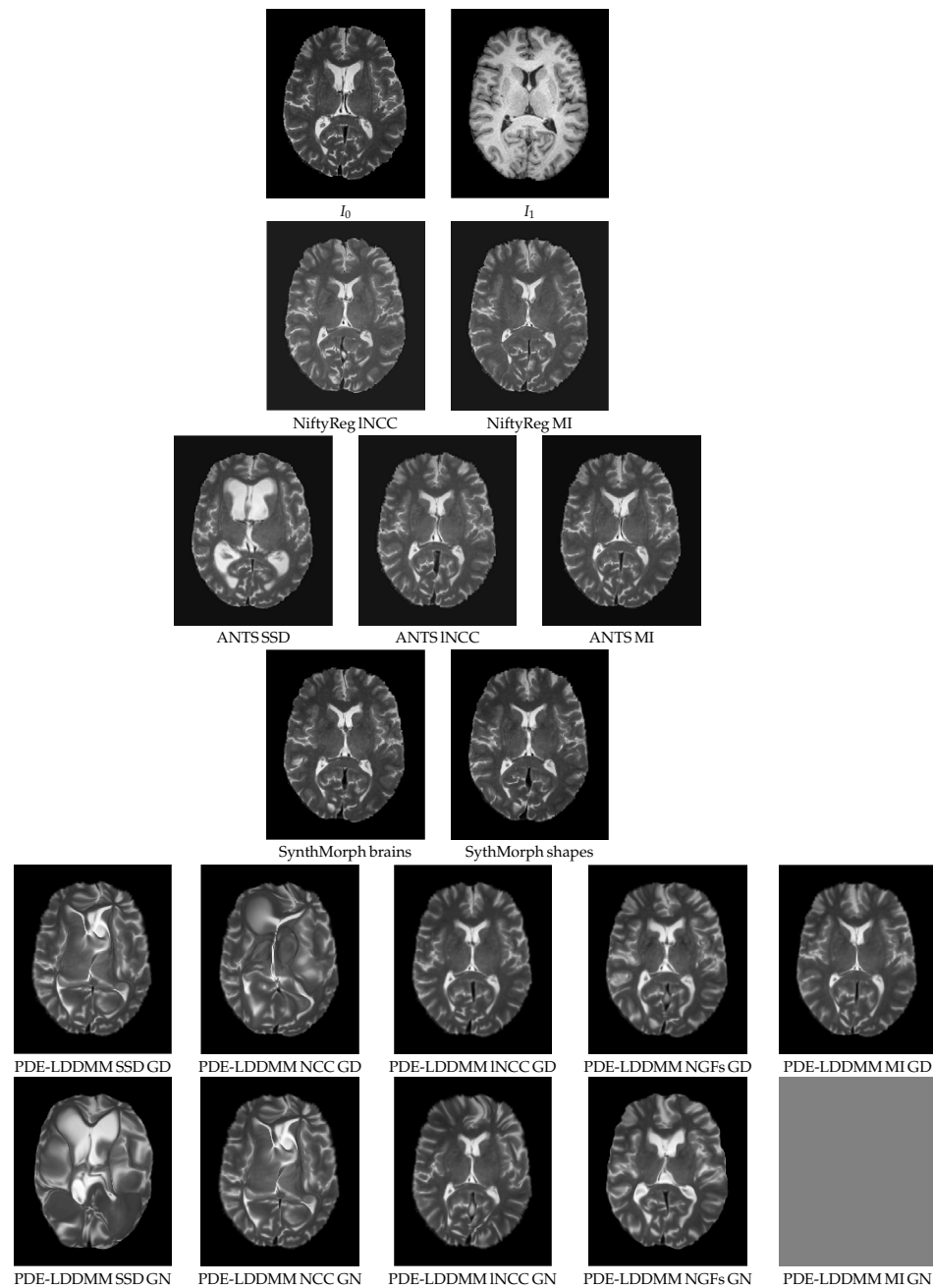
Figure 8. Oasis. BL PDE-LDDMM Variant II. Warped sources obtained for the methods considered in the multimodal simulated experiment.

7. Discussion and Conclusions

In this work, we have presented a unifying framework for introducing different image similarity metrics in the two best-performing variants of PDE-LDDMM with Gauss–Newton–Krylov optimization [22,33]. From the Lagrangian variational problem, we have identified that the change in the image similarity metric involves changing the initial adjoint and the initial incremental adjoint variables. We derived the equations of these variables for NCC, its local version (lNCC), NGFs, and MI. PDE-LDDMM with Gauss–Newton–Krylov optimization being successfully extended from SSD to NCC, and lNCC image similarity metrics. For NGFs, the method was not able to overpass gradient-descent optimization. With MI, the computation of the Hessian-matrix product required the product of dense matrices that requested more than 5000 GBs of memory, thus becoming far from feasible. Therefore, we obtained varying degrees of success in our initial objective.

The evaluation performed in NIREP16 database has shown the superiority of Variant II with respect to Variant I, as happened in [22,33]. In addition, the results reported for the BL version of Variant II were statistically indistinguishable from the SP (spatial) version. For any image similarity metric, BL Variant II overpassed the baseline established by ANTS. For BL Variant II, NCC and its local version were the best-performing metrics, closely followed by the gradient-descent version of NGFs. The superiority of these metrics was statistically significant. The outperformance of lNCC was quantified for the first time for ANTS diffeomorphic registration with gradient-descent and LPBA40 in [36]. Our best-performing variants overpassed QuickSilver, a supervised deep-learning method for diffeomorphic registration. In addition, they provided competitive results when compared with VoxelMorph with the added value of PDE-LDDMM being agnostic to the evaluation metric and providing purely diffeomorphic solutions.

The MSE_{rel} values were in agreement with the DSC distributions obtained with NCC and SSD. However, for lNCC, NGFs and MI, the correlation between the MSE_{rel} values and the DSC seen usually for SSD in previous works does not hold anymore.

The experiments with Klein databases corroborated the superiority of Variant II over Variant I for almost all the metrics. The evaluation in LPBA40 has shown how PDE-LDDMM based on the deformation state equation performs similarly to SyN for the majority of metrics with a reduced number of outliers. The evaluation in IBSR18, CUMC12, and MGH10 datasets has consistently shown lNCC as the best-performing metric for PDE-LDDMM. It is striking that the optimum DSC values greatly vary depending on the dataset used for evaluation. For example, SyN obtains an average DSC value greather than 0.7 for LPBA40 while the average DSC value is close to 0.5 for IBSR18, CUMC12, and MGH10 data. We believe that the disparity of the obtained DSC values depends on the geometry of the anatomies involved in the dataset which may downgrade the overall accuracy.

Although not being able to report functional Gauss–Newton–Krylov PDE-LDDMM methods for NGFs and MI has been disappointing, it encourages us to embed PDE-LDDMM into different optimization methods competing with gradient-descent as Gauss–Newton–Krylov does for the SSD, NCC, and lNCC metrics. In future work, we will address the problem with limited-memory BFGS or, in the framework of Krylov subspace methods, with the generalized minimal residual method (GMRES).

Our method has shown visually acceptable registration results on a challenging multi-modal intra-subject experiment. The results were competitive with SynthMorph, a deep-learning method that uses a loss function based on DSC from image segmentations. The experiment pointed out the differences between the combination of different optimization methods and metrics. In future work, we will explore in depth the influence of metric and optimization selection in the accuracy of multi-modal registration.

Despite the methodological improvements that have been subsequently proposed in PDE-LDDMM for efficiency (Gauss–Newton–Krylov optimization, band-limited parameterization, and Semi-Lagrangian Runge–Kutta integration), our PDE-LDDMM methods are able to compute a diffeomorphism in a volume of size $180 \times 210 \times 180$ in one to five minutes, depending on the variant and the metric. This may be considered a non-acceptable

amount of time in comparison with modern deep-learning approaches where the inference takes about one second. However, the time and resources needed for training are not usually considered in the comparison while they should be at least apportioned. In addition, deep-learning methods are not memory efficient while our proposed methods run in a commodity graphics card with a VRAM of less than 4 GBs.

BL Variant II with SSD, NCC, and lNCC has been recently included in the diffeomorphic normalization step into the pipeline of Spasov et al. [69] for the prediction of stable vs. progressive mild cognitive impairment (MCI) conversion in Alzheimer's disease with multi-task learning and Convolutional Neural Networks [70]. PDE-LDDMM overpassed ANTS-lNCC for this task, in terms of accuracy, sensitivity, and specificity. ANTS-lNCC obtained a median accuracy value of 84%, a sensitivity of 88% and specificity of 81%. Variant II with NCC achieved the best performing accuracy, with a median value of 89%, and sensitivity and specificity values among the best ones, with a median value of 94% and 91%, respectively. Indeed, NCC overpassed lNCC metric in this task, despite the comparable performance achieved by both metrics in the template-based evaluation presented in this work. As future work, we will perform a comprehensive study to find out the whys behind the improved performance of a given configuration with respect to the others.

Our PDE-LDDMM method may serve as a benchmark method for the exploration of different image similarity metrics in the loss function of deep-learning methods. In addition, it may be a good candidate in applications where there are not enough data to generate accurate learning-based models. Even more, it may be used as the backbone of hybrid approaches that combine traditional with modern learning-based models which are being pointed out as one promising research direction [55].

Appendix A. Stationary LDDMM

Stationary LDDMM was proposed in [67,71] as a lightweight approach to the original LDDMM problem proposed in [56]. The method replaces the non-stationary parameterization of diffeomorphisms by steady velocity fields and it uses Gauss–Newton optimization with a block-diagonal approximation of the Hessian matrix. The result is a method close to the LDDMM foundations and competitive with Diffeomorphic Demons [71,72].

In a nutshell, St-LDDMM is formulated from the minimization of the energy functional

$$E(v) = \frac{1}{2}\|v\|_V^2 + \frac{1}{\sigma^2}\|I_0 \circ \varphi^{-1} - I_1\|_{L^2}^2, \tag{A1}$$

where $v \in V$ is an steady velocity field and φ is the solution of the transport equation

$$\frac{d\varphi}{dt} = -v \circ \varphi, \tag{A2}$$

that is solved using scaling and squaring [73].

The gradient can be obtained from the first-order Gâteaux derivative of $E(v)$ and the identity

$$\partial_\eta E(v) = < \nabla_v E(v), \eta >_V . \tag{A3}$$

Thus, for the SSD image similarity metric,

$$\nabla_v E(v) = v - \frac{2}{\sigma^2} K\Big((I_0 \circ \varphi^{-1} - I_1) D(I_0 \circ \varphi^{-1}) \Big). \tag{A4}$$

The energy gradient can be similarly derived for the different E_{img} metrics considered in this work yielding

$$\nabla_v E(v) = v - \frac{2}{\sigma^2} K\Big(\alpha_g D(I_0 \circ \varphi^{-1}) \Big), \tag{A5}$$

where α_g is an scalar field corresponding with the differential of the metric with respect to the residual $I_0 \circ \varphi^{-1} - I_1$. Table A1 gathers the expressions of α_g, which can be easily identified from the derivations of the PDE-LDDMM methods proposed in this work.

The Hessian can be obtained from the second-order Gâteaux derivative of $E(v)$ and the identity

$$\partial_{\eta\eta} E(v) = < \eta, H_v E(v)\eta >_V . \tag{A6}$$

For Gauss–Newton optimization, the Hessian is approximated by a block-diagonal positive-definite matrix yielding

$$H_v E(v) \approx Id_d + \frac{2}{\sigma^2} K\Big(D(I_0 \circ \varphi^{-1}) \Big) \cdot K\Big(D(I_0 \circ \varphi^{-1}) \Big). \tag{A7}$$

for the SSD metric. For the other E_{img} metrics, the energy Hessian is given by

$$H_v E(v) \approx Id_d - \frac{2}{\sigma^2} K\Big(\alpha_h D(I_0 \circ \varphi^{-1}) \Big) \cdot K\Big(D(I_0 \circ \varphi^{-1}) \Big), \tag{A8}$$

where α_h is a scalar field corresponding with the differential of the scalar field α_g. Table A1 gathers the expressions of α_h.

St-LDDMM has been included in our comparison in order to demonstrate the superiority of PDE-LDDMM over unconstrained approaches and the correctness of the derivations of α_g and α_h for NGFs and Gauss–Newton optimization.

Table A1. Scalar fields from the gradients and Hessians of St-LDDMM with the different E_{img} metrics considered in this work. The values of A, B, C, and Θ are analogs of the ones derived in the main manuscript for PDE-LDDMM replacing $m(1)$ with $I_0 \circ \varphi^{-1}$. The notation $(\cdot)_\nu$ in lNCC indicates the restriction of the computation to a neighborhood of size ν.

Metric	α_g	α_h
SSD	$(I_0 \circ \varphi^{-1} - I_1)$	Id
NCC	$\frac{2}{C}\left(\frac{A}{B}\bar{I}_1 - \left(\frac{A}{B}\right)^2 \overline{I_0 \circ \varphi^{-1}} \right)$	$\frac{2}{C}\left(\Theta \bar{I}_1 - 2\frac{A}{B}\Theta \overline{I_0 \circ \varphi^{-1}} - \left(\frac{A}{B}\right)^2 \right)$
lNCC	$\left(\frac{2}{C}\left(\frac{A}{B}\bar{I}_1 - \left(\frac{A}{B}\right)^2 \overline{I_0 \circ \varphi^{-1}} \right) \right)_\nu$	$\left(\frac{2}{C}\left(\Theta \bar{I}_1 - 2\frac{A}{B}\Theta \overline{I_0 \circ \varphi^{-1}} - \left(\frac{A}{B}\right)^2 \right) \right)_\nu$
NGFs	$-2\nabla \cdot \left(\frac{A}{(BC)^2}\nabla I_1 - \left(\frac{A}{B^2C}\right)^2 \nabla(I_0 \circ \varphi^{-1}) \right)$	$2\alpha^T \cdot \alpha$
MI	$-\sum_{r,s} \frac{1}{N_x} \sum_x \frac{\partial \xi}{\partial r}\xi(s - \bar{I}_1(x))\left(1 + \log \frac{p_{I_0 \circ \varphi^{-1}, I_1}(r,s)}{p_{I_0 \circ \varphi^{-1}}(r)} \right)$	n.a.

Appendix B. Remarks on Normalized Gradient Fields (NGFs) Metric

Unlike NCC and lNCC metrics, the differentiation of λ for NGFs needed in Gauss–Newton–Krylov optimization yields a complex differential expression for $\delta\lambda(1)$. In addition, the computation of $\delta\lambda(1)$ in differential form may yield numerical problems during the optimization. For these reasons, the differentiation of λ for the computation of $\delta\lambda(1)$ has been approached in this work writing the differential operators in matrix form after problem discretization, following the ideas in [32].

Thus, let G_x, G_y, and G_z be the matrices associated with the discretization of the partial differential operators ∂_x, ∂_y, and ∂_z. Recall the expression of $\lambda(1)$ in Equation (59)

$$\lambda(1) = -\frac{2}{\sigma^2}\nabla \cdot \left(\frac{A}{(BC)^2}\nabla I_1 - \left(\frac{A}{B^2C}\right)^2 \nabla m(1) \right). \tag{A9}$$

Using the matrix form of the gradient ∇ and the divergence $\nabla\cdot$ operators, $\lambda(1)$ can be written as

$$
\begin{aligned}
\lambda(1) = \quad & \left(\frac{A}{(BC)^2}I_1 \odot Gx - \left(\frac{A}{B^2C}\right)^2 m(1) \odot Gx \right) \odot Gx \\
& + \left(\frac{A}{(BC)^2}I_1 \odot Gy - \left(\frac{A}{B^2C}\right)^2 m(1) \odot Gy \right) \odot Gy \\
& + \left(\frac{A}{(BC)^2}I_1 \odot Gz - \left(\frac{A}{B^2C}\right)^2 m(1) \odot Gz \right) \odot Gz
\end{aligned}
\tag{A10}
$$

where $\odot$ represents the Hadamard matrix product. Then, denoting by Λ the matrix representation of $\lambda(1)$, the Gauss–Newton approximation of $\delta\lambda(1)$ can be computed from

$$\delta\lambda(1) = dm(1)^T \cdot (\Lambda^T \cdot 2Id \cdot \Lambda). \tag{A11}$$

Figure A1 shows the block-diagonal pattern of the matrix $\Lambda^T \cdot 2Id \cdot \Lambda$, involved in the computation of $\delta\lambda(1)$. The matrix is a band matrix with twenty-five not-null diagonals. The distribution of not-null diagonals involves one central pentadiagonal structure, two symmetrically located tridiagonal structures, and the remaining symmetrically located diagonals up to completing the twenty-five diagonals of the band structure. The complexity of the diagonal pattern gives an idea of the complexity of the stencil associated with the differential expression for $\delta\lambda(1)$.
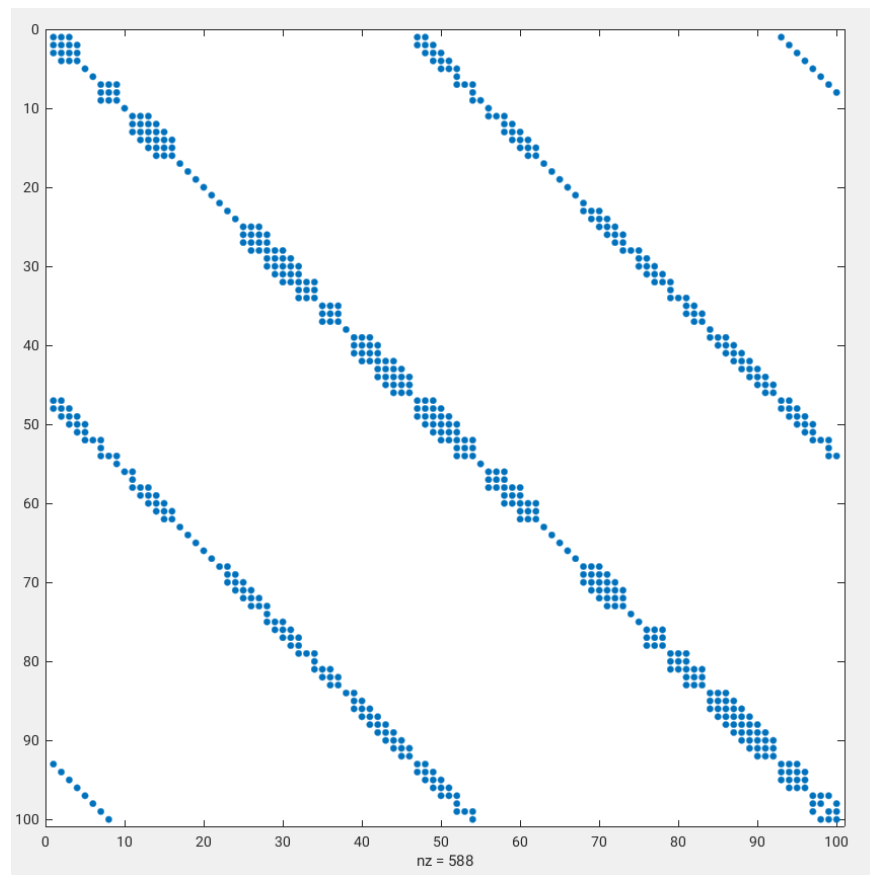
Figure A1. Block-diagonal pattern of the first 100×100 elements of the matrix $(\Lambda^T \cdot 2Id \cdot \Lambda)$.

Inspired on the simpler block-diagonal structure of the Gauss–Newton version of St-LDDMM proposed in [67], we have compared the performance of NGFs PDE-LDDMM using diagonal, penta-diagonal, and all-diagonal truncations of $(\Lambda^T \cdot 2Id \cdot \Lambda)$. Figure A2 shows the DSC metrics obtained by the different methods in NIREP16. The figure shows that the diagonal approximation slightly outperforms the penta- and full-diagonal approximations for both Variant I and II. The performance of the method with the diagonal approximation shows an identical distribution between the spatial variants and approximately identical between the BL variants. For Variant I, the performance of Gauss–Newton–Krylov optimization is similar to gradient-descent optimization. However, for Variant II, the performance of gradient-descent optimization is strikingly superior.

Table A2 shows the mean and the standard deviation of MSE_{rel}, $\|g\|_{\infty,rel}$, and the extrema of the Jacobian determinant obtained with the methods. For the spatial methods, gradient descent obtains the lowest MSE_{rel} values. For the BL methods, Gauss–Newton with the diagonal truncation shares the lowest MSE_{rel} values with gradient descent. The Jacobians reach more extreme values for the Gauss–Newton methods. These results break with the trend that we have observed for SSD and NCC-like metrics and open the possibility that, in some cases, gradient-descent finds better performing minima than Gauss–Newton–Krylov and it is, therefore, a preferable option.
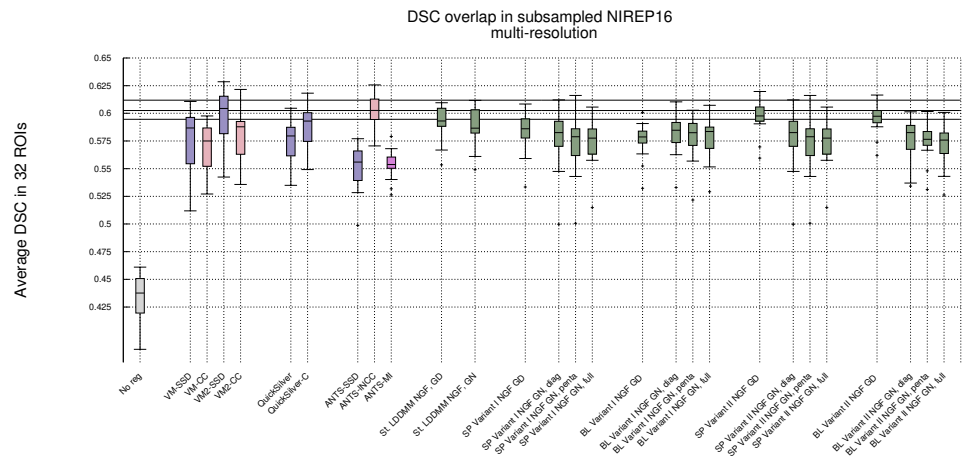
Figure A2. NIREP16. Volume overlap obtained by the NGFs PDE-LDDMM methods considered in this work. Methods differ on the approximation used for matrix ($\Lambda^T \cdot 2Id \cdot \Lambda$): diagonal, penta-diagonal, and full-diagonal.

Table A2. NIREP16. Convergence results. Mean and standard deviation of the relative image similarity error expressed in % (MSE_{rel}), the relative gradient magnitude ($\|g\|_{\infty,rel}$), and the Jacobian determinant extrema associated with the transformation $(\phi_1^v)^{-1}$.

	Multi-Resolution			
Metric	MSE_{rel}	$\|g\|_{\infty,rel}$	$\min(J(\phi_1^v)^{-1})$	$\max(J(\phi_1^v)^{-1})$
SP Variant I, NGFs, GD	22.98 ± 2.54	0.56 ± 0.66	0.14 ± 0.05	4.57 ± 0.46
SP Variant I, NGFs, GN, FD	29.08 ± 2.55	0.66 ± 0.26	0.05 ± 0.04	6.99 ± 1.74
SP Variant I, NGFs, GN, PD	28.86 ± 3.32	0.64 ± 0.17	0.04 ± 0.02	8.95 ± 2.32
SP Variant I, NGFs, GN, D	27.22 ± 2.31	0.68 ± 0.24	0.03 ± 0.02	8.96 ± 1.75
SP Variant II, NGFs, GD	20.63 ± 2.78	0.57 ± 0.28	0.23 ± 0.05	6.88 ± 1.59
SP Variant II, NGFs, GN, FD	24.20 ± 2.41	0.84 ± 0.15	0.23 ± 0.03	4.76 ± 1.08
SP Variant II, NGFs, GN, PD	23.92 ± 3.02	0.94 ± 0.08	0.20 ± 0.03	7.04 ± 2.20
SP Variant II, NGFs, GN, D	23.42 ± 2.74	0.85 ± 0.16	0.19 ± 0.03	7.38 ± 2.40
BL Variant I, NGFs, GD	24.38 ± 2.65	0.08 ± 0.02	0.21 ± 0.06	5.41 ± 0.85
BL Variant I, NGFs, GN, FD	24.10 ± 2.96	0.30 ± 0.15	0.09 ± 0.02	9.62 ± 1.90
BL Variant I, NGFs, GN, PD	23.45 ± 2.47	0.33 ± 0.19	0.09 ± 0.02	12.93 ± 3.35
BL Variant I, NGFs, GN, D	23.31 ± 2.16	0.37 ± 0.23	0.09 ± 0.02	12.07 ± 3.16
BL Variant II, NGFs, GD	21.32 ± 2.80	0.09 ± 0.05	0.22 ± 0.05	5.82 ± 0.96
BL Variant II, NGFs, GN, FD	23.04 ± 2.25	0.51 ± 0.20	0.07 ± 0.02	7.72 ± 2.24
BL Variant II, NGFs, GN, PD	22.39 ± 2.28	0.47 ± 0.23	0.08 ± 0.02	7.66 ± 1.49
BL Variant II, NGFs, GN, D	21.98 ± 2.14	0.53 ± 0.32	0.07 ± 0.02	8.16 ± 1.50

References

1. Sotiras, A.; Davatzikos, C.; Paragios, N. Deformable Medical Image Registration: A Survey. *IEEE Trans. Med. Imaging* **2013**, *32*, 1153–1190. [CrossRef] [PubMed]
2. Hua, X.; Leow, A.D.; Parikshak, N.; Lee, S.; Chiang, M.C.; Toga, A.W.; Jack, C.R.; Weiner, M.W.; Thompson, P.M.; Alzheimer's Disease Neuroimaging Initiative. Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: An MRI study of 676 AD, MCI, and normal subjects. *Neuroimage* **2008**, *43*, 458–469. [CrossRef] [PubMed]
3. Miller, M.I.; Qiu, A. The emerging discipline of Computational Functional Anatomy. *Neuroimage* **2009**, *45*, 16–39. [CrossRef] [PubMed]
4. Qiu, A.; Albert, M.; Younes, L.; Miller, M.I. Time sequence diffeomorphic metric mapping and parallell transport track time-dependent shape changes. *Neuroimage* **2009**, *45*, 51–60. [CrossRef]
5. Durrleman, S.; Pennec, X.; Trouve, A.; Braga, J.; Gerig, G.; Ayache, N. Toward a Comprehensive Framework for the Spatiotemporal Statistical Analysis of Longitudinal Shape Data. *Int. J. Comput. Vision* **2013**, *103*, 22–59. [CrossRef]
6. Zhang, M.; Fletcher, P.T. Bayesian Principal Geodesic Analysis for estimating Intrinsic diffeomorphic image variability. *Med. Image Anal.* **2015**, *25*, 37–44. [CrossRef]

7. HadjHamou, M.; Lorenzi, M.; Ayache, N.; Pennec, X. Longitudinal analysis of image time series with diffeomorphic deformations: A computational framework based on stationary velocity fields. *Front. Neurosci.* **2016**, *10*, 236.

8. Liu, Y.; Li, Z.; Ge, Q.; Lin, N.; Xiong, M. Deep feature selection and causal analysis of Alzheimer's disease. *Front. Neurosci.* **2019**, *13*, 1198. [CrossRef]

9. Schnabel, J.A.; Heinrich, M.P.; Papiez, B.W.; Brady, J.M. Advances and challenges in deformable image registration: From image fusion to complex motion modelling. *Med. Image Anal.* **2016**, *33*, 145–148. [CrossRef]

10. Uneri, A.; Goerres, J.; de Silva, T. Deformable 3D-2D Registration of Known Components for Image Guidance in Spine Surgery. In Proceedings of the 19th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'16), Athens, Greece, 17–21 October 2016; pp. 124–132.

11. Girija, J.; Krishna, G.N.; Chenna, P. 4D medical image registration: A survey. In Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS17), Palladam, India, 7–8 December 2017.

12. Mansi, T.; Pennec, X.; Sermesant, M.; Delingette, H.; Ayache, N. iLogDemons: A Demons-Based Registration Algorithm for Tracking Incompressible Elastic Biological Tissues. *Int. J. Comput. Vis.* **2011**, *92*, 92–111. [CrossRef]

13. Mang, A.; Biros, G. An inexact Newton-Krylov algorithm for constrained diffeomorphic image registration. *SIAM J. Imaging Sci.* **2015**, *8*, 1030–1069. [CrossRef] [PubMed]

14. Mang, A.; Biros, G. Constrained H1 regularization schemes for diffeomorphic image registration. *SIAM J. Imaging Sci.* **2016**, *9*, 1154–1194. [CrossRef] [PubMed]

15. Mang, A.; Gholami, A.; Davatzikos, C.; Biros, G. PDE-constrained optimization in medical image analysis. *Optim. Eng.* **2018**, *19*, 765–812. [CrossRef]

16. Hart, G.L.; Zach, C.; Niethammer, M. An optimal control approach for deformable registration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'09), Miami, FL, USA, 20–25 June 2009.

17. Vialard, F.X.; Risser, L.; Rueckert, D.; Cotter, C.J. Diffeomorphic 3D Image Registration via Geodesic Shooting using an Efficient Adjoint Calculation. *Int. J. Comput. Vision* **2011**, *97*, 229–241. [CrossRef]

18. Polzin, T.; Niethammer, M.; Heinrich, M.P.; Handels, H.; Modersitzki, J. Memory efficient LDDMM for lung CT. In Proceedings of the 19th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'16), Athens, Greece, 17–21 October 2016; pp. 28–36.

19. Mang, A.; Gholami, A.; Davatzikos, C.; Biros, G. Claire: A distributed-memory solver for constrained large deformation diffeomorphic image registration. *SIAM J. Sci. Comput.* **2019**, *41*, C548–C584. [CrossRef]

20. Hernandez, M. Band-Limited Stokes Large Deformation Diffeomorphic Metric Mapping. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 362–373. [CrossRef]

21. Hernandez, M. PDE-constrained LDDMM via geodesic shooting and inexact Gauss-Newton-Krylov optimization using the incremental adjoint Jacobi equations. *Phys. Med. Biol.* **2019**, *64*, 025002. [CrossRef]

22. Hernandez, M. A comparative study of different variants of Newton-Krylov PDE-constrained Stokes-LDDMM parameterized in the space of band-limited vector fields. *SIAM J. Imaging Sci.* **2019**, *12*, 1038–1070. [CrossRef]

23. Hogea, C.; Davatzikos, C.; Biros, G. Brain-tumor interaction biophysical models for medical image registration. *SIAM J. Imaging Sci.* **2008**, *30*, 3050–3072. [CrossRef]

24. Mang, A.; Toma, A.; Schuetz, T.A.; Becker, S.; Eckey, T.; Mohr, C.; Petersen, D.; Buzug, T.M. Biophysical modeling of brain tumor progression: From unconditionally stable explicit time integration to an inverse problem with parabolic PDE constraints for model calibration. *Med. Phys.* **2012**, *39*, 4444–4460. [CrossRef]

25. Gholami, A.; Mang, A.; Biros, G. An inverse problem formulation for parameter estimation of a reaction-diffusion model for low grade gliomas. *J. Math. Biol.* **2016**, *72*, 409–433. [CrossRef] [PubMed]

26. Khanal, B.; Lorenzi, M.; Ayache, N.; Pennec, X. A biophysical model of brain deformation to simulate and analyze longitudinal MRIs of patients with Alzheimer's disease. *Neuroimage* **2016**, *134*, 35–52. [CrossRef] [PubMed]

27. Yang, X.; Niethammer, M. Uncertainty quantification for LDDMM using a low-rank Hessian approximation. In Proceedings of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'15), Lecture Notes in Computer Science, Munich, Germany, 5–9 October 2015; Volume 9350, pp. 289–296.

28. Wang, J.; Wells, W.M.; Golland, P.; Zhang, M. Registration uncertainty quantification via low-dimensional characterization of geometric deformations. *Magn. Reson. Imaging* **2019**, *64*, 122–131. [CrossRef] [PubMed]

29. Brunn, M.; Himthani, N.; Biros, G.; Mehl, M. Fast GPU 3D diffeomorphic image registration. *J. Parallel Distrib. Comput.* **2020**, *149*, 149–162. [CrossRef] [PubMed]

30. Kutten, K.S.; Charon, N.; Miller, M.I.; Ratnanather, J.T.; Deisseroth, K.; Ye, L.; Vogelstein, J.T. A Diffeomorphic Approach to Multi-modal Registration with Mutual Information: Applications to CLARITY Mouse Brain Images. *arXiv* **2016**, arXiv:1612.00356v2.

31. Modersitzki, J. *Numerical Methods for Image Registration*; Oxford University Press: Oxford, UK, 2004.

32. Modersitzki, J. *FAIR: Flexible Algorithms for Image Registration*; SIAM: Philadelphia, PA, USA, 2009.

33. Hernandez, M. Combining the Band-Limited Parameterization and Semi-Lagrangian Runge–Kutta Integration for Efficient PDE-Constrained LDDMM. *J. Math. Imaging Vis.* **2021**, *63*, 555–579. [CrossRef]

34. Rueckert, D.; Sonoda, L.I.; Hayes, C.; Hill, D.L.; Leach, M.O.; Hawkes, J. Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images. *IEEE Trans. Med. Imaging* **1999**, *18*, 712–721. [CrossRef]

35. Haber, E.; Modersitzki, J. Beyond Mutual Information: A Simple and Robust Alternative. In *Bildverarbeitung Für Die Medizin*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 350–355.

36. Avants, B.B.; Tustison, N.J.; Song, G.; Cook, P.A.; Klein, A.; Gee, J.C. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* **2011**, *54*, 2033–2044. [CrossRef]

37. Myronenko, A.; Song, X. Intensity-based Image Registration by Minimizing Residual Complexity. *IEEE Trans. Med. Imaging* **2010**, *29*, 1882–1891. [CrossRef]

38. Lorenzi, M.; Ayache, N.; Frisoni, G.; Pennec, X.; ADNI. LCC-Demons: A robust and accurate symmetric diffeomorphic registration algorithm. *Neuroimage* **2013**, *81*, 470–483. [CrossRef]

39. Hernandez, M. Efficient momentum conservation constrained PDE-LDDMM with Gauss–Newton–Krylov optimization, Semi-Lagrangian Runge–Kutta solvers, and the band-limited parameterization. *J. Comput. Sci.* **2021**, *55*, 101470. [CrossRef]

40. Avants, B.B.; Epstein, C.L.; Grossman, M.; Gee, J.C. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* **2008**, *12*, 26–41. [CrossRef] [PubMed]

41. Christensen, G.E.; Geng, X.; Kuhl, J.G.; Bruss, J.; Grabowski, T.J.; Pirwani, I.A.; Vannier, M.W.; Allen, J.S.; Damasio, H. Introduction to the Non-rigid Image Registration Evaluation Project (NIREP). In Proceedings of the 3rd International Workshop on Biomedical Image Registration (WBIR'06), Utrecht, The Netherlands, 9–11 July 2006; Volume 4057, pp. 128–135.

42. Klein, A.; Andersson, J.; Ardekani, B.A.; Ashburner, J.; Avants, B.; Chiang, M.C.; Christensen, G.E.; Collins, D.L.; Gee, J.; Hellier, P.; et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* **2009**, *46*, 786–802. [CrossRef] [PubMed]

43. Dosovitskiy, A.; Fischer, P.; Ilg, E.; Hausser, P.; Hazirbas, C.; Golkov, V.; Van Der Smagt, P.; Cremers, D.; Brox, T. FlowNet: Learning Optical Flow with Convolutional Networks. In Proceedings of the 15th IEEE International Conference on Computer Vision (ICCV'15), Santiago, Chile, 11–18 December 2015.

44. Boveiri, H.; Khayami, R.; Javidan, R.; Mehdizadeh, A. Medical image registration using deep neural networks: A comprehensive review. *Comput. Electr. Eng.* **2020**, *87*, 106767. [CrossRef]

45. Rohe, M.M.; Datar, M.; Heimann, T.; Sermesant, M.; Pennec, X. SVF-Net: Learning deformable image registration using shape matching. In Proceedings of the 20th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'17), Quebec City, QC, Canada, 11–13 September 2017; pp. 266–274.

46. Yang, X.; Kwitt, R.; Styner, M.; Niethammer, M. Quicksilver: Fast Predictive Image Registration—A Deep Learning Approach. *Neuroimage* **2017**, *158*, 378–396. [CrossRef]

47. Dalca, A.V.; Blakrishnan, G.; Guttag, J.; Sabuncu, M. Unsupervised learning for fast probabilistic diffeomorphic registration. In Proceedings of the 21th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'18), Granada, Spain, 16–20 September 2018; pp. 729–738.

48. Krebs, J.; Delingetter, H.; Mailhe, B.; Ayache, N.; Mansi, T. Learning a probabilistic model for diffeomorphic registration. *IEEE Trans. Med. Imaging* **2019**, *38*, 2165–2176. [CrossRef]

49. Fan, J.; Cao, X.; Yap, P.; Shen, D. BIRNet: Brain image registration using dual-supervised fully convolutional networks. *Med. Image Anal.* **2019**, *54*, 193–206. [CrossRef]

50. Balakrishnan, G.; Zhao, A.; Sabuncu, M.R.; Guttag, J.; Dalca, A.V. VoxelMorph: A Learning Framework for Deformable Medical Image Registration. *IEEE Trans. Med. Imaging* **2019**, *38*, 1788–1800. [CrossRef]

51. Wang, J.; Zhang, M. DeepFLASH: An efficient network for learning-based medical image registration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'20), Seattle, WA, USA, 13–20 June 2020.

52. Mok, T.; Chung, A. Large Deformation Diffeomorphic Image Registration with Laplacian Pyramid Networks. In Proceedings of the 23th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'20), Lima, Peru, 4–8 October 2020.

53. Hoffmann, M.; Billot, B.; Greve, D.; Iglesias, J.; Fischl, B.; Dalca, A. SynthMorph: Learning image registration without images. *IEEE Trans. Med. Imaging* **2021**. [CrossRef]

54. Hoopes, A.; Hoffmann, M.; Fischl, B.; Guttag, J.; Dalca, A. HyperMorph: Amortized Hyperparameter Learning for Image Registration. In Proceedings of the International Conference on Information Processing and Medical Imaging (IPMI'21), Virtual, 28–30 June 2021; Volume 12729, pp. 3–17.

55. Hering, A.; Hansen, L.; Mok, T.C.; Chung, A.; Siebert, H.; Häger, S.; Lange, A.; Kuckertz, S.; Heldmann, S.; Shao, W.; et al. Learn2Reg: Comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning. *arXiv* **2021**, arXiv:2112.04489.

56. Beg, M.F.; Miller, M.I.; Trouve, A.; Younes, L. Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms. *Int. J. Comput. Vis.* **2005**, *61*, 139–157. [CrossRef]

57. Younes, L.; Qiu, A.; Winslow, R.L.; Miller, M.I. Transport of Relational Structures in Groups of Diffeomorphisms. *J. Math. Imaging Vis.* **2008**, *32*, 41–56. [CrossRef] [PubMed]

58. Zhang, M.; Fletcher, T. Fast diffeomorphic image registration via Fourier-Approximated Lie algebras. *Int. J. Comput. Vis.* **2019**, *127*, 61–73. [CrossRef]

59. Zhang, M.; Fletcher, P.T. Finite-Dimensional Lie algebras for fast diffeomorphic image registration. In Proceedings of the International Conference on Information Processing and Medical Imaging (IPMI'15), Isle of Skye, UK, 28 June–3 July 2015.

60. Zhang, M.; Liao, R.; Dalca, A.V.; Truk, E.A.; Luo, J.; Grant, P.E.; Golland, P. Frequency Diffeomorphisms for Efficient Image Registration. In Proceedings of the International Conference on Information Processing and Medical Imaging (IPMI'17), Boone, NC, USA, 25–30 June 2017; pp. 559–570.

61. Dame, A.; Marchand, E. Second-order optimization of mutual information for real-time image registration. *IEEE Trans. Image Process.* **2012**, *21*, 4190–4203. [CrossRef]

62. Ou, Y.; Akbari, H.; Bilello, M.; Da, X.; Davatzikos, C. Comparative evaluation of registration algorithms in different brain databases with varying difficulty: Results and insights. *IEEE Trans. Med. Imaging* **2014**, *33*, 2039–2065. [CrossRef]

63. Hernandez, M. Primal-Dual convex optimization in Large Deformation Diffeomorphic Metric Mapping: LDDMM meets robust regularizers. *Phys. Med. Biol.* **2017**, *62*, 9067–9098. [CrossRef]

64. Hernandez, M. Primal-dual optimization strategies in Huber-L1 optical flow with temporal subspace constraints for non-rigid sequence registration. *Image Vis. Comput.* **2018**, *69*, 44–67. [CrossRef]

65. Ruijters, D.; Thevenaz, P. GPU Prefilter for Accurate Cubic B-Spline Interpolation. *Comput. J.* **2012**, *55*, 15–20. [CrossRef]

66. Trefethen, L.N. *Spectral Methods in Matlab*; SIAM: Philadelphia, PA, USA, 2000.

67. Hernandez, M. Gauss-Newton inspired preconditioned optimization in large deformation diffeomorphic metric mapping. *Phys. Med. Biol.* **2014**, *59*, 6085–6115. [CrossRef]

68. Rohlfing, T. Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable. *IEEE Trans. Med. Imaging* **2012**, *31*, 153–163. [CrossRef]

69. Spasov, S.E.; Passamonti, L.; Duggento, A.; Lio, P.; Toschi, N.; Alzheimer's Disease Neuroimaging Initiative. A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. *Neuroimage* **2019**, *189*, 276–287. [CrossRef] [PubMed]

70. Ramon-Julvez, U.; Hernandez, M.; Mayordomo, E.; Alzheimer's Disease Neuroimaging Initiative. Analysis of the influence of diffeomorphic normalization in the prediction of stable vs. progressive MCI conversion with convolutional neural networks. In Proceedings of the 17th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'20), Iowa City, IA, USA, 3–7 April 2020.

71. Hernandez, M.; Olmos, S. Gauss-Newton optimization in diffeomorphic registration. In Proceedings of the 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'08), Paris, France, 14–17 May 2008.

72. Vercauteren, T.; Pennec, X.; Perchant, A.; Ayache, N. Diffeomorphic Demons: Efficient non-parametric image registration. *Neuroimage* **2009**, *45*, S61–S72. [CrossRef] [PubMed]

73. Arsigny, V.; Commonwick, O.; Pennec, X.; Ayache, N. Statistics on Diffeomorphisms in a Log-Euclidean framework. In Proceedings of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'06), Lecture Notes in Computer Science, Copenhagen, Denmark, 1–6 October 2006; Volume 4190, pp. 924–931.