



A Physics-Informed Deep Learning Deformable Medical Image Registration Method Based on Neural ODEs

Amirhossein Amiri-Hezaveh¹ · Shelly Tan¹ · Qing Deng¹ · David Umulis¹ · Lauren Cunniff² ·
Johannes Weickenmeier³ · Adrian Buganza Tepole¹

Received: 9 June 2024 / Accepted: 3 May 2025 / Published online: 8 June 2025

© The Author(s) 2025

Abstract

An unsupervised machine learning method is introduced to align medical images in the context of the large deformation elasticity coupled with growth and remodeling biophysics. The technique, which stems from the principle of minimum potential energy in solid mechanics, consists of two steps: Firstly, in the predictor step, the geometric registration is achieved by minimizing a loss function composed of a dissimilarity measure and a regularizing term. Secondly, the physics of the problem, including the equilibrium equations along with growth mechanics, are enforced in a corrector step by minimizing the potential energy corresponding to a Dirichlet problem, where the predictor solution defines the boundary condition and is maintained by distance functions. The features of the new solution procedure, as well as the nature of the registration problem, are highlighted by considering several examples. In particular, registration problems containing large non-uniform deformations caused by extension, shearing, and bending of multiply-connected regions are used as benchmarks. In addition, we analyzed a benchmark biological example (registration for brain data) to showcase that the new deep learning method competes with available methods in the literature. We then applied the method to various datasets. First, we analyze the regrowth of the zebrafish embryonic fin from confocal imaging data. Next, we evaluate the quality of the solution procedure for two examples related to the brain. For one, we apply the new method for 3D image registration of longitudinal magnetic resonance images of the brain to assess cerebral atrophy, where a first-order ODE describes the volume loss mechanism. For the other, we explore cortical expansion during early fetal brain development by coupling the elastic deformation with morphogenetic growth dynamics. The method and examples show the ability of our framework to attain high-quality registration and, concurrently, solve large deformation elasticity balance equations and growth and remodeling dynamics.

Keywords Deformable image registration · Physics-informed neural networks · Brain registration · Growth and remodeling · Brain development · Brain atrophy · Zebrafish biophysics · Neural ordinary differential equations · Medical image analysis

1 Introduction

Image registration aims to obtain a smooth map that aligns the source and target images. There are two main categories of registrations: affine (or rigid) and deformable (non-rigid)

Communicated by Ziyue Xu.

✉ Adrian Buganza Tepole
abuganza@purdue.edu

¹ School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA

² Department of Mechanical Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA

³ Department of Engineering Science, University of Oxford, Oxford OX3 7DQ, UK

registrations. In the earlier, the transformation consists of specific modes of deformations, including scaling, shearing, translation, and rotation. These transformations can be constructed using homogeneous coordinates and matrix transformation of size 4×4 and 3×3 for 3d and 2d images, respectively. Due to its finite parameters, this type of transformation is a good candidate when anatomical structures are mostly preserved during deformation. On the other hand, the latter is accompanied by more degrees of freedom, which can capture anatomical changes, such as embryo development and brain variations in different individuals, among other examples.

A subcategory of the non-rigid registration is physics-based models. Some of these models are based on the principles of continuum mechanics. The balance of internal and

external forces dictates the transformation in this framework. The external forces are typically defined through similarity measures, e.g., cross-correlation of intensity of target and moving images (Pawar et al., 2018). The underlying balance laws are written either according to infinitesimal strain theory (Broit, 1981) or fluid flow algorithms (Christensen et al., 1994, 1996). The latter was specifically proposed to introduce large deformation theory in the context of the registration. In parallel, the utilization of nonlinear solid mechanics, which was manifested by using hyper-elastic constitutive equations and including nonlinear terms in the strain measure, can be seen, for instance, in Rabbitt et al. (1995), Pennec et al. (2005), Yanovsky et al. (2008), Pawar et al. (2022) (see more references in Sotiras et al. (2013)). As mentioned in Sotiras et al. (2013), from the biomedical application perspective, it is significant that the resulting transformation is invertible, and both the transformation and its inversion are differentiable (diffeomorphism). This group of transformations can be generated when the deformation is modeled as a (pseudo) flow, which is governed by the transport equation (Miller et al., 2002). This framework was originally proposed in the context of continuum flow algorithms (Christensen et al., 1996) and further developed in as large deformation diffeomorphic metric mapping (LDDMM) approaches (Joshi & Miller, 2000; Beg et al., 2005; Cao et al., 2005; Ceritoglu et al., 2009; Oishi et al., 2009; Hernandez et al., 2009; Ceritoglu et al., 2010), among others.

Apart from the above methods, the widespread use of deep learning methods in medical registration has been increasingly emerging. These methods can be mainly categorized into (Zou et al., 2022; Haskins et al., 2020) 1) deep interactive methods, 2) fully supervised methods, 3) unsupervised methods, and 4) weakly supervised. Among them, the development of unsupervised methods is significant due to the difficulty in providing a sufficient amount of ground truth data for the supervised methods (Zou et al., 2022). In the case of similarity-based unsupervised methods (Zou et al., 2022), the transformation is a neural network whose parameters are determined from the minimization of the loss function that consists of similarity measure and regularization term. De Vos et al. (2017) and Balakrishnan et al. (2019) are two examples of unsupervised learning approaches, among others (López et al., 2023). In contrast to classical methods, the registration in VoxelMorph is accomplished globally. That is, instead of aligning each pair of images separately, a global transformation is defined and parameterized by training a convolution neural network where the inputs are pairs of images. While this global training leads to a fast registration for unseen pairs of images, this method may not give good results for large deformations and is not guaranteed to provide diffeomorphic maps (Shen et al., 2019; Kuang & Schmah, 2019). However, there is a new trend to develop

deep learning methods that learn diffeomorphic maps such as (Mok & Chung, 2020; Wang et al., 2024).

The resulting transformation of the classical frameworks lies in the solution of governing partial differential equations, which were mainly solved by means of numerical methods such as finite difference and finite element methods (FEM). However, the universal approximator property of feed-forward deep neural networks (Hornik et al., 1989) motivated researchers to employ neural networks as solvers for physical boundary value problems (so-called physics-informed neural networks (PINN)). Early attempts in this regard can be seen in Lagaris et al. (1997), Lagaris et al. (1998), Lagaris et al. (2000), McFall (2006), McFall and Mahan (2009). However, with the advent of libraries such as Tensorflow, Pytorch, and JAX, where the differentiation is computed through automatic differentiation in an efficient manner, there is a marked tendency to exploit deep neural networks as function approximators in various physical problems. These methods are mostly meshless, and there exist variants of it such as collocation (Sirignano & Spiliopoulos, 2018; Raissi et al., 2019), or integral forms (Kharazmi et al., 2019, 2021; Jagtap & Karniadakis, 2021; Rezaei et al., 2022). One long-standing problem with the meshless methods is the satisfaction of the Dirichlet boundary condition since the function expansion does not satisfy partial unity. In the classical version of PINN, these boundary conditions are enforced by adding an extra penalty term in the loss function, which leads to an inexact imposition of the boundary conditions. This approximation is negatively reflected in the accuracy of the solution (see (Sukumar & Srivastava, 2022)). Hence, Sukumar and Srivastava (2022), by using the approximate distance function—which can be generated from either theory of R functions (Rvachev, 1982; Biswas & Shapiro, 2004) or theory of mean value potential fields, defined a new format of solution and exactly fulfilled the displacement boundary conditions.

In this study, we propose a PINN-like approach in the context of large deformation solid mechanics for the registration of deformable (biomedical) images. The method consists of two parts: 1) predictor and 2) corrector. In the predictor part, the central focus is on obtaining a map that maximizes the similarity measure between moving and source images. To this end, we employ the following conventional multi-objective:

$$\arg \min_{\theta} (sim(T(\varphi(\mathbf{X}; \theta)), I(\mathbf{X})) + \beta R(\varphi(\mathbf{X}; \theta))). \quad (1.1)$$

in which T and I denote the target (moving) and initial (source) images, $\varphi(\mathbf{X}; \theta)$ is the corresponding map with parameters θ that aligns the target and initial images, R is a regularizing term, and β is a relatively small positive hyper-parameter. It is noted that the terminology 'predictor'

is employed to emphasize that, due to the presence of the regularizer term in (1.1), this step constructs a map that is ideally not far from a physical solution.

Subsequently, in the corrector part, the strain energy is minimized while the similarity obtained earlier remains *the same* by applying the distance function. As shall be discussed, the resulting solution corresponds to an equivalent Dirichlet boundary value problem where the displacement boundary conditions are dictated based on the predictor map.

Moreover, to guarantee the invertibility of the (predictor) corrector map, we employ neural ordinary differential equations (NODEs) and define the transformation in terms of its (pseudo) velocity, making the method conceptually in line with diffeomorphism algorithms (Miller et al., 2002). We note that the registration problem is ill-posed in its nature in the sense that the map aligning the target and initial images is not unique. Hence, the resulting solution of the present predictor-corrector algorithm represents one of the infinitely many admissible transformations for the registration problem. However, in contrast to previous methods where a multi-objective optimization was only adapted, the main features of the present method are: 1) the method is an integral variant of PINNs, and consequently, it avoids calculation of second spatial derivatives, 2) the predictor part minimizes the mismatch loss in a better fashion because it does not need to solve the physics problem concurrently, 3) the corrector part solves a physics problem without losing the registration, allowing for accurate modeling of complex biophysics e.g. growth and remodeling.

As shown in Pawar et al. (2022), the solution of a multi-objective function converges to the one that balances the mismatch and the regularizer losses, which is a function of the hyper-parameter β . This interplay means that neither problem is solved adequately. Hence, here in the predictor part, we decrease the hyper-parameter β as much as possible so that we can obtain a diffeomorphic map that minimizes the mismatch loss in a more accurate manner. The resulting transformation from the corrector systematically satisfies the equilibrium equations of continuum mechanics, including the growth equations, and thus may represent underlying physics in a better fashion. Furthermore, many biomedical applications require a multi-physics coupling, e.g., growth and remodeling (Ambrosi et al., 2011). We also introduce such a coupling in this manuscript to capture embryo development and brain atrophy. Fig 1 shows the main steps of the new method.

The remainder of the paper is organized as follows: firstly, we enumerate the main equations of solid mechanics. Then, the concept of NODE combined with the forward Euler method is briefly reviewed, and the underlying reason to consider this architecture is explained. Next, the theory of the predictor-corrector algorithm is explained, where the incorporation of the growth problem is omitted for clarity.

Specifically, it is shown how the new solution procedure directly embodies the energy (variational) method, which is the foundation of any FEM, and accordingly, the new method represents a physically meaningful deformation defined as an invertible map between source and target images that systematically fulfills an equilibrium equation. Subsequently, the theory that incorporates growth is developed, where a first-order evolutionary ODE is introduced to model the addition of mass. The numerical results for two- and three-dimensional models with rather simple geometry are defined to highlight features of the method and registration problem. Subsequently, three biomedical practical problems with the presence of growth are investigated: 1) growth in the embryo of zebrafish tail, 2) shrinkage of the brain with aging, and 3) fetal brain development. Finally, an ablation study is performed to better understand capability of the method.

2 Background

2.1 Fundamental Equations

Let us assume the initial and target images define domains containing, respectively, undeformed and deformed configurations of an elastic continuum, denoted by \mathcal{B}^I and \mathcal{B}^T . We also show the whole domain of integral with Ω . Then the equilibrium equations read as:

$$\nabla_{\mathbf{X}} \cdot \mathbf{P} + \rho_0 \mathbf{b} = 0, \quad \mathbf{P} \mathbf{F}^T = \mathbf{F} \mathbf{P}^T \quad \text{on } \mathcal{B}^I, \quad (2.1)$$

in which, $\nabla_{\mathbf{X}} \cdot$ is the divergence operator with respect to material coordinates and \mathbf{P} , $\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}}$, and ρ_0 are the first Piola Kirchhoff stress, deformation gradient, and reference density, respectively. Moreover, by assuming hyper-elasticity, we can write:

$$\mathbf{S} = 2\rho_0 \frac{\partial \psi(\mathbf{C})}{\partial \mathbf{C}} \quad \text{on } \mathcal{B}^I, \quad \mathbf{S} = \mathbf{F}^{-T} \mathbf{P}, \quad \mathbf{C} = \mathbf{F}^T \mathbf{F}, \quad (2.2)$$

with ψ , \mathbf{C} , and \mathbf{S} are a convex potential function, the Cauchy strain tensor, and the second Piola Kirchhoff stress, respectively. The above equilibrium and constitutive equations define a mixed boundary value problem by considering the following prescribed boundary conditions:

$$\begin{aligned} \mathbf{u} &= \mathbf{G}(\mathbf{X}) && \text{on } \partial \mathcal{B}_u^I, \\ \mathbf{P}\mathbf{N} &= \mathbf{H}(\mathbf{X}) && \text{on } \partial \mathcal{B}_t^I, \end{aligned} \quad (2.3)$$

where $\partial \mathcal{B}_u^I$ and $\partial \mathcal{B}_t^I$ are disjoint and complementary sets of boundary points with prescribed displacement and traction boundary conditions \mathbf{G} and \mathbf{H} , and \mathbf{N} is the unit normal vector on the boundary of $\partial \mathcal{B}^I$.

A systematic approach to solve the above mixed-boundary value problem is to satisfy the localized (partial differential) equations weakly through an integral sense. Indeed, the solution of the problem is the stationary point of the following integral form:

$$\begin{aligned} \arg \min_{\mathbf{u}} \Pi(\mathbf{u}) &= \int_{\mathcal{B}^I} \psi(\mathbf{C}) d\Omega - \int_{\mathcal{B}^I} \rho_0 \mathbf{b} \cdot \mathbf{u} d\Omega - \int_{\partial \mathcal{B}_t^I} \mathbf{P} \mathbf{N} \cdot \mathbf{u} d\Gamma, \\ \mathbf{u} &= \mathbf{x} - \mathbf{X}, \\ \mathbf{u} &= \mathbf{G}(\mathbf{X}) \quad \text{on } \partial \mathcal{B}_u^I. \end{aligned} \quad (2.4)$$

Now, in the case of the purely Dirichlet boundary condition and absence of body force, the above equation reduces to:

$$\arg \min_{\mathbf{u}} \Pi(\mathbf{u}) = \int_{\mathcal{B}^I} \psi(\mathbf{C}) d\Omega, \quad (2.5)$$

which is central to our registration method. In the FEM, the satisfaction of the above equations is achieved uniformly in the domain by discretization and approximately vanishing the first variation of the above functional, while here, we define (2.5) as the loss function and directly process the minimization of the total potential energy, which is equivalent.

2.2 NODE Architecture

In continuum mechanics, admissible solutions are smooth and invertible. The invertibility of the deformation field leads to $J = \det(\mathbf{F}) > 0$. In the classical FEM, the displacement of a point is interpolated based on the displacement of surrounding nodal points. Hence, the above-mentioned non-negativity condition on Jacobian is identically fulfilled as long as the interpolation is regular. In contrast, in neural network architectures, a stack of nonlinear and linear transformations results in a map that is not necessarily injective.

One significant property of invertible maps is that they are closed under composition. As a result, it is possible to build complex invertible transformations by composing invertible building blocks, establishing a new family of invertible maps that are called flow methods (Papamakarios et al., 2021), which found applications in generative models. NICE (Dinh et al., 2014), and real NVP (Dinh et al., 2016) are notable examples of discrete flow algorithms that have a predefined structure in each building block. The invertible residual networks (Behrmann et al., 2019; Chen et al., 2019) is another example that is defined as

$$\mathbf{x} = \mathbf{X} + NN(\mathbf{X}; \mathbf{W}, \mathbf{b}), \quad (2.6)$$

where the function is a neural network with weights \mathbf{W} and biases \mathbf{b} . As mentioned in Behrmann et al. (2019), Chen et

al. (2019), a sufficient condition for (2.6) to be invertible is when NN is contractive in the sense of:

$$d_x(NN(\mathbf{X}_2), NN(\mathbf{X}_1)) \leq K d_X(\mathbf{X}_2, \mathbf{X}_1), \quad K < 1, \quad (2.7)$$

where K is a constant independent of X_1 and X_2 , and d_x and d_X are metrics defined in space \mathbf{x} and \mathbf{X} , respectively. Hence, one can define a stack of residual networks with a parameter Δt :

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \Delta t \, NN(\mathbf{x}^{(i)}) \quad (2.8)$$

in which Δt is small enough such that $\Delta t \, NN(\mathbf{x}^{(i)})$, $i = 1, \dots, n$ is contractive.

Apparently, the above formulation is the forward Euler time integration of the following ODE:

$$\frac{d\mathbf{x}}{dt} = NN(\mathbf{x}; \mathbf{w}, \mathbf{b}), \quad (2.9)$$

which is equivalent to the concept of Neural ODE (Chen et al., 2018). Additionally, (2.7) is analogous to diffeomorphism algorithms (Christensen et al., 1994, 1996). Hence, in the present method, to obtain invertible transformation, we employ (2.8) and define Δt such that we guarantee (2.7).

3 Method

3.1 Predictor

In the predictor part, the main focus is to attain a map that ideally aligns two images. To this end, we follow the classical formulation of registration, where the transformation is obtained from the minimization of a loss function consisting of similarity and regularization terms. Given two (binary) images discretized in pixels denoted by T and I , following (Pawar et al., 2022), we define two functions $S_1(\mathbf{X})$ and $S_2(\mathbf{x})$ as follows:

$$\begin{aligned} S_1(\mathbf{X}, \mathcal{B}^I) &= \sum_i c_i \xi_i(\mathbf{X}), \quad c_i = \begin{cases} 1, & \mathbf{X}_i \in \mathcal{B}^I \\ 0, & \text{otherwise} \end{cases}, \\ S_2(\mathbf{x}, \mathcal{B}^T) &= \sum_i d_i \xi_i(\mathbf{x}), \quad d_i = \begin{cases} 1, & \mathbf{x}_i \in \mathcal{B}^T \\ 0, & \text{otherwise} \end{cases}, \end{aligned} \quad (3.1)$$

where ξ_i are (linear Lagrange) interpolating functions, and \mathbf{x}_i and \mathbf{X}_i are nodal coordinates corresponding to d_i and c_i ,

respectively. Then, the predictor map is obtained from the minimization of the following loss function:

$$\begin{aligned} \arg \min_{\mathbf{w}_p, \mathbf{b}_p} & \left\{ \int_{\Omega} \left(S_1(\mathbf{X}, \mathcal{B}^I) - S_2 \left(\mathbf{x}_p^{(n_p)} (\mathbf{w}_p, \mathbf{b}_p), \mathcal{B}^T \right) \right)^2 d\Omega \right. \\ & \left. + \beta \int_{\Omega} R \left(S_1(\mathbf{X}, \mathcal{B}^I) \left(\mathbf{x}_p^{(n_p)} (\mathbf{w}_p, \mathbf{b}_p) - \mathbf{X} \right) \right) d\Omega \right\}, \end{aligned} \quad (3.2)$$

where

$$\begin{aligned} \mathbf{x}_p^{(i+1)} &= \mathbf{x}_p^{(i)} + \Delta\tau \mathbf{NN}(\mathbf{x}_p^{(i)}; \mathbf{w}_p, \mathbf{b}_p), \\ i &= 0, \dots, n_p - 1, \quad \Delta\tau = \frac{1}{n_p}, \\ \mathbf{x}_p^{(0)} &= \mathbf{X}, \end{aligned} \quad (3.3)$$

in which \mathbf{w}_p and \mathbf{b}_p denote parameters of the neural network for predictor transformation, which are determined by minimization of the functional (3.2), R is a regularization term that typically depends on \mathbf{F} , and n_p is defined to guarantee invertibility of the resulting transformation. In this study, for synthetic and biological examples, we utilize the strain energy density function as the regularization term, i.e., $R = \psi(\mathbf{F})$. However, in the ablation study part, we show that the method is applicable when the regularization term is different than the strain energy potential. This fact implies that the method can successfully adapt admissible geometrical registration data obtained from other methods and attach physics by converting the predictor deformation into a distance function and applying the corrector part with this constraint. Different strain energy models could be used, but we opt for a neo-Hookean model or the St Venant Kirchhoff model. Additionally, in (3.2), β is a hyperparameter that may vary from problem to problem. Nevertheless, in the new method, a small value for β is considered such that the predictor deformation field produces an admissible transformation in the sense of non-vanishing Jacobians while a satisfactory registration can be achieved. This is in contrast to classical methods, where the registration is entirely carried out by (3.2), and hence, the choice of the hyperparameter is crucial to find a good map. For us, as long as the predictor map is injective, regularity in the sense of equilibrium of large deformation elasticity is systematically validated and enforced in the corrector step.

3.2 Corrector

By employing (3.2), one can *optimally* obtain a smooth transformation that maps the boundary points of the initial image into the boundary points of the target image. However, the aforementioned transformation does not necessarily satisfy the equilibrium equations for interior points, even if the reg-

ularizing term is selected as the strain potential energy. Thus, the physics of the problem needs to be validated and enforced as an additional step. In doing so, we present an approach that, similar to FEMs, seeks to obtain a stationary point of total potential energy while maintaining the alignment of the predictor.

In doing so, having obtained the *predictor* solution from (3.2), we define the following Dirichlet boundary value problem:

$$\begin{aligned} \nabla_{\mathbf{X}} \cdot \mathbf{P} &= \mathbf{0}, \\ \mathbf{u}(\mathbf{X}) &= \mathbf{x}_p^{(n_p)} - \mathbf{X} \quad \text{on } \mathcal{B}^I \end{aligned} \quad (3.4)$$

Then, to satisfy (3.4)₁, we alternatively minimize the potential energy (2.5):

$$\arg \min_{\mathbf{w}_c, \mathbf{b}_c} \Pi = \int_{\Omega} \psi \left(S_1(\mathbf{X}, \mathcal{B}^I) \left(\mathbf{x}_c^{(n_c)} (\mathbf{w}_c, \mathbf{b}_c) - \mathbf{X} \right) \right) d\Omega, \quad (3.5)$$

where we consider the following form for the correction:

$$\begin{aligned} \mathbf{x}_c^{(i+1)} &= \mathbf{x}_c^{(i)} + \Delta t D_p(\mathbf{x}_c^{(i)}; \mathbf{x}_{\mathcal{B}^I}^{p(n_p)}) \mathbf{NN}(\mathbf{x}_c^{(i)}; \mathbf{w}_c, \mathbf{b}_c), \\ i &= 0, \dots, n_c - 1, \quad \Delta t = \frac{1}{n_c}, \\ \mathbf{x}_c^{(0)} &= \mathbf{x}_c^{(n_p)}. \end{aligned} \quad (3.6)$$

D_p , in (3.6), can be any function that fixes the boundary condition imposed by the predictor solution while the interior points are transformed to minimize (3.2). A suitable choice, in this regard, is the distance function:

$$D_p(\mathbf{x}; \mathbf{x}_{\mathcal{B}^I}^{p(n)}) = \min(\|\mathbf{x} - \mathbf{x}_{\mathcal{B}^I}^{p(n)}\|) \quad (3.7)$$

in which $\|\cdot\|$ indicates the Euclidean norm, and $\mathbf{x}_{\mathcal{B}^I}^{p(n)}$ denotes the set of transformed boundary points by predictor map. While there are several methods available to approximate the above distance functions (see, for example, Berg and Nystöm (2018) and Sukumar and Srivastava (2022)), we shall directly utilize (3.7) since in the numerical part we deal with complex 3d geometry such as brains. Thus, in the numerical implementation, we approximate the distance function based on a finite representation of $\mathbf{x}_{\mathcal{B}^I}^{p(n)}$. Note that there is no need to use the same potential for the predictor and corrector steps. For the corrector step, one could adopt more complex material behavior specific to a particular soft tissues, e.g. brain (Mihai et al., 2015), skin Limbert (2014), or myocardium Holzapfel and Ogden (2009), to name a few examples. Multi-physics coupling can also be introduced. For instance, in the following section, we present the formulation for finite volume growth biophysics as part of the corrector (Ambrosi et al., 2011).

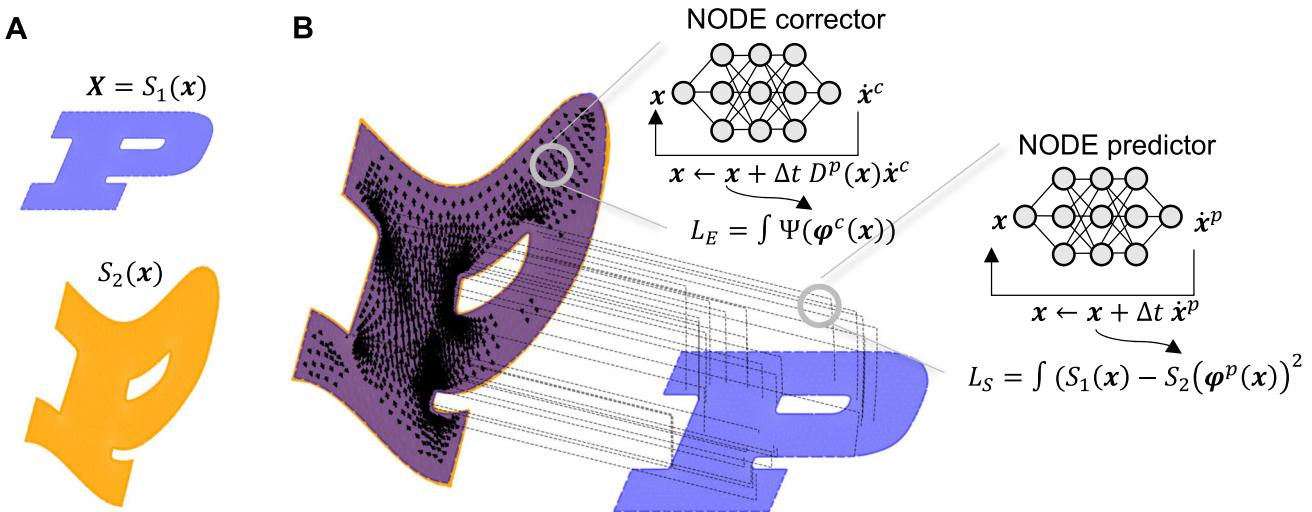


Fig. 1 Method overview. A) Source and target images. B) The method consists of two steps. In a predictor step, a neural ordinary differential equation (NODE) is used to map the source to the target by minimizing a mismatch loss with a regularization term. The NODE guarantees an invertible map provided a small enough time step is used. In a corrector step, the map found by the predictor is enforced through distance functions, and the loss is that of a physics problem, e.g., minimization of total elastic energy

3.3 Finite Growth Biophysics

All living biological tissues show behavior beyond large deformation elasticity, including coupled electrophysiology (Göktepe & Kuhl, 2009), reaction-diffusion (Tepole, 2017), and ability to grow and remodel (Ambrosi et al., 2011). Here, we tackle that last one by introducing the multiplicative split of the deformation gradient

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^g, \quad (3.8)$$

where \mathbf{F}^e and \mathbf{F}^g indicate elastic and inelastic or growth deformation, respectively. The first consequence of such a split is that the energy in the corrector is assumed to depend only on the elastic deformation $\psi(\mathbf{F}^e)$. In FEM implementations of growth, or if the strong form was desired, one should be careful with the computation of the stress with the appropriate pull-back or push-forward operations from the intermediate state (Himpel et al., 2005). We do not encounter such a problem here because the energy is minimized directly. In addition to $\psi(\mathbf{F}^e)$, the second change to the corrector is the introduction of a rate equation with an initial condition for the growth tensor \mathbf{F}^g .

Assuming isotropic homogeneous growth:

$$\mathbf{F}^g = \theta^g \mathbf{I}, \quad (3.9)$$

where the scalar θ^g captures the addition of volume at constant reference density ρ_0 (Himpel et al., 2005). Other options for growth tensor are to model fiber or area growth (Eskandari & Kuhl, 2015). In many cases, the kinetic equation for

growth update depends on the elastic deformation $\dot{\theta}^g(\mathbf{F}^e)$ (Eskandari & Kuhl, 2015). However, particularly for brain development or atrophy shown here, the growth rate can be assumed to be morphogenetic or dependent on the concentration of certain chemicals that drive growth or atrophy, and the dependence on \mathbf{F}^e can be ignored (Weickenmeier et al., 2018; Wang et al., 2021). Assuming exponential growth independent of deformation and initial condition $\theta^g(0) = 1$, one obtains

$$\theta^g(t) = \exp(\kappa t). \quad (3.10)$$

where κ is the growth rate parameter. Hence, for the predictor part in the biological growth examples, we solve the same problem as before (3.2), while in the corrector, we solve the following problem:

$$\arg \min_{\mathbf{w}_c, \mathbf{b}_c, \kappa} \Pi = \int_{\Omega} \psi \left(S_1 \left(\mathbf{X}, \mathcal{B}^I \right) (\mathbf{F}^e), \kappa \right) d\Omega \quad (3.11)$$

where it can be seen that we optimize for the deformation as well as for the growth κ such that elastic energy is minimized.

3.4 Sequential Approach Coupled With Growth

In this part, we explain the method when a stack of images as a function of time is considered. In this regard, we assume that changes through time are slow enough to neglect the inertial force in our analysis. Considering the algorithm mentioned earlier for single registration, there are several

possibilities for carrying out the registration for sequential image data. In this paper, given an initial image and sequence of images, which are indicated by I_0 and I_j , $j = 1, \dots, n_s - 1$, with the pertinent predictor shown by $\mathbf{x}_{p(j)}^{(n_p)}$, $j = 1, \dots, n_s - 1$, $\mathbf{x}_{p(0)}^{(n_p)} = \mathbf{X}$, we define the forward composition function:

$$\mathbf{x}_{p(0 \rightarrow j)}^{(n_p)} \equiv \begin{cases} \mathbf{X}, & j = 0 \\ \mathbf{x}_{p(j)}^{(n_p)} \circ \mathbf{x}_{p(j-1)}^{(n_p)} \circ \dots \circ \mathbf{x}_{p(1)}^{(n_p)}(\mathbf{X}), & \text{otherwise} \end{cases} \quad (3.12)$$

Analogously, we define the backward composition function, $\mathbf{X}_{p(k \rightarrow 0)}^{(n_p)}$, as follows

$$\mathbf{X}_{p(j \rightarrow 0)}^{(n_p)} \equiv \begin{cases} \mathbf{X}, & j = 0 \\ \mathbf{x}_{p(1)}^{(n_p)} \circ \mathbf{x}_{p(2)}^{(n_p)} \circ \dots \circ \mathbf{x}_{p(j)}^{(n_p)}(\mathbf{X}), & \text{otherwise} \end{cases} \quad (3.13)$$

in which $\mathbf{X}_{p(k)}^{(n_p)}(\mathbf{X})$ denotes the inverse map of the k^{th} step, which can be approximated from the following recursive equation by setting $i = n_p - 1$:

$$\begin{aligned} \tilde{\mathbf{X}}_p^{(0)} &= \mathbf{X} \\ \tilde{\mathbf{X}}_p^{(i+1)} &= \tilde{\mathbf{X}}_p^{(i)} - \Delta\tau NN(\tilde{\mathbf{X}}_p^{(i)}; \mathbf{w}_{p(k)}, \mathbf{b}_{p(k)}); \\ i &= 0, \dots, n_p - 1; \quad \Delta\tau = \frac{1}{n_p}. \end{aligned} \quad (3.14)$$

Then, the predictor part may be constructed in the sequential format of (3.2):

$$\begin{aligned} &\arg \min_{\mathbf{w}_{p(j)}, \mathbf{b}_{p(j)}} \\ &= \left\{ \int_{\Omega} \left(S_1 \left(\mathbf{X}_{p(j-1 \rightarrow 0)}^{(n_p)}, \mathcal{B}^{I_0} \right) - S_2 \left(\mathbf{x}_{p(j)}^{(n_p)}, \mathcal{B}^{I_j} \right) \right)^2 d\Omega \right. \\ &\quad \left. + \beta \int_{\Omega} R(S_1 \left(\mathbf{X}, \mathcal{B}^{I_0} \right) \left(\mathbf{x}_{p(0 \rightarrow j)}^{(n_p)} - \mathbf{X} \right)) d\Omega \right\}, \\ &\mathbf{x}_{p(0)} = \mathbf{X}, \quad j = 1, \dots, n_s - 1, \end{aligned} \quad (3.15)$$

The first term of the above integral can be simplified, provided that an acceptable registration is performed. That is, assuming a good registration, one can use the following approximation in (3.15):

$$S_1 \left(\mathbf{X}_{p(j-1 \rightarrow 0)}^{(n_p)}, \mathcal{B}^{I_0} \right) \approx S_1 \left(\mathbf{X}, \mathcal{B}^{I_{j-1}} \right) \quad j = 1, \dots, n_s - 1, \quad (3.16)$$

leading to the following minimization:

$$\begin{aligned} &\arg \min_{\mathbf{w}_{p(j)}, \mathbf{b}_{p(j)}} \\ &= \left\{ \int_{\Omega} \left(S_1 \left(\mathbf{X}, \mathcal{B}^{I_{j-1}} \right) - S_2 \left(\mathbf{x}_{p(j)}^{(n_p)}, \mathcal{B}^{I_j} \right) \right)^2 d\Omega \right. \\ &\quad \left. + \beta \int_{\Omega} R(S_1 \left(\mathbf{X}, \mathcal{B}^{I_0} \right) \left(\mathbf{x}_{p(0 \rightarrow j)}^{(n_p)} - \mathbf{X} \right)) d\Omega \right\}, \\ &\mathbf{x}_{p(0)} = \mathbf{X}, \quad j = 1, \dots, n_s - 1. \end{aligned} \quad (3.17)$$

In this paper, we used (3.17) to obtain the predictors. Subsequently, for the corrector part, we define:

$$\mathbf{x}_{pc(k)}^{(n_p, n_c)} \equiv \mathbf{x}_{c(k)}^{(n_c)} \circ \mathbf{x}_{p(k)}^{(n_p)}, \quad k = 1, \dots, n_s - 1. \quad (3.18)$$

Next, we modify (3.12) as:

$$\mathbf{x}_{pc(0 \rightarrow j)}^{(n_p, n_c)} \equiv \mathbf{x}_{pc(j)}^{(n_p, n_c)} \circ \mathbf{x}_{pc(j-1)}^{(n_p, n_c)} \circ \dots \circ \mathbf{x}_{pc(1)}^{(n_p, n_c)}, \quad j = 1, \dots, n_s - 1. \quad (3.19)$$

The corrector solution for each pair of images is then obtained consecutively by using the following:

$$\arg \min_{\mathbf{w}_{c(j)}, \mathbf{b}_{c(j)}, \kappa(j)} \Pi = \int_{\Omega} \psi \left(S_1 \left(\mathbf{X}, \mathcal{B}^{I_0} \right) \left(\mathbf{x}_{pc(0 \rightarrow j)}^{(n_p, n_c)} - \mathbf{X} \right) \right) d\Omega, \quad (3.20)$$

such that the minimization of the energy in the corrector steps always keeps track of the deformation with respect to the initial configuration even though the predictor and corrector steps are done sequentially.

4 Numerical Results

To show the potential of the method, we provided several examples, including synthetic benchmarks, followed by three practical applications. In synthetic and biological examples, for constitutive equations, we consider either Saint Venant-Kirchhoff material with the following strain energy potential:

$$\psi(\mathbf{C}) = \frac{\lambda}{2} (\text{tr}\mathbf{E}) + \mu \text{tr}(\mathbf{E}^2), \quad \mathbf{E} = \left(\frac{\mathbf{C} - \mathbf{I}}{2} \right), \quad (4.1)$$

or neo-Hookean material:

$$\psi(\mathbf{C}) = \frac{\mu}{2} (\text{tr}\mathbf{C} - 3) - \mu \log(J) + \frac{\lambda}{2} (\log(J))^2, \quad (4.2)$$

where for simplicity, we assume $\mu = \lambda = 1 \text{ Pa}$ in all examples. For the ablation study, we consider the following strain

potential function:

$$\begin{aligned}\psi(\mathbf{C}) &= 3(J-1)^2 + (I_1 - 3) + (I_2 - 3)/12 + (I_1 - 3)^2 \\ I_1 &= \text{tr}\mathbf{C}, \quad I_2 = \frac{1}{2} \left((\text{tr}\mathbf{C})^2 - \text{tr}\mathbf{C}^2 \right), \quad J = \sqrt{\det \mathbf{C}}.\end{aligned}\quad (4.3)$$

During the learning process, we employed the Adam optimizer with a learning rate of 0.00005. The choice of β varied depending on the problem, with specific values to be mentioned accordingly. Additionally, the architecture of the neural network used in the NODE consists of three hidden layers, each containing 40 neurons for 2d and 60 neurons for 3d problems, which is in the range taken for similar problems in the literature (see the neural network architectures taken in Sukumar and Srivastava (2022), Bai et al. (2023)). Furthermore, we take 15 steps in the NODE, i.e., $\Delta t = \Delta\tau = 1/15$ for both predictor and corrector. We evaluated the integrals by discretizing the domain into 200×200 elements for 2d problems and $45 \times 45 \times 45$ elements for 3d problems. In each direction, we considered two integration points for every element, i.e., four points for 2d elements and eight points for 3d elements. This process involved using a mini-batch of 4000 uniformly randomly distributed elements in the predictor part for each epoch. In the corrector part, we utilized mini-batches of 2000 and 1500 uniformly randomly distributed elements in each epoch for 2d and 3d examples, respectively.

4.1 Synthetic Examples

For synthetic examples, we explore non-uniform deformations, including extension, shearing, and bending modes of deformation, applied to a multiple-connected plate. Furthermore, we compare the results obtained when using analytical expressions for S_1 and S_2 against discretization in pixels. We further investigate the impact of using binary data versus RGB image data with 256-bit information per channel, thereby providing more detailed information than binary data.

For the first example, we analyze the extension of a Saint Venant-Kirchhoff plate with a hole. The displacement boundary conditions are as follows:

$$\begin{aligned}u_1(-1, X_2) &= -1, \quad u_1(1, X_2) = 1, \\ u_2(X_1, -1) &= u_2(X_1, 1) = 0, \\ u_1(X_1, X_2) &= u_2(X_1, X_2) = 0, \quad X_1^2 + X_2^2 = 0.25\end{aligned}\quad (4.4)$$

and traction-free boundary conditions are assumed for the Neumann boundary conditions. The FEM results are used as a reference solution. For the *PINN inexact* approach, the hyper-parameter for displacement boundary conditions has been set to one, while 0.01 was selected as the hyper-parameter pertinent to strain energy. Additionally, in the predictor part of the new approach, $\beta = 1/6000$ was considered. Results perti-

nant to the energy and Jacobian have been plotted in Fig 2-C labeled with *FE*. Fig 2-C, *PINN exact* column, are the solutions obtained from a PINN-like by minimizing the strain energy with the imposition of exact boundary conditions. In contrast to classical PINNs (Raissi et al., 2019), or variational PINNs (Kharazmi et al., 2019), we minimize the energy and not the residual of the weak form or the strong form. As can be seen, there is a good agreement between the energy form of the PINN with the exact imposition of boundary conditions and the FEM solution. The results of the integral energy form of the PINN with enforcing boundary conditions as a penalty term have been depicted in Fig 2-C, labeled with *PINN inexact*. In comparison to the FEM results, there are some discrepancies in the results of PINN with penalty terms. In particular, the displacement boundary conditions have not been exactly satisfied over the boundary of the hole, while the resulting energy and Jacobian fields have some disagreement with those corresponding to FEM. This discrepancy can be attributed to the fact that PINN when boundary conditions are enforced with penalty terms, is a multi-objective optimization, and thus, the resulting solution generally satisfies both objects to some extent.

We next analyze the registration problem, where we do not have complete information regarding the boundary conditions. We employ an analytical format for S_1 and S_2 (see Fig. 2-B). The results for the predictor part have been reported in Fig. 2-C, labeled with *Predictor*. As can be seen, for the registration case, incomplete information regarding the boundary condition and small value of the regularization β leads to discrepancies in the energy and Jacobian fields with respect to FEM, but the predictor achieves accurate registration with a smooth invertible map. Subsequently, the results after applying the corrector part have been illustrated in Fig. 2-C labeled with *Corrector*. The difference between the results of the corrector and predictor is shown. The path of minimization of the total potential energy in the corrector part has also been shown in Fig. 2-D. It is worth noting that, as the name suggests, the corrector part is designed to guarantee the minimization of the potential energy corresponding to the Dirichlet boundary value problem specified by the predictor. In addition, we quantified how the corrector maintains the quality of registration by calculating the Dice index across the entire domain. In segmentation analysis, the quality of registration is typically measured using the Dice index, which is defined as follows:

$$\text{Dice}(S_2^k, S_1^k \circ \varphi^{-1}) = 2 \frac{|S_2^k \cap (S_1^k \circ \varphi^{-1})|}{|S_2^k| + |S_1^k \circ \varphi^{-1}|}, \quad (4.5)$$

where superscript k denotes the k^{th} region of interest and $S_1^k \circ \varphi^{-1}$ stands for the composition of S_1 with the inverse of transformation. This function basically constructs the predicted values of S_2 for direct comparison to the ground truth

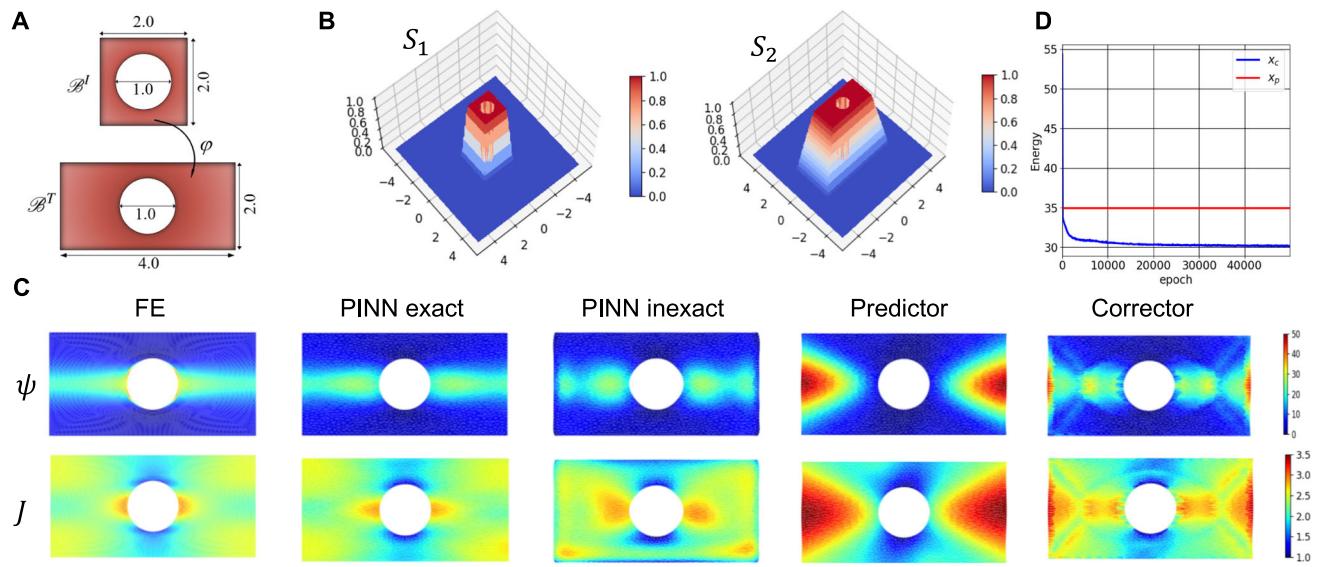


Fig. 2 Comparison of registration, physics-informed neural networks, and finite element analysis: A) extension of a plate with the whole; undeformed and deformed configurations. B) 3d views of analytic S_1 and S_2 functions. C) strain energy density and Jacobian fields obtained from

different methods, including finite element method, integral form of the PINN when the Dirichlet boundary conditions are imposed exactly, integral form of PINN when the boundary conditions are imposed as penalty terms, predictor part of the registration, and corrector solution

S_2 by mapping the S_1 intensities to the S_2 coordinate grid using the NODE map. In this example, the Dice index for predictor and corrector parts are 0.95 and 0.948, indicating good maintenance of the registration in the corrector part.

As explained earlier, the existence of the corrector part in the method is essential. The multi-objective format of the predictor does not necessarily lead to a physically meaningful solution. Especially if a relatively low regularization β is required to achieve a good alignment between initial and target images, particularly in large deformations, the transformation can be far from equilibrium. Another fact that should be noticed is that even if the predictor produces an excellent alignment, it does not necessarily lead to the same boundary value problem with similar Dirichlet boundary conditions defined in FEM. In fact, there are infinitely many predictor solutions that minimize the misalignment but lead to a different Dirichlet problem. This is not a limitation of the method, though, but rather a difference between the nature of registration and classical well-defined boundary value problems in physics where complete boundary information is available. Thus, the corrector solution in Fig. 2-C solves a Dirichlet problem, just not the same problem as the one solved with FEM or PINNs.

As the next example, we consider the shear deformation of the neo-Hookean plate with the hole while the hole is under contraction. In this regard, the following displacement boundary conditions have been imposed:

$$\begin{aligned} u_1(-1, X_2) &= u_1(1, X_2) = 0.25(X_2 + 1), \\ u_2(-1, X_2) &= u_2(1, X_2) = 0, \\ u_2(X_1, -1) &= u_2(X_1, 1) = 0, \\ u_r(X_1, X_2) &= 0.2, \quad u_\theta(X_1, X_2) = 0, \quad X_1^2 + X_2^2 = 0.25, \end{aligned} \quad (4.6)$$

and the traction boundary conditions are vanishing.

We utilize this example to compare the performance of the general neural network with the NODE in generating an admissible deformation field in continuum mechanics. For the NODE, we utilized (3.2) with $\beta = 1/6000$. We do not consider the corrector part in this example.

It is worth mentioning that although there are popular and fast methods in the literature, the resulting transformation is not necessarily invertible (e.g., Balakrishnan et al. (2019)), which is in contradistinction with the admissible deformation field in classical continuum mechanics. Similar to the previous example, we use an analytical format of S_1 and S_2 (see Fig. 3-B). Fig. 3-C shows the resulting transformation for neural network and NODE architectures for the entire domain of integrals and not only \mathcal{B}^T . The corresponding loss functions during training for both regular neural networks and the NODE have also been reported in Fig. 3-D. In Fig. 3-C, we indicated points with negative Jacobian with red color. While the transformation map for the NODE structure generates positive Jacobian for the entire domain, there are points inside of the hole with negative Jacobian in the map attained from the regular neural network. This example confirms that in contrast to regular neural networks that are not necessarily invertible, the NODE is a good candidate

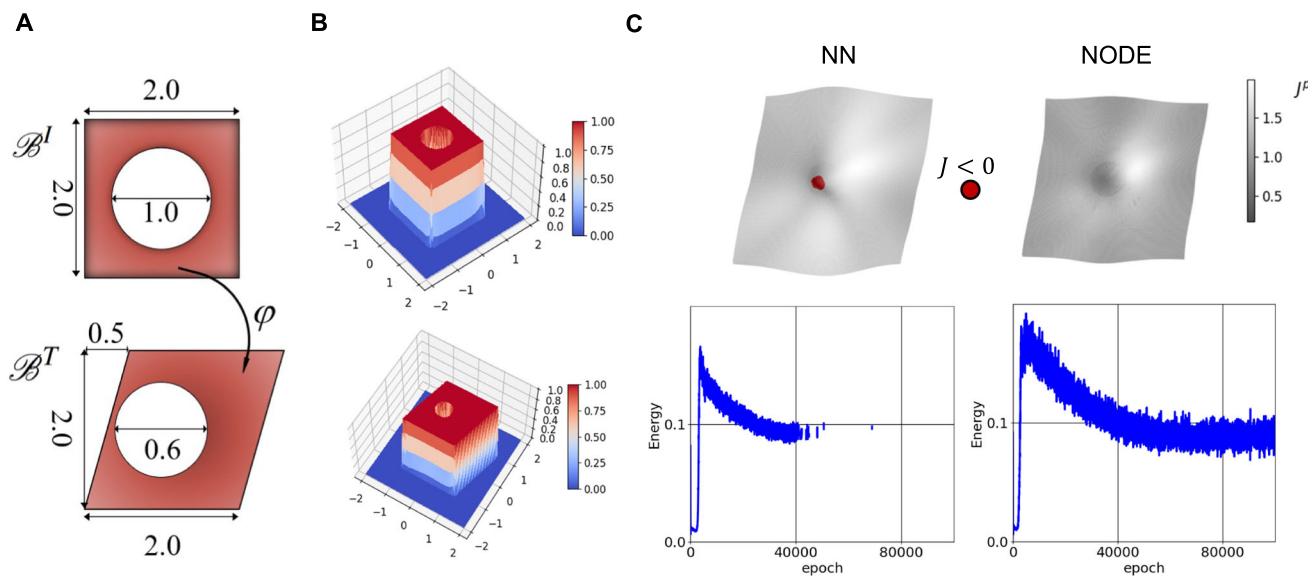


Fig. 3 The performance of the resulting deformation field was evaluated when the displacement field is approximated using neural networks, compared to when the (pseudo) velocity is constructed based on neural networks (Neural ODE registration) A) extensive shear deformation of the neo-Hookean plate with the whole; undeformed and deformed configurations. B) 3d views of analytic S_1 and S_2 functions. C) Deforma-

tion field when displacement is expanded in terms of the neural network along with the corresponding similarity loss function. Red Dots indicate points with negative Jacobian. D) the same outputs when NODE is employed, showing the resulting transformation is physically meaningful ($J > 0$ everywhere) (Color figure online)

for invertible transformations, although it is computationally more expensive in the back-propagation step as well as in the calculation of the deformation gradient. Similarly to the previous example, we computed the Dice score for the entire domain, which is 0.986 for the NODE method.

Using the next synthetic example, we intend to illustrate how the resolution of image data might affect the registration problem. In this regard, we analyzed the bending of the plate containing three ellipsoidal holes. To consider the effect of the resolution, we solved the problem for two sets of data: images with high resolution, which have been defined by using a fine mesh discretization, and low-resolution images that are 90×90 pixels. Neo-Hookean material has been selected for the constitutive equations. In this example, we performed the registration by selecting $\beta = 1/6000$, with a total of 75000 epochs for the predictor. The resulting strain energy (i.e., ψ) for the predictor solutions has been shown in Figs. 4-E & 4-F, respectively, for high- and low-resolution data. From these figures, it is clear that the resolution may affect the final results. In particular, the computed Dice scores for the entire domain are 0.972 and 0.981 for the low and high resolutions, respectively. Additionally, the figure shows that a lower elastic energy level was achieved in the solution corresponding to the high-resolution data.

The last synthetic example is employed to understand how more information in image data can change the results of a registration problem. In this regard, we solved the extension of a neo-Hookean plate with three ellipsoidal holes. We ana-

lyzed the problem using two image datasets, each with a size of 512×512 pixels: 1) Binary images containing information about the boundary points (Fig. 5-A), and 2) RGB images having information about the interior points in addition to boundary points (Fig. 5-C). To create the RGB images, we plotted $D(\mathbf{X})$. This format provides us with more information about the transformation of interior points. In addition, to measure the accuracy of the registration, we created undeformed and deformed versions of a synthetic labeled image as shown in Fig. 5-I. Furthermore, we compared the results of the registration in terms of the Dice index with AntsPy Elastic (xxx,yyy), which is a composition transformation of affine and deformable maps.

Fig. 5-A and 5-C illustrate the results for binary and RGB registrations, respectively. From these figures, one can notice that in terms of the outer boundary, both image data sets lead to almost the same accuracy, better than AntsPy Elastic (Fig. 5-K). Moreover, Figs. 5-B and 5-D indicate similar minimization paths for both datasets in both the predictor and corrector parts. However, as can be seen in Fig 5-I, the map resulting from the RGB data set has obtained better dice indices compared to those of binary and AntsPy Elastic. This result implies that when the image pair has more information, a better registration can be expected. Also, we have shown energy and Jacobian fields relevant to binary and RGB data in the corrector solutions in Figs 5-E to 5-H, which are almost the same for both data sets. Hence, in general, better geomet-

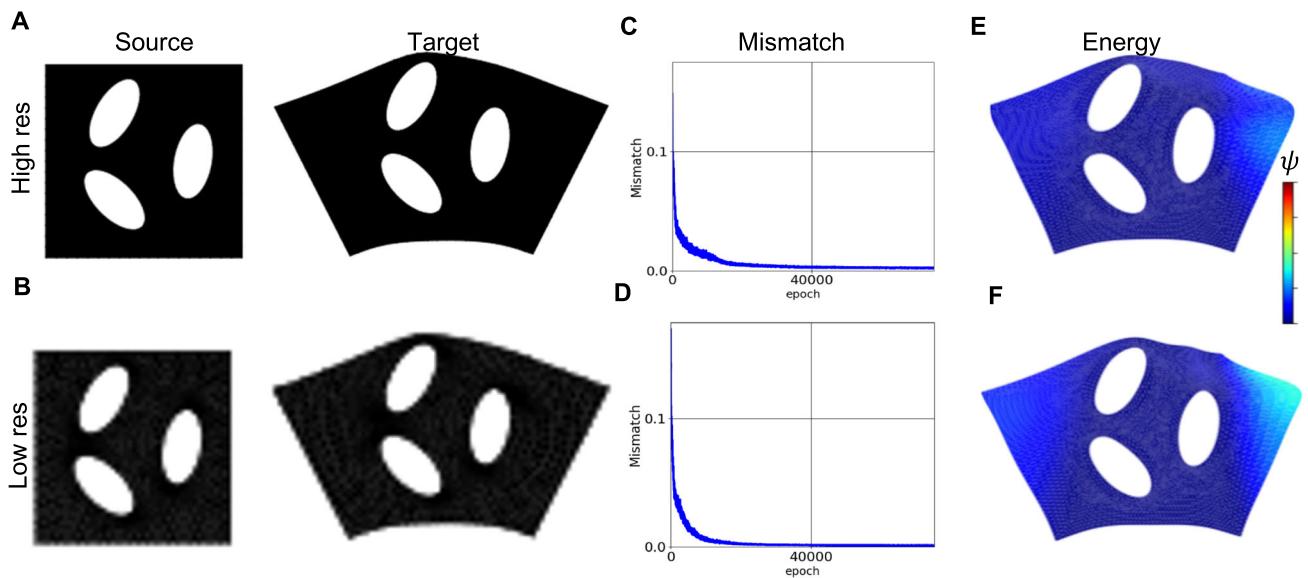


Fig. 4 Resolution of image (discretization of source and target functions) impacts registration quality: A) source and target high-resolution images. B) source and target low-resolution images. C & D) mismatch

energy loss for high- and low-resolutions, respectively. E & F) strain energy field corresponding to high- and low-resolutions

rical registration can be achieved when more information is given in the dataset, as expected in medical image analysis.

4.2 Biological Examples

In this part, we apply the new method to address practical questions. In particular, we utilized the method to quantify mechanical stresses during the regeneration of a zebrafish tail, brain shrinkage induced by cerebral atrophy associated with normal brain aging, and cortical expansion during fetal brain development during gestational weeks 21 through 36. However, before introducing the examples with growth, we start by showcasing a benchmark registration problem relevant for medical image analysis without growth biophysics.

4.2.1 Benchmarking Brain Registration

Prior to considering the coupling of growth and deformation with our method, we first show that the method is able to obtain high-quality registration for a complex benchmark such as brain registration, which had received widespread attention in the literature (Toga & Thompson, 2001; Hernandez et al., 2009; Lombaert et al., 2014). In this regard, we utilize NIREP data, a publically available data set, that has been skull-stripped and annotated to a total of 33 regions of interest (ROIs) by neuroradiologists (for more information, we refer the reader to Christensen et al. (2006)). The MRI data, along with the segmented nifty files, can be found in <https://github.com/andreasmang/nirep>. We intentionally showcase two brain scans that exhibit substantial differences

with respect to intracranial volume and orientation of the brain to assess how effectively the new setup can register them. To that end, we selected images **na10.nii.gz** and **na02.nii.gz** as the initial and target images, respectively. we selected $\beta = 0.0001$.

The different views of the brain have been plotted to show the quality of the registration. In particular, Fig. 6-A summarizes the results of the problem by including a 3d view of the initial, registered, and target MRIs along with some random selection of the corresponding sagittal, coronal, and axial cross sections. For better comparison, we also included the merged version of cross sections in Fig. 6-C. The results show a good alignment between the warped-out initial and target images. We also quantitatively compared the result of our method with that of **xxx (yyy)**, a Python library based on ANTs, one of the most commonly used registration algorithms for brain data (Klein et al., 2009; Ou et al., 2014). Both methods were compared based on the Dice similarity coefficient defined in (4.5). As previously stated, the NIREP data contains segmented data, a total of 33 labels. Thus, we evaluated (4.5) for each label and reported the results alongside those obtained using the method "SyNOnly" in Fig 6-B. This result suggests that our new method produces comparable registration results in comparison to one of the well-established algorithms widely used by the neuroscience community. Additionally, we analyzed the effect of discretization in the predictor calculation by considering grids of $20 \times 20 \times 20$, $45 \times 45 \times 45$, and $100 \times 100 \times 100$, which demonstrates the convergence of the calculation. To further ensure the independence of the comparison from the choice

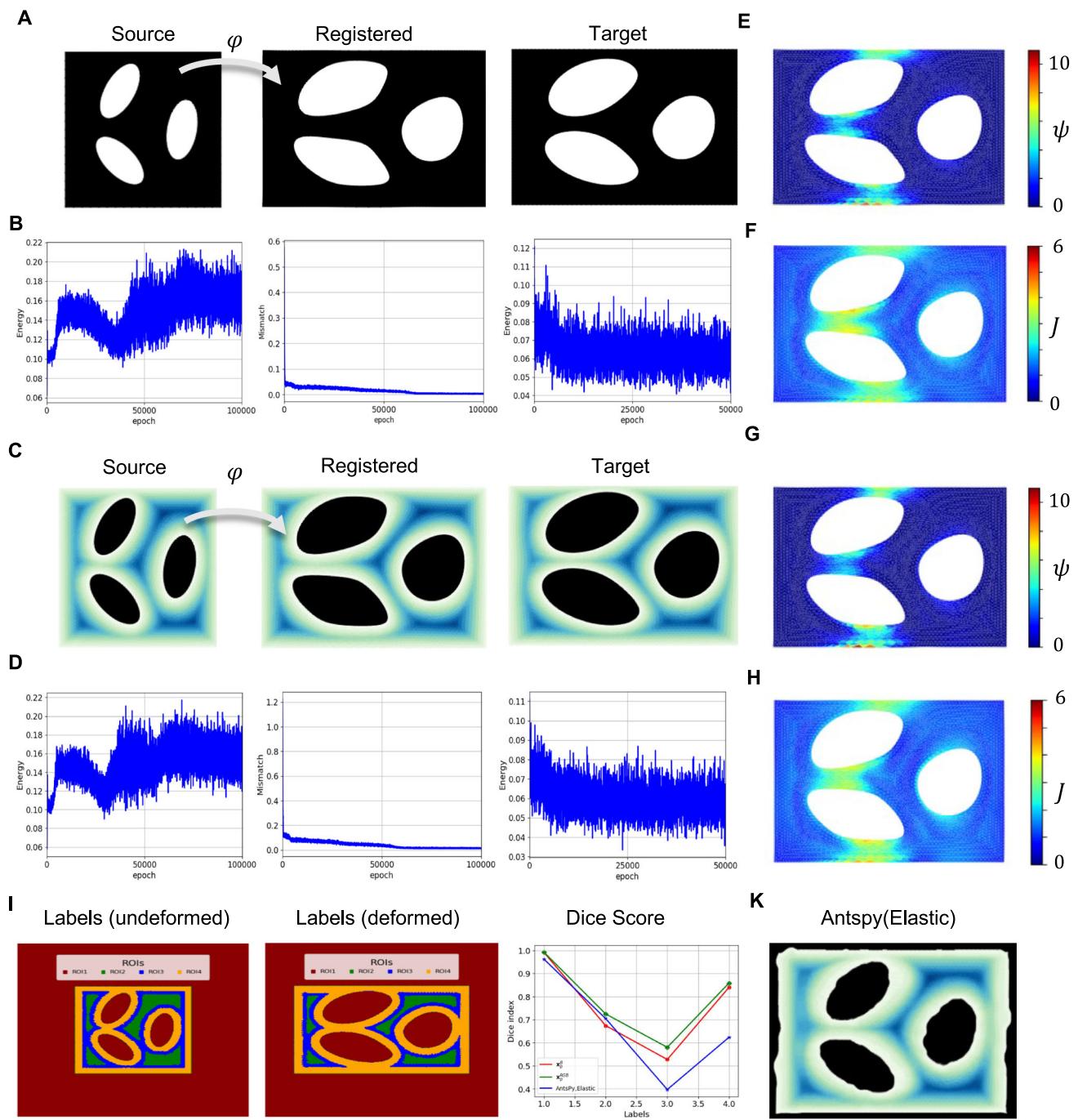


Fig. 5 Gray scale or RGB information improve registration compared to binary approaches: A) Source, registered, and target images for the binary dataset. B) Mismatch energy, strain potential energy loss for predictor, and strain potential energy loss for corrector pertinent to binary datasets. C) Source, registered, and target images for the RGB dataset. D) Mismatch energy, strain potential energy loss for predictor, and strain potential energy loss for corrector pertinent to RGB datasets.

E & F) Strain potential energy & Jacobian fields resulting from the binary dataset analysis. G & H) Strain potential energy & Jacobian fields resulting from the RGB dataset analysis. I) Initial and deformed configurations of a synthetic segmentation, along with the resulting Dice indices for both binary and RGB cases, and K) the registered configuration through AntspyElastic

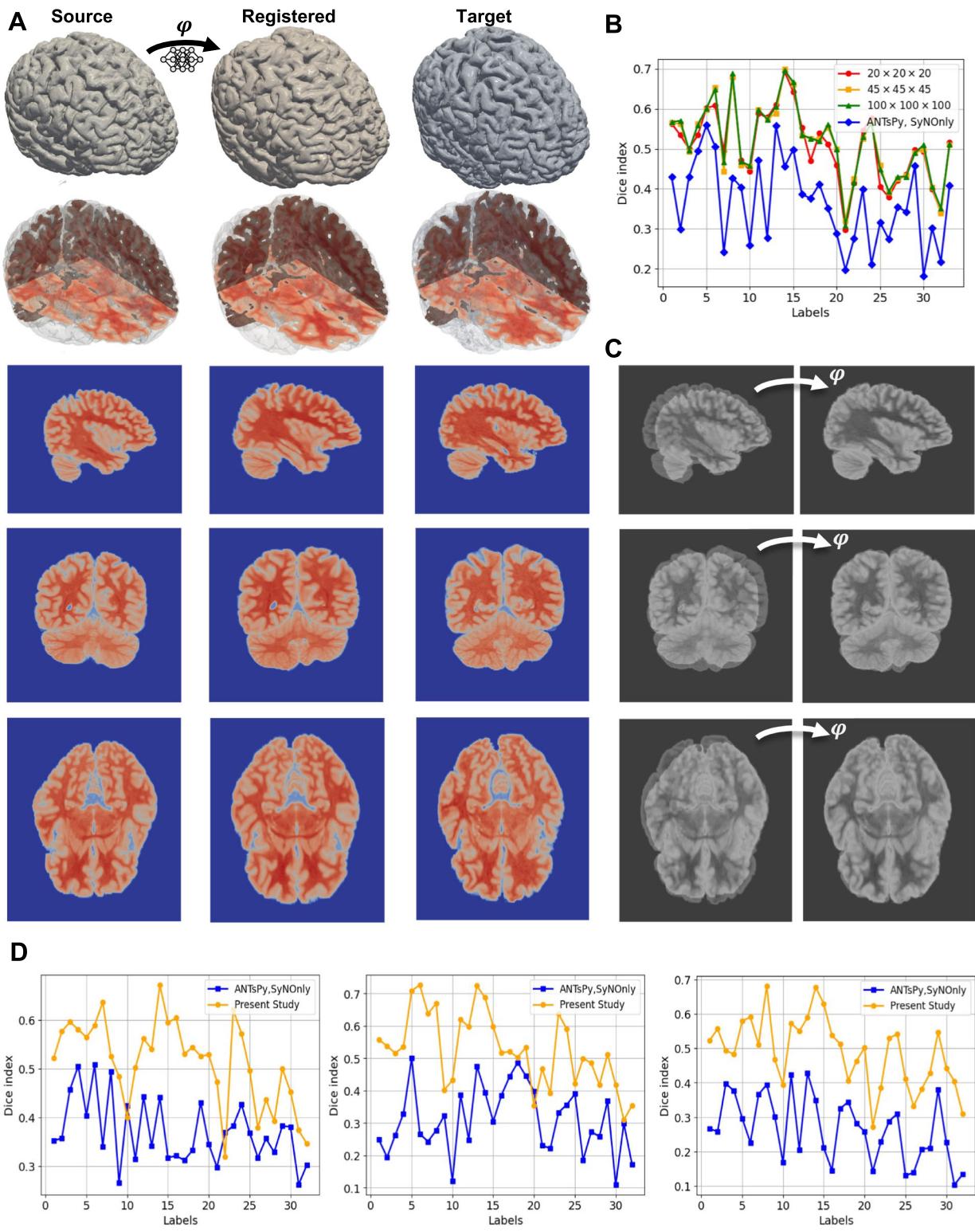


Fig. 6 Benchmark problem for brain registration. A) Iso-surfaces along sample sagittal, coronal, and axial cross sections corresponding to source, registered, and target of NIREP MRI datasets MRIs (na10.nii.gz to na02.nii.gz). B) Comparison of registration methods in terms of the Dice index between the current method across different discretization sizes and ANTsPy (SyNOnly) for all ROIs in the dataset. C) Over-

lay of initial and target intensities as well as the target and deformed geometry from the registration process, showing good alignment was achieved. D) Comparison of the result of registrations for three more pairs chosen from NIREP data sets(left: na16.nii.gz to na05.nii.gz, middle: na11.nii.gz to na03.nii.gz, and right: na12.nii.gz to na01.nii.gz)

of samples, we repeated the same calculation for three pairs of datasets, which shows a similar trend in the quality of registration.

4.2.2 Growth Examples

In the first example, we perform an analysis of the regrowth of the zebrafish tail. After amputation, the tail starts growing, and the exact signals that activate progenitor cells to cue regeneration are unknown. A hypothesis, in this regard, is that the mechanical forces within the tissue cue regeneration. To explore this assumption, it is required to approximate the spatial distribution of deformation over the tissue given a series of images of the regenerating tail. Physics-based registration can thus provide a better insight into this process. Experimentally, we cut the tail using a scalpel and imaged the fish for 24 hours using a cell cycle indicator (Sugiyama et al., 2009) to see where replication is initiated during the process of regrowing lost tissue (see Fig 7-A).

For registration analysis, we followed the sequential approach mentioned above. In this regard, we considered a total of six frames from the image data (i.e., frames at 10 min, 1hr 50 min, 3hr 30 min, 5hr 10min, 6hr 50 min, and 8hrs 30 min) in such a way that the selected images capture the main deformation over 24 hrs. We performed the analysis in a time-reversed manner for simplicity. Specifically, in our sequential approach, we started from frame 6 and sequentially returned to frame 1. Our reasoning for doing so is to counteract the limitations of the experimental dataset, in which the majority of the initial deformation happens within the first fifteen minutes. Because the tail fin relaxes gradually from the peak of that deformation over the course of many hours, correctly mapping the movement of the tissue from the relaxed to the contracted state by starting with a later shape and working backward through a slower shape change was more intuitive than attempting to derive the entire tissue movement from a very fast initial contraction. In the predictor part, we trained for 50000 epochs overall, with $\beta = 0.1$. For the corrector part, we trained the first three steps for 50000 epochs each and the last two steps for 100000 epochs each. We also set initial growth parameters as follows: $\kappa = -17, -17, -17, -11, -5$ for the first to fifth steps, respectively. Fig 7-B shows the result of the registration. The first two rows in Fig 7-B shows directly the image data obtained from experiments, while the third row is the result of the registration, i.e., the deformation of the initial frame into the shapes of the subsequent frames. To facilitate clearer observation, we merged the second and third rows and reported in the fourth one. As can be seen a good registration obtained in all steps. To show this fact, the dice score across the entire domain for each step of registration has been plotted in Fig. 7-C.

The key outcome from the registration is, however, not the alignment between data and deformed images but the implications for tissue growth. The first row of Fig 7-C shows the contours of the maximum principal stretch after running the corrector to satisfy mechanical equilibrium. Initially, the deformation is minimal, which is intuitive since the images do not seem to indicate large shape changes over the last 6 hrs. However, in the last two time frames, between the first two hours, from 10 minutes to 1 hour 50 minutes, there is significant deformation, with larger stretches toward the tip of the regenerating tail. Also depicted in Fig 7-C are the elastic contributions to the deformation. Recall that the corrector also allows for tissue growth to dissipate mechanical energy. If the deformation were fully plastic, that is, if growth were able to completely absorb the shape changes, then the resulting elastic stretch contour would be 1. In reality, Fig 7-C shows that while growth allows for most of the energy to be dissipated, the resulting elastic deformation is not exactly 1 but rather varies slightly with respect to 1. In the last frame of Fig 7-C, it can be seen that the largest elastic deformation is expected at distal end of the regenerating tail. We anticipate that this new information will shed light on the biological mechanisms regulating zebrafish tail wound healing.

In the next biological case, we quantify brain volume loss in a cognitively subject between the ages of 71 and 75 whose scans were obtained from the Alzheimer's Disease Neuroimaging Initiative (Jack et al., 2008; Petersen et al., 2010). In this example, we are primarily interested in brain deformation, and not just registration between images alone. To that end, we first used ANTsPy to rigidly align the source and target image. Then, we used our method while prescribing homogeneous and isotropic shrinkage, i.e., negative growth, as detailed in Section 3.3. In our analysis, we optimized for the (negative) growth rate κ for the given MRI data while maintaining equilibrium. Growth is optimized such that elastic energy is dissipated. Because this deformation is shrinking or negative growth rather than positive growth or addition of mass, we switch the notation from θ^s to θ^s , but this is just a slight change in notation, the split of the total Jacobian is the same as Eq. (3.8), $J = J^e J^s$, $J^s = (\theta^s)^3$. Fig. 8-A shows multiple views of the brain's displacement magnitude field as well as a limited number of nodal displacement vectors that illustrate how the brain deforms. We clearly observe a uniform contraction of the cortical surface representative of brain shrinkage. Our method allows us to quantify the degree of volume loss, as shown in Fig. 8-B. As can be seen, there is a five percent permanent reduction in volume between the baseline and follow-up scan $\kappa = 0.013111$, $J^s = (\theta^s)^3 = 0.961431$, which accords with the average value of the total Jacobian, i.e., $\langle J \rangle \approx 0.96$. Also, it is observable that a good convergence of the shrinkage parameter was achieved, implying that the major part of the deformation is due to permanent brain shrinkage,

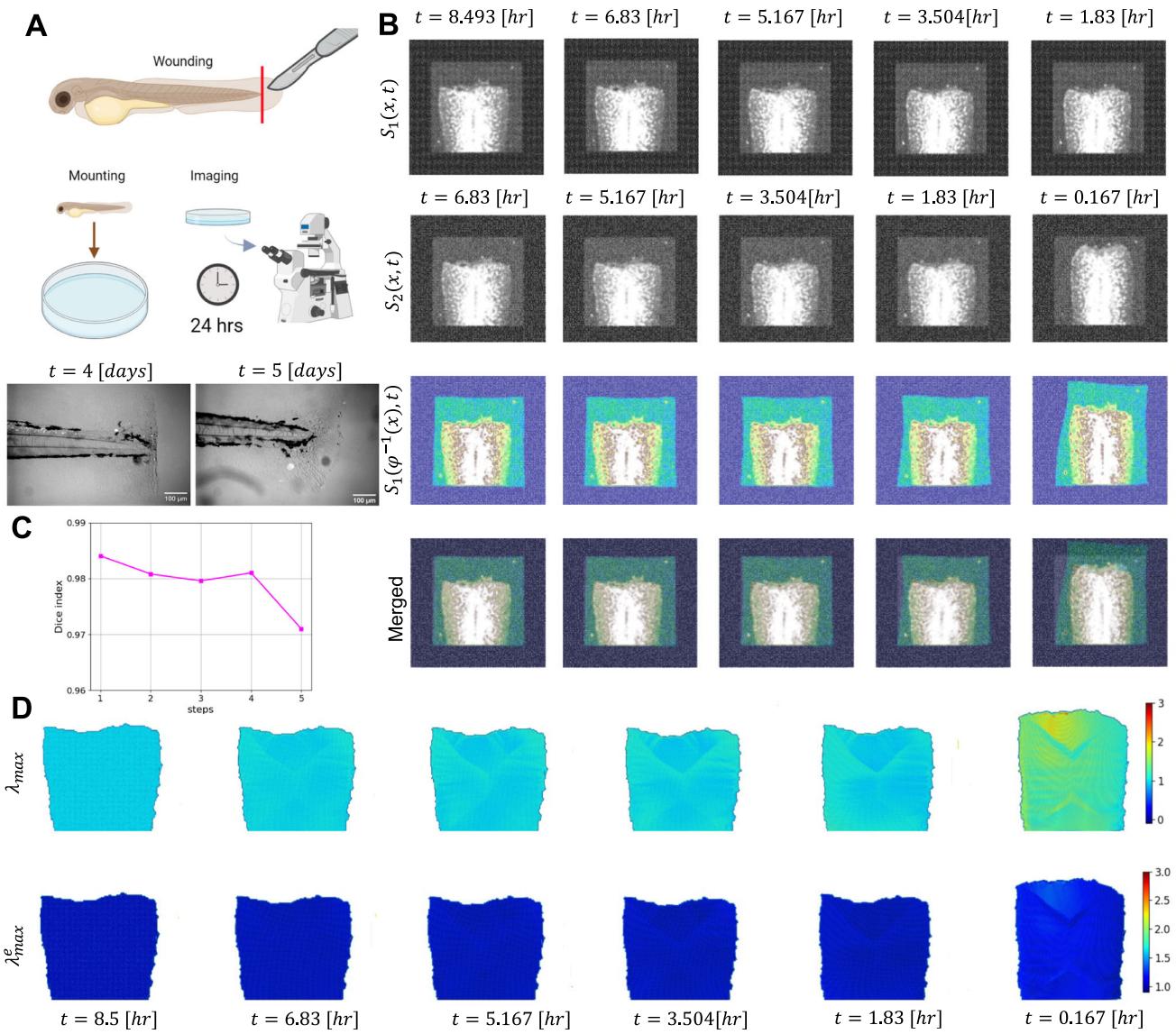


Fig. 7 Zebrafish tail wound. A) A schematic view of the experimental steps carried out to obtain the dataset and focal images corresponding to four and five days after the amputation. B) Initial ($S_1(x, t)$), target ($S_2(x, t)$), and registration results ($S_1(\varphi^{-1}(x), t)$) have been represented in the first, second, and third rows, respectively, while the

merged of them have been demonstrated in the fourth one. C) Dice score for all steps of the registration. D) Comparison of the total and elastic deformation gradients in terms of maximum principal stretches after applying the corrector step with growth dissipation

with only a small amount of residual elastic deformation left. As shown in Fig. 8-B, the major deformation occurs in cerebellum. Therefore, to assess the quality of geometric registration, we intentionally selected a slice of the brain in this region and compared the results with ANTsPy, as illustrated in Fig. 8-C. We present results for three zoomed-in slices: one from the cerebellum and two randomly selected. For each slice, we provide the corresponding target image and the results obtained using ANTsPy (SyNOnly). To facilitate comparison, we separately compare each slice by presenting the initial image, the results from ANTsPy, and the results from the present method against the target image. As observed, the geometric registration obtained using the

NODE method is comparable to the results achieved with ANTsPy. Additionally, the resulting Dice score for the entire geometry is 0.965, indicating a good registration.

For the last biological example, we chose to study the extensive cortical expansion associated with early fetal brain development. To that end, we analyzed the unbiased, deformable, spatiotemporal atlas of the fetal brain provided by The Computational Radiology Laboratory at Harvard University (Gholipour et al., 2017). The atlas consists of weekly images between gestational weeks 21 and 37. To demonstrate the capabilities of the algorithm, we selected the images three weeks apart, i.e., weeks 21, 24, 27, 30, 33, and 36. Analogous to the zebrafish example, we implemented the sequential ver-

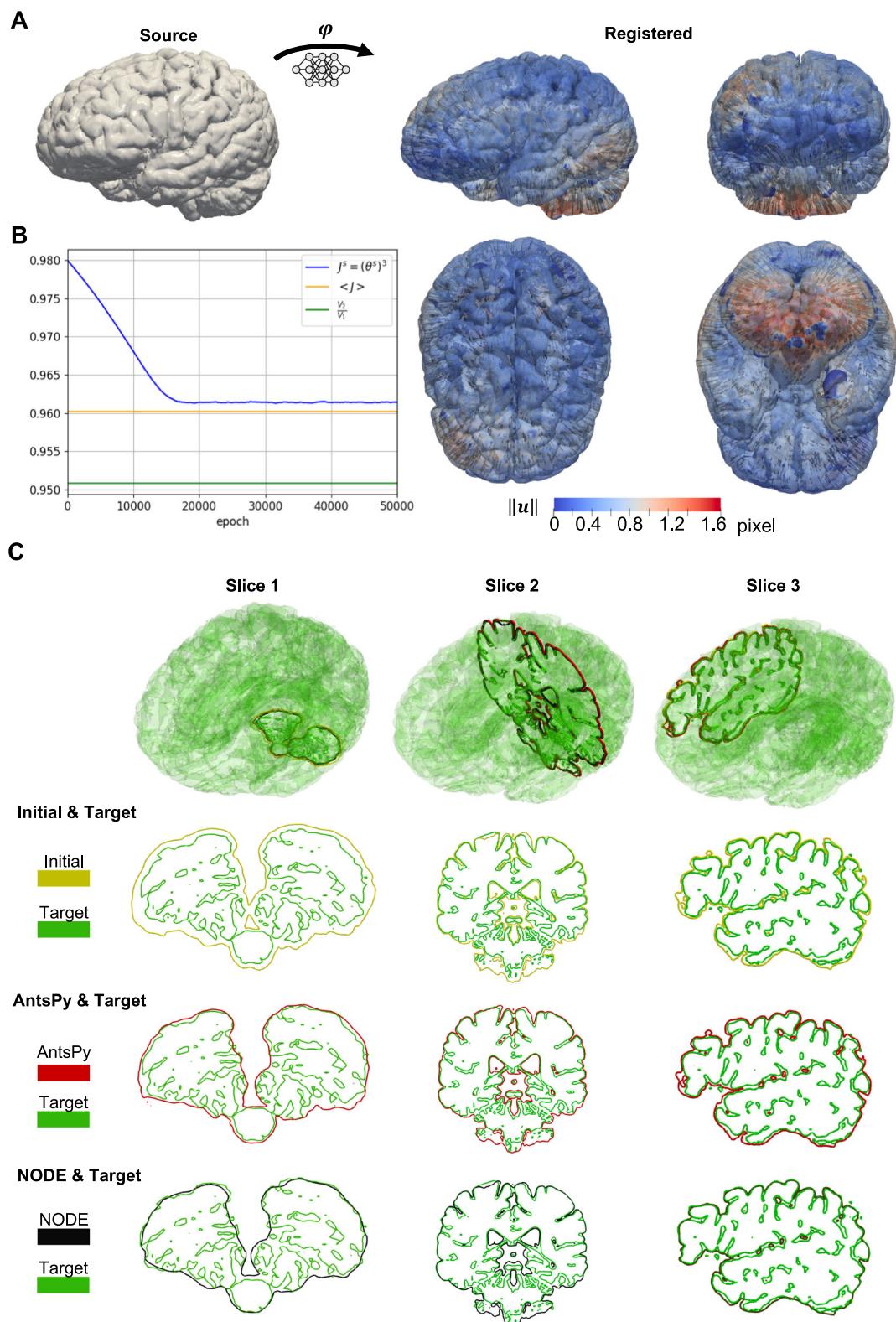


Fig. 8 Brain atrophy example. A) Iso-surface of the baseline and the follow-up obtained from registration along with quiver plots of displacement fields on selected points. The direction of the arrows confirms the contraction of the brain. B) The convergence of the new solution procedure in the corrector part. The major part of contraction is due to

permanent shrinkage, $J^s = (\theta^s)^3 = 0.961431$, denoting a reduction of 5%. C) Comparison of the registration between the present study and ANTsPy (SyNOnly) in terms of the iso-surface of the brain with zoomed-in slices

Table 1 Central parameters for ablation study

Steps	Architecture	Time Step	Regularizing Term	Strain Potential
Predictor	[2,40,40,40,2]	1/15	$R = \ \mathbf{F}\ ^2 + (\log J)^2$	N/A
Corrector	[2,40,40,40,2]	1/15	N/A	(4.3)

sion of the method. We used $\beta = 0.1$ and 50000 as the epoch number in the predictor part. In the corrector part, to obtain a good convergence for the growth parameter, we considered 100000 epochs in the first step and 50000 for others. Fig. 9-A shows the displacement magnitude fields and displacement vector field. We generally observe a relatively uniform expansion of the cortical surface. Additionally, early weeks are characterized by very similar displacement magnitudes across the whole surface, while later weeks exhibit increasingly heterogeneous magnitudes. Specifically, we see a localization of maximum displacement magnitudes in the frontal and temporal lobes. In Fig. 9-B, we statistically summarized the distribution of elastic, total, and growth Jacobian in the whole volume. As can be seen, the average Jacobian due to growth is the same as the average of the one relevant to total deformation during all weeks, while the average elastic Jacobian increases slightly more than unity as a function of time. Strikingly, the brain's volume increases nearly 5-fold during the mid-to-late gestational stages. The total Jacobian (top row) and elastic Jacobian (bottom row) shown in Fig. 9-C indicate regions of highest volume growth. It is evident from the coronal sections that growth predominantly occurs in deep tissue structures as well as the outer cortical layer. This confirms that deformations during these gestational weeks can be explained by tissue growth, which can minimize the total elastic energy without inducing large residual stresses. Analogous to previous examples, we compared the registered images with target ones and computed the Dice score for all steps of registration, as shown in Fig. 9-D and Fig 9-E, respectively.

5 Ablation Study

In this part, we present a comprehensive ablation and hyperparameter study of the proposed algorithm for the shear-contraction problem defined in Fig. 3. This problem is one that proves difficult for conventional deep neural networks and required the NODE architecture to guarantee invertibility. Considering that the method has two major parts, i.e., the predictor and corrector, in which we approximate the deformation field with a flow algorithm (i.e., NODEs as the deformation approximator), we analyze the following: 1) for the predictor part, we consider the effect of different regularization terms, different values of hyperparameter β , the number of neurons of each layer for the core neural network,

the number of its layers, and pseudo time steps $\Delta\tau$; 2) in the corrector part, we test different choices of regularizing terms and various β assumed in the predictor, the number of neurons in each layer of the core neural network, the number of its layers, and the number of pseudo time step Δt . For varying the parameters we consider those in Table 1. Then, we vary the target parameter while keeping all other parameters fixed at their baseline values.

Figs. 10-A & -B show the mismatch loss and the elastic energy loss associated with the ablation study of the predictor part by varying β , the number of neurons, and the number of pseudo time steps. As mentioned in Pawar et al. (2022), the mismatch loss and the regularizing energy term are two competing factors balanced by β . Larger values of β lead to a smooth mapping that partially minimizes the mismatch loss, while lower values result in effective registration but potentially loss of regularity. Hence, the value of β should be small enough to achieve good registration while still resulting in a smooth transformation. The trade-off of the multi-objective optimization as a function of the hyper-parameter β is also illustrated in the Appendix.

Furthermore, similar behavior is observed when changing the number of neurons and layers: the higher values for number of neurons and layers lead to faster and better convergence. However, this choice is accompanied by more computational costs during learning and back propagation. Hence, in this setting, it is required to strike a balance between the accuracy and computational costs. Changing the $\Delta\tau$ has less effect on the path of the minimization. However, as shown in the shear contraction example, the definition of the flow algorithm is crucial since it guarantees the invertibility of the resulting transformation.

In the corrector part, depending on the choice of the regularizing term in the predictor, one expects different minimization paths in the corrector. This fact has been illustrated in Figs. 10 -C & -D. Fig 10-C shows the strain energy path minimization when the regularizing term is $R = \|\mathbf{F}\|^2 + (\log J)^2$, while Fig. 10-D indicates the minimization of the elastic energy by the corrector when the regularizing term is selected to be the same as the strain energy potential. As can be seen, the corrector part reached to a stationary point of total potential energy in all cases, but with different initial conditions given by the predictor. Also, Fig 10-D shows that when $\beta = \frac{1}{6000}$, the predictor path leads to a strain energy that is then further minimized by the corrector. In contrast, for $\beta = \frac{1}{1000}$, the corrector solution is close to the predictor

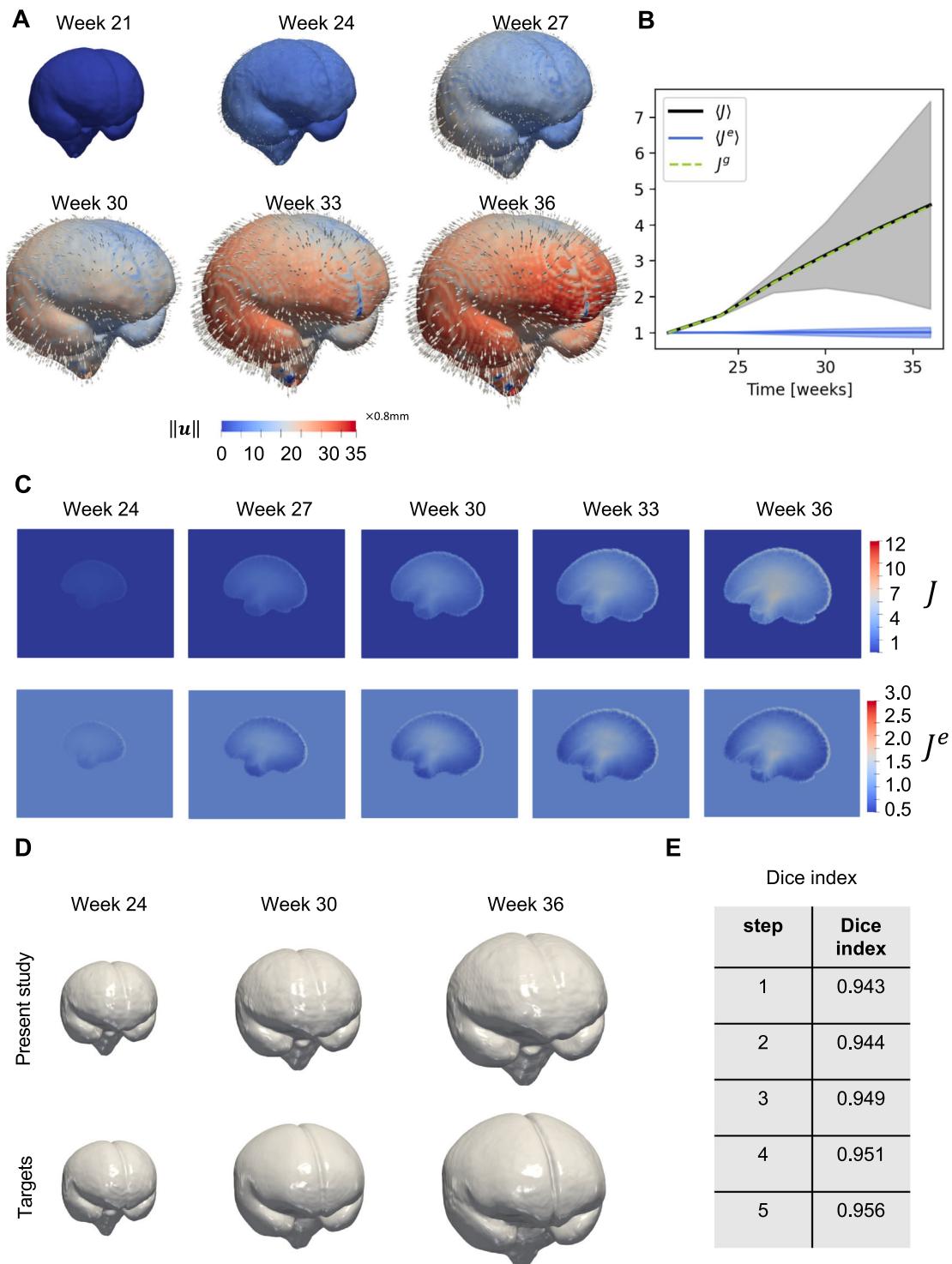


Fig. 9 The extensive cortical expansion associated with early fetal brain development. A) Brain iso-surfaces for all six steps together with quiver displacement fields. B) statistical representation of the Jacobian field for total, elastic, and growth parts. The average growth Jacobian is closely equal to the total Jacobian for all steps, while the elastic Jacobian remained close to unity. C) Comparison of the total and elastic Jacobians for a given sagittal cross section for all steps, confirming the

deformation is mainly non-elastic, and the growth deformation is able to dissipate most elastic energy and residual stresses. The results suggest that growth rather than elastic deformation is likely the cause of the large shape and volume changes during embryonic development. D) Comparison of three steps of the registration between the present method and the AttnPy(SyNOnly). E) The dice score indices corresponding to registration of the whole geometry of the brain

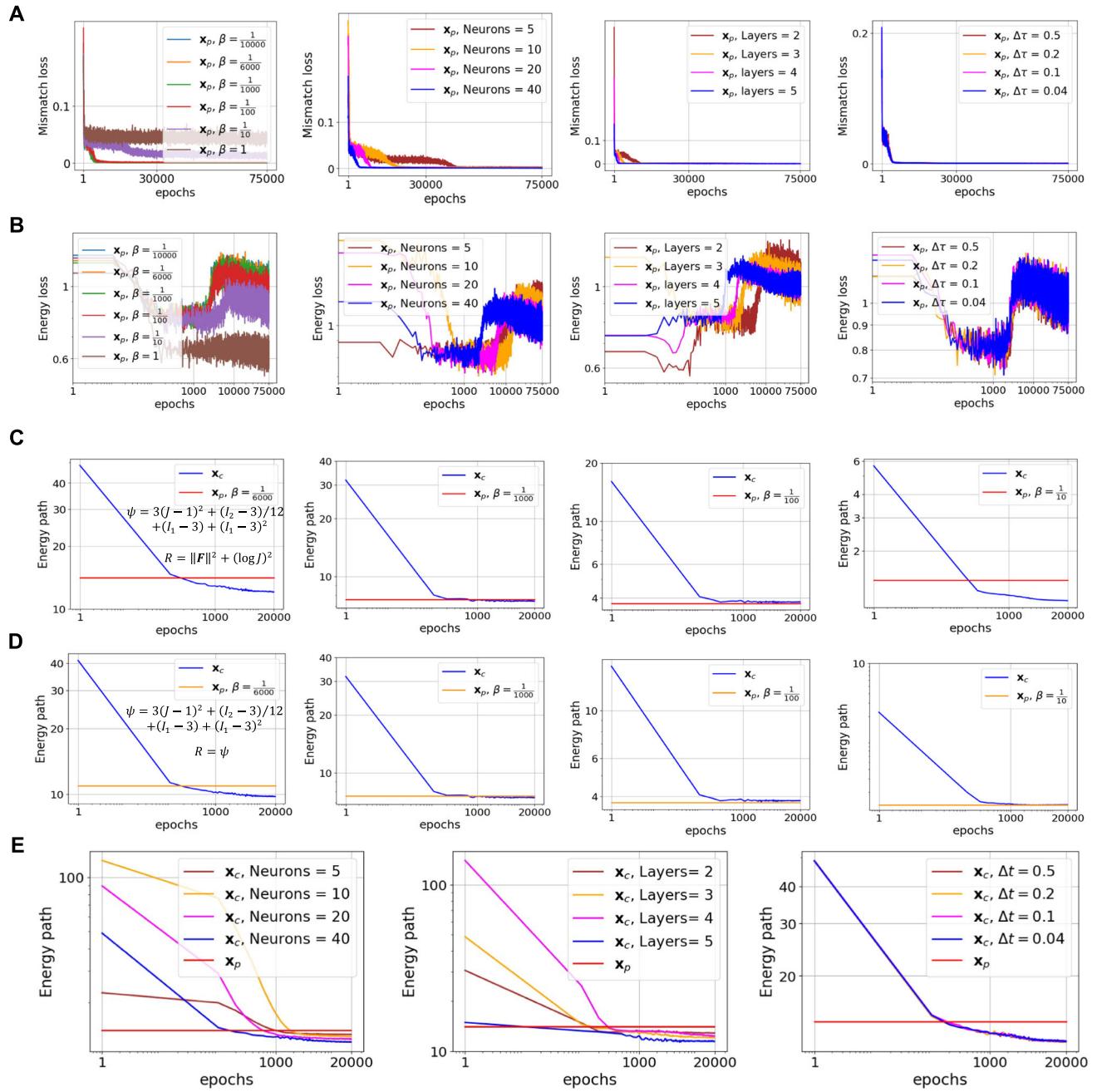


Fig. 10 Ablation study for shear-constriction problem defined in Fig. 3: A) Mismatch loss versus epoch (log-log scale) in the predictor part for different values of β , the number of neurons, the number of layers, and pseudo time steps. B) Energy loss versus epoch (log-log scale) in the predictor part for different values of β , the number of neurons, the number of layers, and pseudo time steps. C) Minimization of energy paths in the corrector part for different values of β when the regularizer

result, implying the existence of some specific values of β where the predictor solution may be close to satisfaction of the equilibrium equations. Nevertheless, the choice of β such that it provides good registration and good minimization of the elastic energy is difficult to obtain and might not exist. In contrast, our predictor-corrector split provides robust regis-

is selected to be different from the strain potential energy. D) Minimization of energy paths in the corrector part for different values of β when the regularizer is selected to be the same as the strain potential energy. E) Minimization of energy paths in the corrector part for different numbers of neurons, layers, and pseudo time steps with the regularizer $R = (\log J)^2 + \|\mathbf{F}\|^2$

tration as well as minimization of the strain energy (indicative of equilibrium), as shown in Fig. 10-D.

Fig. 10-E shows the results pertinent to sensitivity analysis of corrector solution with respect to the number of neurons, the number of layers, and Δt have similar effect in the path of minimization of the corrector as in the predictor part. In this

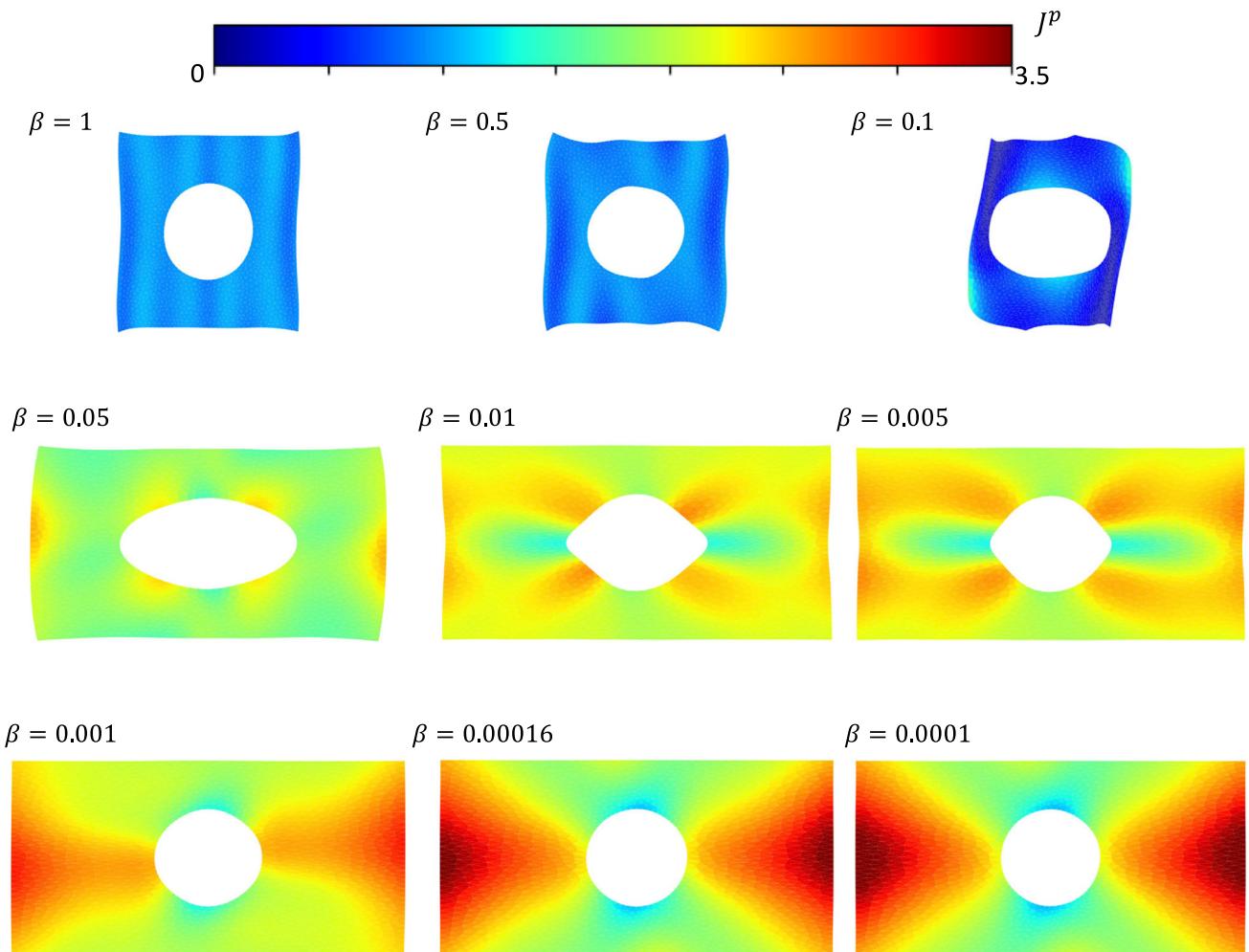


Fig. 11 Effect of the hyperparameter β on registration quality in the extension of the plate with a hole considered in Fig. 2. As can be seen, this parameter should be selected small enough to achieve proper registration

analysis, we selected the results corresponding to $\beta = \frac{1}{100}$ and $R = \|\mathbf{F}\|^2 + (\log J)^2$.

6 Discussion and Conclusion

In this article, we presented a new algorithm for registering 2d and 3d images while concomitantly solving physics-based boundary value problems. The approach is composed of two parts: 1) predictor and 2) corrector. In the predictor, the algorithm minimizes the mismatch loss, while a regularizer is defined to preserve the smoothness of the resulting transformation. In the corrector, a Dirichlet boundary value problem is then defined and solved in a weak sense, similar to the finite element method. The boundary condition in the corrector is determined by the map obtained from the predictor transformation, which is preserved via distance functions. This corrector part results in finding a stationary (minimizer)

point of the total potential energy, leading to a solution that fulfills the equilibrium equation of continuum mechanics in a weak sense. The predictor-corrector split crucially allows not just for a better result than multi-objective optimization registration methods (Pawar et al., 2022) but enables complex multi-physics coupling such as growth and remodeling (Ambrosi et al., 2011).

To show the plausibility of the method, we considered several synthetic and real biological examples. Specifically, by using synthetic examples, we showed that the integral format of a PINN (direct potential energy minimization) where the prescribed displacement boundary condition is satisfied exactly, obtains a solution close to the solution of the FEM. This fact is because the method follows the same paradigm as FEM, which is the minimization of the potential energy. In contrast, strong form or variational PINNs commonly impose boundary conditions via a penalty and result in multi-objective optimization, and it may result in a solution that only partially satisfies equilibrium and boundary conditions

Fig. 12 Comparison of registration obtained from the present method with those of ANTsPy for NIREP data: A) na16.nii.gz to na05.nii.gz, B) na11.nii.gz to na03.nii.gz, and C) na12.nii.gz to na01.nii.gz.

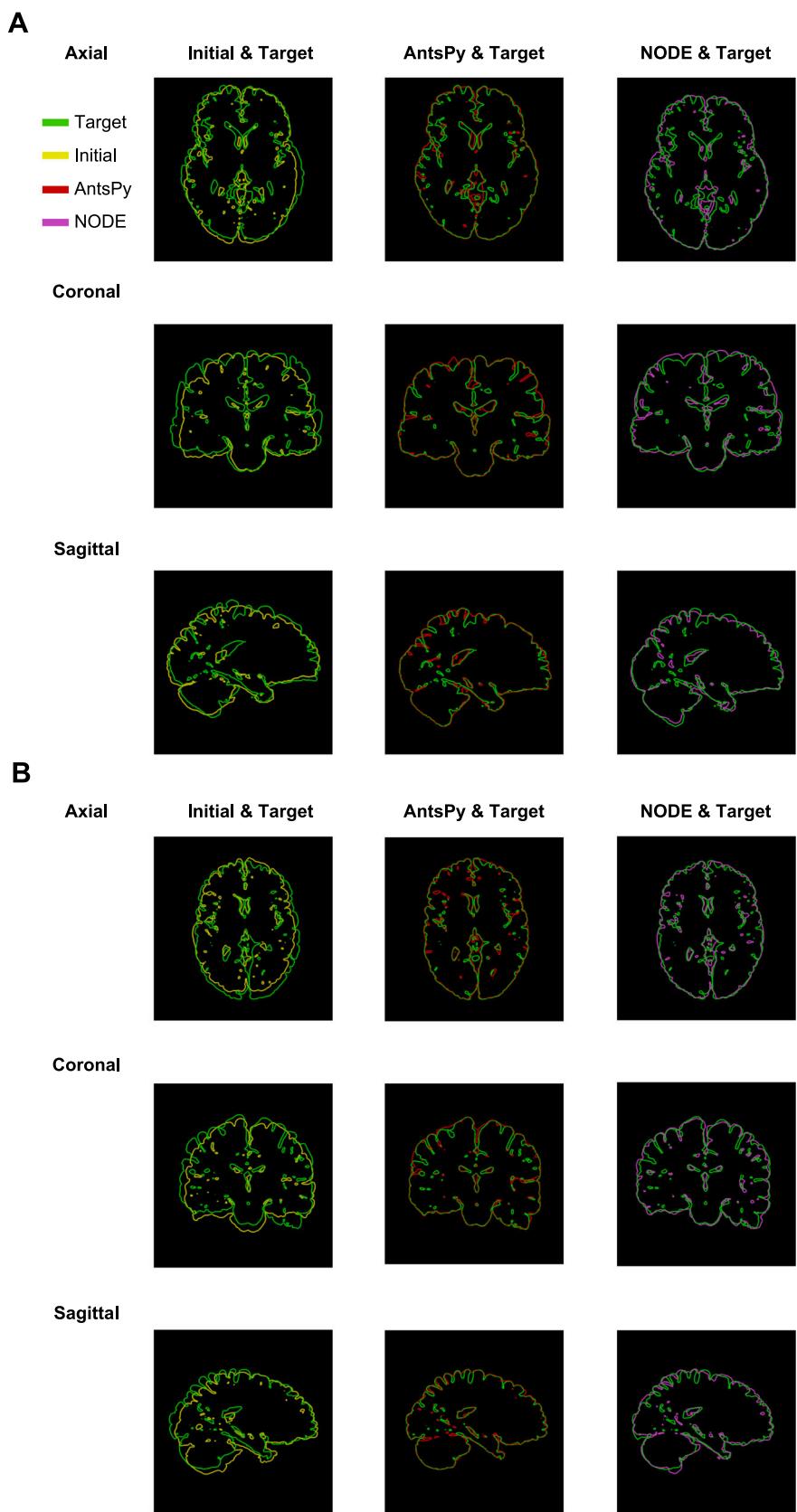
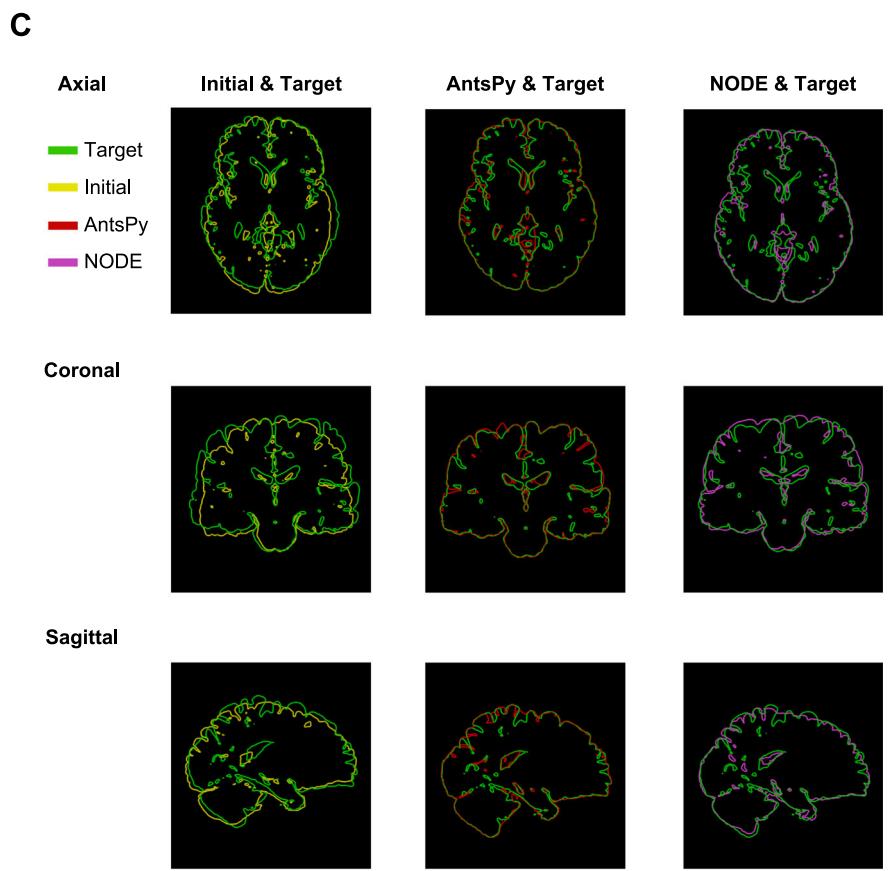


Fig. 12 continued

(Sukumar & Srivastava, 2022). The registration method, in particular the corrector part, satisfies a Dirichlet problem, i.e., it solves a problem with exact boundary conditions, where these boundary conditions are the result of the predictor. The boundary conditions are imposed through distance functions, a strategy also used in the broader context of PINNs to impose boundary conditions (Sukumar & Srivastava, 2022). In contrast to the usual PINN problems, the registration problem is inherently ill-posed as there are infinitely many predictor solutions that vanish the mismatch loss but define a different Dirichlet problem for the corrector. Nevertheless, this characteristic is not considered an anomaly but rather lies in the nature of registration problems.

Among existing deep learning architectures proposed for image registration (see, for example, Balakrishnan et al. (2019)), we showed that the NODE is a good candidate to produce invertible transformations while having strong expressibility. It is worth mentioning that other invertible architectures, such as NICE (Dinh et al., 2014) and Real NVP (Dinh et al., 2016), were tested. Still, obtaining a good deformation field was challenging, which can be attributed to the predesigned invertible architectures of the main building block of such methods. We further observed that the resolution of the image may result in different solutions. Hence, input data with high resolution leads to better registration.

By performing registration on binary and RGB images, we also noticed that better registration may be obtained when the image has more information in terms of intensity.

To benchmark the method against relevant applications, we applied it to perform the registration of brain MR images. In this regard, we selected two MRI data with noticeable differences from NIREP data (Christensen et al., 2006), conducted the registration (only predictor part), and compared the results with ANTsPy (Avants et al., 2009a,b). The results indicate that the NODE structure is comparable and can even outperform state-of-the-art.

The main motivation of the proposed method is to go beyond registration and demonstrate the ability to solve biophysics problems within the same framework, which typically involve complex mechanics and multi-physics coupling and additional external software (López et al., 2023; Min et al., 2023; Wang et al., 2020).

In particular, in this paper, we contribute to the field by introducing the predictor-corrector split, which gives us much more flexibility for the types of multi-physics problems we can tackle while retaining the same deep learning framework. Three biological examples with tissue growth were considered to underscore the unique abilities of the proposed approach. Firstly, through the sequential format of the new technique, we managed to map total, elastic, and

growth deformations across zebrafish wound healing. Additionally, we modeled brain atrophy. The problem of fetal brain development was also studied. The algorithm was able to find a continuous deformation map from the geometry at week 21 up to week 36 during pregnancy. The volume changes observed in this time frame are substantial, with total Jacobians averaging above 4-fold volume changes. Thus, the method can handle extreme deformation. Growth biophysics explains that most of the deformation can be attributed to growth, leading to the dissipation of mechanical energy.

The biological examples contain several assumptions regarding the material behavior, growth ODE form, growth tensor form, and assumption of homogeneous material. Each of these problems can be further refined in future work to account for the complex biological mechanisms behind zebrafish wound healing (De Vos et al., 2017), brain atrophy in neurological disease (Blinkouskaya et al., 2021; Blinkouskaya & Weickenmeier, 2021), and fetal brain development (Wang et al., 2020). Furthermore, in this work, we enforced the physics with a method similar to variational PINNs, while the method of the neural operator networks can be adapted, which generalizes the solution to other PDEs (He et al., 2024, ?). This is an interesting avenue of future research when the PDE is unknown or to accelerate the solution in the case that similar geometries are processed, e.g., in the case of brain atrophy (Visser et al., 2023). Extension to thin membranes is also a problem of interest for medical image analysis we intend to tackle (Pouch et al., 2014, 2020). Moreover, in this study, we demonstrated the application of the new method by testing it on a few brain examples. However, to establish statistically significant knowledge regarding brain biophysics, the new method should be applied to larger datasets (Marcus et al., 2007; Ou et al., 2014).

Furthermore, the ablation study provided deeper insights into the parameters of the new algorithm. The numerical results indicated that altering the number of neurons and layers in the neural network has a similar effect on the accuracy of the method in both the predictor and corrector, suggesting an optimal architecture to balance computational cost and accuracy. Additionally, the ablation study further confirmed that, unlike other methods, our proposed approach can obtain the stationary point of the total potential energy similar to the finite element method in the corrector step, regardless of the chosen regularizing term. Hence, as a future work, it is desirable to develop a hybrid method where the predictor part is done with current registration methods available in literature while the corrector part is carried out by the present algorithm. The key requirement for this approach is ensuring the invertibility of the predictor.

We investigated the computational cost of the method with two predictor cases, both with epoch=50000. For a 2d analysis, which we performed on the plate extension with a hole, the computational cost was about an hour with Nvidia

A100 Tensor Core GPU (80 GB). For a 3d problem, which we performed for the predictor of shrinkage problem, we observed a computational cost of 3 hrs with Nvidia A10 Tensor Core GPU (24 GB). The major computational cost can be attributed to the NODE's training. However, as illustrated in Fig 3, obtaining a general function based on an evolutionary first-order dynamical system is a rich structure, although it might be computationally demanding. A strategy that can reduce the computational cost of the method, which we plan to implement in future work, is to adopt an adaptive scheme by applying an early stop algorithm. In particular, one can provide a criterion that prevents the algorithm from further training when the desired convergence is achieved. Another strategy, in this regard, is by employing refinement strategies to sample in regions with higher error, similar adaptive re-meshing in finite element or spline-based registration (Pawar et al., 2016). Finally, since the displacement field is constructed from the (pseudo) velocity field, the order of the problem can be reduced by leveraging reduced-order methods, such as principal component analysis, which allows it to be solved with fewer degrees of freedom.

As the theory and numerical results suggest, the new method is robust. In particular, the geometric registration in the predictor part is comparable with existing methods. In addition, the new solution procedure attaches physics to the underlying transformation by finding the minimizer of the total potential energy for an equivalent Dirichlet problem in the corrector part. The predictor-corrector split enables rich biophysics modeling in the corrector portion, for example, modeling tissue growth, a hallmark phenomenon of living matter. We thus anticipate that this method can be widely used in various applications, especially medical image analysis.

Appendix

Since the calculation of the deformation gradient is required in the new method, we present an approach for computing it in this section. The approach is the consecutive application of the chain rule that can be implemented concurrently with the NODE calculation. This method can be particularly effective when a small (pseudo) time step is necessary to produce an invertible transformation, as it circumvents the one-time calculation of automatic differentiation for the entire forward Euler method employed in NODE calculation.

For the predictor, one can compute the deformation gradient alongside of (3.3) as follows:

$$\begin{aligned} \mathbf{F}_p^{(i+1)} &= \left(\mathbf{I} + \Delta\tau \nabla_{\mathbf{x}_p^{(i)}} \mathbf{NN}(\mathbf{x}_p^{(i)}; \mathbf{w}_p, \mathbf{b}_p) \right) \mathbf{F}_p^{(i)}, \\ i &= 0, \dots, n_p, \quad \Delta\tau = \frac{1}{n_p}, \\ \mathbf{F}_p^{(0)} &= \mathbf{I}, \end{aligned} \tag{A.1}$$

where $\nabla_{\mathbf{x}_p^{(i)}}$ denotes the gradient of the neural network with respect to $\mathbf{x}_p^{(i)}$, which can be carried out with automatic differentiation. Apparently, the calculation of (A.1) can be carried out concurrently with (3.3).

In the corrector part, a similar recursive equation can be obtained in parallel with (3.6):

$$\begin{aligned} \mathbf{F}_c^{(i+1)} &= \tilde{\mathbf{F}}_c^{(i)} \mathbf{F}_c^{(i)}, \quad i = 0, \dots, n_c, \quad \Delta t = \frac{1}{n_c}, \\ \tilde{\mathbf{F}}_c^{(i)} &= \mathbf{I} + \Delta \tau \nabla_{\mathbf{x}_c^{(i)}} D_p(\mathbf{x}_c^{(i)}; \mathbf{x}_{\mathcal{B}^I}^{p(n)}) \\ &\otimes \mathbf{NN}(\mathbf{x}_p^{(i)}; \mathbf{w}_p, \mathbf{b}_p) + D_p(\mathbf{x}_c^{(i)}; \mathbf{x}_{\mathcal{B}^I}^{p(n)}) \\ &\quad \nabla_{\mathbf{x}_c^{(i)}} \mathbf{NN}(\mathbf{x}_c^{(i)}; \mathbf{w}_c, \mathbf{b}_c), \\ \mathbf{F}_c^{(0)} &= \mathbf{F}_p^{(n_p)}, \quad \tilde{\mathbf{F}}_c^{(0)} = \mathbf{I}, \end{aligned} \quad (\text{A.2})$$

where \otimes stands for the dyadic product.

Appendix B

In this part, the tuning process of the hyperparameter β for the extension of the plate with a hole (see Fig. 2) is shown. To ensure the convergence of the mismatch loss for all values of β , the number of epochs was set to 100,000 for all cases. As can be seen in Fig 11, the large values of β lead to almost zero deformation while the quality of the registration improves with the reduction of β , where good registration is obtained by setting $\beta = \frac{1}{6000}$.

Appendix C

In this section, we additionally present axial, coronal, and sagittal views of the registration results for the three image pairs used to generate the results in Fig. 6-D. As observed, the results of the proposed method are comparable to those of ANTsPy.

Acknowledgements This work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institute of Health, United States under award R01AR074525. This work is also based upon efforts supported by EMBRIO Institute, contract #2120200, a National Science Foundation (NSF) Biology Integration Institute.

Data Availability The code to replicate this manuscript's example is available at https://github.com/AHA-H1987/NODE_Registration.

Declarations

Conflicts of Interest There are no conflicts of interest to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the

source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- “ANTsPy: Python abstraction layer for advanced normalization tools.” Ambrosi, D., Ateshian, G. A., Arruda, E. M., Cowin, S., Dumais, J., Goriely, A., Holzapfel, G. A., Humphrey, J. D., Kemkemer, R., Kuhl, E., et al. (2011). Perspectives on biological growth and remodeling. *Journal of the Mechanics and Physics of Solids*, 59(4), 863–883.
- Avants, B. B., Tustison, N. J., Song, G., & Gee, J. C. (2009). Ants: Open-source tools for normalization and neuroanatomy. *Heanetle*, 10, 1–11.
- Avants, B. B., Tustison, N., Song, G., et al. (2009). Advanced normalization tools (ants). *Insight j*, 2(365), 1–35.
- Bai, J., Jeong, H., Batuwatta-Gamage, C. P., Xiao, S., Wang, Q., Rathnayaka, C., Alzubaidi, L., Liu, G.-R., & Gu, Y. (2023). An introduction to programming physics-informed neural network-based computational solid mechanics. *International Journal of Computational Methods*, 20(10), 2350013.
- Balakrishnan, G., Zhao, A., Sabuncu, M. R., Guttag, J., & Dalca, A. V. (2019). Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8), 1788–1800.
- Beg, M. F., Miller, M. I., Trouvé, A., & Younes, L. (2005). Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International journal of computer vision*, 61, 139–157.
- Behrmann, J., Grathwohl, W., Chen, R. T., Duvenaud, D., & Jacobsen, J.-H. (2019). “Invertible residual networks,” in *International conference on machine learning*, pp. 573–582, PMLR.
- Berg, J., & Nyström, K. (2018). A unified deep artificial neural network approach to partial differential equations in complex geometries. *Neurocomputing*, 317, 28–41.
- Biswas, A., & Shapiro, V. (2004). Approximate distance fields with non-vanishing gradients. *Graphical Models*, 66(3), 133–159.
- Blinkouskaya, Y., & Weickenmeier, J. (2021). Brain shape changes associated with cerebral atrophy in healthy aging and alzheimer’s disease. *Frontiers in Mechanical Engineering*, 7, Article 705653.
- Blinkouskaya, Y., Caçollo, A., Gollamudi, T., Jalalian, S., & Weickenmeier, J. (2021). Brain aging mechanisms with mechanical manifestations. *Mechanisms of ageing and development*, 200, Article 111575.
- Broit, C. (1981). *Optimal registration of deformed images*. University of Pennsylvania.
- Cao, Y., Miller, M. I., Winslow, R. L., & Younes, L. (2005). Large deformation diffeomorphic metric mapping of vector fields. *IEEE transactions on medical imaging*, 24(9), 1216–1230.
- Ceritoglu, C., Oishi, K., Li, X., Chou, M.-C., Younes, L., Albert, M., Lyketos, C., van Zijl, P. C., Miller, M. I., & Mori, S. (2009). Multi-contrast large deformation diffeomorphic metric mapping for diffusion tensor imaging. *Neuroimage*, 47(2), 618–627.
- Ceritoglu, C., Wang, L., Selemón, L. D., Csernansky, J. G., Miller, M. I., & Ratnanather, J. T. (2010). Large deformation diffeomorphic metric mapping registration of reconstructed 3d histological section images and in vivo mr images. *Frontiers in human neuroscience*, 4, 895.

- Chen, R. T., Behrmann, J., Duvenaud, D. K., & Jacobsen, J.-H. (2019). “Residual flows for invertible generative modeling,” *Advances in Neural Information Processing Systems*, vol. 32.
- Chen, R. T., Rubanova, Y., Bettencourt, J., & Duvenaud, D. K. (2018). “Neural ordinary differential equations,” *Advances in neural information processing systems*, vol. 31.
- Christensen, G. E., Rabbitt, R. D., & Miller, M. I. (1994). 3d brain mapping using a deformable neuroanatomy. *Physics in Medicine & Biology*, 39(3), 609.
- Christensen, G. E., Rabbitt, R. D., & Miller, M. I. (1996). Deformable templates using large deformation kinematics. *IEEE transactions on image processing*, 5(10), 1435–1447.
- Christensen, G. E., Geng, X., Kuhl, J. G., Bruss, J., Grabowski, T. J., Pirwani, I. A., Vannier, M. W., Allen, J. S., Damasio, H. (2006) “Introduction to the non-rigid image registration evaluation project (nirep),” in *Biomedical Image Registration: Third International Workshop, WBIR 2006, Utrecht, The Netherlands, July 9–11, 2006. Proceedings* 3, pp. 128–135, Springer.
- De Vos, B. D., Berendsen, F. F., Viergever, M. A., Staring, M., & Işgum, I. (2017). “End-to-end unsupervised deformable image registration with a convolutional neural network,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings* 3, pp. 204–212, Springer.
- Dinh, L., Krueger, D., & Bengio, Y. (2014). “Nice: Non-linear independent components estimation,” *arXiv preprint arXiv:1410.8516*.
- Dinh, L., Sohl-Dickstein, J., & Bengio, S. (2016). “Density estimation using real nvp,” *arXiv preprint arXiv:1605.08803*.
- Eskandari, M., & Kuhl, E. (2015). Systems biology and mechanics of growth. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 7(6), 401–412.
- Göktepe, S., & Kuhl, E. (2009). Computational modeling of cardiac electrophysiology: a novel finite element approach. *International journal for numerical methods in engineering*, 79(2), 156–178.
- Gholipour, A., Rollins, C. K., Velasco-Annis, C., Ouaalam, A., Akhondi-Asl, A., Afacan, O., Ortinau, C. M., Clancy, S., Limperopoulos, C., Yang, E., et al. (2017). A normative spatiotemporal mri atlas of the fetal brain for automatic segmentation and analysis of early brain growth. *Scientific reports*, 7(1), 476.
- Haskins, G., Kruger, U., & Yan, P. (2020). Deep learning in medical image registration: a survey. *Machine Vision and Applications*, 31, 1–18.
- He, J., Koric, S., Abueidda, D., Najafi, A., Jasiuk, I. (2024). “Geom-deepnet: A point-cloud-based deep operator network for field predictions on 3d parameterized geometries,” *arXiv preprint arXiv:2403.14788*.
- He, J., Kushwaha, S., Park, J., Koric, S., Abueidda, D., & Jasiuk, I. (2024). Sequential deep operator networks (s-deepnet) for predicting full-field solutions under time-dependent loads. *Engineering Applications of Artificial Intelligence*, 127, Article 107258.
- Hernandez, M., Bossa, M. N., & Olmos, S. (2009). Registration of anatomical images using paths of diffeomorphisms parameterized with stationary vector field flows. *International Journal of Computer Vision*, 85, 291–306.
- Himpel, G., Kuhl, E., Menzel, A., Steinmann, P., et al. (2005). Computational modelling of isotropic multiplicative growth. *Comp Mod Eng Sci*, 8, 119–134.
- Holzapfel, G. A., & Ogden, R. W. (2009). Constitutive modelling of passive myocardium: a structurally based framework for material characterization. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1902), 3445–3475.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feed-forward networks are universal approximators. *Neural networks*, 2(5), 359–366.
- Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P. J., Whitwell, J. L., Ward, C., et al. (2008). The alzheimer’s disease neuroimaging initiative (adni): Mri methods. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 27(4), 685–691.
- Jagtap, A. D., & Karniadakis, G. E. (2021). “Extended physics-informed neural networks (xpinns): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations,” in *AAAI spring symposium: MLPS*, vol. 10.
- Joshi, S. C., & Miller, M. I. (2000). Landmark matching via large deformation diffeomorphisms. *IEEE transactions on image processing*, 9(8), 1357–1370.
- Kharazmi, E., Zhang, Z., & Karniadakis, G. E. (2021). hp-vpinns: Variational physics-informed neural networks with domain decomposition. *Computer Methods in Applied Mechanics and Engineering*, 374, Article 113547.
- Kharazmi, E., Zhang, Z., & Karniadakis, G. E. (2019). “Variational physics-informed neural networks for solving partial differential equations,” *arXiv preprint arXiv:1912.00873*.
- 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. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *Neuroimage*, 46(3), 786–802.
- Kuang, D., Schmah, T. (2019). “Faim—a convnet method for unsupervised 3d medical image registration,” in *Machine Learning in Medical Imaging: 10th International Workshop, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings* 10, pp. 646–654, Springer.
- López, P. A., Mella, H., Uribe, S., Hurtado, D. E., & Costabal, F. S. (2023). Warppinn: Cine-mr image registration with physics-informed neural networks. *Medical Image Analysis*, 89, Article 102925.
- Lagaris, I. E., Likas, A. C., & Papageorgiou, D. G. (2000). Neural-network methods for boundary value problems with irregular boundaries. *IEEE Transactions on Neural Networks*, 11(5), 1041–1049.
- Lagaris, I. E., Likas, A., & Fotiadis, D. I. (1997). Artificial neural network methods in quantum mechanics. *Computer Physics Communications*, 104(1–3), 1–14.
- Lagaris, I. E., Likas, A., & Fotiadis, D. I. (1998). Artificial neural networks for solving ordinary and partial differential equations. *IEEE transactions on neural networks*, 9(5), 987–1000.
- Limbert, G. (2014). “State-of-the-art constitutive models of skin biomechanics,” *Computational biophysics of the skin*, vol. 558.
- Lombaert, H., Grady, L., Pennec, X., Ayache, N., & Cheriet, F. (2014). Spectral log-demons: diffeomorphic image registration with very large deformations. *International journal of computer vision*, 107, 254–271.
- Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2007). Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience*, 19(9), 1498–1507.
- McFall, K. S., & Mahan, J. R. (2009). Artificial neural network method for solution of boundary value problems with exact satisfaction of arbitrary boundary conditions. *IEEE Transactions on Neural Networks*, 20(8), 1221–1233.
- McFall, K. S. (2006). *An artificial neural network method for solving boundary value problems with arbitrary irregular boundaries*. Georgia Institute of Technology.

- Mihai, L. A., Chin, L., Janmey, P. A., & Goriely, A. (2015). A comparison of hyperelastic constitutive models applicable to brain and fat tissues. *Journal of The Royal Society Interface*, 12(110), 20150486.
- Miller, M. I., Trouvé, A., & Younes, L. (2002). On the metrics and euler-lagrange equations of computational anatomy. *Annual review of biomedical engineering*, 4(1), 375–405.
- Min, Z., Baum, Z. M., Saeed, S. U., Emberton, M., Barratt, D. C., Taylor, Z. A., & Hu, Y. (2023). “Non-rigid medical image registration using physics-informed neural networks,” in *International Conference on Information Processing in Medical Imaging*, pp. 601–613, Springer.
- Mok, T.C., & Chung, A. C. (2020). “Large deformation diffeomorphic image registration with laplacian pyramid networks,” in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part III* 23, pp. 211–221, Springer.
- Oishi, K., Faria, A., Jiang, H., Li, X., Akhter, K., Zhang, J., Hsu, J. T., Miller, M. I., van Zijl, P. C., Albert, M., et al. (2009). Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: application to normal elderly and alzheimer’s disease participants. *Neuroimage*, 46(2), 486–499.
- Ou, Y., Akbari, H., Bilello, M., Da, X., & Davatzikos, C. (2014). Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights. *IEEE transactions on medical imaging*, 33(10), 2039–2065.
- Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., & Lakshminarayanan, B. (2021). Normalizing flows for probabilistic modeling and inference. *The Journal of Machine Learning Research*, 22(1), 2617–2680.
- Pawar, A., Li, L., Gosain, A. K., Umulis, D. M., & Tepole, A. B. (2022). PDE-constrained shape registration to characterize biological growth and morphogenesis from imaging data. *Engineering with Computers*, 38(5), 3909–3924.
- Pawar, A., Zhang, Y., Jia, Y., Wei, X., Rabczuk, T., Chan, C. L., & Anitescu, C. (2016). Adaptive fem-based nonrigid image registration using truncated hierarchical b-splines. *Computers & Mathematics with Applications*, 72(8), 2028–2040.
- Pawar, A., Zhang, Y. J., Anitescu, C., Jia, Y., & Rabczuk, T. (2018). “Dthb3d_reg: dynamic truncated hierarchical b-spline based 3d nonrigid image registration,” *Communications in Computational Physics*, vol. 23, no. 3.
- Pennec, X., Stefanescu, R., Arsigny, V., Fillard, P., & Ayache, N. (2005). “Riemannian elasticity: A statistical regularization framework for non-linear registration,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 943–950, Springer.
- Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., Jack, C., Jr., Jagust, W. J., Shaw, L. M., Toga, A. W., et al. (2010). Alzheimer’s disease neuroimaging initiative (adni) clinical characterization. *Neurology*, 74(3), 201–209.
- Pouch, A. M., Wang, H., Takabe, M., Jackson, B. M., Gorman, J. H., & III, Gorman, R. C., Yushkevich, P. A., & Sehgal, C. M. (2014). Fully automatic segmentation of the mitral leaflets in 3d trans-esophageal echocardiographic images using multi-atlas joint label fusion and deformable medial modeling. *Medical image analysis*, 18(1), 118–129.
- Pouch, A. M., Yushkevich, P. A., Aly, A. H., Woltersom, A. H., Okon, E., Aly, A. H., Yushkevich, N., Parameshwaran, S., Wang, J., & Oguz, B. et al., (2020). “Automated meshing of anatomical shapes for deformable medial modeling: Application to the placenta in 3d ultrasound,” in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 1842–1846, IEEE.
- Rabbitt, R. D., Weiss, J. A., Christensen, G. E., & Miller, M. I. (1995). “Mapping of hyperelastic deformable templates using the finite element method,” in *Vision Geometry IV*, vol. 2573, pp. 252–265, SPIE.
- Raiissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378, 686–707.
- Rezaei, S., Harandi, A., Moeineddin, A., Xu, B.-X., & Reese, S. (2022). A mixed formulation for physics-informed neural networks as a potential solver for engineering problems in heterogeneous domains: comparison with finite element method. *Computer Methods in Applied Mechanics and Engineering*, 401, Article 115616.
- Rvachev, V. (1982). “Theory of r-functions and some applications,”
- Shen, Z., Han, X., Xu, Z., & Niethammer, M. (2019). “Networks for joint affine and non-parametric image registration,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4224–4233.
- Sirignano, J., & Spiliopoulos, K. (2018). Dgm: A deep learning algorithm for solving partial differential equations. *Journal of computational physics*, 375, 1339–1364.
- Sotiras, A., Davatzikos, C., & Paragios, N. (2013). Deformable medical image registration: A survey. *IEEE transactions on medical imaging*, 32(7), 1153–1190.
- Sugiyama, M., Sakae-Sawano, A., Iimura, T., Fukami, K., Kitaguchi, T., Kawakami, K., Okamoto, H., Ichi Higashijima, S., & Miyawaki, A. (2009). Illuminating cell-cycle progression in the developing zebrafish embryo. *Proceedings of the National Academy of Sciences*, 106(49), 20812–20817.
- Sukumar, N., & Srivastava, A. (2022). Exact imposition of boundary conditions with distance functions in physics-informed deep neural networks. *Computer Methods in Applied Mechanics and Engineering*, 389, Article 114333.
- Tepole, A. B. (2017). Computational systems mechanobiology of wound healing. *Computer methods in applied mechanics and engineering*, 314, 46–70.
- Toga, A. W., & Thompson, P. M. (2001). The role of image registration in brain mapping. *Image and vision computing*, 19(1–2), 3–24.
- Visser, V. L., Caçoilo, A., Rusinek, H., & Weickenmeier, J. (2023). “Mechanical loading of the ventricular wall as a spatial indicator for periventricular white matter degeneration,” *journal of the mechanical behavior of biomedical materials*, vol. 143, p. 105921.
- Wang, H., Ni, D., & Wang, Y. (2024). “Recursive deformable pyramid network for unsupervised medical image registration,” *IEEE Transactions on Medical Imaging*.
- Wang, S., Demirci, N., & Holland, M. A. (2021). Numerical investigation of biomechanically coupled growth in cortical folding. *Biomechanics and Modeling in Mechanobiology*, 20(2), 555–567.
- Wang, Z., Martin, B., Weickenmeier, J., & Garikipati, K. (2020). “An inverse modelling study on the local volume changes during early growth of the fetal human brain,” *BioRxiv*.
- Weickenmeier, J., Kuhl, E., & Goriely, A. (2018). Multiphysics of pri-onlike diseases: Progression and atrophy. *Physical review letters*, 121(15), Article 158101.
- Yanovsky, I., Le Guyader, C., Lew, A., Toga, A. W., Thompson, P. M., & Vese, L. (2008). “Unbiased volumetric registration via nonlinear elastic regularization,” in *2nd MICCAI workshop on mathematical foundations of computational anatomy*.
- Zou, J., Gao, B., Song, Y., & Qin, J. (2022). A review of deep learning-based deformable medical image registration. *Frontiers in Oncology*, 12, 1047215.