On a Hybrid Joint Segmentation/Multimodal Registration Model via Finite Distortion Mappings and Implicit Neural Representations

Diane Chan Sock Line, Noémie Debroux, Carole Le Guyader, Caroline Petitjean

On a Hybrid Joint Segmentation/Multimodal Registration Model via Finite Distortion Mappings and Implicit Neural Representations*

Diane Chan Sock Line[1], Noémie Debroux[2][0000−0002−5329−5485], Carole Le Guyader[1][0000−0003−2192−6630], and Caroline Petitjean[3][0000−0003−0013−5370]

[1] INSA Rouen Normandie, Normandie Univ, LMI UR 3226, F-76000 Rouen, France
diane.chan_sock_line,carole.le-guyader@insa-rouen.fr
[2] Université Clermont Auvergne, CNRS, SIGMA Clermont, Institut Pascal, France
noemie.debroux@uca.fr
[3] Normandie Univ, Université Le Havre Normandie, INSA Rouen Normandie, Normandie Univ, LITIS UR 4108, F-76000 Rouen, France
caroline.petitjean@univ-rouen.fr

Abstract. Image segmentation and registration are pivotal preliminary steps in image analysis exhibiting a dual nature, especially in a multimodal context where salient features are to be matched. As such, intertwining them in a unified framework reduces uncertainty propagation and yields positive mutual influence. Registration compensates for weak boundary definition and encodes intrinsic topological requirements like preserving contextual relations between objects. In return, accurate segmented structures foster relevant registration, breeding reliable estimations of the deformation pairing the encoded structures. These observations underpin the proposed contribution blending variational techniques (versatility/adaptability/interpretability) and deep-learning-based approaches (more proficient at handling computationally intensive tasks) through coordinate Multi-Layer Perceptron (MLP) with periodic sinusoidal activation functions. More precisely, in a hybrid nonlinear-elasticity-grounded framework, an unsupervised pairwise joint segmentation/registration 2D model is introduced. Working in the class of mappings with finite distortion enables one to guarantee that the engendered deformation is a homeomorphism. Also, combining a directional-total-variation-based term promoting gradient alignment with the segmentation task, viewed as a substitute for classical intensity-based data terms in registration, enlarges the scope of applications to multimodal images. The existence of a minimiser for the optimisation problem constitutes the core of the paper. Eventually, we test and evaluate our model on both synthetic and medical images to exhibit the accuracy and relevance of our model.

* This research was funded by the French Research National Agency ANR via the AAP CE23 MEDISEG ANR project. The third author would like to thank Dr. A. Molchanova for enlightening discussions and for pointing references [15,16].

Keywords: Finite distortion mappings · Homeomorphic image registration · Directional total variation · Implicit Neural Representations · Hybridisation variational methods/deep learning.

1 Introduction and Motivations

1.1 Intertwining segmentation and registration: interest and proposed nested variational/deep learning-based framework

Segmentation ([1, Chapter 4]) aims to reflect the proficiency of human beings to pinpoint relevant constituents for quantitative and visual purposes. Meanwhile, deformable image registration ([14]) intends to align salient structures visible in an image R —called Reference —into their counterpart in another one T —named Template —through a homeomorphic mapping φ mimicking the true physical transformation involved. As relevant structure pairing and intensity homogeneity criteria dictate registration, incorporating both tasks in a single optimisation formulation seems pertinent. Registration can be seen as the inclusion of topological priors in the segmentation process. Reciprocally, well-segmented elements drive the registration process, providing some nonlocal geometrical information (transfer of edges, etc.). A large body of literature exemplifies the interest in joint segmentation/registration frameworks, ranging from purely variational approaches ([6], [7], [21], [23]) to deep-learning-inspired strategies ([9], [13], [25], [18]). By contrast, the novelty of our proposed work resides in the design of a unified loss function addressing both segmentation and registration in a hybrid setting and relying on the theory of finite distortion mappings. The segmentation task is achieved by the variational component whilst an unsupervised learned-based approach via coordinate MLP for pairwise registration estimates the expected mapping pairing R and T. It thus leverages the strengths of each paradigm variational/learned-based. For the former formalism: (i) increased versatility and adaptability in the mathematical phrasing (in practice, any segmentation tool could be plugged into the learning-based registration component without entailing substantial modifications), (ii) higher predictability of the parameter effect and subsequently interpretability. For the latter one: (iii) adjustability for the deformation regularisation with a high degree of freedom to represent it, and subsequently, ability to learn intricate deformations encoding global and more local variations. Only the pair of images to be aligned is required as input (no pre-training) without relying on (pseudo-)ground-truth deformations, (iv) proficiency to execute computationally intensive tasks.

1.2 Design of a suitable loss function handling multimodality

Several criteria motivate the design of our unifying loss function. As for the segmentation part, the underlying premise is that the geometrical structure of $T \circ \varphi$ resembles that of R, exhibiting a shared prescribed number L of homogeneous regions with thus possibly distinct intensities due to different modalities. This

goal is achieved through a piecewise-constant-Mumford-Shah-like term ([17]) enforcing a single partitioning $u = (u_l)_{1 \leq l \leq L}$ of $T \circ \varphi$ and R, and enlarging thus the scope to possibly multimodal imaging. This segmentation part can be thought of as a substitute for standard data fidelity terms in registration. To process multimodal images and to promote geometrical structure alignment, it is complemented by a term involving first-order information, namely the directional total variation $DTV(T \circ \varphi, M)$ (see [8] and [**?**] for related works), fostering the alignment of ∇R and $\nabla(T \circ \varphi)$ (same orientation but with possibly opposite directions if the contrasts are reversed).[4] As highlighted in Section 2, M denotes a field of symmetric positive definite matrices depending on the anisotropy degree of R. Also, physical and theoretical requirements rule the selection of an accurate model to represent the deformations: smoothness, mechanical plausibility of the resulting deformations faithfully reflecting the natural behaviour and prohibiting matter self-penetration, and theoretical guarantees. The framework of finite distortion mappings ([19]) meets these conditions, ensuring that a homeomorphic deformation is engendered with mild assumptions on φ unlike prior works by Ball [2] requiring that $\varphi \in W^{1,p}(\Omega, \mathbb{R}^2)$ with $p > n(= \text{dimension}) = 2$.

1.3 Structure of the paper and contributions

The overall structure of the remainder of the paper takes the form of two sections. Section 2 zooms in on the mathematical modelling yielding a non-smooth non-convex optimisation problem solved in an alternating scheme. It then elaborates on the mathematical soundness of the modelling, providing the main theoretical result, *i.e*, a minimiser existence result, ensuring, in particular, that the engendered deformation is a homeomorphism. A numerical algorithm is analysed in Section 3, based on a splitting strategy and reflecting the intertwining between the variational component and the learned-based one. Preliminary experiments are then provided, focusing on three main topics: qualitative and quantitative analysis of the results (mechanically meaningful deformations, compliance with real-life scenarios, etc.), ability to generate large deformations, and capability to process multimodal images. Our contribution thus takes on several forms: (i) of a methodological nature through the design of a melded variational and learned-based framework for joint segmentation and multimodal registration, (ii) of a theoretical nature by placing ourselves in the finite distortion mapping setting (not explored so far to the best of our knowledge), thus requiring weakened theoretical assumptions on φ but ensuring nevertheless its homeomorphic character, (iii) of a numerical nature, by developing a nested algorithm involving Chambolle-Pock algorithm ([3]) and an Implicit Neural Representation ([24]) approach. We end this section by emphasising that the paper focuses on the methodological and mathematical aspects. The numerical experiments are preliminary but exemplify the feasibility and relevance of this hybrid approach.

[4] In our terminology, gradients with opposite signs share the same orientation but point in opposite directions.

2 Mathematical Modelling and Main Theoretical Result

2.1 Notations

Let Ω be a connected bounded open subset of $\mathbb{R}^2$ of class $\mathcal{C}^1$. Let $\varphi : \bar{\Omega} \to \mathbb{R}^2$ be the unknown deformation pairing R and T. The deformation gradient is $\nabla\varphi : \bar{\Omega} \to M_2(\mathbb{R})$, with $M_2(\mathbb{R})$, the set of 2×2 real matrices. We denote by $A : B = \text{tr}\left(A^T B\right)$, the classical inner product and by $\|A\| = \sqrt{A : A}$, the related Frobenius norm. We define in a standard way, the functional space $\mathcal{C}_c^\infty(\Omega)$, space of smooth functions with compact support, the Lebesgue spaces $L^p(\Omega)$ and $L^p_{\text{loc}}(\Omega)$ of integrable functions, and the Sobolev spaces $W^{1,p}(\Omega)$ and $W^{1,p}_{\text{loc}}(\Omega)$, $1 \leq p \leq \infty$. The homogeneous Sobolev space $L^p_1(\Omega)$ consists of locally integrable functions $f : \Omega \to \mathbb{R}$ having generalised gradient ∇f and finite semi-norm $\|\nabla f\|_{L^p(\Omega)}$. At last, $f * g$ denotes convolution, while $|\cdot|$ may refer to the Euclidean or operator norm. The loss function combines a segmentation-dedicated part and a registration counterpart and gathers three main terms discussed now.

2.2 Design of a loss function blending segmentation and registration

Segmentation component We postulate that the geometrical structure of R resembles that of $T \circ \varphi$, whether they are of the same modality or not, and that they display L (prescribed number) homogeneous regions with respective intensities $a = (a_l)_{1 \leq l \leq L} \in \mathbb{R}^L$ and $b = (b_l)_{1 \leq l \leq L} \in \mathbb{R}^L$. To reach this objective, a piecewise-constant-Mumford-Shah-like term ([17]) enforcing a single partitioning $u = (u_l)_{1 \leq l \leq L}$ of $T \circ \varphi$ and R is introduced. It reads:

$$\begin{aligned}\mathcal{E}_{R,T}(u, \varphi, a, b) &= \tfrac{1}{2}\sum_{l=1}^{L} \int_\Omega u_l\,(a_l - R)^2\,dx + \tfrac{1}{2}\sum_{l=1}^{L} \int_\Omega u_l\,(b_l - T\circ\varphi)^2\,dx + i_{\mathbb{A}}(u) + \gamma \sum_{l=1}^{L} TV(u_l),\\ &= \mathcal{D}_{R,T}(u, \varphi, a, b) + i_{\mathbb{A}}(u) + \mathcal{Q}(u),\end{aligned}$$

with $\gamma > 0$, a tuning parameter, and $\mathbb{A}$, the admissible set for u defined by

$$\mathbb{A} = \left\{ u = (u_l)_{1 \leq l \leq L} \in BV(\Omega, [0,1])^L \,|\, \sum_{l=1}^{L} u_l = 1 \text{ a.e. on } \Omega \right\}.$$

Registration component: data attachment term To enlarge the scope of the model to potentially multimodal image pairs, a first-order constituent, namely a directional-total-variation-based element denoted by $DTV(T \circ \varphi, M)$, complements the segmentation component. It aims to exploit the similar location of R and $T \circ \varphi$ edges and to favour parallel level sets, using a smooth weighting field M identified to a set of symmetric positive definite matrices spatially varying. For theoretical purposes, $M(\cdot)^{-1}$ is assumed to be smooth. In other words, the $DTV(T \circ \varphi, M)$ term leverages the alignment of ∇R and $\nabla(T \circ \varphi)$. The considered directional total variation is defined as ([**?**])

$$DTV(f, M) = \sup\left\{ \int_\Omega f\,\text{div}(M\Psi)\,dx \,|\, \Psi \in \mathcal{C}^1_c(\Omega, \mathbb{R}^2),\, |\Psi(x)|_2 \leq 1 \text{ for all } x \in \Omega \right\},$$

while the normed space of $\mathbb{R}$-valued functions of bounded directional total variation with respect to the field M reads

$$BDTV(\Omega, M) = BDTV(\Omega, M, \mathbb{R}) = \left\{ f \in L^1(\Omega, \mathbb{R}) \,|\, DTV(f, M) < \infty \right\},$$

$$\|f\|_{BDTV(\Omega,M)} = \|f\|_{L^1(\Omega,\mathbb{R})} + DTV(f, M).$$

The field of matrices M involved in the directional total variation is built upon the amount of anisotropy of R, following [22]. The structure tensor (symmetric matrix) $J_\rho(\nabla R_\sigma) = K_\rho * \left(\nabla R_\sigma \, \nabla R_\sigma^T\right)$ is first computed with K_ρ, a Gaussian kernel, and $R_\sigma = K_\sigma * R$. Let then $\mu_1 \geq \mu_2$ be its ordered eigenvalues and let $\{w_1, w_2\}$ denote the corresponding orthonormal set of eigenvectors. The eigenvector $w_2(x)$ encodes the local direction of anisotropy and corresponds to an approximation of $\nabla^\perp R/|\nabla R|$. A measure reflecting the amount of anisotropy is defined by $\kappa = (\mu_1 - \mu_2)^2$. When large, expressing that the eigenvalues strongly differ, the local anisotropy is high, while it tends to zero when R exhibits isotropic structures. To promote the alignment of ∇R and $\nabla (T \circ \varphi)$, the field $M = M(x)$ is thus constructed as follows. It has $w_1(x)$ and $w_2(x)$ as eigenvectors, whilst the eigenvalues are defined by $\mu_1 = \alpha \in (0, 1)$ and $\mu_2 = \begin{vmatrix} \alpha & \text{if } \kappa = 0 \\ \alpha + (1-\alpha)\exp\left(-\frac{C}{\kappa}\right) & \text{otherwise} \end{vmatrix}$, $C > 0$ serving as a threshold parameter. Note that a structure theorem is available for $BDTV(\Omega, M)$ derived from [10, Theorem 1, Section 5.1] and that $DTV(\cdot, M)$ exhibits fine properties such as lower-semicontinuity in $L^1(\Omega, \mathbb{R})$.

Registration component: regularisation of $\boldsymbol{\varphi}$ We place ourselves within the framework of mappings with finite distortion in nonlinear elasticity ([15,16], [19]): on the one hand, for the biological/clinical plausibility of the deformations generated and, on the other, for theoretical reasons. Indeed, in comparison to Ball's results, we relax summability conditions for admissible deformations to $W^{1,2}(\Omega, \mathbb{R}^2)$ ($W^{1,p}(\Omega, \mathbb{R}^2)$ with $p > n = 2$ the dimension in [2]). To compensate for that, we require the condition $\frac{\|\nabla\varphi\|^2}{\det(\nabla\varphi)} \leq G \in L^s(\Omega)$, $s > 1$ on the characteristic of distortion (G given), still ensuring, in the end, that the produced deformations are homeomorphisms. Following [12], Ω being an open connected set in $\mathbb{R}^n$ and considering a mapping $W^{1,n}_{\text{loc}}(\Omega, \mathbb{R}^n) \ni f : \Omega \to \mathbb{R}^n$ with non-negative Jacobian, $\det(\nabla f(x)) \geq 0$ almost everywhere, one says that f has finite dilatation or distortion, written $f \in FD(\Omega, \mathbb{R}^n)$, whenever

$$|\nabla f(x)|^n \leq K(x) \det(\nabla f(x)) \text{ for almost all } x \in \Omega,$$

where $1 \leq K(x) < \infty$ for almost every $x \in \Omega$ and $|\nabla f(x)|$ denotes the norm of the differential[5] $\nabla f(x) : \mathbb{R}^n \to \mathbb{R}^n$. Basically, it means that $\nabla f(x) = 0$ for almost all $x \in \Omega$ such that $\det(\nabla f(x)) = 0$. The distortion coefficient at the

[5] The norm equivalence in finite dimension enables one to work alternatively with the Frobenius norm.

points $x \in \Omega$ with $\det(\nabla f(x)) \neq 0$ is defined by

$$K(x,f) = \frac{|\nabla f(x)|^n}{\det(\nabla f(x))} \geq 1.$$

If $\det(\nabla f(x)) = 0$, then $\nabla f(x) = 0$, and in this case, we set $K(x,f) = 1$ a.e. Therefore, the distortion function $K(\cdot,f) : \Omega \to [1,+\infty)$ is defined almost everywhere in Ω. A generalisation of distortion coefficients more conducive to handling composition operators may be considered. For a mapping $f : \Omega \to \Omega'$ of class $W^{1,1}_{\text{loc}}(\Omega)$, one defines the distortion operator function as $K_{f,p}(x) = \begin{cases} \dfrac{|\nabla f(x)|}{|\det(\nabla f(x))|^{\frac{1}{p}}} & \text{for} \quad x \in \Omega \setminus Z \\ 0 & \text{otherwise} \end{cases}$, where Z is a zero set of the Jacobian $\det(\nabla f(x))$. Note that $K(x,f) = K^n_{f,n}(x)$ if $x \in \Omega \setminus Z$. Hence $K(\cdot,f) \in L^s(\Omega)$ yields $K_{f,n} \in L^{ns}(\Omega)$, while $\|K(\cdot,f)\|_{L^s(\Omega)} \leq \|G\|_{L^s(\Omega)}$ leads to $\|K_{f,n}(\cdot)\|_{L^{ns}(\Omega)} \leq \|G\|^{\frac{1}{n}}_{L^s(\Omega)}$. Equipped with this material, the proposed regularisation on φ, penalising changes in length and area via weights $a_1 > 0$ and $a_2 > 0$, is chosen as

$$\mathcal{R}(\varphi) = \int_\Omega W(\nabla\varphi)\,dx = \int_\Omega \left[a_1 \|\nabla\varphi\|^2 + a_2\,(\det(\nabla\varphi) - 1)^2\right] dx,$$

over the admissible functional space

$$\mathcal{A} = \big\{\varphi \in W^{1,2}(\Omega,\mathbb{R}^2) \cap FD(\Omega,\mathbb{R}^2),\, K(\cdot,\varphi) \leq G \in L^s(\Omega),\, s > 1,$$

$$\varphi_{|\partial\Omega} = \mathrm{Id}_{|\partial\Omega} \text{ a.e. on } \partial\Omega,\, \det(\nabla\varphi(x)) \geq 0 \text{ a.e. in } \Omega\big\}.$$

Overall loss function Aggregating those components yields the following unified loss function $\mathcal{L}_{R,T}$ and the related optimisation problem

$$\inf_{(u,\varphi,a,b)} \mathcal{L}_{R,T}(u,\varphi,a,b) = \mathcal{D}_{R,T}(u,\varphi,a,b) + i_{\mathbb{A}}(u) + \mathcal{Q}(u) + \mu\, DTV(T\circ\varphi, M) + \mathcal{R}(\varphi). \qquad (\mathcal{P})$$

It is straightforwardly observed that problem $(\mathcal{P})$ is equivalent to the reduced optimisation problem $(\overline{\mathcal{P}})$

$$\inf_{(u,\varphi)} \left\{\overline{\mathcal{L}_{R,T}}(u,\varphi) = \inf_{(a,b)} [\mathcal{D}_{R,T}(u,\varphi,a,b)] + i_{\mathbb{A}}(u) + \mathcal{Q}(u) + \mu\, DTV(T\circ\varphi, M) + \mathcal{R}(\varphi)\right\}, \quad (\overline{\mathcal{P}})$$

so that we eventually theoretically [6] tackle the problem $(\overline{\mathcal{P}})$ with

$$a_l = \left(\int_\Omega u_l R\,dx\right) / \left(\int_\Omega u_l\,dx\right) \text{ (0 if } u_l = 0) \qquad (\mathcal{C})$$
$$\text{and } b_l = \left(\int_\Omega u_l T\circ\varphi\,dx\right) / \left(\int_\Omega u_l\,dx\right) \text{ (0 if } u_l = 0).$$

[6] Functional $\mathcal{L}_{R,T}(\cdot,\varphi,\cdot,\cdot)$ is biconvex —convex in the variable u as well as in the variable (a,b) but not jointly convex —. Nevertheless, problem $(\mathcal{P})$ is the one implemented in practice through a splitting strategy.

2.3 Main result

Theorem 1. *Existence of minimisers*

Let us assume that $R \in L^\infty(\Omega)$ and $T \in W^{1,m}(\Omega)$ with $m > 2$ [7]*. Then problem $(\overline{\mathcal{P}})$ admits minimisers (u_0, φ_0) on $\mathbb{A} \times \mathcal{A}$. Moreover, $\varphi_0 : \bar{\Omega} \to \bar{\Omega}$ is a homeomorphism.*

Proof. The proof is motivated by [15,16]. Due to the limited number of pages, we restrict ourselves to emphasising the main stages and focus primarily on the φ-depending part. Classical arguments and Fatou's lemma enable one to handle the part in u. Functional $\overline{\mathcal{L}_{R,T}}$ is proper and the infimum is finite. Let $((u_k)_k, (\varphi_k)_k)$ be a minimising sequence of $(\overline{\mathcal{P}})$. Using the fact that $(a-b)^2 \geq \frac{1}{2}a^2 - b^2$, a standard coercivity inequality of the type $\overline{\mathcal{L}_{R,T}}(u,\varphi) \geq a_1 \|\nabla\varphi\|^2_{L^2(\Omega,M_2(\mathbb{R}))} + \frac{a_2}{2}\|\det(\nabla\varphi)\|^2_{L^2(\Omega)} - a_2\,\mathrm{meas}(\Omega)$ can be established, which enables one to conclude (invoking the generalised Poincaré inequality) that there exist a subsequence still denoted by $(\varphi_k)_k$ and $\varphi_0 \in W^{1,2}(\Omega,\mathbb{R}^2)$ such that $\varphi_k \underset{k\to+\infty}{\rightharpoonup} \varphi_0$ weakly in $W^{1,2}(\Omega,\mathbb{R}^2)$ and $\det(\nabla\varphi_k) \underset{k\to+\infty}{\rightharpoonup} \det(\nabla\varphi_0)$ weakly in $L^2(\Omega)$ ([5, Theorem 1.14]). By continuity of the trace, $\varphi_{0_{|\partial\Omega}} = \mathrm{Id}_{|\partial\Omega}$.

• **$\det(\nabla\boldsymbol{\varphi_0}) \geq 0$ a.e.** By definition, $\forall k \in \mathbb{N}$, $\det(\nabla\varphi_k) \geq 0$ a.e. on Ω and $\det(\nabla\varphi_k) \underset{k\to+\infty}{\rightharpoonup} \det(\nabla\varphi_0)$ weakly in $L^2(\Omega)$, which means that $\forall\phi \in L^2(\Omega)$, $\int_\Omega \det(\nabla\varphi_k)\ \phi\,dx \underset{k\to+\infty}{\longrightarrow} \int_\Omega \det(\nabla\varphi_0)\ \phi\,dx$. We aim to prove that $\det(\nabla\varphi_0) \geq 0$ a.e., *i.e.*, $\mathrm{meas}(\{x \mid \det(\nabla\varphi_0) < 0\}) = 0$, the set $\{x \mid \det(\nabla\varphi_0) < 0\}$ being measurable since $\det(\nabla\varphi_0)$ is measurable. Assume that the set $\{x \mid \det(\nabla\varphi_0) < 0\}$ has positive measure. Let us take $\phi = 1_{\{\det(\nabla\varphi_0)<0\}}$ in the previous relation. Then $\underbrace{\int_{\{\det(\nabla\varphi_0)<0\}} \det(\nabla\varphi_k)\ dx}_{\geq 0} \underset{k\to+\infty}{\longrightarrow} \int_{\{\det(\nabla\varphi_0)<0\}} \det(\nabla\varphi_0)\,dx$, which raises a contradiction and achieves this element of the proof.

• **$\boldsymbol{\varphi_k} : \bar{\Omega} \to \bar{\Omega}$ homeomorphism and $\boldsymbol{\Psi_k = \varphi_k^{-1}}$, uniformly bounded in $\boldsymbol{W^{1,2}(\Omega,\mathbb{R}^2)}$**: $(\varphi_k)_k \in \mathcal{A}$ and as $\mathrm{Id} : \bar{\Omega} \to \bar{\Omega}$ is a homeomorphism, according to [15, Lemma 4.2], so are the φ_k*'s*: $\bar{\Omega} \to \bar{\Omega}$ —we denote by Ψ_k the inverse mapping $\Psi_k = \varphi_k^{-1}$—. In particular, $\forall k \in \mathbb{N}$, φ_k induces a bounded composition operator ([15, Theorem 3.9]) $\varphi_k^* : L_1^2(\Omega) \to L_1^q(\Omega)$ with $q = \frac{2s}{s+1}$ and where $\varphi_k^*(f) = f\circ\varphi_k$ for $f \in L_1^2(\Omega)$. Furthermore, the estimate $\|\varphi_k^*\| \leq \|K_{\varphi_k,2}\|_{L^{2s}(\Omega)} \leq C\,\|\varphi_k^*\|$ holds, for some constant C. Also, according to [15, Theorem 3.10], $\forall k \in \mathbb{N}$, the inverse mapping Ψ_k (uniformly bounded w.r.t. the norm $\|\cdot\|_{\mathcal{C}^0(\bar{\Omega},\mathbb{R}^2)}$ since Ω is bounded) induces a bounded composition operator $\Psi_k^* : L_1^{q'}(\Omega) \to L_1^2(\Omega)$, where $q' = \frac{2s}{s-1}$, and has finite distortion. Moreover, it holds that $\forall k \in \mathbb{N}$, $\|\Psi_k^*\| = \|\left(\varphi_k^{-1}\right)^*\| \leq \|K_{\varphi_k^{-1},q'}\|_{L^\rho(\Omega)} \leq \|K_{\varphi_k,2}\|_{L^{2s}(\Omega)}$, with $\rho = 2s$. ($\Psi_k^*(f) = f\circ\Psi_k$ for $f \in L_1^{q'}(\Omega)$ and with $\Psi_k = \left(\Psi_k^1, \Psi_k^2\right)$). Taking successively the coordinate functions for f —denoted by f_i thereafter —, $i \in \{1,2\}$, it yields $\|\Psi_k^i\|_{L_1^2(\Omega)} =$

[7] This choice of functional space is motivated in the proof.

$\|\Psi_k^*(f_i)\|_{L_1^2(\Omega)} = \|\nabla\Psi_k^i\|_{L^2(\Omega)} \leq \|\Psi_k^*\|\|f_i\|_{L_1^{q'}(\Omega)} \leq C(i,\Omega,q')\|K_{\varphi_k,2}\|_{L^{2s}(\Omega)} \leq C(\Omega,q')\|G\|_{L^s(\Omega)}^{\frac{1}{2}}$, this latter quantity being uniformly bounded with respect to k. Combining this preliminary result with the generalised Poincaré inequality yields that $\forall k \in \mathbb{N}$, Ψ_k is uniformly bounded in $W^{1,2}(\Omega,\mathbb{R}^2)$.

• $\boldsymbol{T \circ \varphi_k \in BDTV(\Omega, M)}$ At this stage, the assumptions on T can be slackened to $T \in L^\infty(\Omega) \cap W^{1,2}(\Omega)$. According to [11, Theorem 1.3], as $\forall k \in \mathbb{N}$, $\Psi_k \in W^{1,2}(\Omega,\mathbb{R}^2)$ and is of finite dilatation, $\forall k \in \mathbb{N}$, Ψ_k satisfies Lusin's condition.[8] Let $\mathcal{N}$ be such that $\text{meas}(\mathcal{N}) = 0$ and for all $x \in \Omega \setminus \mathcal{N}$, $T(x) \leq c$, with c a positive given constant (according to the assumptions, $T \in L^\infty(\Omega)$). Let $\mathcal{N}_k' = \Psi_k(\mathcal{N})$. Since Ψ_k satisfies Lusin's condition, $\text{meas}(\mathcal{N}_k') = 0$. For every $y \notin \mathcal{N}_k'$, $T \circ \varphi_k(y) = T(x)$ with $x \in \Omega \setminus \mathcal{N}$. Consequently, for almost every $y \in \Omega$, $T \circ \varphi_k(y) \leq c$, yielding the uniform boundedness of $\|T \circ \varphi_k\|_{L^1(\Omega)}$. In addition, as previously highlighted, φ_k induces a bounded composition operator

$$\varphi_k^* : L_1^2(\Omega) \to L_1^q(\Omega)$$

with $q = \frac{2s}{s+1}$ and where $\varphi_k^*(f) = f \circ \varphi_k$ for $f \in L_1^2(\Omega)$, showing that $T \circ \varphi_k \in W^{1,1}(\Omega)$ (even in $W^{1,q}(\Omega)$), and is thus an element of $BV(\Omega)$. Let $\Psi \in \mathcal{C}_c^1(\Omega,\mathbb{R}^2)$ be such that $\forall x \in \Omega$, $|\Psi(x)|_2 \leq 1$. Let now $\phi \in \mathcal{C}_c^1(\Omega,\mathbb{R}^2)$ be defined by $\phi(x) = M(x)\Psi(x)$. For every $x \in \Omega$, $|\phi(x)|_2 \leq 1$ since $\mu_1(x) \leq 1$ and by definition of the space $BV(\Omega)$, $\int_\Omega T \circ \varphi_k \,\text{div}\,(\phi)\, dx = \int_\Omega T \circ \varphi_k \,\text{div}\,(M\Psi)\, dx \leq TV(T \circ \varphi_k) < +\infty$, showing that $T \circ \varphi_k \in BDTV(\Omega, M)$ by passing to the supremum over Ψ.

• **Up to a subsequence, $\boldsymbol{(\Psi_k)_k}$ uniformly converges to $\boldsymbol{\Psi_0}$ up to the boundary of Ω.** Based on [19, Chapter III, Section 4, Lemma 4.3, p. 338], the equicontinuity of the family $(\Psi_k)_{k\in\mathbb{N}}$ on any compact subset K of Ω is proved. The equicontinuity of the family $(\Psi_k)_{k\in\mathbb{N}}$ near $\partial\Omega$ can be proved following the steps of [15, Lemma 4.3, p.23]. Then, by applying Arzelà-Ascoli theorem, there exists a subsequence —still denoted by $(\Psi_k)_{k\in\mathbb{N}}$ —and Ψ_0 such that Ψ_k uniformly converges to Ψ_0 on $\bar{\Omega}$.

• **Almost everywhere injectivity of $\boldsymbol{\varphi_0}$ and finite distortion feature** The limit of homeomorphisms needs not be a homeomorphism or even an injective mapping ([16]). We thus question the almost everywhere injectivity property of φ_0 on $\Omega \setminus \varphi_0^{-1}(\partial\Omega)$[9]. Up to a subsequence, φ_k converges pointwise almost everywhere to φ_0. Let S be a measurable set of measure $|S| = 0$ such that $\forall x \in \Omega \setminus S$, $\varphi_k(x) \underset{k\to+\infty}{\longrightarrow} \varphi_0(x)$. As Ψ_k uniformly converges to Ψ_0 on Ω and as $\forall k \in \mathbb{N}, \forall x \in \Omega$, $\Psi_k \circ \varphi_k(x) = x$, by passing to the limit when $k \to +\infty$, it yields, $\forall x \in \Omega \setminus S$, $\lim_{k\to+\infty} \Psi_k \circ \varphi_k(x) = \Psi_0 \circ \varphi_0(x) = x$. Let $(x_1, x_2) \in \Omega \setminus S \times \Omega \setminus S$ be such that $\varphi_0(x_1) = \varphi_0(x_2) \in \Omega$. Then $\Psi_0 \circ \varphi_0(x_1) = \Psi_0 \circ \varphi_0(x_2)$, which entails $x_1 = x_2$, demonstrating the almost everywhere injectivity feature of φ_0 on $\Omega \setminus \varphi_0^{-1}(\partial\Omega)$. The question that remains to be addressed is whether the set of points of Ω such that $\varphi_0(x) \in \partial\Omega$ is negligible. In that purpose, an adaptation of [15, Lemma 4.7] is proved. In addition, [16, Lemma 10] enables one to conclude

[8] $E \subseteq \Omega, \text{meas}(E) = 0 \Rightarrow \text{meas}(\Psi_k(E)) = 0$.

[9] Note that since φ_0 is the pointwise a.e. limit of homeomorphisms φ_k, the images of some points $x \in \Omega$ may lie on the boundary $\partial\Omega$.

that $\det(\nabla\varphi_0) > 0$ a.e..

• **Weak lower semicontinuity of the DTV term.** Using Rellich-Kondrachov theorem, we prove that the mapping φ_0 induces a bounded composition operator $\varphi_0^* : L_1^2(\Omega) \cap W^{1,m}(\Omega) \to L_1^q(\Omega)$ with $m > 2$ and $q = \dfrac{2s}{s+1}$. In particular, $T \circ \varphi_k \in BDTV(\Omega, M)$ strongly converges to $T \circ \varphi_0$ in $L^1(\Omega)$ and $TDV(\cdot, M)$ being lower semi-continuous in $L^1(\Omega, \mathbb{R})$, it yields $DTV(T \circ \varphi_0, M) \leq \liminf_{k\to+\infty} DTV(T \circ \varphi_k, M)$.

To summarise, at this stage, $\varphi_0 \in W^{1,2}(\Omega)$ induces a bounded composition operator $\varphi_0^* : L_1^2(\Omega) \cap W^{1,m}(\Omega) \to L_1^q(\Omega)$, is a.e. injective on Ω and $\det(\nabla\varphi_0) > 0$ a.e.. According to [16, Lemma 1], φ_0 has thus finite distortion.

• **Behaviour of the dilatation coefficient** Due to [15, Lemma 4.13], as $\varphi_k \underset{k\to+\infty}{\rightharpoonup} \varphi_0$ in $W^{1,2}(\Omega)$ (the φ_k's being with finite distortion), as $\det(\nabla\varphi_k) \geq 0$ a.e. in Ω and as there exists $G \in L^s(\Omega)$ with $s > 1$ such that $K_{\varphi_k,2}(x) \leq G(x)^{\frac{1}{2}}$ a.e., then at the limit[10], $K(x, \varphi_0)^{\frac{1}{2}} = K_{\varphi_0,2}(x) \leq G(x)^{\frac{1}{2}}$ for almost every $x \in \Omega$, resulting in the homeomorphic nature of φ_0 ([15, Lemma 4.2]). This closes the study of the functional behaviour in terms of the functions φ_k's, the weak lower semicontinuity of $\overline{\mathcal{L}_{R,T}}$ deriving, among others, from the polyconvexity ([5]) of W.

This last step enables us to conclude that a global minimiser to problem $(\overline{\mathcal{P}})$ exists, which is the main result of the paper.

3 Numerical Algorithm, Experiments and Analysis

3.1 Splitting strategy for solving problem $(\mathcal{P})$

Problem $(\mathcal{P})$ falls into the category of non-smooth-non-convex problems. As classically done, a splitting strategy is applied, relying on an alternating scheme that breaks the problem into two main subproblems: (i) the former is non-smooth and convex (partial problem obtained when one fixes the deformation φ and the constants (a, b)). This subproblem involves solving a generic saddle-point problem for u tackled using the Chambolle-Pock algorithm ([3]); (ii) the latter (partial problem obtained when one fixes u and (a, b)) is nonlinear non-convex, related to the updating of the deformation φ. This step leverages an Implicit Neural Representation approach (INR) ([24]), where a coordinate-based multilayer perceptron (MLP) is optimised to represent the deformation. To capture fine-grained details and subtle variations, we adopt a strategy inspired by SIREN ([20]) employing periodic sinusoidal functions to model high-frequency content. Algorithm 1. summarises the complete numerical resolution strategy. Note that, for the moment, the distortion coefficient is monitored via a soft penalisation incorporated in $\mathcal{L}_{R,T}$ in the form $a_3 \int_\Omega \left(\|\nabla\varphi\|^2/\det(\nabla\varphi)\right)^s dx$ with $s = 2$. Further analyses will be devoted to a numerical scheme more consistent with the (hard) constraints on the distortion coefficient and to the convergence of such an alternating strategy.

[10] since $\text{meas}(\{\det(\nabla\varphi_0)\}) = 0$.

Algorithm 1 Alternating optimisation scheme

Input: Define $k = 1$, image Reference R, image Template T
Initialisation: Define $nbIter$, γ, μ, L, u^0, a^0, b^0, $\varphi^0 = Id$
While $k < nbIter$ **and** $\det(\nabla\varphi) > 0$ **do:**

- update $u^k \leftarrow \text{argmin}_u \mathcal{L}_{R,T}(u, \varphi^{k-1}, a^{k-1}, b^{k-1})$ by employing Chambolle-Pock algorithm,
- update $a^k = (a_l)_{1\leq l\leq L}^k \in \mathbb{R}^L$ and $b^k = (b_l)_{1\leq l\leq L}^k \in \mathbb{R}^L$ as in $(\mathcal{C})$,
- update $\varphi^k \leftarrow \text{argmin}_\varphi \mathcal{L}_{R,T}(u^k, \varphi, a^k, b^k)$ by training the INR network with $\mathcal{L}_{R,T}(u^k, \cdot, a^k, b^k)$ as loss function.

end while
Output: return $T \circ \varphi$, u, a, b

3.2 Implementation

Our method was implemented in PyTorch and computations were performed on an NVIDIA RTX A3000 Laptop GPU. In all our experiments, the registration process was carried out using an INR architecture, as described in [24], which is resolution-independent. As a result, for each image pair processed with the same network hyperparameters —a 3-layer MLP with 256 hidden units and a spectral bias of 32 —, the registration network component was optimised over 250 epochs per temporal iteration, with all points in the image domain treated during each epoch. Optimisation was performed using the Adam optimiser with a fixed learning rate of 10^{-4}. The parameters γ, μ together with the regularisation terms a_1 and a_2, and the soft penalisation term a_3 were manually fine-tuned.

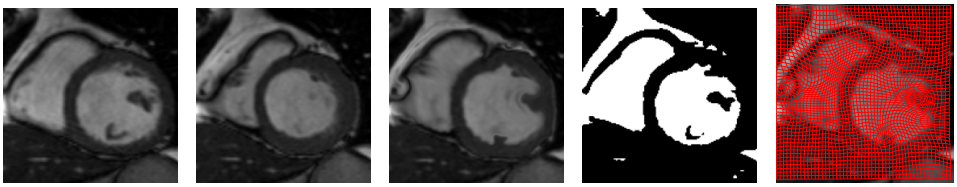
3.3 Numerical experiments and results

We present both synthetic and medical examples to demonstrate the model performance. First, a toy geometric example (200×200), referenced as 1 in Tab. 1, showcases the model ability to handle noisy data and sharp edges/corners and to generate smooth and large deformations (Fig. 1).

(a) R (b) T (c) $T \circ \varphi$ (d) u (e) Deformed Grid

Fig. 1: Mapping of a toy example: $\gamma = 1$, $\mu = 5$, $a_1 = 50$, $a_2 = 1000$, $a_3 = 10$, $\min \det \nabla\varphi = 0.1795$, execution time: $99s$.

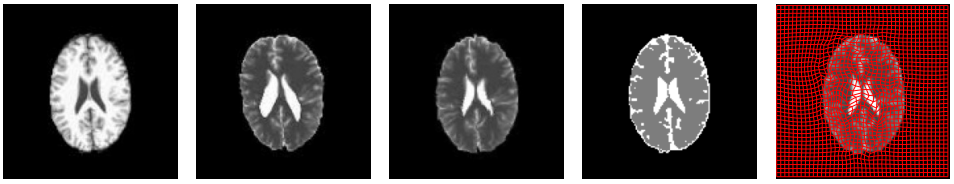
Next, to assess the algorithm accuracy in managing significant deformations while addressing small-scale features, MRI images of a cardiac cycle (150×150) are considered, referenced as 2 in Tab. 1. The reference corresponds to the end of diastole, while the template mirrors the end of systole (Fig. 2).



(a) R (b) T (c) $T \circ \varphi$ (d) u (e) Deformed Grid

Fig. 2: Mapping of a cardiac cycle MRI example $\gamma = 0.001$, $\mu = 5$, $a_1 = 75$, $a_2 = 1000$, $a_3 = 5$, min $\det \nabla\varphi = 0.4499$, execution time: $108s$.

Finally, brain tumour images (128×128) from different modalities, referenced as 3 in Tab. 1, highlight the model capacity to address complex topologies, including thin structures and concavities, while performing multimodal image registration (Fig. 3). In all cases, the deformations remain mechanically



(a) R (b) T (c) $T \circ \varphi$ (d) u (e) Deformed Grid

Fig. 3: Mapping of T1- and T2-Weighted MRI brain tumour example (from BrainWeb [4]) $\gamma = 0.0001$, $\mu = 40$, $a_1 = 50$, $a_2 = 2500$, $a_3 = 10$, min $\det \nabla\varphi = 0.4773$, execution time: $81s$.

meaningful (positivity of the Jacobian determinant). For the first two cases, the results are visually comparable, maybe more relevant than those reported in [6] — purely variational model close in essence to ours and based on nonlocal shape descriptors —, without requiring the regridding step strategy involved in [6] and necessary to preserve topology. In the toy example, our approach achieves a similar registration accuracy to [6], with comparable execution time, though it results in slightly lower segmentation relevance. Nevertheless, in our view, one cannot reduce the model assessment to a mere analysis of metrics, which, for most of them, take on an averaging effect and presuppose the introduction of parameters (threshold for the Dice coefficient, etc.). Visual inspection, combined with the mechanical plausibility of the deformation obtained, the management of fine details, and a good compromise between execution time/computational investment and the absence of sustaining methods like regridding, also seems to us to be essential. For the cardiac MRI images, the model captures fine details within the myocardium, although the Dice coefficient and mutual information

are marginally lower than in [6], but still satisfactory [11] Moreover, our method is 1.5/11 times faster than the counterparts in [6] [12]. In the brain tumour example, visual results demonstrate that the algorithm effectively handles complex topologies, yielding satisfactory outcomes in both segmentation and registration for multimodal image registration (see Tab. 1).

Images	DICE(R,T)	DICE(R,$T \circ \varphi$)	DICE(R,u)	MI(R,T)	MI(R,$T \circ \varphi$)
1	0.76517	**0.97831**,-,0.9783	**0.98068**,-,0.98016	0.51401	0.7035,-,**0.7433**
2	0.54366	0.90884-0.93833-**0.95304**	0.98219-**0.99137**-0.99131	0.93134	1.12444-1.1857-**1.1930**
3	0.73394	0.81306	0.93234	1.0545	1.1168

Table 1: Dice coefficients and Mutual Information. For each metric, the obtained values, when available, are provided in the following order: our method, local method from [6] and nonlocal counterpart from [6]).

References

1. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations. Applied Mathematical Sciences, Springer-Verlag, New York (2001)
2. Ball, J.M.: Global invertibility of Sobolev functions and the interpenetration of matter. P. Roy. Soc. Edin. A **88**(3-4), 315–328 (1981)
3. Chambolle, A., Pock, T.: A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. J. Math. Imaging Vis. **40**(1), 120–145 (2011)
4. Cocosco, C.A., Kollokian, V., Kwan, R.K.S., Evans, A.C.: BrainWeb: Simulated Brain Database (2006), available at `http://www.bic.mni.mcgill.ca/brainweb/`
5. Dacorogna, B.: Direct Methods in the Calculus of Variations, Second Edition. Springer, New York (2008)
6. Debroux, N., Le Guyader, C.: A Joint Segmentation/Registration Model Based on a Nonlocal Characterization of Weighted Total Variation and Nonlocal Shape Descriptors. SIAM J. Imaging Sci. **11**(2), 957–990 (2018)
7. Droske, M., Rumpf, M.: Multiscale joint segmentation and registration of image morphology. IEEE Trans. Pattern Anal. Mach. Intell. **29**(12), 2181–2194 (2007)
8. Ehrhardt, M.J., Betcke, M.M.: Multicontrast MRI Reconstruction with Structure-Guided Total Variation. SIAM J. Imaging Sci. **9**(3), 1084–1106 (2016)
9. Estienne, T., Vakalopoulou, M., Christodoulidis, S., Battistela, E., Lerousseau, M., Carre, A., Klausner, G., Sun, R., Robert, C., Mougiakakou, S., Paragios, N., Deutsch, E.: U-ReSNet: Ultimate Coupling of Registration and Segmentation with Deep Nets. In: MICCAI 2019. pp. 310–319. Springer International Publishing, Cham (2019)
10. Evans, L.C., Gariepy, R.F.: Measure Theory and Fine Properties of Functions. Studies in Advanced Mathematics, Taylor & Francis, N.W., Boca Raton, Florida (1991)
11. Heinonen, J., Koskela, P.: Sobolev mappings with integrable dilatations. Arch. Ration. Mech. Anal. **125**(1), 81–97 (1993)

[11] A Dice coefficient closer to 1 and higher mutual information indicate a more reliable match between two given sets.

[12] To be interpreted with caution since we did not reproduce the experiments from [6] on our own hardware.

12. Iwaniec, T., Sverak, V.: On Mappings with Integrable Dilatation. Proc. Am. Math. Soc. **118**(1), 181–188 (1993)
13. Li, B., Niessen, W.J., Klein, S., Groot, M., Ikram, M.A., Vernooij, M.W., Bron, E.E.: A hybrid deep learning framework for integrated segmentation and registration: evaluation on longitudinal white matter tract changes. In: MICCAI 2019. pp. 645–653 (2019)
14. Modersitzki, J.: FAIR: Flexible Algorithms for Image Registration. Society for Industrial and Applied Mathematics (SIAM), USA (2009)
15. Molchanova, A.O., Vodop'yanov, S.K.: Variational problems of nonlinear elasticity theory in certain classes of mappings with finite distortion (2015)
16. Molchanova, A.O., Vodop'yanov, S.K.: Injectivity almost everywhere and mappings with finite distortion in nonlinear elasticity. Calc. Var. Partial Differ. Equ. **59**(17) (2019)
17. Mumford, D., Shah, J.: Optimal approximation by piecewise smooth functions and associated variational problems. Comm. Pure Appl. Math. **42**(5), 577–685 (1989)
18. Qiu, L., Ren, H.: RSegNet: A Joint Learning Framework for Deformable Registration and Segmentation. IEEE Trans. Autom. Sci. Eng. **19**(3), 2499–2513 (2022)
19. Reshetnyak, Y.G.: Space mappings with bounded distortion, vol. 73. Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, USA (1989)
20. Sitzmann, V., Martel, J.N.P., Bergman, A.W., Lindell, D.B., Wetzstein, G.: Implicit Neural Representations with Periodic Activation Functions. In: NeurIPS 2020. Curran Associates Inc., Red Hook, NY, USA (2020)
21. Swierczynski, P., Papież, B., Schnabel, J.A., Macdonald, C.: A level-set Approach to Joint Image Segmentation and Registration with Application to CT Lung Imaging. Comput. Med. Imaging Graph. **65**, 58–68 (2017)
22. Weickert, J.: Coherence-Enhancing Diffusion Filtering. Int. J. Comput. Vis. **31**(2), 111–127 (1999)
23. Wirth, B.: On the Gamma-limit of joint image segmentation and registration functionals based on phase fields. Interfaces Free Bound. **18**(4), 441–477 (2016)
24. Wolterink, J.M., Zwienenberg, J.C., Brune, C.: Implicit neural representations for deformable image registration. In: MIDL 2022. pp. 1349–1359 (2022)
25. Xu, Z., Niethammer, M.: DeepAtlas: Joint Semi-Supervised Learning of Image Registration and Segmentation. In: MICCAI 2019. pp. 420–429 (2019)