

Supplementary Materials

Alexandre Bône, Paul Vernhet, Olivier Colliot, and Stanley Durrleman

ARAMIS Lab, ICM, Inserm U1127, CNRS UMR 7225, Sorbonne University, Inria,
Paris, France

A Metamorphosis framework

This section aims at providing an introduction to our metamorphosis framework. Although we present the specific case of image metamorphosis, this theory also applies to the larger class of f -shapes (which encompass both shapes and signals on them).

A.1 Continuous theory

Diffeomorphic framework

Common grounds for diffeomorphic registration methods In generic diffeomorphic image registration framework (stationary velocity fields, large deformation diffeomorphic metric mapping, ...), two images can be registered, and thus compared, by assuming that there exists a diffeomorphism φ mapping one into the other. The associated action Φ_φ is given by $\Phi_\varphi \star I = I \circ \varphi$. Common practice is to restrict to specific classes of diffeomorphisms ϕ_1 obtained as integration, over a constant amount of time $[0, 1]$, of the flow of smooth velocity fields, for instance with $v_t \in L^2([0, 1], V)$ where V is a subspace of smooth vanishing fields, *ie* $V \subset C_0^\infty(\Omega, \mathbb{R}^d)$, in the case of LDDMM¹, and $v_t = v$ in the case of SVF². For the former, the flow is therefore given by (1) :

$$\partial_t \phi_t = v_t \circ \phi_t \quad \text{from the identity} \quad \phi_0 = \text{Id}_\Omega. \quad (1)$$

Associating an energy functional - akin to a deformation cost - to a diffeomorphism through the V -norm of its associated infinitesimal velocity field v leads to a geodesic property of optimal path between two images: minimizing diffeomorphisms are exactly geodesic paths. In practice, a relaxed energy functional \mathcal{E} is introduced (as there is no reason for a set of real-world noisy images to be exactly related by a diffeomorphism) and minimized:

$$\mathcal{E}(v) = \int_0^1 \|v_t\|_V^2 dt + \gamma d(\Phi(v) \star I, I') \quad \text{s. t.} \quad \partial_t \phi_t = v_t \circ \phi_t \quad (2)$$

¹ efficient framework especially for large deformations with complex deformation paths.

² fit for smaller deformation fields, as it leverages Lie groups tools and yields powerful theoretical properties and computational algorithms such as the *Baker-Campbell-Hausdorff* formula or *scaling-and-squaring* integration.

where we denote $\Phi(v) = \Phi_{\phi_1}$ the mapping which associates the obtained action diffeomorphism from the vector fields v_t (resp. v), and $d(\cdot, \cdot)$ a data attachment cost such as euclidean mean square error, cross-correlations or optimal transport penalty between I and I' .

Reproducing Kernel Hilbert Space An additional hypothesis is that V is continuously embedded in a reproducing kernel Hilbert space (RKHS) \mathcal{H}_V of reproducing kernel K_V . v then admits dual representation $v^* \in V^*$ called momentum, which verifies:

$$v(\cdot) = \int_{\Omega} K_V(x, \cdot) \cdot v^*(x) \cdot d\Omega(x) \quad (3)$$

The induced norm on V finally writes:

$$\|v\|_V = \iint_{\Omega^2} v^*(x')^\top \cdot K_V(x, x') \cdot v^*(x) \cdot d\Omega(x) \cdot d\Omega'(x') \quad (4)$$

where $(\cdot)^\top$ denotes the transposition operator.

This essential step provides the *reduction* property, which ensures that that the minimizing velocity fields of \mathcal{E} (which are geodesics) are restricted to a finite-dimensional subspace of V , thus reducing optimization to a finite dimensional space.

Distance on images Binding all previous hypotheses together, these frameworks give rise to a distance between images, which boils down to a finite-dimensional optimization problem over specific velocity fields (or their associated momenta). It is in particular possible to extend euclidean statistics to the manifold of images, and define objects such as the (Fréchet) mean of images, *aka* an atlas, which we use in this paper.

Extension to metamorphosis Metamorphosis extends such work by keeping the geometric group action, but replacing diffeomorphic deformations ϕ_t by a transform that acts on both the geometric support (also in a diffeomorphic fashion) and intensities.

Metamorphosis transform

Let $\Omega \subset \mathbb{R}^d$ with $d \in \{2, 3\}$ be a spatial domain, on which is defined the image $I : \Omega \rightarrow \mathbb{R}$ seen as an intensity signal on Ω . We can therefore rewrite the image as $I \equiv (\Omega, I)$ where Ω can be thought of as the geometrical domain (typically a rectangular grid for 2d images). Metamorphosis framework states that this image representation can be acted upon by a mapping Φ such that, with a slight abuse of notation, $\Phi \star I = (\phi_1 \circ \Omega, (I + h_1) \circ \phi_1^{-1})$. Unlike previously, the diffeomorphism ϕ_1 acts on the geometry of the image while an intensity map h_1 - which depends on the geometrical variations of the support - alters the levels of intensity. In particular, this action has the ability to change the initial topology,

for instance by creating holes (by lowering the intensity value) or adding matter (by increasing it).

It can be shown that under regularity assumptions of smooth velocity fields $v \in V$ and intensity maps $\delta \in D$ similar to the previous ones, the action on images is described by a flow of the form (5) :

$$\begin{cases} \partial_t \phi_t = v_t \circ \phi_t \\ \partial_t h_t = \delta_t \\ \phi_0 = \text{Id}_\Omega \\ h_0 = 0 \end{cases} \quad (5)$$

This defines a new family of metrics on images d_m , stemming from the minimization of a corresponding relaxed energy path \mathcal{E}_m under flow assumption (5):

$$\mathcal{E}_m(v, \delta) = \alpha \int_0^1 \|v_t\|_V^2 dt + \beta \int_0^1 \|\delta_t \circ \phi_t^{-1}\|_D^2 dt + \gamma d(\Phi \star I, I') \quad (6)$$

where $(\alpha, \beta, \gamma) \in (\mathbb{R}_+^*)^3$ are weights balancing the contribution of each factor.

A.2 Discretization

We now explain the discretization of previous continuous framework. We considered a one-step approximation of intensity path $h_1 \equiv \delta$, and assumed a SVF framework to model diffeomorphisms, such that the action on images reduces to:

$$\Phi(v, \delta) \star I = \Phi(v) \star (I + \delta) = (I + \delta) \circ \Phi(v) \quad (7)$$

This action can be easily interpreted as the combination of an additive intensity increment δ of the original intensity values I in the reference domain Ω , followed by the more classical deformation action $\Phi(v)$. With this approximation, because the local inversion theorem ensures that $\Phi : V \times D \rightarrow \mathcal{G} = \{\Phi(v, \delta) | v \in V, \delta \in D\}$ is invertible around Id_Ω , there exists a neighbourhood of each image I_0 in the orbit space $\mathcal{G} \star I_0$ (space of all images which are reachable by a metamorphic transform starting from I_0) where we can define a metamorphic distance metric d_{I_0} :

$$d_{I_0}(I, I')^2 = \alpha \|v' - v\|_V^2 + \beta \|\delta' - \delta\|_D^2 \quad \text{where } (v, \delta) = \Phi^{-1}(I), (v', \delta') = \Phi^{-1}(I') \quad (8)$$

As mentioned in the main paper, our MAE outputs momenta \underline{v}^* , $\underline{\delta}^*$ and atlas $\underline{\mathcal{A}}$ discretized on a reference grid g . The discrete velocity field \underline{v} and intensity increment $\underline{\delta}$ are first computed according to the filtering formulae (3) $\underline{v} = \underline{K}_V \cdot \underline{v}^*$ and $\underline{\delta} = \underline{K}_D \cdot \underline{\delta}^*$ where for any grid index k_0 :

$$[\underline{K}_V]_{k_0} = \sum_k \exp\left(\frac{-\|g_k - g_{k_0}\|_{\ell^2}^2}{\rho_V^2}\right) \quad \text{and} \quad [\underline{K}_D]_{k_0} = \sum_k \exp\left(\frac{-\|g_k - g_{k_0}\|_{\ell^2}^2}{\rho_D^2}\right) \quad (9)$$

which corresponds to the discrete version of the reproducing Hilbert norm given in (4), which writes $\|\underline{v}\|_V = (\underline{v}^*)^\top \cdot \underline{K}_V \cdot \underline{v}^*$ (and similarly on D). The choice of Gaussian kernel variances ρ_V^2, ρ_D^2 is driven by the scale of action for each contributions and can be fine-tuned: in our case, we set $\rho_V^2 \geq \rho_D^2$ as deformations are acting on a larger scale than intensities.

The diffeomorphism $\Phi(v) \equiv \phi_1 \circ \Omega$ from (7) is then obtained by integration along the streamlines of v defined by Equation 1, which is discretely carried out with the scaling-and-squaring algorithm which consists in applying T times:

$$\underline{x}_{t+1} = \underline{x}_t + \mathcal{I}(\underline{x}_t - g, \underline{x}_t) \quad \text{from} \quad \underline{x}_0 = g + \underline{v} / 2^T \quad (10)$$

where $\mathcal{I}(\underline{x}_t - g, \underline{x}_t)$ simply denotes the bilinear interpolation \mathcal{I} of the displacement field $\underline{x}_k - g$ at the physical locations \underline{x}_k . The metamorphosis of $\mathcal{A} \equiv (\Omega, \mathcal{A})$ (i.e. Equation 7) is finally approximated as:

$$\Phi(v, \delta) \star \mathcal{A} \approx \Phi(\underline{v}, \underline{\delta}) \star \underline{\mathcal{A}} = \mathcal{I}(\underline{\mathcal{A}} + \underline{\delta}, \underline{x}_T) \quad (11)$$

B Chan-Vese localization from δ maps

BraTs segmentation annotations comprise the GD-enhancing tumor (ET — label 4), the peritumoral edema (ED — label 2), and the necrotic and non-enhancing tumor core (NCR/NET — label 1). We focused on the ET labels, and used the level-sets Chan-Vese segmentation scheme on the intensity increments δ decoded for our T1 contrast-enhanced MRI modality. We estimated hyper-parameters on a small held-out dataset of 32 MRIs, and obtained the following score distribution (Figure 1) on remaining data. Although not sufficient for immediate segmentation of tumors, this supports the hypothesis of an intrinsic disentanglement of intensity information (in δ) and geometric knowledge (in v).

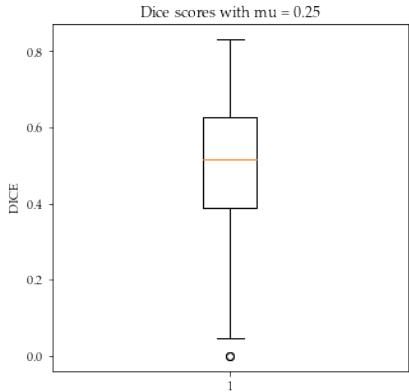


Fig. 1: Boxplot of DICE scores