On Finite Difference Jacobian Computation in Deformable Image Registration

Yihao Liu[1] · Junyu Chen[2] · Shuwen Wei[1] · Aaron Carass[1] · Jerry Prince[1]

Abstract

Producing spatial transformations that are diffeomorphic is a key goal in deformable image registration. As a diffeomorphic transformation should have positive Jacobian determinant $|J|$ everywhere, the number of pixels (2D) or voxels (3D) with $|J| < 0$ has been used to test for diffeomorphism and also to measure the irregularity of the transformation. For digital transformations, $|J|$ is commonly approximated using a central difference, but this strategy can yield positive $|J|$'s for transformations that are clearly not diffeomorphic—even at the pixel or voxel resolution level. To show this, we first investigate the geometric meaning of different finite difference approximations of $|J|$. We show that to determine if a deformation is diffeomorphic for digital images, the use of any individual finite difference approximation of $|J|$ is insufficient. We further demonstrate that for a 2D transformation, four unique finite difference approximations of $|J|$'s must be positive to ensure that the entire domain is invertible and free of folding at the pixel level. For a 3D transformation, ten unique finite differences approximations of $|J|$'s are required to be positive. Our proposed *digital diffeomorphism* criteria solves several errors inherent in the central difference approximation of $|J|$ and accurately detects non-diffeomorphic digital transformations. The source code of this work is available at https://github.com/yihao6/digital_diffeomorphism.

1 Introduction

The goal of deformable image registration is to establish a nonlinear spatial transformation that aligns two images.

✉ Yihao Liu
   yliu236@jhu.edu

   Junyu Chen
   jchen245@jhmi.edu

   Shuwen Wei
   swei14@jhu.edu

   Aaron Carass
   aaron_carass@jhu.edu

   Jerry Prince
   prince@jhu.edu

1 Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA

2 Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins Medical Institutes, Baltimore, MD, USA

For many tasks, it is reasonable to assume that the anatomy in the two images share the same topology. Therefore, for deformable registration algorithms, the ability to produce a topology preserving transformation is preferred. Following the work of (Christensen, 1994), many registration algorithms either penalize (Avants et al., 2008; Chen et al., 2017; Mok & Chung, 2020) or constrain (Beg et al., 2005; Chen et al., 2015; Dalca et al., 2018) their output transformations to be diffeomorphic and preserve topology. In the continuous domain, a diffeomorphic transformation is a smooth and invertible mapping with a smooth inverse that is guaranteed to maintain the topology of the anatomy being transformed. A diffeomorphic transformation should have positive Jacobian determinant $|J|$ everywhere,[1] and testing if a transformation is diffeomorphic involves local computation of $|J|$. When a transformation is not diffeomorphic, it is common to use the number of pixels (2D) or voxels (3D) with negative $|J|$ (Balakrishnan et al., 2019; Chen et al., 2022, 2021; Liu et al., 2022) or the standard deviation of the logarithmic trans-

---

1 We do not consider the case of all negative $|J|$, which would produce a reflection of the entire image.

formed $|J|$ (Hering et al., 2022; Kabus et al., 2009; Johnson & Christensen, 2002) to measure the irregularity of the transformation.

Given a digital transformation that is defined on a regular grid, a widely accepted practice for computing the Jacobian is to use finite difference approximations of spatial derivatives. While certain parametric and B-spline based models offer alternative methods for estimating the Jacobian (Narayanan et al., 2005; Rohlfing et al., 2003) for their resultant deformation field, these techniques are tailored for specific models and are not universally applicable. For the scope of our discussion and its relevance to a broader array of applications, including traditional and deep learning based registration algorithms, we focus on finite difference methods. There are three standard methods for computing finite differences along each axis. We denote the forward, backward, and central differences that operate along the x axis (and similarly for the y and z axes) as $\mathcal{D}^{+x}$, $\mathcal{D}^{-x}$, and $\mathcal{D}^{0x}$, respectively. To approximate $|J|$ in 2D or 3D transformations, either the same or a different type of finite difference can be used for different axes. We denote the central difference approximation of $|J|$ in 2D and 3D as $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|$ and $\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J|$, respectively.

Despite its current popularity, the central difference approximation of the Jacobian does not always work as expected in the evaluation of the diffeomorphism property of spatial transformations. For example, Fig. 1a shows a transformation around center point $\boldsymbol{p}$ that is not diffeomorphic despite the fact that $\mathcal{D}^{0x}\mathcal{D}^{0y}|J| = 1$. In fact, the transformation at $\boldsymbol{p}$ has no effect on the computation of $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|(\boldsymbol{p})$, even if $\boldsymbol{p}$ moves outside the field of view. We call this the *checkerboard problem* because the central difference based Jacobian computations on the transformations of the "black"
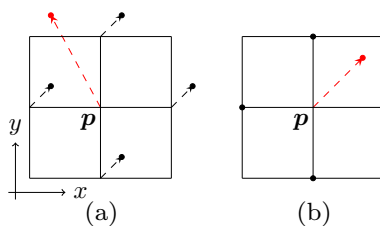

Fig. 1 Two transformations around the center point $\boldsymbol{p}$ are shown. The displacements of pixels around $\boldsymbol{p}$ are visualized as dotted arrows pointing toward solid dots; the displacement at these pixels are used for computing $D^{0x}D^{0y}|J|$ for $\boldsymbol{p}$. The displacement at $\boldsymbol{p}$ itself is highlighted in red. **a** highlights the *checkerboard problem* inherent to $D^{0x}D^{0y}|J|$. Here, $D^{0x}D^{0y}|J|(\boldsymbol{p}) = 1$ due to the translational movement of the neighboring pixels. However, as the displacement of $\boldsymbol{p}$ is *not* involved in computing $D^{0x}D^{0y}|J|(\boldsymbol{p})$, it can take on any value. Consequently, folding caused by the displacement at $\boldsymbol{p}$ remains undetected by $D^{0x}D^{0y}|J|(\boldsymbol{p})$. **b** demonstrates the inconsistency issue arising from the application of a single finite difference approximation in Jacobian determinant computation. Here, $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p})$ and $\mathcal{D}^{+x}\mathcal{D}^{+y}|J|(\boldsymbol{p})$ have opposite signs, resulting in conflicting interpretations

and "white" pixels (of a checkerboard) are independent of each other.

A possible solution to this problem is to use $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|$ or $\mathcal{D}^{+x}\mathcal{D}^{+y}|J|$ instead. However, Fig. 1b shows an example in which $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|$ and $\mathcal{D}^{+x}\mathcal{D}^{+y}|J|$ have opposite signs, leading to contradictory conclusions. These examples illustrate the problem with naive application of finite differences in this application. We seek a better approach that still involves finite differences, but avoids these types of contradictory situations.

In this work, we first investigate the geometric meaning of finite difference approximations of $|J|$. We show that when using forward or backward differences, the sign of $|J|(\boldsymbol{p})$ determines if the underlying transformation $\mathcal{T}$ is invertible and orientation-preserving in a triangle (2D) or tetrahedron (3D) adjacent to $\boldsymbol{p}$. Reversing the orientation indicates folding in space. We formally define digital transformations that are globally invertible and free of folding as digital diffeomorphisms. In order to determine if a transformation is a digital diffeomorphism, at each point it is necessary to consider four finite difference approximated $|J|$'s for 2D transformations and ten $|J|$'s for 3D transformations. We also demonstrate that because of the *checkerboard problem* and other errors that are inherent in the central difference based $|J|$, the number of non-diffeomorphic voxels it reports is always less than or equal to the actual number. Finally, we propose to use non-diffeomorphic area (2D) and volume (3D) as more meaningful measurements of irregularity in computed transformations.

2 Methodology

2.1 Backward Difference Based Jacobian Determinant in 2D

Consider the standard Euclidean space $\mathbb{R}^2$ that follows the right-hand rule. Let $\mathcal{T}$ be a digital transformation for $\mathbb{R}^2$ that is defined for every grid point $\boldsymbol{p}$. When using backward differences on both x and y axes for approximating $|J|$ at point $\boldsymbol{p}$, we have the formulation

$$\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) = \begin{vmatrix} \mathcal{D}^{-x}\mathcal{T}_x(\boldsymbol{p}) & \mathcal{D}^{-y}\mathcal{T}_x(\boldsymbol{p}) \\ \mathcal{D}^{-x}\mathcal{T}_y(\boldsymbol{p}) & \mathcal{D}^{-y}\mathcal{T}_y(\boldsymbol{p}) \end{vmatrix}, \qquad (1)$$

where $\mathcal{T}_x(\boldsymbol{p})$ and $\mathcal{T}_y(\boldsymbol{p})$ are the x and y components of $\mathcal{T}(\boldsymbol{p})$.

We denote the 4-connected neighbors of $\boldsymbol{p}$ as $\boldsymbol{p}^{-x}$, $\boldsymbol{p}^{+x}$, $\boldsymbol{p}^{-y}$, and $\boldsymbol{p}^{+y}$, as shown in Fig. 2a. Their transformed locations are denoted with subscripts t, as shown in Fig. 2b. For example, $\boldsymbol{p}_t^{+x} := \mathcal{T}(\boldsymbol{p}^{+x})$. We denote the triangular region defined by the vectors $\boldsymbol{p}\boldsymbol{p}^{-x}$ and $\boldsymbol{p}\boldsymbol{p}^{-y}$ as $\triangle \boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}$, and we assume that the 2D transformation $\mathcal{T}$ is linearly interpolated on $\triangle \boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}$.
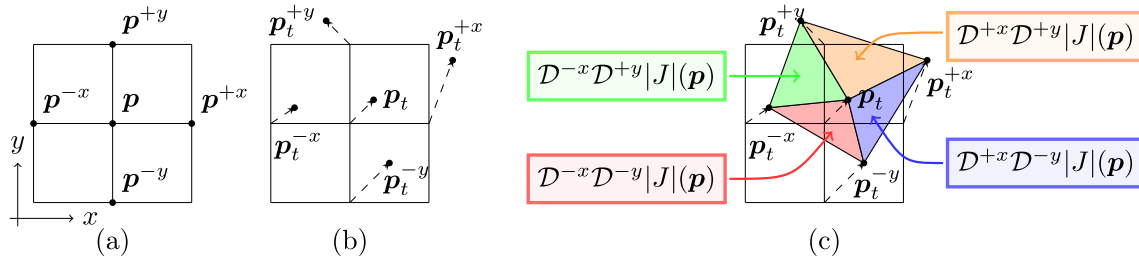
Fig. 2 **a** shows the notations for $\boldsymbol{p}$ and its 4-connected neighbors. **b** shows $\boldsymbol{p}_t$, the transformed version of $\boldsymbol{p}$ as well as the transformed position of the 4-connected neighbors. **c** shows the triangular regions and their corresponding forward and backward difference based $|J|$'s

Proposition 1 *A 2D transformation $\mathcal{T}$ is invertible for $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$ if and only if $\mathcal{D}^{-x}\mathcal{D}^{-y}\,|J|\,(\boldsymbol{p})$ for $\mathcal{T}$ is nonzero.*

Proof Since $\mathcal{T}$ is linearly interpolated on $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$, $\mathcal{T}$ is linear for $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$ and can be written as a 2×2 matrix with $\boldsymbol{p}_t\,\boldsymbol{p}_t^{-x}$ and $\boldsymbol{p}_t\,\boldsymbol{p}_t^{-y}$ as its columns. Thus, $\mathcal{T}$ is invertible if and only if $\boldsymbol{p}_t\,\boldsymbol{p}_t^{-x}$ and $\boldsymbol{p}_t\,\boldsymbol{p}_t^{-y}$ are linearly independent (not colinear). $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p})$ can be written as a triple product:

$$\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) = (\boldsymbol{p}_t\,\boldsymbol{p}_t^{-x} \times \boldsymbol{p}_t\,\boldsymbol{p}_t^{-y}) \cdot \boldsymbol{n}, \tag{2}$$

where $\boldsymbol{n}$ is the unit vector perpendicular to vectors $\boldsymbol{p}_t\,\boldsymbol{p}_t^{-x}$ and $\boldsymbol{p}_t\,\boldsymbol{p}_t^{-y}$ and is positively oriented following the right-hand rule. Therefore, $\mathcal{T}$ is invertible if and only if $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) \neq 0$. □

Definition 1 A 2D transformation $\mathcal{T}$ is said to *cause folding* of $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$ if the orientation of $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$ is reversed by $\mathcal{T}$.

Proposition 2 *A 2D transformation $\mathcal{T}$ is free of folding for $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$ if and only if $\mathcal{D}^{-x}\mathcal{D}^{-y}\,|J|\,(\boldsymbol{p})$ for $\mathcal{T}$ is positive.*

Proof By the right-hand rule, $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$ is positively oriented.
($\Rightarrow$) When $\mathcal{T}$ is free of folding, the orientation of $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$ is preserved by $\mathcal{T}$. Because of linear interpolation, $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}$ is transformed to $\triangle \boldsymbol{p}_t\,\boldsymbol{p}_t^{-x}\,\boldsymbol{p}_t^{-y}$, which is also positively oriented. Equation 2 shows that $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p})$ equals twice the signed area of $\triangle \boldsymbol{p}_t\,\boldsymbol{p}_t^{-x}\,\boldsymbol{p}_t^{-y}$. Therefore, $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) > 0$.
($\Leftarrow$) When $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) > 0$, from Eq. 2 $\triangle \boldsymbol{p}_t\,\boldsymbol{p}_t^{-x}\,\boldsymbol{p}_t^{-y}$ is positively oriented. Because of linear interpolation, $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}$ is transformed to $\triangle \boldsymbol{p}_t\,\boldsymbol{p}_t^{-x}\,\boldsymbol{p}_t^{-y}$ and both of them are positively oriented. Therefore, $\mathcal{T}$ is free of folding for $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$. □

Definition 2 A 2D transformation $\mathcal{T}$ is *digitally diffeomorphic* for the region $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$ if $\mathcal{T}$ is invertible and free of folding for $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$.

Proposition 3 *A 2D transformation $\mathcal{T}$ is digitally diffeomorphic for the region $\triangle \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}$ if and only if $\mathcal{D}^{-x}\mathcal{D}^{-y}\,|J|\,(\boldsymbol{p}) > 0$.*

**Proof** Proposition 3 is a direct consequence of Propositions 1 and 2. □

In conclusion, $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|$ for the point $\boldsymbol{p}$ informs us about the digitally diffeomorphic property of a triangle adjacent to $\boldsymbol{p}$, under the assumption that the transformation is linearly interpolated.

2.2 Digital Diffeomorphism in 2D

Similar to $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p})$, we can replicate Proposition 3 to establish that any $|J|(\boldsymbol{p})$ approximated using any combination of forward and backward differences is testing if the transformation is digitally diffeomorphic for a triangular region around $\boldsymbol{p}$ (see Fig. 2c), assuming that the region is linearly interpolated. Since these $|J|(\boldsymbol{p})$'s cover different regions, their signs are independent of each other. For example, $\mathcal{T}$ can have a positive $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p})$ and a negative $\mathcal{D}^{+x}\mathcal{D}^{+y}|J|(\boldsymbol{p})$ at the same time (see Fig. 1b). For each of these approximations, the implied triangular regions for all $\boldsymbol{p}$'s taken together cover only half of the space (*e.g.*, $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|$ only considers the red tiles in Fig. 3a). Consequently, even if a transformation $\mathcal{T}$ has $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) > 0$ for all $\boldsymbol{p}$'s, half the space is not considered and can potentially exhibit folding or be non-invertible. This is also the case for any other forward and backward difference computations of $|J|$.

Although combining the triangular regions of $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|$ and $\mathcal{D}^{+x}\mathcal{D}^{+y}|J|$ can cover the entire space, positive $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|$ and $\mathcal{D}^{+x}\mathcal{D}^{+y}|J|$ only guarantee that the transformation is digitally diffeomorphic when the transformation is piecewise linearly interpolated as shown in Fig. 3a. A different choice, for example that shown in Fig. 3b, which corresponds to $\mathcal{D}^{-x}\mathcal{D}^{+y}|J|$ and $\mathcal{D}^{+x}\mathcal{D}^{-y}|J|$ can give contradictory conclusions. Since there are two ways of dividing the square-size area between grid points, there are many more piecewise linear transformations that correspond to the same digital transformation, *e.g.*, Fig. 3c. To anticipate all possi-

Fig. 3 Illustration of the combinations of Jacobian determinants and their corresponding triangular regions. **(a)** $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p})$ and $\mathcal{D}^{+x}\mathcal{D}^{+y}|J|(\boldsymbol{p})$. **(b)** $\mathcal{D}^{-x}\mathcal{D}^{+y}|J|(\boldsymbol{p})$ and $\mathcal{D}^{+x}\mathcal{D}^{-y}|J|(\boldsymbol{p})$. **(c)** Each pixel uses a different combination of Jacobian determinants
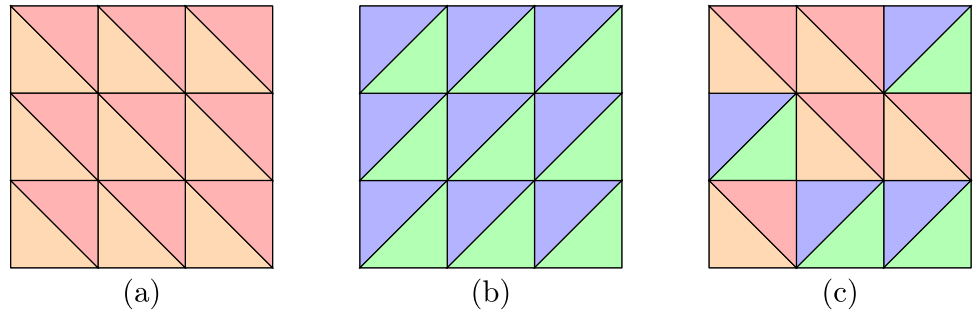
ble choices of triangulation—each of which leads to its own invertible transformation on the plane—we should therefore consider all finite difference approximations in determining whether a transformation is digitally diffeomorphic. This leads naturally to the following definition:

Definition 3 A 2D digital transformation $\mathcal{T}$ is a *digital diffeomorphism* if for every grid point $\boldsymbol{p}$ it satisfies $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) > 0$, $\mathcal{D}^{-x}\mathcal{D}^{+y}|J|(\boldsymbol{p}) > 0$, $\mathcal{D}^{+x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) > 0$, and $\mathcal{D}^{+x}\mathcal{D}^{+y}|J|(\boldsymbol{p}) > 0$.

The proposed digital diffeomorphism definition guarantees the transformation to be free of folding and invertible regardless of the piecewise linear transformation (on triangles that divide the squares between pixel centers) that is used.

2.3 Central Difference Based Jacobian Determinant

In this section, we analyze the central difference approximation of $|J|$ given the previous analysis. We first ask where $\boldsymbol{p}$ can be positioned to yield a digital diffeomorphism when its neighbors have fixed transformations. We start with the following definitions:

Definition 4 For grid point $\boldsymbol{p}$ with fixed $\boldsymbol{p}_t^{-x}$, $\boldsymbol{p}_t^{-y}$, $\boldsymbol{p}_t^{+x}$, and $\boldsymbol{p}_t^{+y}$, $\mathcal{R}(\boldsymbol{p})$ is defined to be the region in $\mathbb{R}^2$ such that $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) > 0$, $\mathcal{D}^{-x}\mathcal{D}^{+y}|J|(\boldsymbol{p}) > 0$, $\mathcal{D}^{+x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) > 0$, and $\mathcal{D}^{+x}\mathcal{D}^{+y}|J|(\boldsymbol{p}) > 0$.

Definition 5 Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be points in $\mathbb{R}^2$. The half-plane $\mathcal{H}(\boldsymbol{a}, \boldsymbol{b})$ is defined as $\boldsymbol{p} \in \mathbb{R}^2$ such that $\triangle \boldsymbol{a}\boldsymbol{b}\boldsymbol{p}$ is positively oriented.

Proposition 4 $\mathcal{R}(\boldsymbol{p})$ *is the intersection of the four half-planes* $\mathcal{H}(\boldsymbol{p}_t^{-x}, \boldsymbol{p}_t^{-y})$, $\mathcal{H}(\boldsymbol{p}_t^{-y}, \boldsymbol{p}_t^{+x})$, $\mathcal{H}(\boldsymbol{p}_t^{+x} \boldsymbol{p}_t^{+y})$, *and* $\mathcal{H}(\boldsymbol{p}_t^{+y}, \boldsymbol{p}_t^{-x})$.

Proof Eq. 1 can be written as:

$$\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) = (\boldsymbol{p}_t^{-x}\boldsymbol{p}_t^{-y} \times \boldsymbol{p}_t^{-x}\boldsymbol{p}_t) \cdot \boldsymbol{n}.$$

Therefore, when $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p}) > 0$, $\triangle \boldsymbol{p}_t^{-x}\boldsymbol{p}_t^{-y}\boldsymbol{p}_t$ is positively oriented. Thus $\boldsymbol{p}_t$ is in the half-plane $\mathcal{H}(\boldsymbol{p}_t^{-x}, \boldsymbol{p}_t^{-y})$. Analogous statements can be made for the other forward

and backward difference based $|J|$'s. When all the four $|J|$'s are positive, $\boldsymbol{p}_t$ must be inside the four half-planes $\mathcal{H}(\boldsymbol{p}_t^{-x}, \boldsymbol{p}_t^{-y}), \mathcal{H}(\boldsymbol{p}_t^{-y}, \boldsymbol{p}_t^{+x}), \mathcal{H}(\boldsymbol{p}_t^{+x} \boldsymbol{p}_t^{+y})$, and $\mathcal{H}(\boldsymbol{p}_t^{+y}, \boldsymbol{p}_t^{-x})$. Therefore, $\mathcal{R}(\boldsymbol{p})$ is the intersection of these four half-planes. $\square$

Figure 4 shows examples of $\mathcal{R}(\boldsymbol{p})$. In Figs. 4c and d, no matter where $\boldsymbol{p}_t$ is located, at least one forward or backward difference based $|J|(\boldsymbol{p})$ is guaranteed to be negative and, thus, $\mathcal{R}(\boldsymbol{p})$ is empty. This makes sense since the transformations in Fig. 4c and d cause folds in space.

Proposition 5 *Assume* $\boldsymbol{p}_t^{-x}$, $\boldsymbol{p}_t^{-y}$, $\boldsymbol{p}_t^{+x}$, *and* $\boldsymbol{p}_t^{+x}$ *forms a simple polygon (without self-intersection). Then* $\mathcal{R}(\boldsymbol{p})$ *is non-empty if and only if* $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|(\boldsymbol{p}) > 0$.

Proof $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|(\boldsymbol{p})$ can be written as:

$$\mathcal{D}^{0x}\mathcal{D}^{0y}|J| = \frac{1}{2}\left(\mathcal{D}^{-x}\mathcal{D}^{-y}|J| + \mathcal{D}^{+x}\mathcal{D}^{-y}|J| \right.$$
$$\left. + \mathcal{D}^{-x}\mathcal{D}^{+y}|J| + \mathcal{D}^{+x}\mathcal{D}^{+y}|J|\right). \quad (3)$$

($\Rightarrow$) If $\mathcal{R}(\boldsymbol{p})$ is non-empty, for every $\boldsymbol{p}_t \in \mathcal{R}(\boldsymbol{p})$ all its forward and backward difference based $|J|(\boldsymbol{p})$'s are positive by Definition 4, and therefore from Eq. 3, $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|(\boldsymbol{p}) > 0$. ($\Leftarrow$) On the other hand, if $\mathcal{D}^{0x}\mathcal{D}^{0y}|J| > 0$, the polygon $\boldsymbol{p}_t^{-x}\boldsymbol{p}_t^{-y}\boldsymbol{p}_t^{+x}\boldsymbol{p}_t^{+y}$ is positively oriented (Braden, 1986) (*i.e.* interior to the left). Thus, if there exists a $\boldsymbol{p}_t$ that is visible to all vertices of the polygon, then $\boldsymbol{p}_t \in \mathcal{R}(\boldsymbol{p})$ by Proposition 4. Following (Chvátal, 1975), such $\boldsymbol{p}_t$ always exists for simple polygons with five or fewer vertices. Therefore, $\mathcal{R}(\boldsymbol{p})$ is non-empty. $\square$

Although a positive $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|(\boldsymbol{p})$ ensures that $\mathcal{R}(\boldsymbol{p})$ is non-empty, $\boldsymbol{p}$ can still be digitally non-diffeomorphic when $\boldsymbol{p}_t \notin \mathcal{R}(\boldsymbol{p})$, which we see in the *checkerboard problem*. In addition to the *checkerboard problem*, $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|$ also fails to provide a meaningful interpretation when the polygon $\boldsymbol{p}_t^{-x}\boldsymbol{p}_t^{-y}\boldsymbol{p}_t^{+x}\boldsymbol{p}_t^{+y}$ exhibits self-intersection (see Fig. 4d). In conclusion, $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|(\boldsymbol{p}) \leq 0$ always indicates that $\mathcal{T}$ is digitally non-diffeomorphic but $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|(\boldsymbol{p}) > 0$ does not mean it is digitally diffeomorphic because of the checkerboard or self-intersection problems. Therefore, use of central differences alone to estimate $|J|$ is an inadequate characterization of digital diffeomorphism.

Fig. 4 Examples of $\mathcal{R}(\boldsymbol{p})$ (blue shaded region) for $\boldsymbol{p}_t$ such that $\mathcal{T}$ is digitally diffeomorphic. **c** and **d** show examples of $\mathcal{T}$'s that are causing folds in space and thus $\mathcal{R}(\boldsymbol{p})$ is an empty set
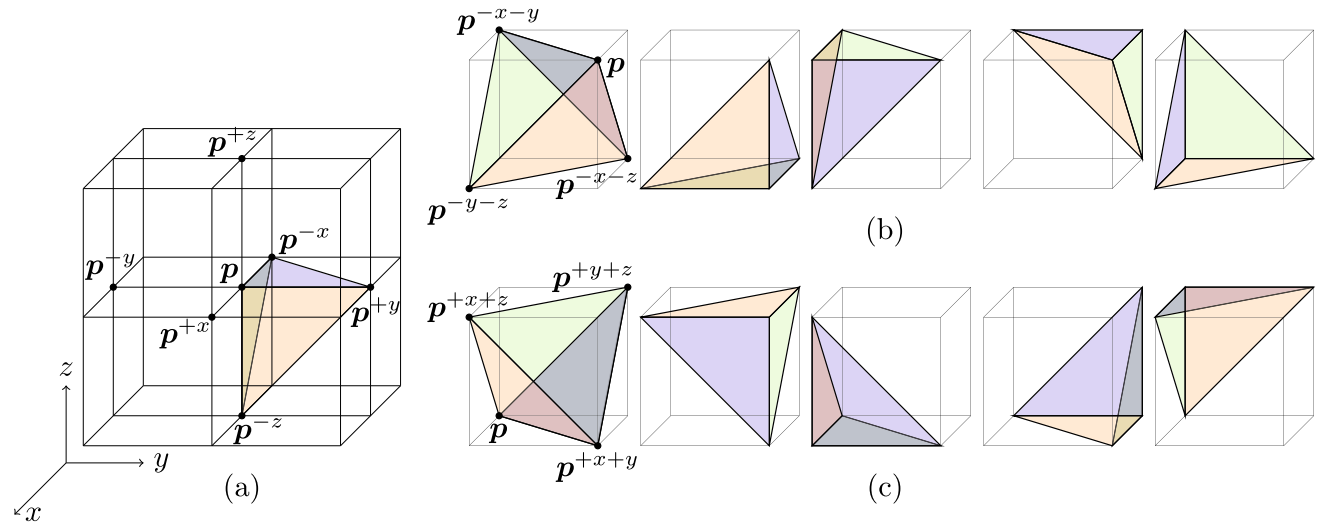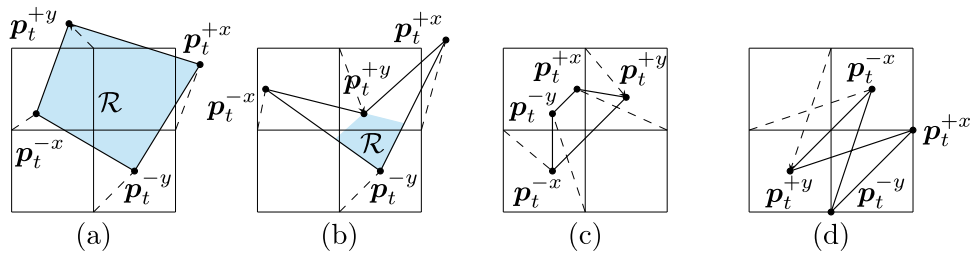

Fig. 5 **a** shows the notations for $\boldsymbol{p}$ and its 6-connected neighbors in 3D and the tetrahedron considered by $\mathcal{D}^{-x}\mathcal{D}^{+y}\mathcal{D}^{-z}|J|$. **b** and **c** are illustrations of the two schemes to divide the cube volume in-between grid points in 3D

2.4 Digital Diffeomorphism in 3D

Consider the standard Euclidean space $\mathbb{R}^3$ that follows the right-hand rule. Let $\mathcal{T}$ be a digital transformation of $\mathbb{R}^3$ that is defined for every grid point $\boldsymbol{p}$. We denote the 6-connected neighbors of $\boldsymbol{p}$ as $\boldsymbol{p}^{\pm x}$, $\boldsymbol{p}^{\pm y}$, $\boldsymbol{p}^{\pm z}$ (see Fig. 5a) and their transformed locations are denoted with the subscript t. We denote the tetrahedron defined by the vectors $\boldsymbol{p}\boldsymbol{p}^{-x}$, $\boldsymbol{p}\boldsymbol{p}^{-y}$, and $\boldsymbol{p}\boldsymbol{p}^{-z}$ as $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$, and we assume that the 3D transformation $\mathcal{T}$ is linearly interpolated on $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$.

Proposition 6 *A 3D transformation $\mathcal{T}$ is invertible for $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$ if and only if $\mathcal{D}^{-x}\mathcal{D}^{-y}\mathcal{D}^{-z}|J|$ for $\boldsymbol{p}$ is nonzero.*

Proof Since $\mathcal{T}$ is linearly interpolated on $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$, $\mathcal{T}$ is linear for $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$ and can be written as a 3×3 matrix with $\boldsymbol{p}_t\boldsymbol{p}_t^{-x}$, $\boldsymbol{p}_t\boldsymbol{p}_t^{-y}$, and $\boldsymbol{p}_t\boldsymbol{p}_t^{-z}$ as its columns. Thus, $\mathcal{T}$ is invertible if and only if $\boldsymbol{p}_t\boldsymbol{p}_t^{-x}$, $\boldsymbol{p}_t\boldsymbol{p}_t^{-y}$, and $\boldsymbol{p}_t\boldsymbol{p}_t^{-z}$ are linearly independent (not colinear). $\mathcal{D}^{-x}\mathcal{D}^{-y}\mathcal{D}^{-z}|J|(\boldsymbol{p})$ can be written as a triple product:

$$\mathcal{D}^{-x}\mathcal{D}^{-y}\mathcal{D}^{-z}|J|(\boldsymbol{p}) = -(\boldsymbol{p}_t\boldsymbol{p}_t^{-x} \times \boldsymbol{p}_t\boldsymbol{p}_t^{-y}) \cdot \boldsymbol{p}_t\boldsymbol{p}_t^{-z}. \quad (4)$$

Therefore, $\mathcal{T}$ is invertible if and only if $\mathcal{D}^{-x}\mathcal{D}^{-y}\mathcal{D}^{-z}|J|(\boldsymbol{p}) \neq 0$. □

Definition 6 A 3D transformation $\mathcal{T}$ is said to cause *folding* for $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$ if the orientation of $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$ is reversed by $\mathcal{T}$.

Proposition 7 *A 3D transformation $\mathcal{T}$ is free of folding for $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$ if and only if $\mathcal{T}$ has $\mathcal{D}^{-x}\mathcal{D}^{-y}\mathcal{D}^{-z}|J|(\boldsymbol{p}) > 0$.*

Proof ($\Rightarrow$) When $\mathcal{T}$ is free of folding, the orientation of $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$ (negatively oriented by the right-hand rule) is preserved by $\mathcal{T}$. Because of linear interpolation, $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$ is transformed to $\triangle\boldsymbol{p}_t\boldsymbol{p}_t^{-x}\boldsymbol{p}_t^{-y}\boldsymbol{p}_t^{-z}$, which is also negatively oriented. Equation 4 shows that $\mathcal{D}^{-x}\mathcal{D}^{-y}\mathcal{D}^{-z}|J|(\boldsymbol{p})$ equals to six times the negative signed volume of $\triangle\boldsymbol{p}_t\boldsymbol{p}_t^{-x}\boldsymbol{p}_t^{-y}\boldsymbol{p}_t^{-z}$. Therefore, $\mathcal{D}^{-x}\mathcal{D}^{-y}\mathcal{D}^{-z}|J|(\boldsymbol{p}) > 0$.

($\Leftarrow$) $\mathcal{D}^{-x}\mathcal{D}^{-y}\mathcal{D}^{-z}|J|(\boldsymbol{p}) > 0$ indicates that $\triangle\boldsymbol{p}_t\boldsymbol{p}_t^{-x}\boldsymbol{p}_t^{-y}\boldsymbol{p}_t^{-z}$ is negatively oriented by Eq. 4. Because of linear interpolation $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$ is mapped to $\triangle\boldsymbol{p}_t\boldsymbol{p}_t^{-x}\boldsymbol{p}_t^{-y}\boldsymbol{p}_t^{-z}$ and both of them are negatively oriented. Therefore, $\mathcal{T}$ is free of folding for $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$. □

Definition 7 A 3D transformation $\mathcal{T}$ is *digitally diffeomorphic* for the region $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$ if $\mathcal{T}$ is invertible and free of folding for $\triangle\boldsymbol{p}\boldsymbol{p}^{-x}\boldsymbol{p}^{-y}\boldsymbol{p}^{-z}$.

Proposition 8 *A 3D transformation $\mathcal{T}$ is digitally diffeomorphic for the region $\triangle\!\!\!\!\triangle\, \boldsymbol{p}\,\boldsymbol{p}^{-x}\,\boldsymbol{p}^{-y}\,\boldsymbol{p}^{-z}$ if and only if $\mathcal{D}^{-x}\mathcal{D}^{-y}\mathcal{D}^{-z}|J|(\boldsymbol{p}) > 0$.*

Proof Proposition 8 is a direct consequence of Propositions 6 and 7. □

Similarly, we can prove that a $|J|(\boldsymbol{p})$ approximated using any combination of forward and backward differences is testing if the transformation is digitally diffeomorphic in a tetrahedron adjacent to $\boldsymbol{p}$. One of these tetrahedra (for $\mathcal{D}^{-x}\mathcal{D}^{+y}\mathcal{D}^{-z}|J|$) is visualized in Fig. 5a. However, unlike the 2D case where all the triangular regions completely cover the entire 2D space, the union of all these adjacent tetrahedra does not fill the entire 3D space. Therefore, even if all forward and backward difference approximations of $|J|$ are positive, the transformation can still cause folding in the spaces not covered by these adjacent tetrahedra.

To solve this issue, we introduce another tetrahedron $\triangle\!\!\!\!\triangle\, \boldsymbol{p}\,\boldsymbol{p}^{-x-y}\,\boldsymbol{p}^{-x-z}\,\boldsymbol{p}^{-y-z}$, as shown in Fig. 5b. When combined with four existing tetrahedra from finite difference based $|J|$'s, they completely cover the entire volume. We define $|J_1^\star|(\boldsymbol{p})$ as

$$|J_1^\star|(\boldsymbol{p}) = (\boldsymbol{p}_t\boldsymbol{p}_t^{-x-y} \times \boldsymbol{p}_t\boldsymbol{p}_t^{-x-z}) \cdot \boldsymbol{p}_t\boldsymbol{p}_t^{-y-z}, \tag{5}$$

where $\boldsymbol{p}_t^{-x-y}$, $\boldsymbol{p}_t^{-x-z}$, and $\boldsymbol{p}_t^{-y-z}$ are the transformed locations of $\boldsymbol{p}^{-x-y}$, $\boldsymbol{p}^{-x-z}$, $\boldsymbol{p}^{-y-z}$. The signed volume of the extra tetrahedron after applying $\mathcal{T}$ equals $\frac{1}{6}|J_1^\star|(\boldsymbol{p})$. Similar to 2D, there are two ways of dividing the cube-size volume in-between grid points into five tetrahedra, as shown in Figs. 5b, c. The signed volume of the extra tetrahedron in Fig. 5(c) can be computed as $\frac{1}{6}|J_2^\star|(\boldsymbol{p})$, where

$$|J_2^\star|(\boldsymbol{p}) = (\boldsymbol{p}_t\boldsymbol{p}_t^{+x+y} \times \boldsymbol{p}_t\boldsymbol{p}_t^{+y+z}) \cdot \boldsymbol{p}_t\boldsymbol{p}_t^{+x+z}. \tag{6}$$

There are other ways to divide a cube into tetrahedra (Carr et al., 2006). These other schemes are not considered here because 1) their computation involve finite difference approximations at different points than those we consider, or 2) their computation requires interpolating the transformation at non-grid point.

Definition 8 A 3D digital transformation $\mathcal{T}$ is a digital diffeomorphism if for every grid point $\boldsymbol{p}$ its forward and backward difference based $|J|$'s of $\boldsymbol{p}$ are all positive and both $|J_1^\star|(\boldsymbol{p})$ and $|J_2^\star|(\boldsymbol{p})$ are positive.

For similar reason as in 2D, our definition of digital diffeomorphism involves both schemes shown in Fig. 5b, c. Specifically, each of the two schemes corresponds to a particular piecewise linear transformation. But for a given digital transformation there are many plausible piecewise linear transformations (see Fig. 3c). By considering both schemes,

we avoid the ambiguity of choosing different schemes for every cube-sized volume.

The central difference approximation of $|J|$ in 3D calculates the signed volume of an octahedron with vertices $\boldsymbol{p}_t^{-x}$, $\boldsymbol{p}_t^{+x}$, $\boldsymbol{p}_t^{-y}$, $\boldsymbol{p}_t^{+y}$, $\boldsymbol{p}_t^{-z}$, and $\boldsymbol{p}_t^{+z}$, when the octahedron is simple. The proof is a straightforward extension of Eq. 3. It is easy to show that $\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J|$ can also have the checkerboard problem or the self-intersection problem as in the 2D case. Note that Proposition 5 cannot be generalized to 3D because $\boldsymbol{p}_t^{-x}\boldsymbol{p}_t^{+x}\boldsymbol{p}_t^{-y}\boldsymbol{p}_t^{+y}\boldsymbol{p}_t^{-z}\boldsymbol{p}_t^{+z}$ may not be tetrahedralizable (O'Rourke et al., 1987) and thus, $\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J|(\boldsymbol{p}) > 0$ cannot guarantee $\mathcal{R}(\boldsymbol{p})$ in 3D is non-empty.

2.5 Non-diffeomorphic Space Measurement

For non-diffeomorphic transformations in 3D, the number and percentage of non-diffeomorphic voxels is often used to measure its irregularity. Specifically, a voxel is considered non-diffeomorphic if its center location $\boldsymbol{p}$ has a central difference based Jacobian determinant that is not positive (*i.e.* in 3D: $\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J|(\boldsymbol{p}) \le 0$). However, as we have demonstrated in Sects. 1 and 2, the central difference based Jacobian determinant underestimates the non-diffeomorphic space, in general. Given the limitations of the central difference-based Jacobian determinant, we seek an alternative way of evaluating the diffeomorphic property of digital transformations. We aim to find a quantitative measure that involves the computation of finite difference-based Jacobians and is consistent with our definition of digital diffeomorphism. As a result, we propose *non-diffeomorphic volume* (NDV) to quantify the size of the non-diffeomorphic space in 3D caused by a digital transformation.

Given a grid point $\boldsymbol{p}$, we denote its eight forward and backward difference based Jacobian determinants as $|J_i|$, $i \in [1, \ldots, 8]$. As shown in Sect. 2.4, each of these determinants is equal to six times the signed volume of a tetrahedron adjacent to $\boldsymbol{p}$. Thus, the non-diffeomorphic volume caused by a given transformation within any tetrahedron is given by $-\min(|J_i|, 0)/6$, which is positive only if the tetrahedron is folded.

Following Def. 8, we use the total volume of all *folded* tetrahedrons from 1) $|J_i|$, $i \in [1, \ldots, 8]$, 2) $|J_1^\star|(\boldsymbol{p})$ given in Eq. 5, and 3) $|J_2^\star|(\boldsymbol{p})$ given in Eq. 6 to compute NDV exhibited in the entire domain as:

$$\begin{aligned} \text{NDV} = -\frac{1}{2}\sum_{\boldsymbol{p}} &\left[\sum_{i=1}^{8} \frac{\min(|J_i|(\boldsymbol{p}), 0)}{6} + \right. \\ &\left. \frac{\min(|J_1^\star|(\boldsymbol{p}), 0)}{6} + \frac{\min(|J_2^\star|(\boldsymbol{p}), 0)}{6} \right]. \end{aligned} \tag{7}$$

This computation is the average of the two tetrahedralization schemes shown in Figs. 5b, c. While it is possible to report the minimum or maximum achievable non-diffeomorphic volume by tetrahedralizing each cube between voxels based on the specific transformation, doing so would inevitably associate the given digital transformation with one particular piecewise linear transformation. Therefore, averaging the two tetrahedralization schemes provides a more comprehensive and unbiased measure for a digital transformation. Further discussion on the impact of spatially varying tetrahedralization schemes can be found in Sect. 4.

The proposed NDV is connected to the ideas of simplex counting (SC) (Yushkevich et al., 2010) and surface propagation (SP) (Pai et al., 2016), which are used to assess the degree of volume change on a per-voxel basis. However, those methods typically use only one scheme to discretize the space. In contrast, NDV considers the two tetrahedralization schemes shown in Fig. 5b, c. This choice is motivated by the definition of digital diffeomorphism but it also offers a rotation-invariant property. Specifically, rotating the transformation by any multiple of 90 degrees does not affect its result. This cannot be achieved using only one scheme. Further discussion on the relationship and distinction between SC, SP, and the proposed NDV can be found in Sect. 4.

For completeness we note the 2D version of non-diffeomorphic space, which we term *non-diffeomorphic area* (NDA), follows from Def. 3,

$$\text{NDA} = -\frac{1}{2} \sum_{\boldsymbol{p}} \sum_{i=1}^{4} \frac{\min(|J_i|(\boldsymbol{p}), 0)}{2}, \tag{8}$$

where $|J_i|, i \in [1, \ldots, 4]$, are the Jacobian determinants approximated using the four possible combinations of forward and backward differences.

A demonstration of NDA is provided in Fig. 6. When using the central difference based $|J|$, all three cases in Fig. 6 would be considered diffeomorphic ($\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J|(\boldsymbol{p}) > 0$) because of the checkerboard problem. The forward difference based $|J|$ is able to identify that Fig. 6b, c exhibit
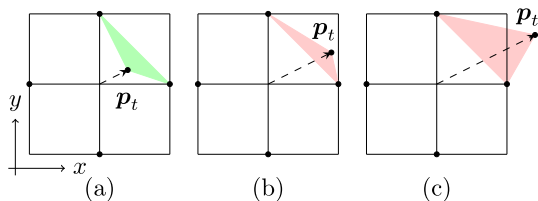

Fig. 6 A demonstration of measuring non-diffeomorphic space in 2D. The transformations are visualized as displacement fields. In each of the three subfigures, only the center points are transformed to their corresponding $\boldsymbol{p}_t$'s and the other grid points remain fixed. The triangular region corresponds to the forward difference based $|J|$ is shown in green if $|J| > 0$, otherwise it is shown in red

folding, but only NDA can provide the observation that the non-diffeomorphic space caused by the transformation shown in Fig. 6c is larger than the non-diffeomorphic space in Fig. 6b.

3 Experiments

We compared the commonly used central difference based Jacobian determinant ($\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J|$) and our proposed non-diffeomorphic volume (NDV) using several deformable registration algorithms on two publicly available datasets:

IXI: A total of 576 T1-weighted brain magnetic resonance (MR) images from the publicly available IXI dataset were used. 403 scans were used in training for the task of atlas-to-subject registration (Kim et al., 2021) and 58 scans were used for validation. The transformations generated from registering an atlas brain MR images to 115 test scans were evaluated.

Learn2Reg OASIS: We also used the brain T1-weighted MR images from the 2021 Learn2Reg challenge (Hering et al., 2022; LaMontagne et al., 2019). Scans were preprocessed using FreeSurfer (Fischl, 2012; Hoopes et al., 2021). All algorithms were trained using the training set of 414 scans and the transformations for the 19 validation pairs were evaluated.

The central difference Jacobian determinant approximation was implemented directly from the 2021 Learn2Reg challenge evaluation script. The implementation details and hyper-parameters for each of the algorithms were adopted from Chen et al. (2022) and Liu et al. (2022). Since, finite differences can be efficiently computed using linear filters with fixed, separable kernels, we implemented the proposed non-diffeomorphic space measurements using the correlation operator. It is worth noting that when a GPU is available, this implementation can be further accelerated by incorporating the convolutional layers in modern deep learning libraries.

A visualization of a result from the IXI dataset is shown in Fig. 7. Since it is difficult to visualize displacements across slices (anterior-to posterior direction in this case), only the displacements within the coronal plane were considered. Thus, we computed the non-diffeomorphic area instead of non-diffeomorphic volume. For each pixel in Fig. 7d, a higher red intensity indicates larger non-diffeomorphic area around that pixel. The non-diffeomorphic pixels computed from $\mathcal{D}^{0x}\mathcal{D}^{0y}|J|$ are highlighted in Fig. 7c for comparison.

The results on the IXI dataset are summarized in Table 1 and for the Learn2Reg OASIS in Table 2. Only the voxels within the brain were considered and the percentages were calculated relative to the brain volume of the fixed image.
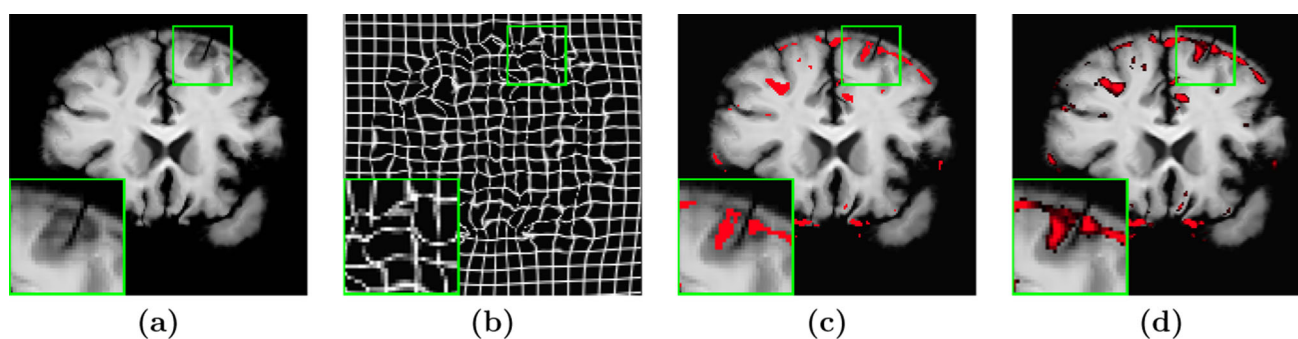
Fig. 7 A visualization of the proposed non-diffeomorphic area. Due to the difficulty in visualizing the deformation across slices (anterior-to posterior direction), only the displacement inside the coronal plane was considered in this example. **a** The warped image. **b** The grid line representation of the transformation (generated using Voxelmorph (Balakrishnan et al., 2019)). **c** The warped image with non-diffeomorphic pixels (marked in red) as measured by the central difference Jaco-bian determinant ($\mathcal{D}^{0x}\mathcal{D}^{0y}|J|$) highlighted in red. **d** The warped image with a map overlay indicating the non-diffeomorphic area with brighter shades of red indicating larger non-diffeomorphic area. We also provide a video demonstrating the NDV in 3D, that we make available at https://github.com/yihao6/digital_diffeomorphism/blob/main/docs/_static/imgs/ndv_demonstration.gif

Table 1 Our proposed non-diffeomorphic volume and several other measures on the IXI dataset and the seven comparison algorithms.

		$\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J| \leq 0$ # of voxels (%)	Any $|J_i| \leq 0$ # of voxels (%)	Proposed NDV (%)
NiftyReg	Modat et al. (2010)	222.2 ± 776.1 (0.01%)	288.9 ± 945.7 (0.02%)	10.9 ± 45.4 (< 0.00%)
deedsBCV	Heinrich et al. (2015)	5704.7 ± 1939.6 (0.37%)	12212.9 ± 3118.1 (0.79%)	1597.4 ± 661.5 (0.10%)
Voxelmorph	Balakrishnan et al. (2019)	41233.1 ± 8091.3 (2.64%)	98241.0 ± 17061.1 (6.26%)	16261.5 ± 2709.3 (1.04%)
Cyclemorph	Kim et al. (2021)	44126.5 ± 8526.4 (2.83%)	99560.0 ± 16833.8 (6.38%)	17923.6 ± 3083.3 (1.15%)
MIDIR	Qiu et al. (2021)	0.0 ± 0.0 (< 0.00%)	0.1 ± 0.7 (< 0.00%)	0.0 ± 0.0 (< 0.00%)
Transmorph	Chen et al. (2022)	35324.1 ± 7887.7 (2.26%)	88263.1 ± 17261.7 (5.65%)	14034.7 ± 2902.5 (0.90%)
im2grid	Liu et al. (2022)	8291.4 ± 5928.4 (0.53%)	20282.7 ± 7853.8 (1.3%)	822.6 ± 248.5 (0.05%)

For '$\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J| \leq 0$' and 'Any $|J_i| \leq 0$' we report the mean number of voxels (voxel #) over the 115 test subjects and the corresponding standard deviation ($\pm$), as well as the percentage (%) with respect to the brain mask. We also report our proposed measure of non-diffeomorphic space—*i.e.* non-diffeomorphic volume (NDV) in 3D—and the corresponding standard deviations and percentages. Methods are listed in the order in which they were published

In both tables, we report the number of non-diffeomorphic voxels (# of voxel) and its percentage (%) based on the $\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J|$. We also report the number (and percentage) of voxels that have at least one $|J_i| \leq 0$, denoted as 'Any $|J_i| \leq 0$' in the tables. We observe that in most of the cases, there are actually more than twice the number of voxels having $|J_i| \leq 0$ for some finite difference than found using only the central difference. The differences between '$\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J| \leq 0$' and 'Any $|J_i| \leq 0$' highlight that errors in using the central difference approximation are very common in practice.

The proposed average non-diffeomorphic volume and its percentage are shown in the last three columns of Tables 1 and 2. While the non-diffeomorphic voxel count and NDV in Tables 1 and 2 show a similar ranking order, we also observe different behaviors. For example, in Table 1, deeds-BCV reports fewer non-diffeomorphic voxels than im2grid, but has a higher NDV. This might be linked to the unique algorithm and regularization techniques that each method employs. As a result, even though im2grid presents a higher number of voxels with folding or non-invertible (higher 'Any $|J_i| \leq 0$'), these folded regions result in a smaller non-diffeomorphic volume. This example highlights the

Table 2 Results for the `Learn2Reg OASIS` dataset.

		$\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J| \leq 0$ voxels (%)	Any $|J_i| \leq 0$ # of voxels (%)	Proposed NDV (%)
SyN	Avants et al. (2008) (0.01%)	205.5 ± 331.1 (0.02%)	278.2 ± 415.3 (0.01%)	85.0 ± 204.1
Voxelmorph	Balakrishnan et al. (2019)	40418.4 ± 8991.3 (2.84%)	94343.8 ± 18452.7 (6.64%)	18448.2 ± 4319.0 (1.30%)
Transmorph	Chen et al. (2022) (2.22%)	31646.6 ± 7609 (5.52%)	78533.7 ± 16113.0 (0.95%)	13502.7 ± 3779.9
im2grid	Liu et al. (2022)	22905.6 ± 4142.3 (1.61%)	161071.8 ± 18271.5 (11.36%)	8774.7 ± 975.6 (0.62%)

For '$\mathcal{D}^{0x}\mathcal{D}^{0y}\mathcal{D}^{0z}|J| \leq 0$' and 'Any $|J_i| \leq 0$' we report the mean number of voxels (voxel #) over the 19 validation subject pairs and the corresponding standard deviation ($\pm$), as well as the percentage (%) with respect to the brain mask. We also report our proposed measure of non-diffeomorphic space—i.e. non-diffeomorphic volume (NDV) in 3D—and the corresponding standard deviations and percentages. Methods are listed in the order in which they were published

capability of NDV to provide more insights into the severity of folding, which is not captured by the existing use of counting the non-diffeomorphic voxels. For algorithms that impose strong regularization on the transformations (*e.g., MIDIR* (Qiu et al., 2021), *deedsBCV* (Heinrich et al., 2015), *NiftyReg* (Modat et al., 2010), and *SyN* (Avants et al., 2008)), a small—even zero—NDV is observed. However, for the deep learning methods that directly output deformation fields (Kim et al., 2021; Balakrishnan et al., 2019; Chen et al., 2022; Liu et al., 2022), we usually have higher NDVs. It is important to note that the results shown in Tables 1 and 2 do not reflect the accuracy of the algorithms, just the proportion of their deformation that is non-diffeomorphic.

4 Discussion

The Jacobian determinant of a transformation is a widely used measure in deformable image registration, but the details of its computation are often overlooked. In this paper, we focused on the finite difference based approximation of $|J|$. Contrary to what one might expect, the commonly used central difference based $|J|$ does not reflect if the transformation is diffeomorphic or not. Our investigation shows that each of the finite difference approximated $|J|$ corresponds to the signed area of a triangle in 2D or the signed volume of a tetrahedron in 3D when the digital transformations are assumed to be piecewise linear. Following this, we propose the definition of a digital diffeomorphism that allows diffeomorphisms—a concept in continuous domain—to be applied to digital transformations. It solves several problems that are inherent in the central difference based $|J|$. We further propose to use non-diffeomorphic volume to measure the irregularity of 3D transformations and non-diffeomorphic area for 2D transformations. As demonstrated in Figs. 6 and

7, our proposed approach measures the *severity* of the irregularity whereas the commonly used central difference based $|J|$ is only a binary indicator of folding (also with errors). The transformation shown in Fig. 6(b) is obviously more favorable than the one shown in Fig. 6(c) in terms of regularity. As such, it is important for us to be able to draw distinctions between these two scenarios.

The non-diffeomorphic measures presented in Eqs. 7 and 8 are averages of two choices of tetrahedralization of the volume. It is possible, however, to tetrahedralize each cube between voxels based on the specific registration outcome, for example, to yield the minimum or maximum achievable non-diffeomorphic volume. These measures must be computed for the entire image domain since the brain mask is defined at voxel level. For the `IXI` dataset, the average non-diffeomorphic volume for the best case is 33227 voxel³ for *Voxelmorph* and 31488 voxel³ for *Transmorph* while the average non-diffeomorphic Volume for the worse case is 41903 voxel³ for *Voxelmorph* and 40964 voxel³ for *Transmorph*. In comparison, our proposed NDV is 37565 voxel³ for *Voxelmorph* and 36226 voxel³ for *Transmorph* for the entire image domain. This result shows that, at least for these two algorithms, the non-diffeomorphic volume cannot be made substantially different from the average value that we specified in Eq. 7 by alternative selection of tetrahedralization.

The idea of discretizing the space using triangles or tetrahedrons has been presented before for regularizing deformable transformations and deformation-based volume change estimation (Pai et al., 2016; Holland et al., 2011; Yushkevich et al., 2010). It was previously believed that computing the volume change in discretized space using tetrahedrons is more accurate than using Jacobian determinants because the latter involves finite difference approximation (Yushkevich et al., 2010). Our analysis shows that

discretizing the space using triangles or tetrahedrons is in fact the consequence of finite difference approximation of Jacobian determinants. Therefore, the two approaches are equivalent when the corresponding finite differences are chosen for a given partition of space. (Haber & Modersitzki, 2004, 2007) and Burger et al. (2013) used the volume of triangles in 2D or tetrahedrons in 3D as a regularization term for their registration algorithm. Initially, Haber et al. proposed a hard equality constraint (Haber & Modersitzki, 2004) that enforces the preservation of the discretized volume (or area in 2D) of every deformed box. Recognizing that this approach could not detect "twists" (i.e., folding), they later proposed an inequality constraint (Haber & Modersitzki, 2007) that calculates the volumes of the tetrahedrons to prevent such twists by imposing positive volumes of the tetrahedrons. However, their methods only account for a single combination of Jacobian determinants (as shown in Fig. 3(a)), which is insufficient to guarantee a digital diffeomorphism. Moreover, these previous works were motivated by the fact that calculating the volume of a "twisted" polygon or octahedron is difficult. Our analysis on the central difference approximated $|J|$ explains why using a polygon or octahedron is inaccurate.

Throughout the manuscript, linear interpolation is assumed for the transformations at non-grid points (*e.g.,* the transformation is assumed to be piecewise linear). This assumption is intrinsically linked to the computation of finite differences. To clarify, Eq. 2 is derived from the triple product definition and is independent of the interpolation method used. As such, regardless of the chosen interpolation for transformations at non-grid points, $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p})$ always equals twice the signed area of $\triangle \boldsymbol{p}_t \boldsymbol{p}_t^{-x} \boldsymbol{p}_t^{-y}$. Under the linear interpolation assumption, $\triangle \boldsymbol{p}_t \boldsymbol{p}_t^{-x} \boldsymbol{p}_t^{-y}$ represents the transformed version of $\triangle \boldsymbol{p} \boldsymbol{p}^{-x} \boldsymbol{p}^{-y}$. The sign of $\mathcal{D}^{-x}\mathcal{D}^{-y}|J|(\boldsymbol{p})$ provides insight into the orientation of $\triangle \boldsymbol{p}_t \boldsymbol{p}_t^{-x} \boldsymbol{p}_t^{-y}$, which, when differing from the orientation of $\triangle \boldsymbol{p} \boldsymbol{p}^{-x} \boldsymbol{p}^{-y}$, indicates folding. An alternative interpolation method would permit $\triangle \boldsymbol{p} \boldsymbol{p}^{-x} \boldsymbol{p}^{-y}$ to transform into unpredictable shapes, rendering the orientation information of $\triangle \boldsymbol{p}_t \boldsymbol{p}_t^{-x} \boldsymbol{p}_t^{-y}$ unreliable for determining diffeomorphic properties. As such, the assumption of linear interpolation becomes crucial to precisely decipher diffeomorphic characteristics through the finite difference approximated Jacobian. In the context of deep learning based registration methods, the scaling-and-squaring approach has become increasingly popular with numerous works ensuring the diffeomorphic nature of their network outputs by incorporating this method in their output layer (Dalca et al., 2019; Chen et al., 2022; Hoopes et al., 2021). However, several works have reported that the scaling-and-squaring approach produces voxels with negative central difference based Jacobian determinant. The issue of interpolation, as we discussed earlier, is the reason why the scaling-and-squaring approach cannot consistently guarantee a positive Jacobian determinant. Specifically, when composing two digital transformations, one of the transformations needs to be interpolated at non-grid locations. The current methods usually adopt bilinear or trilinear interpolation for this process. These interpolation methods, however, are inconsistent with the piecewise linear transformation that is implicitly assumed by the finite difference based Jacobian determinant computation. Even if the original transformation is digitally diffeomorphic, the interpolated transformation may have negative Jacobian determinant when considered by the finite difference approximation.

For recent deep learning-based registration algorithms, the Jacobian determinant is frequently integrated into the loss function during network training. Many algorithms utilize the Jacobian determinant to either penalize negative Jacobian determinant (Mok & Chung, 2020) values—thereby encouraging diffeomorphism—or to penalize deviations from a determinant value of one, thereby encouraging volume preservation or incompressibility (Bian et al., 2023). The space discretization and volume estimation scheme presented in this work can refine these loss functions and better instill the desired properties within the trained networks.

Data availability The datasets analysed during the current study are available in the OASIS repository, https://www.oasis-brains.org/ and the IXI repository, https://brain-development.org/ixi-dataset/

Declarations

Conflict of interest The authors have no Conflict of interest to declare that are relevant to the content of this article

Ethical approval This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the relevant local Institutional Review Boards, and performed in line with the Declaration of Helsinki.

References

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis, 12*(1), 26–41.

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*(2), 139–157.

Bian, Z., Xing, F., Yu, J., Shao, M., Liu, Y., Carass, A., Woo, J., & Prince, J. L. (2023). DRIMET: Deep registration-based 3d incompressible motion estimation in tagged-MRI with application to the tongue. In *Medical imaging with deep learning*.

Braden, B. (1986). The surveyor's area formula. *The College Mathematics Journal, 17*(4), 326–337.

Burger, M., Modersitzki, J., & Ruthotto, L. (2013). A hyperelastic regularization energy for image registration. *SIAM Journal on Scientific Computing, 35*(1), B132–B148.

Carr, H., Moller, T., & Snoeyink, J. (2006). Artifacts caused by simplicial subdivision. *IEEE Transactions on Visualization and Computer Graphics, 12*(2), 231–242.

Chen, J., He, Y., Frey, E. C., Li, Y., & Du, Y. (2021). ViT-v-net: Vision transformer for unsupervised volumetric medical image registration. In *Medical imaging with deep learning*.

Chen, J., Frey, E. C., He, Y., Segars, W. P., Li, Y., & Du, Y. (2022). Transmorph: Transformer for unsupervised medical image registration. *Medical Image Analysis, 82*, 102615.

Chen, M., Lang, A., Ying, H. S., Calabresi, P. A., Prince, J. L., & Carass, A. (2015). Analysis of macular OCT images using deformable registration. *Biomedical Optics Express, 5*(7), 2196–2214.

Chen, M., Carass, A., Jog, A., Lee, J., Roy, S., & Prince, J. L. (2017). Cross contrast multi-channel image registration using image synthesis for MR brain images. *Medical Image Analysis, 36*, 2–14.

Christensen, G. E. (1994). *Deformable shape models for anatomy*. Washington University. in St.

Chvátal, V. (1975). A combinatorial theorem in plane geometry. *Journal of Combinatorial Theory, Series B, 18*(1), 39–41.

Dalca, A. V., Balakrishnan, G., Guttag, J., Sabuncu, M. R. (2018). Unsupervised learning for fast probabilistic diffeomorphic registration. In *International conference on medical image computing and computer-assisted intervention* (pp. 729–738). Springer.

Dalca, A. V., Balakrishnan, G., Guttag, J., & Sabuncu, M. R. (2019). Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. *Medical Image Analysis, 57*, 226–236.

Fischl, B. (2012). Freesurfer. *NeuroImage, 62*(2), 774–781.

Haber, E., & Modersitzki, J. (2004). Numerical methods for volume preserving image registration. *Inverse Problems, 20*(5), 1621.

Haber, E., & Modersitzki, J. (2007). Image registration with guaranteed displacement regularity. *International Journal of Computer Vision, 71*(3), 361–372.

Heinrich, M. P., Maier, O., & Handels, H. (2015). Multi-modal multi-atlas segmentation using discrete optimisation and self-similarities. *VISCERAL Challenge@ ISBI, 1390*, 27.

Hering, A., Hansen, L., Mok, T. C., Chung, A. C., Siebert, H., Häger, S., Lange, A., Kuckertz, S., Heldmann, S., Shao, W., & Vesal, S. (2022). Learn2Reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning. *IEEE Transactions on Medical Imaging, 42*, 697–712.

Holland, D., Dale, A. M., Initiative, Alzheimer's Disease Neuroimaging. (2011). Nonlinear registration of longitudinal images and measurement of change in regions of interest. *Medical Image Analysis, 15*(4), 489–497.

Hoopes, A., Hoffmann, M., Fischl, B., Guttag, J., & Dalca, A. V. (2021). Hypermorph: Amortized hyperparameter learning for image registration. In *International conference on information processing in medical imaging* (pp. 3–17). Springer.

Johnson, H. J., & Christensen, G. E. (2002). Consistent landmark and intensity-based image registration. *IEEE Transactions on Medical Imaging, 21*(5), 450–461.

Kabus, S., Klinder, T., Murphy, K., van Ginneken, B., Lorenz, C., & Pluim, J. P. (2009). Evaluation of 4D-CT lung registration. In *International conference on medical image computing and computer-assisted intervention—MICCAI 2009* (vol. 5761, pp. 747–754). Springer.

Kim, B., Kim, D. H., Park, S. H., Kim, J., Lee, J. G., & Ye, J. C. (2021). CycleMorph: Cycle consistent unsupervised deformable image registration. *Medical Image Analysis, 71*, 102036.

LaMontagne, P. J., Benzinger, T. L., Morris, J. C., Keefe, S., Hornbeck ,R., Xiong, C., Grant, E., Hassenstab, J., Moulder, K., Vlassenko, A. G., & Raichle, M. E. (2019). OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease. *MedRxiv*.

Liu, Y., Zuo, L., Han, S., Xue, Y., Prince, J. L., & Carass, A. (2022). Coordinate translator for learning deformable medical image registration. In: *International workshop on multiscale multimodal medical imaging* (pp. 98–109). Springer.

Modat, M., Ridgway, G. R., Taylor, Z. A., Lehmann, M., Barnes, J., Hawkes, D. J., Fox, N. C., & Ourselin, S. (2010). Fast free-form deformation using graphics processing units. *Computer Methods and Programs in Biomedicine, 98*(3), 278–284.

Mok, T. C., & Chung, A. (2020). Fast symmetric diffeomorphic image registration with convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition* (pp. 4644–4653).

Narayanan, R., Fessler, J. A., Park, H., & Meyer, C. R. (2005). Diffeomorphic nonlinear transformations: A local parametric approach for image registration. In $19^t h$ *Information proceedings in medical imaging (IPMI 2005)* (pp. 174–185). Springer.

O'Rourke, J. (1987). *Art gallery theorems and algorithms* (Vol. 57). Oxford University Press.

Pai, A., Sporring, J., Darkner, S., Dam, E. B., Lillholm, M., Jørgensen, D., Oh, J., Chen, G., Suhy, J., Sørensen, L., & Nielsen, M. (2016). Deformation-based atrophy computation by surface propagation and its application to Alzheimer's disease. *Journal of Medical Imaging, 3*(1), 014005.

Qiu, H., Qin, C., Schuh, A., Hammernik, K., & Rueckert, D. (2021). Learning diffeomorphic and modality-invariant registration using b-splines. In *Medical imaging with deep learning*.

Rohlfing, T., Maurer, C. R., Bluemke, D. A., & Jacobs, M. A. (2003). Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. *Transactions on Medical Imaging, 22*(6), 730–741.

Yushkevich, P. A., Avants, B. B., Das, S. R., Pluta, J., Altinay, M., Craige, C., Alzheimer's Disease Neuroimaging Initiative. (2010). Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: an illustration in ADNI 3T MRI data. *NeuroImage, 50*(2), 434–445.