G-VR: INVERTIBLE GEOMETRIC VOLUME REPRESENTATION FOR LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Investigating geometric structure representation is key to explainable and controllable learning. We introduce G-VR, an invertible geometric representation for 3D volume data that enables lossless reversible encoding and exact restoration. Our framework operates on a standard 3D uniform grid, where the associated scalar field, whether voxel intensities from medical scans or occupancy ratios from solid geometries, serves as the target measure. Grounded in optimal transport and quasiconformal mappings on three-manifolds, G-VR encodes each volume into a canonical geometric signature: a quasiconformal deformation field that captures the intrinsic geometric structure of the data. Together with the extrinsic total mass of the volume, this representation guarantees exact restoration of both the geometry and the original scalar field. The representation naturally preserves geometric structure under interpolation and manipulation, while ensuring full invertibility. Experiments on both image-based and geometry-based volumes validate its fidelity, reversibility, and effectiveness in interpolation.

1 INTRODUCTION

Volume data are pervasive in medical imaging and scientific simulations, encompassing both image-based volumes (regular voxel grids from CT/MRI) and geometry-based volumes (unstructured tetrahedral meshes from CAD or FEM). In learning systems, these two modalities are typically handled by disparate architectures, 3D CNNs for voxel grids (Çiçek et al., 2016; Milletarì et al., 2016) and graph neural networks for meshes (Farazi et al., 2023; Deshpande et al., 2024), leading to modality-specific solutions that lack a unified geometric understanding. A critical limitation is that most existing approaches neglect the intrinsic geometric structure of the volume, such as shape morphology, spatial deformation, and topological features, all of which are essential for explainable and controllable learning. This gap poses fundamental challenges for developing representations that can faithfully encode and restore volumetric data without information loss.

Several recent studies have sought to incorporate geometric structure into volume data processing. One line of work focuses on shape-aware representations, extracting geometric features such as curvature, normals, or signed distance functions from volumetric inputs to augment learning tasks (Xue et al., 2020; Li et al., 2020; Bogensperger et al., 2025). Another line leverages deformation-based embeddings, encoding volumes via diffeomorphic or registration-based mappings that capture spatial deformations as geometric priors (Beg et al., 2005; Dalca et al., 2019; Fleishman & Fletcher, 2025). A third line employs mesh-based or point-based proxies, converting volume data into surface representations to enable geometric deep learning on volumetric domain (Wickramasinghe et al., 2020; Kong et al., 2021; Teng et al., 2025). Despite these efforts, most existing approaches treat geometric structure as an auxiliary feature or an intermediate representation rather than as a first-class citizen in the encoding process. Furthermore, they are predominantly modality-specific, designed for either regular voxel grids or tetrahedral meshes, but not both, and critically, rarely offer invertibility, limiting their applicability in tasks that demand faithful reconstruction or reversible data manipulation.

To address these limitations, we introduce G-VR (Geometric Volume Representation), a framework that encodes 3D volume data into a canonical geometric representation, a quasiconformal deformation field on a uniform grid. The key insight is to treat the volume's density distribution, whether voxel intensities from medical scans or occupancy ratios from solid geometries, as a target measure on a standard 3D uniform grid, and seek an optimal transport map from the uniform measure to this

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

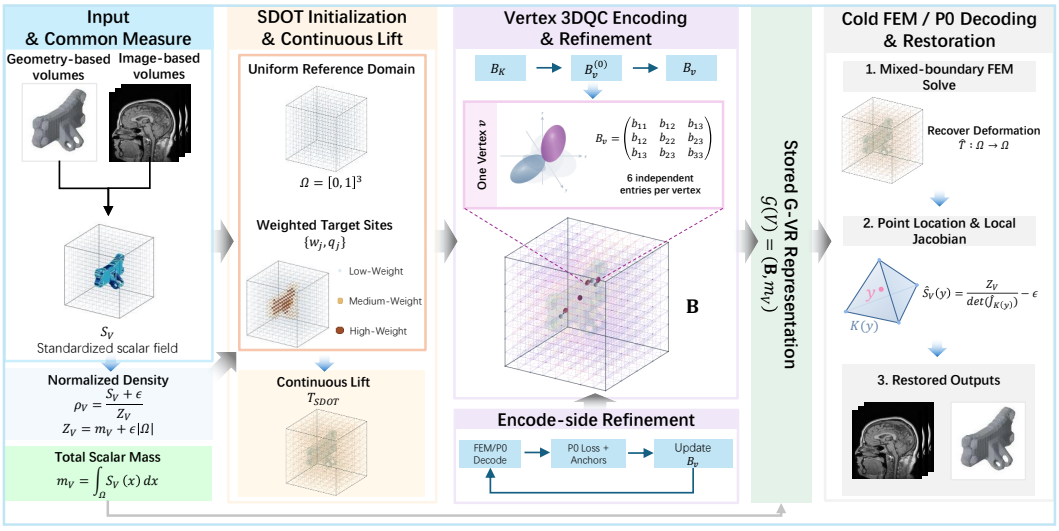


Figure 1: Overview of G-VR. Image- and geometry-based volumes are first converted to a common scalar measure. SDOT provides an orientation-preserving initialization, which is encoded and refined as a vertex 3DQC field. The stored G-VR representation is then decoded by the cold FEM/P0 pipeline to restore the original volumetric data.

target (Brenier, 1991). The resulting map, constrained to be quasiconformal (Zhang et al., 2022), describes how the uniform density must be deformed to match the volume's actual distribution. The deformation field itself serves as the intrinsic geometric signature of the volume, analogous to how conformal parameterization encodes a surface's geometry into a canonical domain. To enable complete restoration, this intrinsic signature is complemented by the extrinsic total mass of the volume, a single scalar that encodes the overall scale of the density field. Together, the intrinsic deformation field and the extrinsic total mass form a complete representation: the original volume can be exactly restored by first applying the inverse quasiconformal map to recover the normalized density distribution, then rescaling by the total mass to recover the original scalar field. This representation is geometric by construction: any operation performed in this canonical space, whether interpolation, generation, or manipulation, directly operates on the intrinsic geometry of the data. The formulation further guarantees invertibility: the original volume can be exactly restored from this geometric signature via the inverse quasiconformal map, ensuring lossless encoding.

The core idea is to characterize each volume by the geometric structure of a deformation map derived from semi-discrete optimal transport (SDOT) (Lévy, 2015). Intuitively, we treat the density at each point as volume mass and compute an optimal transport map that pushes the uniform reference measure onto the volume-induced target measure. We formulate this map as a quasiconformal deformation, whose 3D Beltrami coefficients, i.e., a 3D quasiconformal (3DQC) tensor field (Chen & Lui, 2025), serve as a compact geometric signature encoding both local shape distortions and global spatial organization. This signature, together with the exterior total mass, uniquely determines the original volume and enables exact, lossless restoration. Figure 1 illustrates the G-VR pipeline, demonstrating how both image-based and geometry-based volumes are encoded into a unified geometric representation along with intermediate visulizations. Unlike conventional approaches that process voxel grids and tetrahedral meshes through disparate architectures, G-VR provides a single representation form that supports invertible encoding and consistent downstream learning across heterogeneous volumetric modalities. *In contrast to methods that perform reconstruction from incomplete inputs, G-VR performs exact restoration from a complete geometric representation.*

A key technical challenge is to ensure that the resulting mappings are bijective and orientation-preserving (diffeomorphic in the smooth setting), so that the encoding is geometrically valid and the decoding is well-defined. To this end, we constrain the Beltrami coefficients of the optimal transport map on a 3D uniform grid to satisfy quasiconformality conditions, enforced through an L_2 SDOT solve, a shared-template lift, and element-wise Jacobian positivity checks. To account for numerical discretization errors, we further introduce optimization modules that enforce the fidelity of both the

measure functions and the Beltrami representation, via a vertex-based 3DQC refinement with a P0 reconstruction loss and map and coefficient anchors, keeping the encoding-decoding pipeline robust and accurate.

In summary, the main contributions are summarized as follows:

- **Invertible encoding with theoretical guarantees.** A mechanism grounded in optimal transport and quasiconformal mapping that captures intrinsic geometric structure while ensuring exact, lossless restoration. The formulation operates uniformly on any scalar field defined on a 3D grid, whether from image intensities or occupancy ratios, naturally accommodating heterogeneous volumetric inputs without modality-specific adaptation.
- **Canonical and structure-preserving geometric representation.** G-VR encodes each volume into a quasiconformal deformation field, a geometric signature that is itself the structure of the data, not a learned latent code. Operations in this canonical space naturally preserve geometry, as the representation and the geometry are one and the same.
- **Efficient empirical validation.** Comprehensive experiments on diverse volumetric datasets demonstrate high-fidelity restoration, reversibility, and effectiveness in interpolation tasks.

2 Related Work

Geometric Representations for 3D Data. Geometric representations are fundamental to 3D data understanding. A wide range of representations have been developed, including point clouds (Qi et al., 2017a), polygonal meshes (Hanocka et al., 2019), voxel grids (Wu et al., 2015), and implicit neural representations such as Occupancy Networks (Mescheder et al., 2019), DeepSDF (Park et al., 2019), and NeRF (Mildenhall et al., 2021). Among these, voxel grids and implicit representations are inherently volumetric, encoding information throughout the interior rather than merely on surfaces. However, voxel-based methods are typically limited by memory scalability (Liu et al., 2019), while implicit neural representations require per-scene optimization or large-scale pretraining and lack explicit geometric interpretability and invertibility. Geometric representations have also been studied for 2D images (Chen et al., 2026), though volumetric data presents distinct challenges in three-manifolds with interior structures and heterogeneous modalities. Recent work has introduced explicit quasiconformal representations for 3D bijective mappings (Chen & Lui, 2025). G-VR distinguishes itself by using such geometric map descriptors to encode the interior structure of volumetric data itself, rather than merely representing a given deformation map.

Volume Data Processing in Deep Learning. Volume data in learning systems are predominantly processed in two disjoint tracks. For image-based volumes (e.g., CT/MRI), standard practice employs 3D convolutional neural network (Çiçek et al., 2016; Milletarì et al., 2016) operating on regular voxel grids. For geometry-based volumes (e.g., tetrahedral meshes), learning pipelines typically leverage graph neural networks (Farazi et al., 2023; Pfaff et al., 2020) or point-based architectures (Qi et al., 2017b) on irregular domains. Hybrid approaches have begun to bridge different representation families, for example by combining point- and voxel-based processing (Liu et al., 2019) or coupling volumetric encoders with mesh decoders (Wickramasinghe et al., 2020). However, these methods still rely on representation-specific processing pathways and do not define a common geometric encoding shared across heterogeneous volumetric modalities. G-VR operates on a standard 3D uniform grid with scalar field measures, whether from image intensities or occupancy ratios, naturally accommodating both modalities without modality-specific adaptation.

Invertible Representations and Geometric Mappings. Invertibility has been explored in various 3D data processing contexts. Normalizing flows (Dinh et al., 2017) and invertible neural networks (Ardizzone et al., 2018) learn bijective mappings between data and latent spaces. In medical imaging, diffeomorphic registration methods (Beg et al., 2005; Dalca et al., 2019) explicitly model invertible deformations between volumes. More recently, invertible neural architectures have also been explored for 3D shape representation and reconstruction (Paschalidou et al., 2021).

Beyond explicit invertible architectures, optimal transport (OT) has been widely used to define continuous, measure-preserving mappings between probability measures, with applications in generative modeling (Rubenstein et al., 2018) and domain adaptation (Courty et al., 2017). Meanwhile, quasiconformal (QC) mappings have been employed in shape analysis and surface parameterization

for their ability to capture geometric distortions via compact Beltrami coefficient representations (Lui et al., 2012). While both OT and QC mappings are compatible with invertible formulations. OT maps are measure-preserving and can be bijective under proper conditions (e.g., strictly convex cost and absolutely continuous measures), while QC mappings are homeomorphisms by definition. However, their combination for the canonical geometric representation of heterogeneous 3D volume data has not been explored.

G-VR bridges this gap. To the best of our knowledge, it is the first framework that encodes volumetric data into a canonical geometric representation, a quasiconformal deformation field that is itself the geometric structure of the data, rather than a latent feature from which geometry is only inferred. This representation guarantees lossless invertibility and exact restoration, while naturally preserving geometric properties under interpolation and manipulation.

3 METHOD

G-VR converts a nonnegative 3D scalar volume into a canonical geometric deformation via optimal transport, and represents its intrinsic geometry by a 3D quasiconformal tensor field together with its extrinsic total mass. Optimal transport constructs the canonical mapping; the quasiconformal tensor provides the geometric representation. Figure 1 illustrates the pipeline. The per-volume representation is

$$\mathcal{G}(V) = \big(\mathbf{B},\, m_V\big), \tag{1}$$

where $\mathbf{B} = (B_v)_{v\in\mathcal{X}}$ is the nodal 3DQC code, $\mathcal{X}$ is the vertex set of the template mesh, $B_v \in \mathbb{R}^{3\times3}_{\mathrm{sym},+}$ is the symmetric positive definite 3DQC tensor at vertex v, and m_V is the extrinsic total mass.

3.1 FROZEN PROTOCOL

All volumes share a single frozen protocol $\mathcal{P}$, which fixes the template, connectivity, face-normal boundary convention, density floor, and output-grid convention. These are specified once and never adapted per volume, so that encode and decode share the same discretization and the decoder is truly cold. The protocol is defined by:

- **Template.** A reference tetrahedral mesh with $n = 97$ vertices, shared across all volumes.
- **Boundary convention.** For each coordinate $\ell \in \{1, 2, 3\}$, the normal coordinate is fixed on the face pair $\Gamma^-_\ell, \Gamma^+_\ell$, while tangential components use the natural zero-flux condition.
- **Density floor.** A fixed $\epsilon > 0$ used in the standardization $\rho_V = (S_V + \epsilon)/Z_V$.
- **Output grid.** A fixed 64^3 query grid with fixed axis ordering and spacing.
- **SDOT discretization.** A 16^3 measure discretization for semi-discrete optimal transport.
- **Continuation schedule.** A fixed $2 \times 8 + 4 \times 8$ continuation schedule.
- **Lift rule.** A fixed 12-neighbor moving-least-squares / Shepard lift for vertex initialization.
- **FEM/P0 rule.** Linear tetrahedral FEM for decoding, with P0 reconstruction at query points.
- **Floors.** Fixed thresholds δ_a (initialization), δ (optimization), and η (nodal eigenvalue).
- **Decoder audits.** The decoder outputs $\min \det J$, NegJ, and the tensor consistency E_B as certificate quantities.
- **Evaluation metrics.** SSIM, RMSE, HFEN, and PSNR are used for scalar fidelity in experiments.

The serialized representation consists of the nodal code $\mathbf{B}$ and the scalar m_V, together with the fixed protocol $\mathcal{P}$. No sample-specific shape or preprocessing fields are stored. A cold decode therefore reads only $\mathcal{G}(V)$ and $\mathcal{P}$.

3.2 OVERALL ALGORITHM

Algorithm 1 summarizes the full encoding and decoding pipeline. Encode and Decode share the template, FEM assembly, point-location rule, and P0 decoder. All constants, grids, schedules, and floors are fixed by the protocol $\mathcal{P}$ in Section 3.1. Figure 1 visualizes the intermediate outputs of each stage. The following subsections detail each step.

Algorithm 1 G-VR encoding and decoding

Input: volume V, frozen protocol $\mathcal{P}$
Output: code $\mathcal{G}(V) = (\mathbf{B}, m_V)$ and decoder audits

Encode:
Standardize V: compute S_V, m_V, Z_V, and $\{\omega_j, q_j\}$.
Run L2 SDOT; lift $q_j - c_j$ with MLS/Shepard and reject infeasible lifts.
Compute J_K, $B(J_K)$, T_a, and the volume-weighted $B_v^{(0)}$.
Refine Cholesky parameters with the three terms in Eq. (5); keep feasible finite-step candidates and return $\mathcal{G}(V)$ plus metadata.
Decode:
Construct $\mathcal{B}_h$ from $\mathbf{B}$ and solve the three FEM problems in Eq. (6).
for each query point y **do**
    Locate $K(y)$, form $\widehat{J}_{K(y)}$, and set $\widehat{S}_V(y) = Z_V / \det \widehat{J}_{K(y)} - \epsilon$.
end for
Return scalar metrics, E_B, $\min \det J$, NegJ, and solver/representation certificates.

3.3 Canonical Mapping by Semi-Discrete Optimal Transport

We treat the volume as a nonnegative scalar field S_V on the unit cube Ω, where $S_V(y)$ denotes the intensity (for image-based volumes) or occupancy ratio (for geometry-based volumes) at voxel y. We introduce a density floor $\epsilon > 0$ to ensure a strictly positive target measure, and define

$$m_V = \int_\Omega S_V(y)\, dy, \qquad Z_V = m_V + \epsilon|\Omega|, \qquad \rho_V(y) = \frac{S_V(y) + \epsilon}{Z_V}, \tag{2}$$

where the scalar m_V is the extrinsic total mass, Z_V the mass after adding the floor, and ρ_V the normalized density satisfying $\rho_V > 0$ and $\int_\Omega \rho_V = 1$. The density ρ_V encodes the intrinsic distribution of mass, and is the target measure for optimal transport.

We then seek a map $T : \Omega \to \Omega$ that pushes the uniform reference measure onto ρ_V. We solve this with semi-discrete optimal transport (Gu et al., 2016; Mérigot, 2011; Lévy, 2015), obtaining an orientation-preserving initialization T_{SDOT} of the canonical mapping. This step fixes the global correspondence between the uniform grid and the volume's mass distribution, and provides the anchor for all subsequent geometric encoding.

3.4 Encoding by 3D Quasiconformal Representation

Given the optimized canonical mapping T, let DT denote its spatial Jacobian. The orientation-safe continuation rejects any candidate unless $\det DT \geq \delta_a > 0$ on every template tetrahedron. For each element $K \in \mathcal{K}$, let $J_K = DT_{\mathrm{SDOT}}|_K$. We encode the local deformation by the 3D quasiconformal tensor

$$B_K = B(J_K) = \det(J_K)\,(J_K^\top J_K)^{-1}, \tag{3}$$

where $B(J) = \det(J)(J^\top J)^{-1}$ for any Jacobian J. This tensor form follows the standard 3D quasiconformal representation; see (Chen & Lui, 2025) for a detailed treatment. The tensor B_K is symmetric positive definite for an orientation-preserving nonsingular deformation. Its eigenvalues characterize the local anisotropic deformation, while its eigenvectors give the corresponding principal directions. It is therefore a geometric signature of the volume, not an unconstrained learned latent code. This is the sense in which G-VR is geometric by construction.

3.5 Refinement of the Nodal 3DQC Field

The element-wise tensor B_K is a faithful local descriptor, but the decoder operates on a nodal field. Wetherefore lift B_K to a vertex code $\mathbf{B} = (B_v)_{v \in \mathcal{X}}$ and refine it so that the decoded map reproduces the target measure.

Initialization and interpolation. The initial nodal code is obtained by volume-weighted projection, and the vertex code is linearly lifted with barycentric functions λ_v to a continuous field $\mathcal{B}_h$:

$$B_v^{(0)} = \frac{\sum_{K\ni v} |K|\, B_K}{\sum_{K\ni v} |K|}, \qquad \mathcal{B}_h(x) = \sum_{v\in K} \lambda_v(x)\, B_v, \qquad \overline{B}_K = |K|^{-1} \int_K \mathcal{B}_h\, dx = \frac{1}{4} \sum_{v\in K} B_v, \tag{4}$$

where the sums run over tetrahedra incident to v. The quantity $\overline{B}_K$ is the element-averaged tensor assembled by the FEM decoder, and $\mathcal{B}_h$ is the continuous field it approximates. Both are used in the FEM system of Eq. (6).

Refinement objective. For each target voxel center y_i, let $K_i(\mathbf{B})$ denote its containing element in the mapped mesh, and define $J_K(\mathbf{B}) = D\mathcal{F}_h[\mathbf{B}]|_K$. The refinement combines a P0 reconstruction term, a map anchor toward the FEM map T_a, and a coefficient anchor toward $\mathbf{B}^{(0)}$:

$$\begin{aligned} \mathbf{B} = \arg\min_{\mathbf{B}} \; & \underbrace{\frac{1}{N_q} \sum_{i=1}^{N_q} \left[\frac{Z_V}{\det J_{K_i}(\mathbf{B})} - \epsilon - S_V(y_i) \right]^2}_{\mathcal{L}_{\mathrm{P0}}(\mathbf{B})} \\ & + \lambda_{\mathrm{map}} \frac{1}{3N_v h^2} \sum_{v\in\mathcal{X}} \|\mathcal{F}_h[\mathbf{B}]_v - T_{a,v}\|_2^2 + \lambda_B \frac{1}{9N_v} \sum_{v\in\mathcal{X}} \|B_v - B_v^{(0)}\|_F^2. \end{aligned} \tag{5}$$

Candidates with non-positive sample Jacobians are rejected before evaluating $\mathcal{L}_{\mathrm{P0}}$. An accepted candidate must decrease the complete objective and satisfy $\det J_K \geq \delta > 0$. Since B_v is symmetric, it is fully determined by its six independent entries. The optimizer therefore uses six parameters per vertex through the Cholesky factorization $B_v = L_v L_v^\top$, where L_v is lower triangular and its diagonal is parameterized in the log domain. Candidate steps also pass the nodal eigenvalue floor $\lambda_{\min}(B_v) \geq \eta$. Gradients pass through the FEM solve using an adjoint system. The finite schedule is an approximate solve: the last accepted feasible code is saved, without a claim of global optimality. The final code $\mathbf{B}$ and m_V form $\mathcal{G}(V)$ with the fixed-format normalization and grid metadata.

3.6 Reversible Decoding by Mixed-Boundary FEM

Given the nodal code $\mathbf{B}$, let $\mathcal{B}_h$ denote its interpolated continuous field from Eq. (4). Let $\widehat{T} = (\widehat{T}^1, \widehat{T}^2, \widehat{T}^3)$. For each coordinate $\ell \in \{1,2,3\}$, define $\Gamma_\ell^- = \{x \in \partial\Omega : x_\ell = 0\}$ and $\Gamma_\ell^+ = \{x \in \partial\Omega : x_\ell = 1\}$. The decoder solves the following mixed-boundary FEM system for each ℓ:

$$\begin{cases} \nabla \cdot (\overline{B}\, \nabla \widehat{T}^\ell) = 0 & \text{in } \Omega, \\ \widehat{T}^\ell = 0 \text{ on } \Gamma_\ell^-, \quad \widehat{T}^\ell = 1 \text{ on } \Gamma_\ell^+, \quad n^\top \overline{B}\, \nabla \widehat{T}^\ell = 0 \text{ on } \partial\Omega \setminus (\Gamma_\ell^- \cup \Gamma_\ell^+), \end{cases} \tag{6}$$

where n is the outward unit normal and $\overline{B}|_K = \overline{B}_K$. Only the normal coordinate is fixed on each pair of faces; tangential components use the natural zero-flux condition. No sample-specific full boundary map is required. This is what makes the decoder cold and protocol-shared.

We solve the three systems with linear tetrahedral finite elements, and the P0 decoder recovers the scalar field at each query point through mass conservation:

$$\rho_V(T(x))\, \det DT(x) = 1, \qquad \widehat{\rho}_V(y) = \frac{1}{\det \widehat{J}_{K(y)}}, \qquad \widehat{S}_V(y) = Z_V\, \widehat{\rho}_V(y) - \epsilon, \tag{7}$$

where $K(y)$ is the containing deformed tetrahedron found by point location, and $\widehat{J}_K = D\widehat{T}|_K$. This is the discrete form of the mass-conservation relation, assuming that mapped elements do not overlap in their interiors. Thus the 3DQC tensor field captures the intrinsic deformation geometry, while the total mass preserves its extrinsic scalar magnitude, and the representation recovers the original scalar field under the stated invertibility conditions.

Decoder audits. The decoder outputs three certificate quantities: $\min \det J$, the minimum Jacobian determinant over all elements; NegJ, the number of elements with non-positive determinant; and the tensor consistency

$$E_B = \frac{\left(\sum_K |K|\, \|\overline{B}_K - B_K^{\mathrm{dec}}\|_F^2 \right)^{1/2}}{\left(\sum_K |K|\, \|B_K^{\mathrm{dec}}\|_F^2 \right)^{1/2} + \xi}, \tag{8}$$

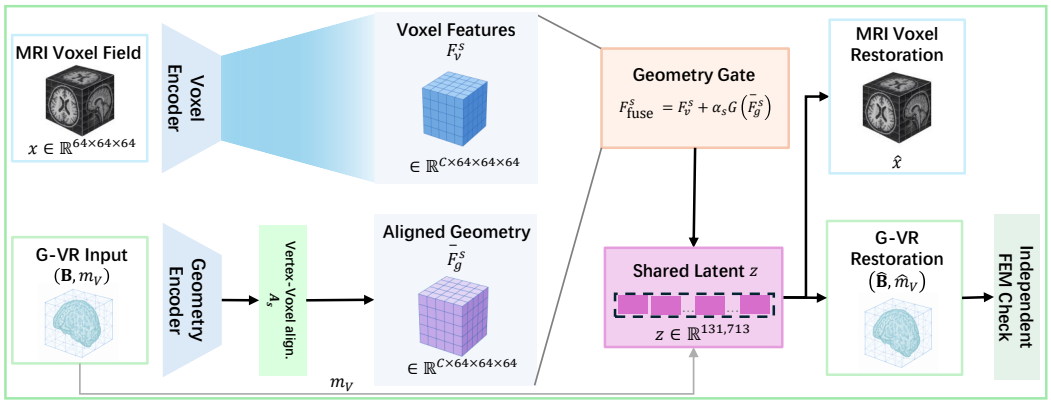


Figure 2: The G-VR-guided Hybrid receives aligned G-VR features through a middle-layer geometry gate, while sharing the same latent payload as the Voxel-only AE.

where $B_K^{\text{dec}} = B(D\widehat{T}|_K)$, $\xi > 0$ prevents division by zero, and $\|\cdot\|_F$ is the Frobenius norm. These quantities are produced by the decoder as part of its output. Element-wise positivity is not taken as a proof of global injectivity, and tensor consistency is evaluated separately from scalar fidelity.

4 Experiments

4.1 Experiment Setup

Datasets. We evaluate G-VR on image-based and geometry-based volume data. For *image-based volumes*, we use 100 subject-level OASIS-1 MR1 MRI volumes and 43 CT-ORG CT volumes. All image-based volumes are standardized and evaluated on a 64^3 grid. For *geometry-based volumes*, we use the verified SimJEB package, consisting of a 10-case development pilot and a 30-case unseen holdout (40 verified cases in total). All cases use the shared normalized-domain protocol and the same frozen vertex/FEM decoding route.

Baselines. We compare the G-VR restoration pipeline on the image-based and geometry-based cohorts with the same case-level protocol. For the learned models, we compare a Voxel-only autoencoder with a G-VR-guided Hybrid at matched capacity. Both models use the same latent payload of 131,713 FP16 values (263,426 bytes); the Voxel-only AE has 112,402 trainable parameters and the Hybrid has 114,034. The Hybrid incorporates the aligned G-VR features through a middle-layer geometry gate (see Figure 2).

Metrics. Scalar fidelity is measured by 3D-SSIM, RMSE, PSNR, and HFEN between the original and restored volumes. We additionally report the decoder audits $\min \det J$, NegJ, and the tensor consistency E_B, as defined in Section 3.6. For geometry-based volumes, we report Dice, surface Dice, ASSD, HD95, mapping safety, and topology status. For the learned Hybrid experiment, we also report parameter count, latent payload, and interpolation metrics.

Implementation Details. All volumes use the 64^3 grid. Restoration uses the Geogram semi-discrete OT bridge followed by the vertex-field and cold FEM decoder. Each case is checked for finite Jacobians, positive tetrahedral orientation, solver success, and the fixed certificate. MRI learning uses a 90/10 subject-level split, with the checkpoints fixed before evaluating the 10 test cases. Means and standard deviations are computed at case level.

4.2 Restoration

Note that we use the term restoration to distinguish our setting from reconstruction: G-VR recovers the original volume from its complete geometric representation via a deterministic inverse map, whereas reconstruction typically refers to inferring missing information from incomplete observations.

Quantitative Results. Table 1 reports restoration results and decoder audits on both data types. G-VR achieves near-lossless restoration under cold decoding, with positive $\min \det J$ in every reported case. All 100 MRI cases and all 43 CT-ORG cases pass the representation and geometric validity checks, and 37 of 43 CT-ORG cases pass the internal candidate-quality gate. On SimJEB, 38 of 40 cases pass the strict gate, all 40 cases are mapping-safe, and the unseen 30-case holdout reaches a 28/30 strict-gate pass rate.

Table 1: Restoration results and decoder audits on image-based and geometry-based volumes. NegJ is zero for all cases.

Dataset	N	3D-SSIM ↑	RMSE ↓	HFEN ↓	$\min \det J$ ↑
OASIS-1 MRI	100	0.998076 ± 0.001973	0.003067 ± 0.000854	0.016179 ± 0.002654	0.250019
CT-ORG	43	0.9905 ± 0.0289	0.00613 ± 0.00653	0.03144 ± 0.05393	0.2653 ± 0.0386
Dataset	N	Dice ↑	surface Dice ↑	ASSD ↓	HD95 ↓
SimJEB	40	0.998723 ± 0.000584	0.997700	0.22280	0.42586

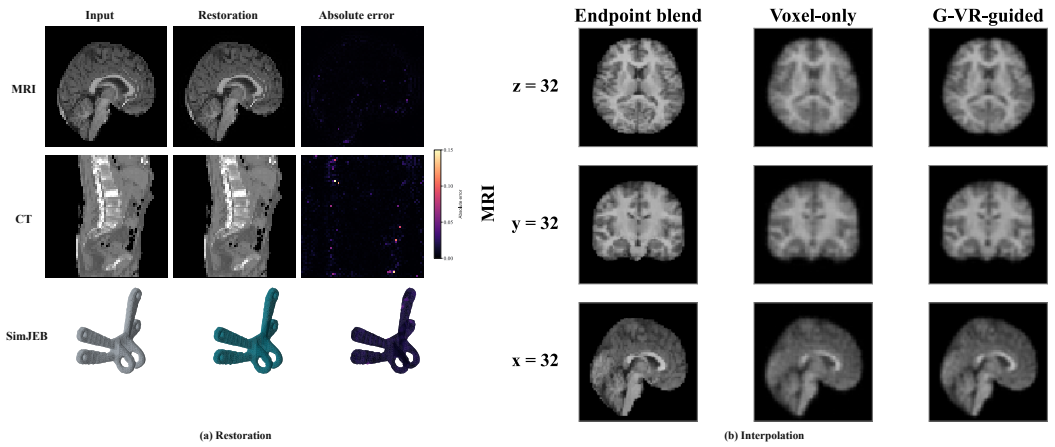


Figure 3: Qualitative results. **(a) Restoration:** MRI, CT, and SimJEB results, showing input, restoration, and absolute error or geometric discrepancy. **(b) Interpolation:** MRI latent interpolation comparing endpoint blending, Voxel-only interpolation, and G-VR-guided interpolation. Endpoint blending serves only as an artifact reference.

Qualitative Results. Figure 3 shows representative restoration results on both data types. The qualitative comparisons use the same intensity range and slice convention for MRI and CT. Each image panel shows the target, the G-VR restoration, and the absolute error. The geometry panel shows the SimJEB target, the restored geometry, and a $\det J$ or mapped-element visualization.

4.3 Interpolation

We evaluate latent interpolation on five fixed held-out MRI pairs at $t \in \{0.25, 0.5, 0.75\}$. The pairs are fixed across all methods to ensure a fair and reproducible comparison. The comparison uses the Voxel-only AE and the G-VR-guided Hybrid (Figure 2) with the matched payload. Both models are first trained for restoration, and interpolation is then performed in the learned latent space.

Quantitative Results. Table 2 reports path second-difference RMS, intermediate RMSE, HFEN, and local out-of-range rate. The G-VR-guided Hybrid reduces path second-difference RMS by 42.5%, intermediate RMSE by 9.7%, HFEN by 11.5%, and local out-of-range rate by 77.7%. Since both models share the same latent payload and have nearly the same parameter count, the improvement comes from the geometry branch rather than from a larger code budget.

Qualitative Results. Figure 3 results. Figure 3 shows a representative MRI interpolation at $t = 0.5$ in three orthogonal slices, comparing the endpoint blend, the Voxel-only AE, and the G-VR-guided Hybrid. The Voxel-only AE produces blurry intermediates whereas the G-VR-guided Hybrid produces

Table 2: Latent interpolation on five fixed held-out MRI pairs and three interior points per pair. Lower values are better. Both models use the same latent payload.

Method	Path 2nd-diff RMS ↓	Intermediate RMSE ↓	HFEN ↓	Local out-of-range ↓
Voxel-only AE	0.001018	0.02746	0.02361	8.77×10^{-5}
G-VR-guided Hybrid	0.000585	0.02480	0.02089	1.96×10^{-5}

smooth, anatomically plausible intermediates that preserve the shape and boundary structure of the endpoints. This improvement is consistent across the five held-out pairs and all three interior times.

4.4 Ablation Study

We use the matched-capacity Voxel-only AE as the core ablation of the G-VR-guided Hybrid. The two models share the same latent payload and nearly the same parameter count, while the Hybrid additionally receives aligned G-VR features through the geometry gate. Under this controlled model-size setting, the Hybrid improves 3D-SSIM from 0.93636 to 0.96610, reduces RMSE from 0.03307 to 0.02721 and HFEN from 0.17885 to 0.12159, and increases Dice from 0.64225 to 0.69704. It also gives lower path second-difference RMS, intermediate RMSE, high-frequency error, and local out-of-range rate for latent interpolation. Both models use the same latent payload, and their parameter counts differ by less than two percent.

4.5 Discussion

The experiments show that G-VR restores both image-based and geometry-based volumes with high fidelity and valid tetrahedral mappings, and that the same frozen protocol works across modalities.

Invertibility. Table 1 reports near-lossless restoration on MRI and CT under cold decoding, with positive $\min \det J$ and zero NegJ in every reported case. The residual error comes from numerical discretization rather than information loss, and the decoder audits confirm this with small E_B.

Structure preservation. Under a matched latent payload, the G-VR-guided Hybrid improves over the Voxel-only AE on all four interpolation metrics, and the matched-capacity ablation shows the same pattern. Since the representation is a quasiconformal deformation field, operations in the canonical space act directly on the geometry, so structure preservation needs no extra loss or constraint.

Unified representation. The same frozen protocol, template, and cold FEM/P0 decoder are used for MRI, CT, and SimJEB without modality-specific adaptation. On SimJEB, the restored geometry passes the strict gate in 38 of 40 cases, and all 40 cases are mapping-safe.

Limitations and future work. The current evaluation uses 64^3 grids and a fixed template with $n = 97$ vertices, and the quasiconformal projection may accumulate small errors under repeated encoding-decoding cycles. Future work will address these directions through more efficient optimal transport solvers, extensions to dynamic volumes, and tighter integration of the QC constraint.

5 Conclusion

In this work, we have presented G-VR, an invertible geometric representation framework for 3D volume data. Grounded in optimal transport and quasiconformal mappings on three-manifolds, G-VR encodes each volume into a canonical geometric signature, a quasiconformal deformation field on a uniform grid, which is itself the geometric structure of the data, rather than a latent feature from which geometry is only inferred. Together with the extrinsic total mass, this representation guarantees exact, lossless restoration. By construction, operations such as interpolation and manipulation in this canonical space preserve the underlying geometry, as the representation and the geometry are one and the same. Experiments on both image-based and geometry-based volumes validate the fidelity, reversibility, and effectiveness of G-VR in interpolation tasks. Future work will explore its potential for generative modeling and shape analysis, as well as extensions to time-varying volumetric data.

AI USE STATEMENT

Generative AI tools were used to assist with literature retrieval and discovery, research ideation and experimental workflow support, code and command generation for author-controlled validation, and the drafting and polishing of portions of the manuscript. All AI-assisted suggestions, code, citations, mathematical statements, experimental procedures, numerical results, and final wording were reviewed and verified by the authors. The authors take full responsibility for the final content of this work. Generative AI was not used as an author.

ETHICS STATEMENT

The MRI experiments use de-identified public OASIS data. The final submission will document the applicable dataset terms, preprocessing, subgroup composition, and limitations of clinical interpretation. The Bunny experiments use public geometric data and synthetic topology modifications.

REPRODUCIBILITY STATEMENT

The appendix records the solver discretization, grids, metrics, seeds, and the artifact paths underlying every number currently shown. The final anonymous supplement will include configuration files, preprocessing scripts, subject-disjoint splits, and commands sufficient to regenerate all tables and figures.

DATA AVAILABILITY

OASIS and the Stanford Bunny are publicly available subject to their respective terms. Derived splits, preprocessing metadata, and non-identifying experiment manifests will be released with the anonymous supplementary material.

REFERENCES

Lynton Ardizzone, Jakob Kruse, Sebastian Wirkert, Daniel Rahner, Eric W Pellegrini, Ralf S Klessen, Lena Maier-Hein, Carsten Rother, and Ullrich Köthe. Analyzing inverse problems with invertible neural networks. *arXiv preprint arXiv:1808.04730*, 2018.

M. Faisal Beg, Michael I. Miller, Alain Trouvé, and Laurent Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005. doi: 10.1023/B:VISI.0000043755.93987.AA.

Lea Bogensperger, Dominik Narnhofer, Alexander Falk, Konrad Schindler, and Thomas Pock. Flowsdf: Flow matching for medical image segmentation using distance transforms. *International Journal of Computer Vision*, 133(7):4864–4876, 2025.

Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on Pure and Applied Mathematics*, 44(4):375–417, 1991. doi: 10.1002/cpa.3160440402.

Qiguang Chen and Lok Ming Lui. A new geometric representation for 3d bijective mappings and applications, 2025.

Xin Chen, Qi Zhao, Wei Zeng, and Zongben Xu. G-ir: Geometric image representation for learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 20307–20315, 2026.

Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II*, pp. 424–432, Berlin, Heidelberg, 2016. Springer-Verlag. ISBN 978-3-319-46722-1. doi: 10.1007/978-3-319-46723-8_49. URL https://doi.org/10.1007/978-3-319-46723-8_49.

Nicolas Courty, Rémi Flamary, Devis Tuia, and Alain Rakotomamonjy. Optimal transport for domain adaptation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(9):1853–1865, September 2017. ISSN 0162-8828. doi: 10.1109/TPAMI.2016.2615921. URL https://doi.org/10.1109/TPAMI.2016.2615921.

Adrian V. Dalca, Guha Balakrishnan, John Guttag, and Mert R. Sabuncu. Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. *Medical Image Analysis*, 57: 226–236, 2019. ISSN 1361-8415. doi: https://doi.org/10.1016/j.media.2019.07.006. URL https://www.sciencedirect.com/science/article/pii/S1361841519300635.

Saurabh Deshpande, Stéphane P.A. Bordas, and Jakub Lengiewicz. Magnet: A graph u-net architecture for mesh-based simulations. *Engineering Applications of Artificial Intelligence*, 133:108055, 2024. ISSN 0952-1976. doi: https://doi.org/10.1016/j.engappai.2024.108055. URL https://www.sciencedirect.com/science/article/pii/S0952197624002136.

Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real nvp, 2017. URL https://arxiv.org/abs/1605.08803.

Mohammad Farazi, Zhangsihao Yang, Wenhui Zhu, Peijie Qiu, and Yalin Wang. Tetcnn: Convolutional neural networks on tetrahedral meshes. In *Information Processing in Medical Imaging: 28th International Conference, IPMI 2023, San Carlos de Bariloche, Argentina, June 18–23, 2023, Proceedings*, pp. 303–315, Berlin, Heidelberg, 2023. Springer-Verlag. ISBN 978-3-031-34047-5. doi: 10.1007/978-3-031-34048-2_24. URL https://doi.org/10.1007/978-3-031-34048-2_24.

Greg M Fleishman and P Thomas Fletcher. Lddmem: Large deformation diffeomorphic metric embedding: Decoupling shape analysis from image registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 323–332. Springer, 2025.

Xianfeng Gu, Feng Luo, Jian Sun, and Shing-Tung Yau. Variational principles for minkowski type problems, discrete optimal transport, and discrete monge-ampere equations. *Asian Journal of Mathematics*, 20(2):383–398, 2016. doi: 10.4310/AJM.2016.v20.n2.a7.

Rana Hanocka, Amir Hertz, Noa Fish, Raja Giryes, Shachar Fleishman, and Daniel Cohen-Or. Meshcnn: a network with an edge. *ACM Trans. Graph.*, 38(4), July 2019. ISSN 0730-0301. doi: 10.1145/3306346.3322959. URL https://doi.org/10.1145/3306346.3322959.

Fanwei Kong, Nathan Wilson, and Shawn Shadden. A deep-learning approach for direct whole-heart mesh reconstruction. *Medical image analysis*, 74:102222, 2021.

Bruno Lévy. A numerical algorithm for L_2 semi-discrete optimal transport in 3d. *ESAIM: Mathematical Modelling and Numerical Analysis*, 49(6):1693–1715, 2015. doi: 10.1051/m2an/2015055.

Shuailin Li, Chuyu Zhang, and Xuming He. Shape-aware semi-supervised 3d semantic segmentation for medical images. In *International conference on medical image computing and computer-assisted intervention*, pp. 552–561. Springer, 2020.

Zhijian Liu, Haotian Tang, Yujun Lin, and Song Han. *Point-voxel CNN for efficient 3D deep learning*. Curran Associates Inc., Red Hook, NY, USA, 2019.

Lok Ming Lui, Tsz Wai Wong, Wei Zeng, Xianfeng Gu, Paul M Thompson, Tony F Chan, and Shing-Tung Yau. Optimization of surface registrations using beltrami holomorphic flow. *Journal of scientific computing*, 50(3):557–585, 2012.

Quentin Mérigot. A multiscale approach to optimal transport. *Computer Graphics Forum*, 30(5): 1584–1592, 2011. doi: 10.1111/j.1467-8659.2011.02032.x.

Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3d reconstruction in function space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: representing scenes as neural radiance fields for view synthesis. *Commun. ACM*, 65(1):99–106, December 2021. ISSN 0001-0782. doi: 10.1145/3503250. URL https://doi.org/10.1145/3503250.

Fausto Milletarì, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, 2016. URL https://api.semanticscholar.org/CorpusID:206429151.

Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. Deepsdf: Learning continuous signed distance functions for shape representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

Despoina Paschalidou, Angelos Katharopoulos, Andreas Geiger, and Sanja Fidler. Neural parts: Learning expressive 3d shape abstractions with invertible neural networks. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3203–3214. IEEE, 2021.

Tobias Pfaff, Meire Fortunato, Alvaro Sanchez-Gonzalez, and Peter W Battaglia. Learning mesh-based simulation with graph networks. *arXiv preprint arXiv:2010.03409*, 2020.

Charles R. Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017a.

Charles R. Qi, Li Yi, Hao Su, and Leonidas J. Guibas. Pointnet++: deep hierarchical feature learning on point sets in a metric space. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, pp. 5105–5114, Red Hook, NY, USA, 2017b. Curran Associates Inc. ISBN 9781510860964.

Paul K Rubenstein, Bernhard Schölkopf, and Ilya O Tolstikhin. Wasserstein auto-encoders: Latent dimensionality and random encoders. In *ICLR (Workshop)*, 2018.

Lin Teng, Shen Zhao, Feng Shi, and Dinggang Shen. A curvature-guided diffeomorphic mesh deformation framework for lifespan brain cortical surface reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 14–23. Springer, 2025.

Udaranga Wickramasinghe, Edoardo Remelli, Graham Knott, and Pascal Fua. Voxel2mesh: 3d mesh model generation from volumetric data. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV*, pp. 299–308, Berlin, Heidelberg, 2020. Springer-Verlag. ISBN 978-3-030-59718-4. doi: 10.1007/978-3-030-59719-1_30. URL https://doi.org/10.1007/978-3-030-59719-1_30.

Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

Yuan Xue, Hui Tang, Zhi Qiao, Guanzhong Gong, Yong Yin, Zhen Qian, Chao Huang, Wei Fan, and Xiaolei Huang. Shape-aware organ segmentation by predicting signed distance maps. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 12565–12572, 2020.

Daoping Zhang, Gary PT Choi, Jianping Zhang, and Lok Ming Lui. A unifying framework for n-dimensional quasi-conformal mappings. *SIAM Journal on Imaging Sciences*, 15(2):960–988, 2022.

A Additional Experimental Protocol

All experiments use the normalized 64^3 volume protocol. MRI and CT inputs are evaluated as image-based volumes, while SimJEB inputs are evaluated as geometry-based volumes. The same vertex representation and cold FEM decoder are used across the reported cohorts.

Table 3: Datasets and evaluation protocols used in the experiments.

Dataset	N	Input type	Protocol
OASIS-1 MRI	100	Image volume	64^3 grid; subject-level split
CT-ORG	43	Image volume	64^3 grid; case-level evaluation
SimJEB	40	Geometry volume	10-case pilot and 30-case holdout

For the learned MRI models, 90 cases are used for training and the remaining 10 cases are held out for testing. The test cases are evaluated only after the training checkpoints are fixed. All reported means and standard deviations are computed at the case level.

B Geometric Validity Audits

In addition to image and geometry fidelity, we audit the decoded maps using the minimum tetrahedral Jacobian determinant, the number of negative-Jacobian tetrahedra (NegJ), certificate validity, and mapping safety. These checks evaluate whether the decoded piecewise-affine map preserves the orientation of the tetrahedral mesh.

Table 4: Geometric validity audits for the reported restoration experiments.

Dataset	N	Valid cases	det$J$ statistic / audit
OASIS-1 MRI	100	100/100	global min 0.250019; NegJ $= 0$
CT-ORG	43	43/43	mean min 0.2653 ± 0.0386; 37/43 gate
SimJEB	40	40/40	38/40 strict gate; 40/40 mapping-safe

For the learned Hybrid, an independent cold FEM audit is performed on the 10 held-out test cases. All 10 cases are valid, with minimum det$J = 0.38565$, mean minimum det$J = 0.46937$, NegJ $= 0$, and zero solver failures.

C Hybrid AE Configuration and Matched-Capacity Ablation

The Voxel-only AE and the G-VR-guided Hybrid use the same latent payload of 131,713 FP16 values (263,426 bytes). Their parameter counts differ by less than two percent. The Hybrid receives aligned G-VR features through a middle-layer geometry gate.

Table 5: Matched-capacity comparison between the Voxel-only AE and the G-VR-guided Hybrid on the held-out MRI cases.

Model	Params	3D-SSIM ↑	RMSE ↓	HFEN ↓	Dice ↑
Voxel-only AE	112,402	0.93636 ± 0.06234	0.03307 ± 0.00859	0.17885 ± 0.03363	0.64225 ± 0.24900
G-VR-guided Hybrid	114,034	$\mathbf{0.96610 \pm 0.00400}$	$\mathbf{0.02721 \pm 0.00481}$	$\mathbf{0.12159 \pm 0.00767}$	$\mathbf{0.69704 \pm 0.27075}$

At matched payload and nearly matched parameter count, the Hybrid improves all four reported restoration metrics. The comparison isolates the effect of the aligned G-VR features and the geometry gate without increasing the code size.

D Additional Latent Interpolation Results

Latent interpolation is evaluated on five fixed held-out MRI pairs at $t \in \{0.25, 0.5, 0.75\}$. Lower values indicate smoother paths and fewer local range violations.

The Hybrid reduces the path second-difference RMS by 42.5%, intermediate RMSE by 9.7%, high-frequency error RMS by 11.5%, and local out-of-range rate by 77.7%.

Table 6: Latent interpolation on five held-out MRI pairs and three interior points per pair.

Metric	Voxel-only AE	G-VR-guided Hybrid
Path second-difference RMS ↓	0.001018	**0.000585**
Intermediate RMSE ↓	0.02746	**0.02480**
High-frequency error RMS ↓	0.02361	**0.02089**
Local out-of-range rate ↓	8.77×10^{-5}	$\mathbf{1.96 \times 10^{-5}}$

E Reproducibility Details

The restoration metrics use the same case-level aggregation for MRI and CT. The geometry metrics use the normalized SimJEB coordinate system. The vertex layout, FEM connectivity, decoder settings, latent payload, and MRI subject split are fixed before evaluating the held-out cases. The supplementary materials contain the preprocessing configuration, split manifest, model configuration, and per-case audit records needed to reproduce the reported tables.