Volume xx (200y), Number z, pp. 1–11

OrganPhys: Scope Captures to CT-Informed, Mesh-Free, Physics Sim-Ready Deformables

Aravind S. Kumar* Vishwesh Nath* Andres-Diaz Pinto Vismay Modi
NVIDIA
* Equal contribution

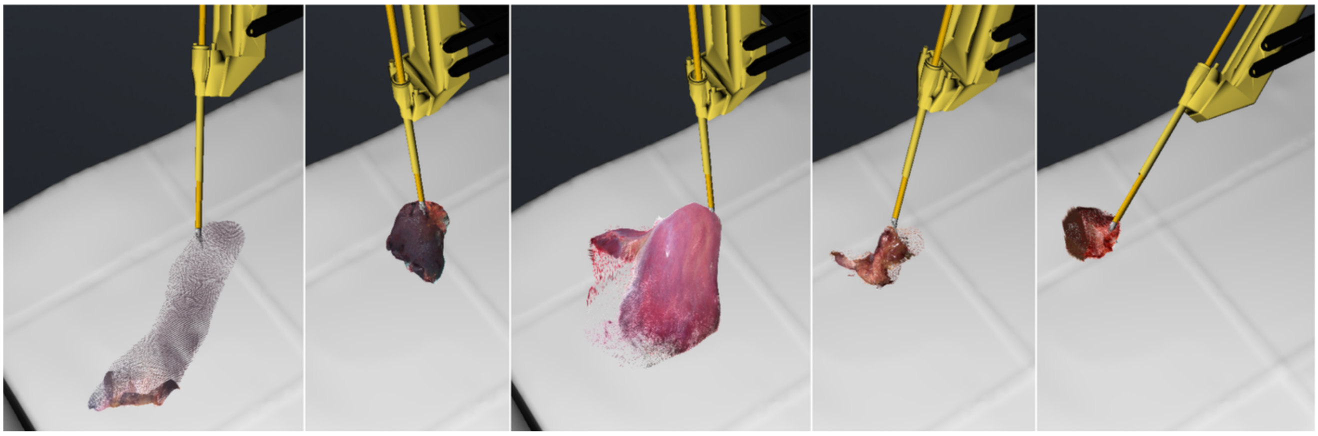

Figure 1: High fidelity deformable simulations requires realistic geometry. We bridge the gap between hyper-realistic 3DGS organ captures and accurate CT-grounded geometry. We show the results of our pipeline via interactive, physics-based deformations, on REAL organ geometry on five separate organs. The first on the far left, we show the gallbladder. Second, we show the spleen. Third, is the liver. Fourth, the pancreas. Fifth, we show the prostate. *Note: we register to the splat's visible surface and never hallucinate unobserved anatomy, so the visual gaps reflect unobserved regions.*

Abstract

Reconstructions of surgical scenes from endoscopic video are useful for surgical training and medical animation, as they accurately capture the clinician's viewpoint. In practice, however, 3D scenes acquired through endoscopy, increasingly represented as 3D Gaussian splats (3DGS), are typically constrained to limited viewpoint ranges. Consequently, even small endoscope camera motions can significantly degrade rendering fidelity and organ shape reconstruction. This gap prevents the direct use of such scenes in physics-based surgical simulation, which requires accurate, shape-consistent geometry. To address this, we propose a pipeline that combines the visual realism of an endoscopic 3DGS scene with simulation-ready organ geometry extracted from pre-operative 3D computed tomography (CT) or MRI scans. Because the pre-operative and endoscopically observed organs often differ in pose, deformation, and geometric completeness, we register the full segmented anatomy onto its partial corresponding region in the splat representation using a mesh-free, elastically regularized alignment method. We perform this alignment in the reduced-order deformation space of Simplicits, an existing state-of-the-art mesh-free simulator based on learned skinning weights, which has not previously been applied to surgical endoscope-captured organs because no CT-grounded alignment pipeline existed to bridge the two modalities. Using its reduced basis, we optimize on the order of 10–20 handle transforms rather than thousands of tetrahedral-mesh vertices. We employ a two-stage procedure: (1) a weighted rigid iterative closest point (ICP), (2) followed by a reduced-order elastic ICP, with a Neo-Hookean elastic energy term and a geometry matching term that uses nearest-neighbour correspondences to the 3DGS target. As shown in Figure 1, our pipeline produces CT-grounded, physics-simulatable organ assets that are visually aligned with the hyper-realistic tissue surfaces captured by endoscopic 3DGS reconstructions, closing the gap between photoreal scene capture and simulation-ready geometry.

CCS Concepts

•***Computing methodologies*** $\rightarrow$ ***Modeling and simulation; Physical simulation;*** *Volumetric models;*

1. Introduction

Minimally invasive surgery is increasingly recorded from the endoscope, and these recordings are used for scene-level reconstruction and visualizations for simulation and educational purposes.

Recently, 3D Gaussian Splatting (3DGS) [KKLD23] offers a practical route to novel-view synthesis and reconstruction for endoscope-captured scenes [LLYY24, XYC*24, YLS*24], but 3DGS reconstructions still suffer from artefacts due to limited endoscope-camera motion and constantly moving, specular tissue with occlusions. All these drawbacks weaken multi-view consistency, so organs are often noisy, incomplete surfaces that can still look plausible from some views. Even with convincingly realistic rendering from certain viewpoints, they remain geometrically inconsistent with the actual organ as only partial surfaces are visible. Figure 2 (liver) contrasts shell-like scope splats (right side).

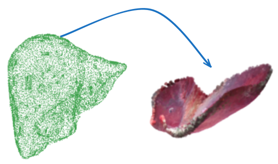

Figure 2: Dense CT organ geometry (left) versus raw splats (right). With limited scope-camera motion, splat reconstructions can look visually compelling yet remain geometrically thin and incomplete for simulation; we align pre-operative CT to splats so physics uses corrected volumetric shape while preserving splat hyper-realism.

Pre-operative CT Figure 2 (left side), by contrast, provides dense volumetric information and 3D segmentation masks from which organ shape can be extracted with well-understood deep learning approaches [WBM*23]. Bridging CT with an endoscopic reconstruction therefore offers a path to the "best of both worlds": photorealistic splat rendering of the scene together with a CT-grounded, simulatable organ model. The central obstacle is geometric alignment: the segmented organ in CT coordinates and the organ implied by splats in the endoscopic reconstruction generally differ in rigid pose, non-rigid deformation, and geometric completeness. As a result, the CT organ must be registered onto the partially visible splat organ before simulation or rendering can share a single coherent representation.

Classical physics-based registration [WMKW07, WJC*10] builds dense tetrahedral meshes from the CT volume and runs FEM-based iterative closest point incorporating an elastic term (Elastic ICP) [BM92]. Instead we adopt Simplicits [MSP*24], a mesh-free, reduced-order deformable model that represents the organ's deformation with a small set of handle transformations applied to the CT-scanned organ via a learned linear blend skinning map.

We present the first pipeline for directly registering an un-meshed CT organ onto a partial 3DGS endoscope reconstruction, yielding a simulation-ready surgical scene with accurate organ geometry. Specifically, our contributions are:

- A weighted, anatomy-aware, rigid registration that resolves the global pose ambiguity on partial geometry via randomized initializations.
- A mesh-free, reduced-order elastic registration formulation using Simplicits [MSP*24], which reduces the optimization from thousands of volumetric DOFs to a few dozen handles, enabling efficient convergence.

2. Related Work

Our method sits at the intersection of classical and elastic point-cloud registration, 3D Gaussian splatting for surgical scenes, pre-operative–to–intra-operative alignment, interactive surgical simulation pipelines, and reduced-order deformable modelling. We review each area below and summarize the landscape in Table 1.

2.1. ICP and Non-Rigid Registration

The iterative closest point (ICP) algorithm [BM92] alternates between establishing nearest-neighbour correspondences and solving for a rigid transform, converging to a local minimum of the mean-square distance. Chetverikov [CSK05] introduced trimmed ICP, which rejects a fraction of the worst correspondences at each iteration, making ICP robust to partial overlap—a property we rely on heavily given the limited endoscopic viewpoint. Amberg [ARV07] generalized ICP to nonrigid registration by introducing locally-affine deformation fields with a stiffness regularizer whose weight is annealed from high to low, yielding a coarse-to-fine deformation. Their stiffness annealing strategy directly inspires our elastic α-schedule. More recently, Semantic-ICP [CZS*25] adds label consistency and a linear elastic energy to ICP for multi-organ point cloud registration. Our work differs from these in two respects: we replace the surface-only, linear regularizer with a *volumetric Neo-Hookean hyperelastic energy* evaluated over the full organ interior, and we operate in a *reduced-order* handle space rather than per-vertex.

Most closely related in spirit is MorphModes [BLGB25], which likewise parameterizes non-rigid registration in a reduced skinning-eigenmode subspace [BZC*23] rather than per-vertex, but targets mesh-based, complete-to-complete shapes and drives alignment with a signed-distance-function (SDF) matching objective. We instead operate mesh-free, on point samples of the CT organ and the splat target, and solve the substantially harder complete-to-partial problem—registering a full CT volume onto a partial, endoscope-visible splat—through an ICP closest-point data term coupled to the volumetric Neo-Hookean prior rather than SDF matching.

In medical imaging, finite element method (FEM) models have long served as regularizers for deformable registration. Wittek [WMKW07] used patient-specific hexahedral FEM with Neo-Hookean tissue to predict brain shift during neurosurgery, and later demonstrated GPU-accelerated real-time performance [WJC*10]. Genet [GSv*18] formulated equilibrium-gap regularization for cardiac image registration. These methods share a common pattern: they drive registration with *image similarity metrics* (mutual information, SSD) on voxel grids, and require explicit volumetric meshing of the domain. Our approach instead combines a FEM-class elastic energy with an *ICP-style closest-point alignment term* on

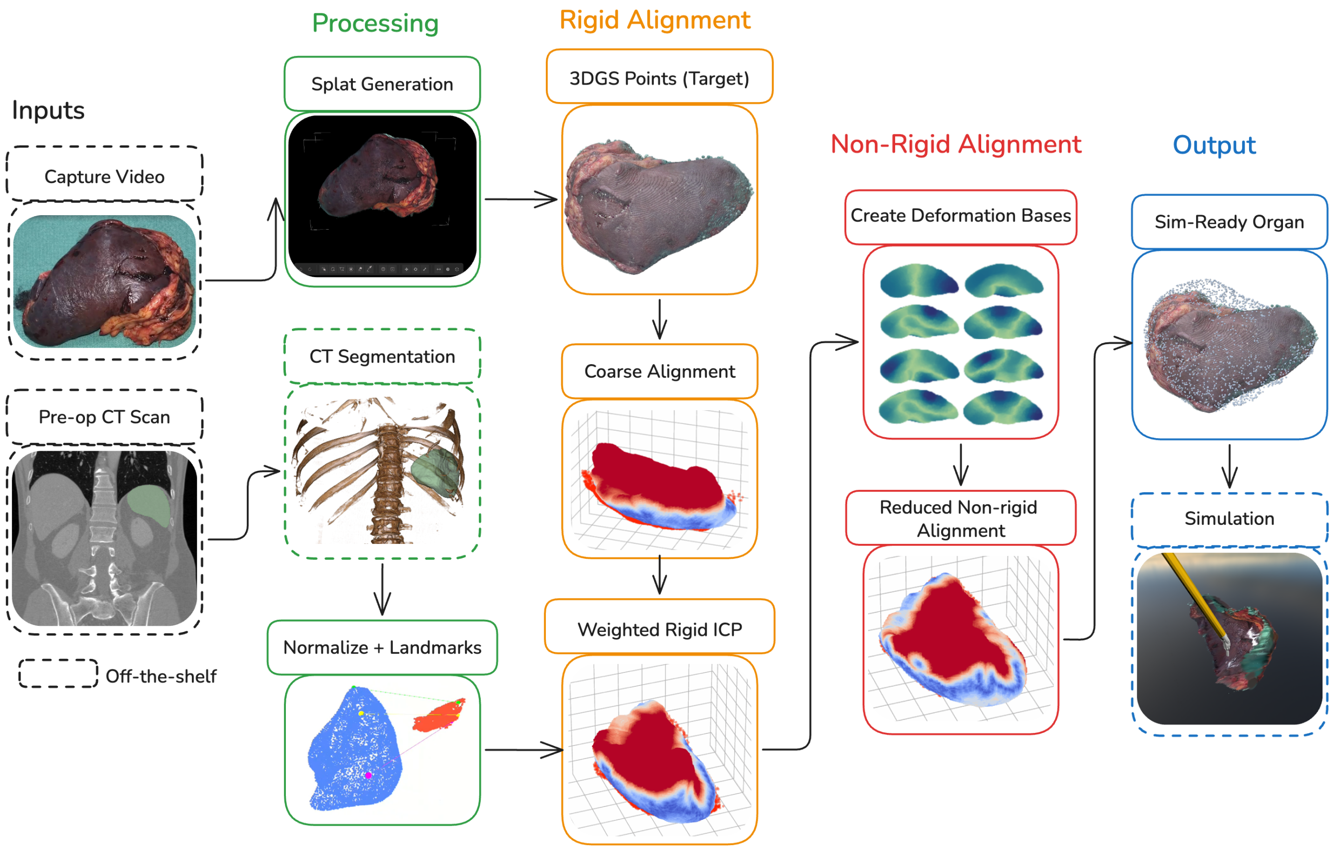

Figure 3: OrganPhys requires two inputs—an intra-operative endoscopic video and a pre-operative CT scan. G-SHARP reconstructs the endoscopic scene as a 3D Gaussian Splat whose Gaussian centres form the target point cloud (rendered in red in subsequent panels), while the CT volume is segmented, normalized, and tagged with anatomical landmarks to yield the source cloud (blue). A two-stage rigid cascade—*coarse landmark alignment* (Section 3.3) followed by *weighted rigid ICP* (Section 3.4)—deforming the full source organ into complete correspondence with the partial, endoscope-visible target. A reduced non-rigid stage then constructs *Simplicits deformation bases* (Section 3.5) and solves an *elastic ICP* (Section 3.6) in the resulting low-dimensional modal subspace, yielding a physics-consistent, simulation-ready organ registered to the endoscopic scene and ready for downstream interactive simulation. Dashed boxes denote off-the-shelf components; solid boxes mark our contributions.

point clouds, bridging the FEM registration and nonrigid ICP communities. Crucially, we avoid volumetric meshing entirely, for both physics and rendering, through the use of Simplicits [MSP*24] and 3D Gaussian Splatting.

2.2. Gaussian Splatting and Surgical Scenes

3D Gaussian Splatting (3DGS) [KKLD23] represents scenes as collections of anisotropic Gaussians with optimizable position, covariance, opacity, and spherical-harmonic colour, achieving real-time rendering via differentiable rasterization. Several methods have adapted 3DGS to endoscopic surgery: EndoGaussian [LLYY24] introduces holistic depth initialization and a spatio-temporal deformation field for dynamic tissue, achieving 195 FPS; SurgicalGaussian [XYC*24] adds forward-mapping MLPs and tool masking; Deform3DGS [YLS*24] achieves one-minute training with per-Gaussian deformation. On the registration side, GaussReg [CXL*24] aligns two Gaussian splat scenes via coarse point-cloud matching followed by image-guided refinement, but is limited to rigid or affine transforms between scenes of the same modality. None of these methods *register* a 3DGS reconstruction to a pre-operative CT segmentation, nor do they produce simulation-ready deformable geometry. Our novel pipeline uses the 3DGS Gaussian centres as the *target* point cloud for ICP, treating the splat scene as a fixed geometric reference against which the pre-operative CT organ is non-rigidly aligned.

2.3. Pre-Operative to Intra-Operative Registration

Bridging pre-operative imaging with intra-operative observations is a central challenge in surgical navigation. Traditional approaches use surface-based or landmark-based rigid registration followed by biomechanical FEM correction [WMKW07, WJC*10], but depend on explicit volumetric meshing and voxel-based similarity metrics. More recently, MIS-NeRF [BFJP*25] reconstructs intra-operative surfaces via NeRF and aligns them to CT with rigid ICP (3.25 mm

Table 1: Summary of pre-op. to intra-op. registration methods.

Method	Key idea
FEM-based [WMKW07, WJC*10]	Rigid + biomechanical FEM on tet-meshes
MIS-NeRF [BFJP*25]	NeRF reconstruction + rigid ICP to CT
PIVOTS [LGS*25]	Learned vol.-to-surf. displacement fields
Land-Reg [CPW*26]	Latent 2D–3D correspondences, rigid + non-rigid
Deep hashing [ZBD*25]	Retrieval-based automatic rigid initialization
Keyhole AR [ERB*24]	Tool-keyhole-aware tumour projection
BridgeSplat [FWH*25]	*Mesh-based*; assumes prior CT–video registration, then rigs splats to a CT mesh
Ours (Simplicits: Reduced, Mesh-free)	**Mesh-free non-rigid CT→partial-3DGS registration (Neo-Hookean prior)**

mean error), while PIVOTS [LGS*25] learns dense volume-to-surface displacement fields with cross-attention, trained on synthetic pneumoperitoneum deformations. Land-Reg [CPW*26] learns latent-grounded 2D–3D landmark correspondences for combined rigid and non-rigid laparoscopic liver registration, and deep hashing [ZBD*25] casts CT-to-video alignment as content-based image retrieval for automatic rigid initialization. Keyhole-aware AR [ERB*24] highlights that standard laparoscopic overlays misproject tumours toward the camera rather than the instrument keyhole, motivating the need for geometrically accurate organ models.

BridgeSplat [FWH*25] is the closest, publicly available method to ours. It assumes an initial CT–video registration, then meshes the CT object and embeds the splats to follow along as the CT deforms. In contrast, we solve the non-rigid CT-to-partial Splat registration problem in addition to embedded splat deformations in a completely mesh-free manner. Table 1 summarizes the key differences.

2.4. Elastic Simulation Methods

Interactive surgical simulation and navigation systems have historically relied on manually created organ geometries and mesh-based deformation pipelines to achieve real-time performance for training, rehearsal, or guidance.The more accurate, gold-standard approach to surgical simulation is Finite Element Method (FEM)-based elastodynamics, but it is computationally intensive and typically yields slow simulations ill-suited to tight interaction loops. Position-based dynamics is therefore the standard technique in interactive settings, but it does not accurately handle material parameters, which can yield implausible deformations despite high frame rates.

Reduced-order models project high-dimensional deformation spaces onto low-dimensional bases for efficient simulation. In physics-based simulation, modal analysis via eigenvalue decomposition is a classic approach to model reduction [BJ05]. Linear blend skinning (LBS), which is traditionally used for skeletal animation [KJP02,JT05,JBK*12], has emerged as an effective tool for reduced-order deformation. Fast Skinning Eigenmodes, uses eigenanalysis to create deformation bases from LBS weights, which are then blended with handle transforms to drive reduced-order elastic deformation of a tetrahedral mesh. Simplicits [MSP*24] generalizes the paradigm of using LBS weights to create a deformation basis to any geometric representation. It can create a reduced-order deformation basis for point clouds, meshes, signed distance fields, NeRFs, and Gaussian splats without volumetric meshing.

In our project, we avoid meshing the CT organ volume and drive non-rigid alignment of the CT organ onto the 3DGS target directly in the Simplicits reduced basis. The Simplicits methodology is applied in two independent stages of the pipeline, each with its own RKPM-trained basis: first on the normalised pre-registration CT organ to parameterise the elastic ICP, and then on the post-registration geometry to drive an interactive elastic simulation (see Section 3.5 for the dual-role breakdown). Our contribution is the registration pipeline itself, which is agnostic to the downstream simulator: it produces a single, consistent, simulation-ready surface and volume that mesh-based or mesh-free pipelines can consume in their native form. Enabled by this registration procedure, we then demonstrate Simplicits – a recent state-of-the-art mesh-free reduced-order simulator – on captured surgical organs for the first time, closing the loop between endoscopic 3DGS capture, CT-grounded geometry, and interactive physics.

3. Method

We present a two-stage registration pipeline that creates simulation-ready endoscope reconstructions with anatomically plausible geometric representations. The proposed method enables mesh-free, physics-informed registration of preoperative CT–derived organ geometry to partially observed endoscopic 3D Gaussian Splat (3DGS) reconstructions, ensuring geometrically consistent alignment under limited viewpoint coverage.

3.1. Overview

The pipeline operates on two inputs: (i) a segmented organ obtained from a preoperative 3D CT scan, and (ii) an endoscopic 3D Gaussian Splat (3DGS) reconstruction of the organ. It produces a splat-aligned, volumetric representation of the CT-derived geometry.

The inputs require pre-processing as outlined in section 3.2. Section 3.3 presents the coarse alignment stage, which establishes scale consistency between the splats and the anatomically accurate CT-derived organ segment. Section 3.5 describes the construction of a reduced deformation basis over the CT volume, together with the extraction of surface point samples using ball morphology. The overall optimization proceeds in two stages. First, 3.4 performs anatomy-aware weighted rigid ICP based on 1-2 manually selected landmarks. This is followed by 3.6, which refines the alignment through non-rigid optimization over the reduced Simplicits degrees of freedom by jointly minimizing an elastic energy term and a CT–splat surface matching term. Figure 3 illustrates the complete pipeline. We deform the complete CT geometry onto the partial splat—and not the reverse—so that an anatomically correct, simulatable volume is brought into the clinician's view.

3.2. Processing Inputs

3.2.1. Endoscope Scene Reconstruction

The intra-operative target geometry is obtained by reconstructing the endoscopic video as a 3D Gaussian Splat scene. The pipeline follows three stages. First, MedSAM3 [LXC*25] extracts individual RGB frames and generates per-frame binary tool-segmentation masks, isolating surgical instruments from tissue. In parallel, the MegaSAM [LTC*25] stack—comprising Depth-Anything [YKH*24] for monocular depth, UniDepth [PSY24] for metric scale, and DROID-SLAM [TD21] for camera tracking—produces per-frame metric depth maps and camera poses. Frames, masks, depth, and poses are assembled into an EndoNeRF-compatible [WLFD22] dataset bundle. Finally, G-Sharp [NTL*25] trains 4D Gaussians with a HexPlane-based temporal deformation field on this dataset, producing a 3DGS PLY file whose Gaussian centres define the target point cloud.

The Gaussian centres in the output PLY approximate the endoscope-visible surface of the organ; we use them as the fixed geometric target for registration. From this PLY we extract the target point set $\mathcal{T} = \{\vec{t}_j\}_{j=1}^{M}$ by retaining only Gaussians with opacity $o_j > 0.1$ and uniformly subsampling to $M = 5{,}000$ points, where M controls the density of the ICP target cloud. A KD-tree over $\mathcal{T}$ is built once and reused in all subsequent ICP stages.

3.2.2. 3D CT Segmentation & Pre-processing

Given a pre-operative CT volume with a multi-organ segmentation mask inferred using TotalSegmentator [WBM*23], we extract the target organ label and clean the binary mask with standard morphological post-processing (largest connected component - removing outliers, opening - for smoothness). From the cleaned segmentation mask we derive two CT-space point sets: **volumetric points** $\mathcal{V}$, one sample per labelled voxel, and **surface points** $\mathcal{S} \subset \mathcal{V}$, the subset on the segmentation boundary. These are distinct from the 3DGS target $\mathcal{T}$ (Section 3.2.1), which is derived from Gaussian splat positions rather than CT voxels. The organ volume is estimated as $V_{\text{organ}} \approx |\mathcal{V}| \cdot v_{\text{voxel}}$, where v_{voxel} is the per-voxel volume from the CT scan metadata.

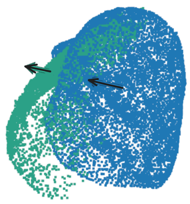

Figure 4: CT organ (blue) and 3DGS target (teal) with their respective front vectors. A Rodrigues rotation aligns the CT front (blue arrow, centroid to landmark midpoint) with the inferred viewer direction (teal arrow, pointing to the endoscopic camera), orienting the organ's visible side toward the camera.

Each of $\mathcal{S}$, $\mathcal{V}$, and the 3DGS target $\mathcal{T}$ is independently centred at its own centroid and scaled to a unit box by dividing by the maximum absolute coordinate. Because each point set is normalised independently, the coarse alignment stage (Section 3.3) must recover the relative scale, rotation, and translation between them.

3.3. Coarse Alignment

Coarse alignment resolves the global scale, rotation, and translation between the normalised CT and 3DGS point clouds via three steps.

Scale. A uniform scale factor s is estimated as the ratio of the splat oriented bounding box extent (PCA-aligned bounding box) to the CT projected diameter (convex-hull diameter in the plane perpendicular to the landmark–centroid axis) (Figure 5).

Primary rotation. The CT "front direction", shown in Figure 4 (from the organ centroid toward the *user-selected* landmark midpoint) is mapped onto the inferred viewer direction $-\hat{\mathbf{c}}_{\text{gs}}$ via a Rodrigues rotation formula.

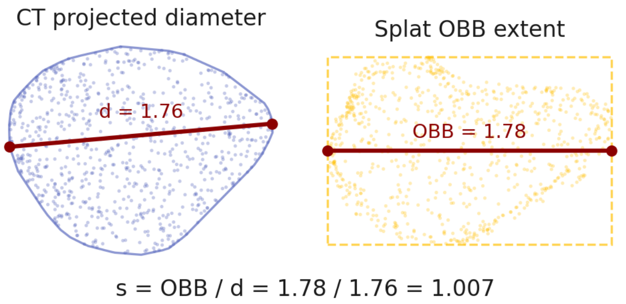

Figure 5: Scale estimation for coarse alignment. Left: the CT organ surface (indigo) is projected onto the plane perpendicular to the landmark–centroid axis, and the convex-hull diameter d is measured. Right: the 3DGS target (amber) is projected onto its PCA-aligned oriented bounding box, and the maximum extent is measured. The uniform scale factor s = OBB/d brings the CT into the same spatial extent as the splat target.

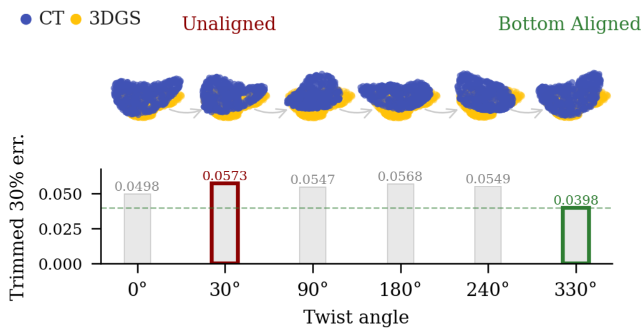

Figure 6: Coarse alignment twist search. The CT organ surface (indigo) is rotated about the viewing axis at six uniformly spaced angles and overlaid on the 3DGS target point cloud (amber). The bar chart shows the trimmed 30 % mean nearest-neighbour distance for each twist angle. The search identifies 330° as the best-aligned orientation (green), minimising the registration error before rigid ICP refinement. At 30° (crimson), the CT and target surfaces are maximally misaligned.

Twist search. The remaining rotational degree of freedom around the viewing axis is resolved by evaluating 12 uniformly spaced twist angles $\theta_k = 2\pi k/12$ and selecting the one minimizing the trimmed (30%) mean nearest-neighbour distance to the target splats $\mathcal{T}$ (Figure 6).

The coarse transform to initialize the is $\mathbf{p}_{\text{coarse}} = s(\mathbf{p}\mathbb{R}^{\top}) + \mathbf{t}$ for each source CT point $\mathbf{p} \in \mathcal{V}$.

3.4. Region-Weighted Rigid ICP

After coarse alignment, we refine the rigid pose using a region-weighted variant of trimmed ICP [BM92, CSK05]. The user selects one or two anatomical landmarks on the CT segmentation surface (e.g the fundus tip and neck of the gallbladder, or a flat landmark on the spleen); placing these requires some anatomical understanding so that the chosen region corresponds to what is visible endoscopically. From these, per-vertex weights $w_i \in [0,1]$ are computed as a raised-cosine (Hann) bell of the form $w_i = \frac{1}{2}\big(1+\cos(\pi d_i/r)\big)$ for $d_i < r$ and $w_i = 0$ otherwise, where d_i is the distance from surface point i to the nearest landmark and r is the inter-landmark distance. We choose the raised cosine rather than a Gaussian because it has *compact support* (exactly zero beyond r), so vertices on the endoscope-occluded back side of the organ contribute nothing to the weighted SVD; its single parameter r is fixed by the geometry of the user-selected landmarks, requiring no additional tuning.

At each ICP iteration, every CT surface point $\vec{p}_i \in \mathcal{S}$ queries the KD-tree of $\mathcal{T}$ for its nearest target point $\vec{q}_i$. Following trimmed ICP [CSK05], we then *sort these correspondence pairs by distance and retain only the closest 50%* – the remaining half are discarded as presumed outliers or non-overlapping correspondences, and contribute nothing to the rigid solve that follows.

Let $(\vec{p}_i, \vec{q}_i)_{i=1}^{N_{\text{trim}}}$ denote the surviving pairs. Their landmark weights w_i are rescaled within this trimmed subset so that $\sum_i w_i = 1$; this is a per-iteration normalisation, not an update across iterations, and ensures the weighted centroids $\bar{\vec{s}} = \sum_i w_i \vec{p}_i$ and $\bar{\vec{t}} = \sum_i w_i \vec{q}_i$ are proper convex combinations of the kept points.

The weighted cross-covariance $H = \sum_i w_i (\vec{p}_i - \bar{\vec{s}})(\vec{q}_i - \bar{\vec{t}})^\top$ is decomposed via SVD to yield the rigid rotation R and translation $\vec{t} = \bar{\vec{t}} - R\bar{\vec{s}}$, which are applied to *all* nodes (surface and volumetric) jointly.

To escape local minima, we run this procedure from $N_{\text{rot}} = 6$ initial twist angles about the viewing axis and retain the result with the lowest trimmed (30%) mean error.

3.5. Simplicits Reduced-Order Model

Rather than tetrahedralizing the organ and optimizing per-vertex displacements, we build this reduced-order deformation model using Simplicits [MSP*24] implemented in the Nvidia Kaolin library [FTS*].

We apply the Simplicits in two distinct places in our pipeline, each with its own independently trained bases: (i) during registration, we train a bases on the normalised CT organ points (surface + volume) in the pre-alignment frame, and its reduced handle space becomes the parametric search space of our non-rigid ICP (Section 3.6); (ii) at inference time, a second bases on the post-registration geometry serves as the substrate for an interactive simulation of the aligned organ under external contacts (eg. a surgical instrument), reported in Section 4.

Simplicits uses linear blend skinning to deform objects. Given $H = \sim 20$ handles, the deformed position of any point i with rest position $\mathbf{x}_i^0$ is

$$\mathbf{x}_i(\mathbf{z}) = B_i\,\mathbf{z} + \mathbf{x}_i^0, \tag{1}$$

where $B_i \in \mathbb{R}^{3\times 12H}$ is the linear blend skinning matrix for point i and $\mathbf{z} \in \mathbb{R}^{12H}$ flattens the twelve 3×4 transformation matrices of all handles (Fig. 3). With $H = 20$ this yields $n_{\text{dof}} = 240$ (including the constant handle).

The linear blend skinning matrix, $B \in \mathbb{R}^{3N\times 12H}$, is a function of $\mathbf{x} \in \mathbb{R}^{N\times 3}$ and skinning weights $W \in \mathbb{R}^{N\times H}$. The function which outputs the skinning weights is learned on the point set $\mathcal{V}$ using the Simplicits training process. After training, separate skinning matrices B_{surf} and B_{vol} are baked for surface and volumetric points respectively.

For elastic energy evaluation, we subsample $N_{\text{qp}} = \sim 2000$ quadrature points from $\mathcal{V}$. The simulation module provides the reduced deformation gradient $F = \partial\mathbf{x}_i(\mathbf{z})/\partial\mathbf{x}_i^0$ at each quadrature point, which we can use to compute the Neohookean elastic energy, given DOFs $\mathbf{z}$

3.6. Non-Rigid Alignment

The physics-based non-rigid alignment alternates between updating correspondences and minimizing the combined elastic and alignment energies over the handles $\mathbf{z}$. The correspondence term maps source CT surface points to nearest target splat points and then the alignment term pulls the source points towards the target.

Correspondences. At each outer iteration, we query the KD-tree of $\mathcal{T}$ for the nearest target point $_i$ of each deformed surface point $\mathbf{x}_i(\mathbf{z})$, with distance d_i. Each correspondence is weighted by a Gaussian kernel:

$$w_i = \exp\left(-\frac{d_i^2}{2\sigma^2}\right), \quad \sigma = 0.05. \tag{2}$$

Combined objective. For a given correspondence set, we minimize

$$E(\mathbf{z}) = \alpha E_{\text{NH}}(\mathbf{z}) + E_{\text{align}}(\mathbf{z}), \tag{3}$$

where α balances elastic regularization against geometric fidelity.

Alignment Term. The weighted closest-point energy and its gradient are

$$E_{\text{align}} = \sum_{i=1}^{|\mathcal{S}|} w_i \,\|\mathbf{x}_i(\mathbf{z}) - \mathbf{u}_i\|^2, \qquad \frac{\partial E_{\text{align}}}{\partial \mathbf{z}} = 2\,B_{\text{surf}}^\top(\mathbf{w}\odot\mathbf{d}), \tag{4}$$

where $\mathbf{d}_i = \mathbf{x}_i - \mathbf{u}_i$.

Elastic Energy Term Given the organ's material parameters, youngs modulus (E), poissons ratio (ν) and deformation gradient $F(\mathbf{z})$ we can use any elastic energy formulation to compute the deformation energy of the object. We compute the Neo-Hookean Elastic energy provided to us in Kaolin's physics module [FTS*]. The reduced gradient, also provided via Kaolin's physics module, is $\partial E_{\text{NH}}/\partial\mathbf{z} = (\partial F/\partial\mathbf{z})^\top(\partial E_{NH}/\partial F)$.

Optimization. The combined objective (3) is minimized with respect to the reduced DOFs $\mathbf{z}$ using L-BFGS-B with analytic gradients (up to 80 inner iterations per correspondence update, convergence tolerance $\|\nabla E\| < 10^{-8}$). Within each α level, correspondences are re-established and the optimization re-run up to 25 times until the maximum vertex displacement falls below 10^{-4}.The stiffness weight α is annealed from high to low following Amberg [ARV07], allowing coarse global alignment before fine local deformation. The output of the optimization is a volumetric, CT-informed, anatomically correct organ representation, aligned onto corresponding surface splats from an endoscope scene.

3.7. Simulation

Using reduced Simplicits bases for non-rigid alignment, our method remains entirely mesh-free. In that same spirit, we use the Simplicits simulator, coupled with Nvidia's Newton IK Solver [New25] to interactively deform the organ with a da-Vinci PSM Gripper [KCD*14, Joh24]. We visualize the combined mesh and splat scene using 3DGRUT [MLMP*24].

Although we use the mesh-free Simplicits simulator for elastodynamcis, our results can also be meshed upon output via marching cubes and subsequently tetrahedralized for use in standard simulators.

4. Results

Armed with CT-grounded organ geometry aligned to visually accurate surface splats, we can achieve an incredible level of realism for surgical simulations. We demonstrate this in Figure 1 and Figure 7 on several distinct human organs, spleen, gallbladder, liver, pancreas and prostate and interactive deformations with an NVIDIA Newton IK-driven PSM gripper [New25, Joh24] can be observed. The entire pipeline, from registration to simulation, runs locally. In our experiments we use a single NVIDIA RTX 6000 Ada GPU for the BFGS objective computation, but the rest of the optimization runs on the CPU. The codebase and data to generate our results will be released publicly.

We evaluate our alignment method on three $alpha$ values: conservative ($\alpha = 0.05$) - which conservatively deforms the organ, moderate ($\alpha = 0.025$), and aggressive($\alpha = 0.005$), each with 35 steps to ensure convergence. Note, based on results, convergence occurs mostly within 8-16 steps. The results are shown in Table Table **??** and Figure 8.

Additionally, we include timings for our rigid and non-rigid alignment steps in Table **??**. For comparison, standard FEM-based elastic ICP, which runs on the CPU, using marching cubes and tetrahedralization for the gallbladder scan, takes on the order of *mins*. Our reduced optimization speeds this up to *secs*.

We also study splat to CT surface distance and volume drift after elastic ICP in Figure 8.The top row reports the trimmed 30% splat→CT surface distance of the mean nearest neighbour residual over the closest 70% of target points, a robust measure of registration fidelity in the well-posed target-to-source direction — and the bottom row reports volume drift, the percentage deviation of the mean deformation Jacobian (mean det F) from unity, which characterizes the net volumetric change the organ undergoes during registration. Surface fidelity improves with the aggressive (lowest) alpha value attaining the lowest residual across all five organs. The associated volume drift increases in magnitude with α aggressiveness, ranging from approximately 1% to 3.5% lower for the liver, gallbladder, prostate, and spleen and up to approximately 12% lower for the pancreas, indicating that more aggressive data-term weighting achieves tighter surface alignment at the cost of volume preservation at poisson ratio 0.49.

Clinicians View: In an expert clinical assessment, a surgeon acknowledged the importance of such deformation methods as critically important as learning tool for surgical simulation. He mentioned, "Achieving high-fidelity tool–tissue interaction remains a critical challenge for surgical simulation, and **progress toward physically grounded, patient-specific reconstructions** could enable more realistic training environments, improved procedural rehearsal, and ultimately safer pathways toward autonomous or semi-autonomous surgical actions."

5. Datasets

Our pipeline requires two inputs per case: a pre-operative volumetric scan with organ segmentation labels, and an endoscopic video of the same patient from which a 3DGS scene is reconstructed.

5.1. Preoperative imaging.

Anatomical models are derived from pre-operative computed tomography (CT) scans. Automatic multi-organ segmentation is performed using the MONAI Auto3DSeg framework [MON23, CLB*22] via the SlicerMONAIAuto3DSeg extension [LS24], which provides pre-trained SegResNet models for over 150 anatomical structures. Each model checkpoint bundles the network architecture, trained weights, and all pre-processing parameters (modality-appropriate normalisation, resampling resolution, and inference configuration), enabling portable inference without external configuration.

For this work we segment abdominal organs—liver, gallbladder, and spleen—from CT scans using the *whole-body segmentation model*. The same pipeline generalises to MRI-derived segmentations and to additional structures covered by publicly available Auto3DSeg checkpoints trained on datasets such as BTCV [har15], the Medical Segmentation Decathlon [ARB*22], and AMOS [JBY*22].

The **Medical Segmentation Decathlon** (MSD) [ARB*22] comprises ten diverse biomedical-image segmentation tasks spanning CT and MRI, including liver tumours, pancreas, colon, and hepatic vessels. The challenge was designed to evaluate generalist segmentation algorithms across modalities and anatomies, and its public training data—2 633 labelled volumes in total—has become a standard benchmark for Auto3DSeg model development.

The **Beyond the Cranial Vault** (BTCV) dataset [har15] provides 50 abdominopelvic CT scans with voxel-level annotations of 13 abdominal organs (spleen, right and left kidneys, gallbladder, oesophagus, liver, stomach, aorta, inferior vena cava, portal and splenic veins, pancreas, and adrenal glands). Originating from the MICCAI 2015 Multi-Atlas Labeling challenge, BTCV is one of the most widely used benchmarks for abdominal organ segmentation.

Both MSD and BTCV are among the training corpora used by the whole-body segmentation model we employ, ensuring robust out-of-the-box performance on the liver, gallbladder, and spleen labels used in our experiments.

5.2. Surgical Videos

Each video is processed by the G-SHARP reconstruction pipeline (Section 3.2.1) to produce a 3DGS PLY whose Gaussian centres serve as the fixed registration target. We evaluate our pipeline on a range of anatomical targets drawn from a mixture of in-house

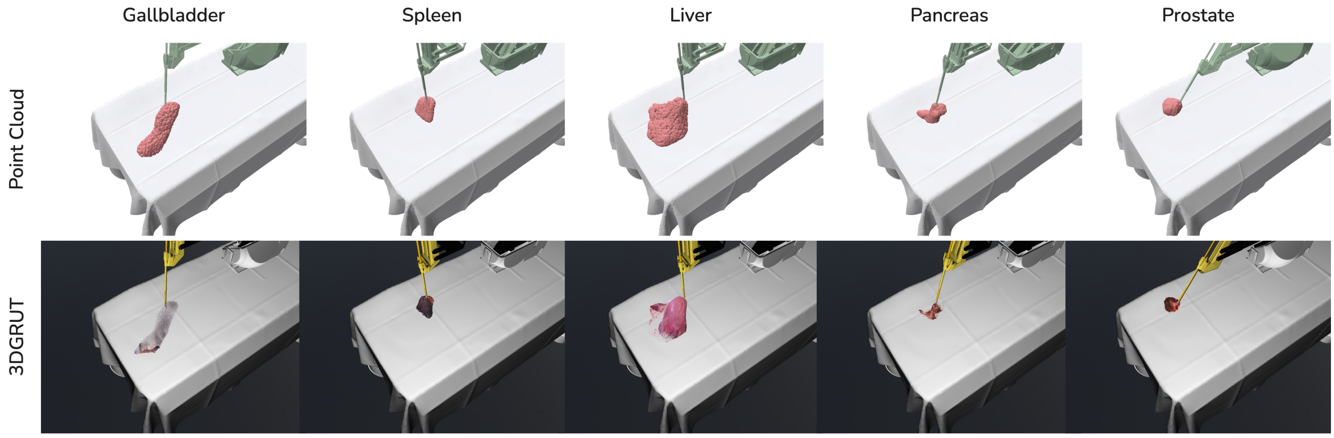

Figure 7: Physics-based deformation of organs with accurate CT geometry using Simplicits. We show each of the deforming organs as a pointcloud representation, and render the aligned splats in 3DGRUT [MLMP*24]

	#Points		**Runtime**		**Pre-rigid (coarse)**		**Post-rigid ICP**		**Post-elastic ICP**	
Organ	CT_{surf}	3DGS	Rigid (s)	Elastic Step Avg (s)	mean	$trim_3 0$	mean	$trim_{30}$	mean	$trim_{30}$
Liver	9992	648345	27	~1.59	0.335	0.076	0.293	0.042	0.132	0.015
Gallbladder	9784	88040	1.5	~1.79	1.079	0.172	0.854	0.078	0.669	0.074
Pancreas	5628	126807	11	~1.1	0.471	0.153	0.190	0.047	0.050	0.018
Prostate	10000	95048	5	~1.5	0.634	0.183	0.478	0.069	0.359	0.034
Spleen	67500	158826	65	~2.9	0.409	0.079	0.308	0.036	0.199	0.026

Table 2: Registration error progression, runtime, and point-cloud sizes per organ. All experiments use a Simplicits model with **20 transformation handles** ($H = 20$, 240 reduced DOFs) and the parameters described in Section 3.5. Errors are nearest-neighbour distances from CT surface points to the 3DGS target cloud, reported in normalised units (CT surface scaled to a unit box). *Mean* is the average nearest-neighbour distance over all CT surface points; $trim_{30}$ discards the worst-matching 30 % of correspondences and averages the remaining 70 %, emphasising the endoscope-visible region where the splat geometry is most trustworthy.

and publicly available laparoscopic recordings. The gallbladder and liver videos come from in-house recordings of routine laparoscopic cholecystectomy and partial-hepatectomy procedures provided by our clinical collaborators. The spleen sequence is from the laparoscopic splenectomy case series of Dimitrov [Dim19]; the prostate sequences are from the robot-assisted radical prostatectomy intraoperative footage published alongside Shin and Lee [SL20]; the pancreas sequences are from the laparoscopic central pancreatectomy case report of Kvashilava [KKG22]. All externally sourced videos were used under the open-access licenses of their respective publications; no patient-identifying information is retained in the resulting 3DGS reconstructions or our derived assets.

6. Conclusion and Limitations

Our project addresses the simulatability gap between view-limited 3D Gaussian splat (3DGS) reconstructions from endoscopy and pre-operative volumetric organ models in medical images. While splats best capture what the camera actually sees, they lack the anatomical grounding present in CT scans for physics-driven simulation. We presented a two-stage pipeline that registers segmented CT organ geometry onto a splat-derived target point cloud—first with anatomy-aware weighted rigid ICP bootstrapped by 2-click, landmark selection, then with a mesh-free, reduced-order elastic ICP built on reduced deformation bases from Simplicits. By optimizing a modest number of handle parameters instead of a full tetrahedral mesh, we obtain a splat-aligned, simulation-ready organ representation that remains physically plausible in the bulk while respecting the visible endoscopic surface.

6.1. Limitations

Data availability. Clean, high-quality endoscopic videos are difficult to obtain at scale, especially for organs beyond the set shown in our experiments. Limited data curtail how broadly we can apply the pipeline and report quantitative trends across different anatomies.

Weighted rigid ICP and user weighting. The rigid stage remains sensitive to local minima: even with trimming and multi-start twists, poor landmark choices or uneven importance weights can yield a globally wrong pose that the elastic stage cannot fully correct. In practice, weights must reflect which tissue is actually visible in the endoscopic view.

Upstream reconstruction and segmentation. End-to-end behaviour depends strongly on 3DGS scene quality (motion, occlusions, specular tissue) and on CT/MRI segmentation and morphol-

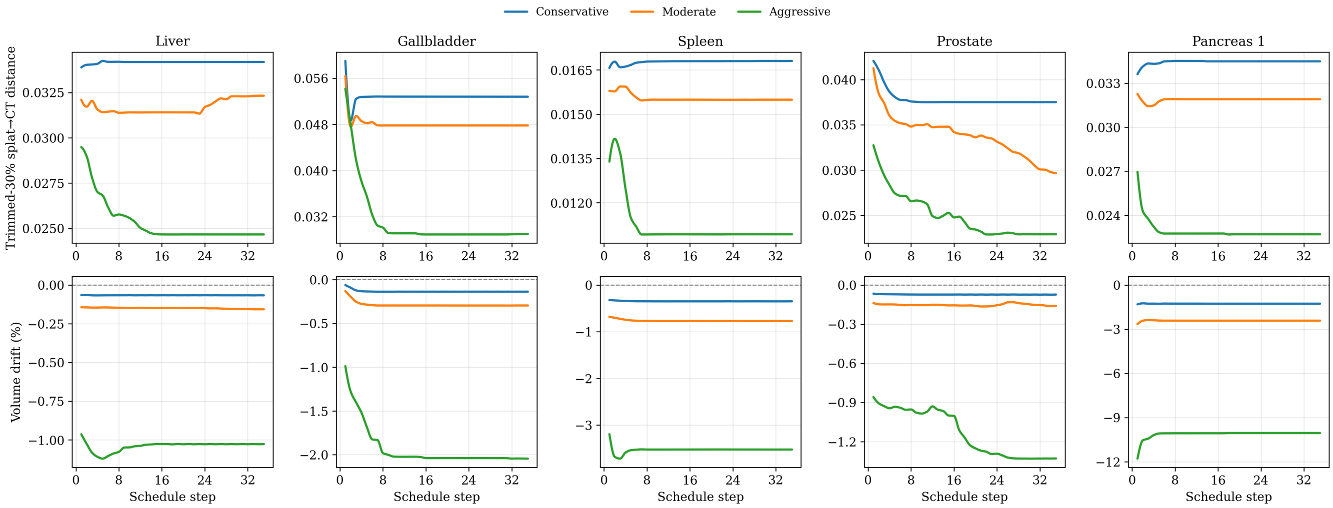

Figure 8: Registration fidelity and volume drift across five organs and three deformation weightings (α) starting at iteration 1. The top row shows the trimmed-30% splat→CT surface distance (registration fidelity) and the bottom row the volume drift — the percentage deviation of the mean deformation Jacobian (mean det F) from unity — both as a function of elastic-ICP step. In each panel the three curves denote the conservative ($\alpha = 0.05$), moderate ($\alpha = 0.025$), and aggressive ($\alpha = 0.005$). Lower values in the top row indicate a closer surface match; negative values in the bottom row indicate net volumetric loss. We ran results for 35 outer steps and 300 inner BFGS max iterations, but convergence occured within 8-16 outer steps.

ogy post-processing. Splat *completion* and *infilling* remain open problems. Producing consistent 3DGS where the scope never observes tissue is nearly impossible. Our method aligns CT to the splat-*visible* organ rather than hallucinating unseen regions.

Tool induced deformation and topology changes. The current pipeline assumes a single 3DGS reconstruction of one organ that is topologically close to the (un-cut) CT geometry. Tools from raw video are masked using MedSAM3 [LXC*25] and occluded regions are recovered implicitly with G-SHARP [NTL*25] for tool-free scenes. For CT alignment, highly local tool-induced tissue deformation, like pinching, is handled implicitly as the elastic stage discards the worst correspondences each iteration (trimmed ICP [CSK05]), naturally rejecting locally tool-pressed tissue as outliers.

6.2. Future work

Promising directions include replacing the manual correspondence step with automatic full-to-partial shape matching. For example, we could use functional-map style spectral correspondences between CT and splat surfaces, which would allow us to seed the correspondence step with more robust correspondences.

Improving perceptual realism through better 3DGS reconstruction and completion, as well as mechanical realism through cutting/fracture simulation will greatly improve the usability of the pipeline.

References

[ARB*22] Antonelli M., Reinke A., Bakas S., Farahani K., Kopp-Schneider A., Landman B. A., Litjens G., Menze B., Ronneberger O., Summers R. M., van Ginneken B., Bilello M., Bilic P., Christ P. F., Do R. K. G., Gollub M. J., Heckers S. H., Huisman H., Jarnagin W. R., McHugo M. K., Napel S., Pernicka J. S. G., Rhode K., Tobon-Gomez C., Vorontsov E., Meakin J. A., Ourselin S., Wiesenfarth M., Arbeláez P., Bae B., Chen S., Daza L., Feng J., He B., Isensee F., Ji Y., Jia F., Kim N., Kim I., Merhof D., Pai A., Park B., Perslev M., Rezaiifar R., Rippel O., Sarasua I., Shen W., Son J., Wachinger C., Wang L., Wang Y., Xia Y., Xu D., Xu Z., Zheng Y., Simpson A. L., Maier-Hein L., Cardoso M. J.: The medical segmentation decathlon. *Nature Communications 13* (2022), 4128. doi:10.1038/s41467-022-30695-9. 7

[ARV07] Amberg B., Romdhani S., Vetter T.: Optimal step nonrigid ICP algorithms for surface registration. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2007), pp. 1–8. doi:10.1109/CVPR.2007.383165. 2, 6

[BFJP*25] Barri Khojasteh S., Fuentes-Jimenez D., Pizarro D., Espinel Y., Bartoli A.: MIS-NeRF: Neural radiance fields in minimally-invasive surgery. *Int. J. Computer Assisted Radiology and Surgery 20* (2025), 1481–1490. doi:10.1007/s11548-025-03429-7. 3, 4

[BJ05] Barbic J., James D. L.: Real-time subspace integration for st.venant-kirchhoff deformable models. *ACM Trans. Graph. 24*, 3 (July 2005), 982–990. 4

[BLGB25] Browne G., Liu M., Grinspun E., Benchekroun O.: MorphModes: Non-rigid registration via adaptive skinning eigenmodes. *arXiv preprint arXiv:2510.18658* (2025). 2

[BM92] Besl P. J., McKay N. D.: A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell. 14*, 2 (1992), 239–256. doi:10.1109/34.121791. 2, 5

[BZC*23] Benchekroun O., Zhang J. E., Chaudhuri S., Grinspun E., Zhou Y., Jacobson A.: Fast complementary dynamics via skinning eigenmodes. *ACM Trans. Graph. 42*, 4 (July 2023). doi:10.1145/3592404. 2

[CLB*22] Cardoso M. J., Li W., Brown R., Ma N., Kerfoot

E., WANG Y., MURREY B., MYRONENKO A., ZHAO C., YANG D., ET AL.: MONAI: An open-source framework for deep learning in healthcare. *arXiv preprint arXiv:2211.02701* (2022). 7

[CPW*26] CUI R., PEI J., WANG H., ZHOU J., TEOH J. Y.-C., HENG P.-A., QIN J.: Preoperative-to-intraoperative liver registration for laparoscopic surgery via latent-grounded correspondence constraints. *arXiv preprint arXiv:2603.01720* (2026). 4

[CSK05] CHETVERIKOV D., STEPANOV D., KRSEK P.: Robust Euclidean alignment of 3D point sets: The trimmed iterative closest point algorithm. *Image and Vision Computing 23*, 3 (2005), 299–309. Originally presented at ICPR 2002. doi:10.1016/j.imavis.2004.05.007. 2, 5, 6, 9

[CXL*24] CHANG J., XU Y., LI Y., CHEN Y., FENG W., HAN X.: GaussReg: Fast 3D registration with Gaussian splatting. In *Proc. European Conference on Computer Vision (ECCV)* (2024). 3

[CZS*25] CHEN W., ZENG Q., STUDDERS C., KWON J. J., PANG E. H., PRISMAN E., SALCUDEAN S. E.: Semantic-icp: Iterative closest point for non-rigid multi-organ point cloud registration. *arXiv preprint arXiv:2503.00972* (2025). 2

[Dim19] DIMITROV D. D.: Laparoscopic splenectomy for solitary splenic metastasis in patients with previous open surgery – case series. *International Journal of Surgery Case Reports 65* (2019), 83–86. doi:10.1016/j.ijscr.2019.10.044. 8

[ERB*24] ESPINEL Y., RABBANI N., BUI T. B., RIBEIRO M., BUC E., BARTOLI A.: Keyhole-aware laparoscopic augmented reality. *Medical Image Analysis 94* (2024), 103161. URL: https://www.sciencedirect.com/science/article/pii/S1361841524000860, doi:https://doi.org/10.1016/j.media.2024.103161. 4

[FTS*] FUJI TSANG C., SHUGRINA M., ET AL.: Kaolin: A pytorch library for accelerating 3d deep learning research. 6

[FWH*25] FEHRENTZ M., WINKLER A., HEILIGER T., HAOUCHINE N., HEILIGER C., NAVAB N.: BridgeSplat: Bidirectionally coupled CT and non-rigid Gaussian splatting for deformable intraoperative surgical navigation. In *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (2025), pp. 44–53. doi:10.1007/978-3-032-05141-7_5. 4

[GSv*18] GENET M., STOECK C. T., VON DEUSTER C., LEE L. C., KOZERKE S.: Equilibrated warping: Finite element image registration with finite strain equilibrium gap regularization. *Medical Image Analysis 50* (2018), 1–22. doi:10.1016/j.media.2018.07.007. 2

[har15] HARRIGR: Segmentation outside the cranial vault challenge, 2015. URL: https://repo-prod.prod.sagebase.org/repo/v1/doi/locate?id=syn3193805&type=ENTITY, doi:10.7303/SYN3193805. 7

[JBK*12] JACOBSON A., BARAN I., KAVAN L., POPOVIĆ J., SORKINE O.: Fast automatic skinning transformations. *ACM Trans. Graph. 31*, 4 (2012), 77:1–77:10. doi:10.1145/2185520.2185573. 4

[JBY*22] JI Y., BAI H., YANG J., GE C., ZHU Y., ZHANG R., LI Z., ZHANG L., MA W., WAN X., LUO P.: AMOS: A large-scale abdominal multi-organ benchmark for versatile medical image segmentation. In *Advances in Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track* (2022). 7

[Joh24] JOHNS HOPKINS UNIVERSITY DVRK TEAM: dvrk_model: URDF models for the da Vinci research kit patient side manipulator. https://github.com/jhu-dvrk/dvrk_model, 2024. 7

[JT05] JAMES D. L., TWIGG C. D.: Skinning mesh animations. *ACM Trans. Graph. 24*, 3 (2005), 399–407. doi:10.1145/1073204.1073206. 4

[KCD*14] KAZANZIDES P., CHEN Z., DEGUET A., FISCHER G. S., TAYLOR R. H., DIMAIO S. P.: An open-source research kit for the da vinci® surgical system. In *2014 IEEE International Conference on Robotics and Automation (ICRA)* (2014), pp. 6434–6439. doi:10.1109/ICRA.2014.6907809. 7

[KJP02] KRY P. G., JAMES D. L., PAI D. K.: Eigenskin: Real time large deformation character skinning in hardware. In *Proc. ACM SIGGRAPH/Eurographics Symp. Computer Animation* (2002), pp. 153–159. doi:10.1145/545261.545286. 4

[KKG22] KVASHILAVA A., KOBALAVA B., GIORGOBIANI G.: Laparoscopic central pancreatectomy with pancreaticogastrostomy: Our initial experience. *Cureus 14*, 4 (2022), e24542. doi:10.7759/cureus.24542. 8

[KKLD23] KERBL B., KOPANAS G., LEIMKÜHLER T., DRETTAKIS G.: 3D Gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph. 42*, 4 (2023), 139:1–139:14. doi:10.1145/3592433. 2, 3

[LGS*25] LIU P., GÜTTNER B., SU Y., LI C., XU J., LIU M., MIN Z., ZHYLKA A., SMIT J., OLTHOF K., FUSAGLIA M., APOLLE R., MIEDERER M., FROHNEBERGER L., RIEDIGER C., WEITZ J., KOLBINGER F., SPEIDEL S., PFEIFFER M.: PIVOTS: Aligning unseen structures using preoperative to intraoperative volume-to-surface registration for liver navigation. *arXiv preprint arXiv:2507.20337* (2025). 4

[LLYY24] LIU Y., LI C., YANG C., YUAN Y.: EndoGaussian: Real-time Gaussian splatting for dynamic endoscopic scene reconstruction. *arXiv preprint arXiv:2401.12561* (2024). 2, 3

[LS24] LASSO A., SLICERMONAIAUTO3DSEG CONTRIBUTORS: SlicerMONAIAuto3DSeg: 3D Slicer extension for MONAI auto3dseg segmentation. https://github.com/lassoan/SlicerMONAIAuto3DSeg, 2024. 7

[LTC*25] LI Z., TUCKER R., COLE F., ET AL.: Megasam: Accurate, fast and robust structure and motion from casual dynamic videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2025). URL: https://arxiv.org/abs/2412.04463. 4

[LXC*25] LIU A., XUE R., CAO X. R., SHEN Y., LU Y., LI X., CHEN Q., CHEN J.: Medsam3: Delving into segment anything with medical concepts. *arXiv preprint arXiv:2511.19046* (2025). 4, 9

[MLMP*24] MOENNE-LOCCOZ N., MIRZAEI A., PEREL O., DE LUTIO R., MARTINEZ ESTURO J., STATE G., FIDLER S., SHARP N., GOJCIC Z.: 3D Gaussian ray tracing: Fast tracing of particle scenes. *ACM Trans. Graph. 43*, 6 (2024). SIGGRAPH Asia 2024 (Journal Track). doi:10.1145/3687934. 7, 8

[MON23] MONAI CONSORTIUM: Auto3DSeg: Automated 3D medical image segmentation framework. https://github.com/Project-MONAI/tutorials/tree/main/auto3dseg, 2023. Part of Project-MONAI; pre-trained SegResNet checkpoints for over 150 anatomical structures. 7

[MSP*24] MODI V., SHARP N., PEREL O., SUEDA S., LEVIN D. I. W.: Simplicits: Mesh-free, geometry-agnostic elastic simulation. *ACM Trans. Graph. 43*, 4 (July 2024). doi:10.1145/3658184. 2, 3, 4, 6

[New25] NEWTON PHYSICS CONTRIBUTORS: Newton: A GPU-accelerated physics simulation engine for robotics. https://github.com/newton-physics/newton, 2025. Open-source project managed by the Linux Foundation; initiated by NVIDIA, Google DeepMind, and Disney Research. 7

[NTL*25] NATH V., TEJERO J. G., LI R., FILICORI F., AZIZIAN M., HUVER S. D.: G-SHARP: Gaussian surgical hardware accelerated real-time pipeline. *arXiv preprint arXiv:2512.02482* (2025). 5, 9

[PSY24] PICCINELLI L., SAKARIDIS C., YU F.: Unidepth: Universal monocular metric depth estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024). 4

[SL20] SHIN T. Y., LEE Y. S.: Robot-assisted radical prostatectomy with clipless intrafascial neurovascular bundle-sparing approach: Surgical technique and one-year functional and oncologic outcomes. *Scientific Reports 10* (2020), 18356. doi:10.1038/s41598-020-74513-y. 8

[TD21] TEED Z., DENG J.: Droid-slam: Deep visual slam for monocular, stereo, and rgb-d cameras. In *Advances in Neural Information Processing Systems (NeurIPS)* (2021). 5

[WBM*23] WASSERTHAL J., BREIT H.-C., MEYER M. T., PRADELLA M., HINCK D., SAUTER A. W., HEYE T., BOLL D. T., CYRIAC J., YANG S., ET AL.: Totalsegmentator: robust segmentation of 104 anatomic structures in ct images. *Radiology: Artificial Intelligence 5*, 5 (2023), e230024. 2, 5

[WJC*10] WITTEK A., JOLDES G., COUTON M., WARFIELD S. K., MILLER K.: Patient-specific non-linear finite element modelling for predicting soft organ deformation in real-time; application to non-rigid neuroimage registration. *Progress in Biophysics and Molecular Biology 103*, 2 (2010), 292–303. Special Issue on Biomechanical Modelling of Soft Tissue Motion. URL: https://www.sciencedirect.com/science/article/pii/S0079610710000696, doi:https://doi.org/10.1016/j.pbiomolbio.2010.09.001. 2, 3, 4

[WLFD22] WANG Y., LONG Y., FAN S. H., DOU Q.: Neural rendering for stereo 3d reconstruction of deformable tissues in robotic surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2022), Springer, pp. 431–441. 5

[WMKW07] WITTEK A., MILLER K., KIKINIS R., WARFIELD S. K.: Patient-specific model of brain deformation: Application to medical image registration. *Journal of Biomechanics 40*, 4 (2007), 919–929. doi:10.1016/j.jbiomech.2006.02.021. 2, 3, 4

[XYC*24] XIE W., YAO J., CAO X., LIN Q., TANG Z., DONG X., GUO X.: SurgicalGaussian: Deformable 3D Gaussians for high-fidelity surgical scene reconstruction. In *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (2024), pp. 617–627. 2, 3

[YKH*24] YANG L., KANG B., HUANG Z., XU X., FENG J., ZHAO H.: Depth anything: Unleashing the power of large-scale unlabeled data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024). 4

[YLS*24] YANG S., LI Q., SHEN D., GONG B., DOU Q., JIN Y.: Deform3DGS: Flexible deformation for fast surgical scene reconstruction with Gaussian splatting. In *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (2024). 2, 3

[ZBD*25] ZHANG H., BULATHSINHALA S., DAVIDSON B. R., CLARKSON M. J., RAMALHINHO J.: Deep hashing for global registration of preoperative CT and video images for laparoscopic liver surgery. *Int. J. Computer Assisted Radiology and Surgery 20* (2025), 1461–1469. doi:10.1007/s11548-025-03418-w. 4