PhysOCT: Physics-Guided Retinal Modeling for OCT Image Synthesis

Yutong Wang, Xiaohui Li, Wenbo Zhang, Yizhe Zhang, Qiang Chen

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China
chen2qiang@njust.edu.cn

Abstract. Optical Coherence Tomography (OCT) is central to retinal disease diagnosis and longitudinal monitoring, yet learning-based OCT analysis is often constrained by limited expert annotations. Conditional generative models can alleviate data scarcity, but prevailing lesion- or segmentation-conditioned methods typically lack physical constraints governing the biomechanical coupling between pathological fluid evolution and retinal-layer deformation. As a result, synthesized scans may look plausible while exhibiting anatomically inconsistent layer configurations near lesions. We propose **PhysOCT**, a physics-guided plug-in framework that augments existing conditional generators with a deformation-consistent structural prior. Given consecutive lesion masks, PhysOCT (i) predicts lesion-induced external forces using a force prediction network, then (ii) propagates retinal geometry via a differentiable mesh-free solver built on enhanced Generalized Moving Least Squares and a stabilized Lagrangian update, producing a predicted next-step layer prior. This physics-derived prior, together with a deformation-aligned warped sketch of the previous scan, is used as additional conditioning for diffusion-based synthesis. Experiments on multiple longitudinal retinal OCT cohorts demonstrate improved one-step deformation prediction and consistent gains in generation fidelity (PSNR/SSIM) and distributional quality (FID) across DDPM and LDM backbones. The code is publicly available at: https://github.com/Terrific0723/PhysOCT.

Keywords: Physics-informed generation · OCT synthesis · Differentiable simulation.

1 Introduction

Optical Coherence Tomography (OCT) is a non-invasive, high-resolution imaging modality that reveals cross-sectional retinal anatomy with micrometer-scale detail [1]. OCT is widely used for early screening, disease monitoring, and treatment planning across major retinal pathologies such as choroidal neovascularization (CNV), diabetic macular edema (DME), and age-related macular degeneration (AMD) [2–4]. In these conditions, the joint configuration of retinal layers and pathological regions (e.g., fluid) provides key diagnostic evidence. However, high-quality pixel-wise annotations of layers and lesions are costly and

time-consuming to obtain, limiting the size and diversity of labeled datasets for training robust segmentation and analysis models.

Controllable medical image synthesis is a promising approach to mitigate annotation scarcity and enable privacy-preserving data sharing. Generative modeling has evolved from GAN-based synthesis [5], to diffusion probabilistic models [6], and more recently to flow-based formulations such as rectified flow [7]. For clinical utility, generation must be conditional and anatomically constrained: recent work explores text and multi-modal conditioning in medical imaging [8, 9], and segmentation-guided diffusion models explicitly enforce mask adherence during sampling [10]. In ophthalmology, diffusion-based OCT synthesis has been used to improve downstream segmentation, including RetiDiff [11] and DDPM-based synthesis guided by coarse retinal-layer sketches [12].

Despite progress, OCT lesion evolution is strongly coupled with retinal-layer deformation over time. Conditioning solely on lesion masks or coarse layer hints can therefore produce anatomically inconsistent outcomes: lesions may expand or regress without correspondingly plausible displacement of surrounding layers, undermining clinical credibility. Purely physics-driven simulation [13, 14] and physics-informed learning [15] offer principled deformation modeling, but they can be computationally demanding and are challenged by uncertain or patient-specific lesion-induced forces. These observations motivate *hybrid* approaches that learn unknown forces from data while retaining a physically structured deformation model, similar in spirit to recent physics-augmented neural rendering frameworks [16, 17].

We introduce **PhysOCT**, a physics-guided plug-in conditioner for longitudinal OCT synthesis. PhysOCT infers lesion-driven forces from temporal changes in lesion masks and advances retinal geometry with a stable differentiable mesh-free solver, yielding a predicted next-step layer prior that can be injected into existing diffusion generators. This results in improved lesion–layer coherence while retaining the flexibility and fidelity of modern generative backbones. The contribution of this work can be summarized as follows: **Physics-guided plug-in for conditional OCT generation.** PhysOCT augments existing diffusion backbones with a physics-derived structural prior, improving anatomical consistency between lesions and retinal layers without requiring large-scale paired structural supervision, as the physics-based deformation model provides a strong prior that substantially reduces the paired samples needed to learn lesion-layer correspondence, compared to purely data-driven image generation. **Learned lesion-induced forces with a differentiable biomechanical solver.** We predict point-wise forces from consecutive lesion masks and propagate deformation with a GMLS-based mesh-free solver informed by a stabilized Lagrangian update. **Consistent gains across cohorts.** PhysOCT improves deformation prediction accuracy, and boosts synthesis metrics (PSNR/SSIM/FID) across multiple datasets and backbones.

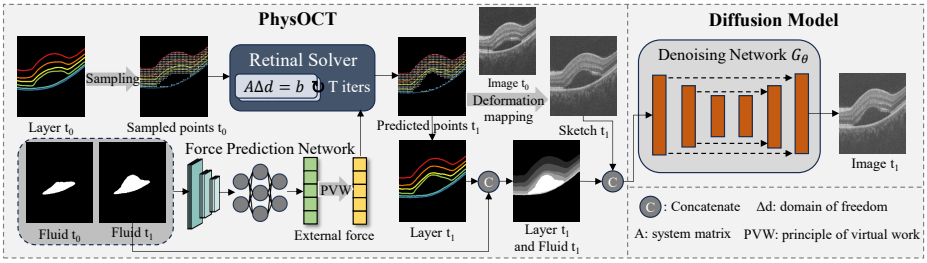

Fig. 1. Overview of PhysOCT. Given current layer structure layer_t and consecutive lesion masks $(\text{fluid}_t, \text{fluid}_{t+1})$, PhysOCT predicts lesion-induced forces and advances retinal geometry using a differentiable solver, producing a deformation mapping and a next-step layer prior. These physics-derived signals are used to form a deformation-aligned sketch and to condition a diffusion-based generator for anatomically consistent synthesis.

2 Method

2.1 Problem Setup and Overview

We consider longitudinal conditional OCT synthesis, where lesion evolution and retinal-layer deformation should remain coherent across time. Given the current OCT B-scan I_t, the corresponding retinal layer structure layer_t, and consecutive lesion (fluid) masks fluid_t and fluid_{t+1}, our goal is to synthesize the next scan $\hat{I}_{t+1}$ such that: (i) the synthesized lesion appearance adheres to fluid_{t+1}, and (ii) retinal layers deform in a physically and anatomically plausible manner consistent with the lesion change. We emphasize that PhysOCT is positioned as a physics-consistent OCT *editing* framework rather than a future-scan predictor: fluid_{t+1} is obtained either from clinician-specified or algorithmically edited masks (e.g., augmenting rare lesion morphologies) or from a separate mask-prediction network, and no real $t+1$ scan is required at inference.

As illustrated in Fig. 1, PhysOCT introduces a physics-guided intermediate representation that bridges mask evolution and image synthesis. PhysOCT consists of three modules: (1) a *force prediction network* that estimates lesion-induced external forces from temporal lesion changes, (2) a *retinal solver* that propagates the retinal configuration using a stable differentiable mesh-free update, and (3) an *image synthesis module* (e.g., DDPM/LDM) that is conditioned not only on the lesion mask but also on the physics-predicted next-step layer prior and a deformation-aligned warped sketch.

2.2 Force Prediction Network

We construct a uniformly sampled point set $\{x_i^t\}_{i=1}^N$ within the retinal region defined by layer_t, excluding the lesion region. Ground-truth targets x_i^{t+1} used in Eq. (2) are obtained by linearly interpolating each point's relative position within

its enclosing inter-layer band between layer_t and layer_{t+1}. Given the current B-scan I_t, we employ a CNN encoder to obtain multi-scale feature maps, augmented with two lesion-aware local cues: the lesion height change between fluid_t and fluid_{t+1}, and a distance cue to lesion-related boundaries. For each point x_i^t, we extract a local feature vector h_i from these maps and predict a 2D point-wise force $f_i^{\text{pred}} \in \mathbb{R}^2$ using a shared regressor. Considering that different inter-layer regions exhibit different elastic characteristics, we further introduce a learnable inter-layer-region-dependent scaling parameter α to modulate the predicted force magnitudes for points belonging to different regions.

To interface with the discretized Lagrangian solver, we assemble point-wise forces into kernel-level (generalized) forces through the principle of virtual work:

$$f_{\text{ker}}^{\text{pred}} = \sum_{i=1}^{N} \Delta x^2 \, N(x_i^t)^{\mathsf{T}} \, f_i^{\text{pred}} \tag{1}$$

where Δx^2 is the cell area induced by uniform sampling and $N(\cdot)$ denotes precomputed Generalized Moving Least Squares (GMLS) shape functions. Injecting $f_{\text{ker}}^{\text{pred}}$ into the retinal solver yields predicted next-step point positions $\{\hat{x}_i^{t+1}\}$. We train the force predictor end-to-end using SmoothL1 supervision on point locations:

$$L_{\text{force}} = \sum_{i=1}^{N} \text{SmoothL1}\big(\hat{x}_i^{t+1},\, x_i^{t+1}\big) \tag{2}$$

Why point-wise forces? Directly regressing kernel-level generalized forces is empirically unstable: GMLS assembly induces strong non-local coupling across kernel degrees of freedom (DoFs), making learning in the generalized-force space ill-conditioned. Predicting point-wise forces decouples learning from the fixed assembly operator in Eq. (1), providing a more stable and transferable interface to the solver.

2.3 Physics-Guided Retinal Solver

Given assembled generalized forces $f_{\text{ker}}^{\text{pred}}$, current sampling points $\{x_i^t\}$, and fixed precomputed GMLS operators, the retinal solver outputs updated point positions $\{\hat{x}_i^{t+1}\}$, a deformation mapping $\varphi_{t\to t+1}$, and a recovered next-step layer prior $\text{layer}_{t+1}^{\text{pred}}$.

Mesh-free deformation parameterization. To model continuous deformation without an explicit mesh, we adopt an enhanced mesh-free GMLS discretization with a quadratic basis and derivative constraints, following PIE-NeRF [16]. The deformation field is parameterized by kernel DoFs d and evaluated by the same GMLS shape functions used in force assembly:

$$\varphi(x) = N(x)\, d, \qquad x \in \Omega \tag{3}$$

where Ω is the 2D retinal domain. Predicted next-step point locations follow $\hat{x}_i^{t+1} = \varphi(x_i^t)$.

Stable differentiable update. We advance the configuration by solving a discretized Lagrangian update in which internal forces are derived from a stabilized Neo-Hookean hyperelastic energy (serving as a physically motivated deformation prior) and external forces are provided by $f_{\text{ker}}^{\text{pred}}$. For stable optimization and end-to-end differentiation, each step is written as a Symmetric Positive Definite (SPD) linear system:

$$A\,\Delta d = b, \qquad d_{t+1} = d_{\text{rest}} + \Delta d \tag{4}$$

where A and b denote the step-wise system matrix and right-hand side, and Δd is the DoF increment from the rest state d_{rest}. The SPD structure yields robust solves and stable gradients through the implicit update.

Boundary constraints. Because consecutive scans are pre-aligned at the bottom-most layer boundary, we impose Dirichlet constraints by fixing sampled points on that boundary. This prevents global drift and improves numerical stability when advancing the deformation.

2.4 Image Synthesis with Physics-Derived Priors

PhysOCT provides two complementary structural signals for synthesis: (1) a deformation mapping $\varphi_{t\to t+1}$ and (2) a predicted next-step layer prior $\text{layer}_{t+1}^{\text{pred}}$. These are used to improve anatomical consistency while allowing the image generator to focus on realistic texture and speckle statistics.

We first warp the current image using the predicted deformation mapping:

$$\tilde{I}_{t\to t+1} = \text{warp}(I_t,\, \varphi_{t\to t+1})\,. \tag{5}$$

The warped sketch $\tilde{I}_{t\to t+1}$ is structurally aligned to time $t{+}1$ while preserving low-frequency appearance from I_t. A conditional generator G_θ then synthesizes $\hat{I}_{t+1}$ using $\tilde{I}_{t\to t+1}$ as input, conditioned on the lesion mask fluid_{t+1} and the physics-predicted layer prior $\text{layer}_{t+1}^{\text{pred}}$. During training, we supervise $\hat{I}_{t+1}$ with the ground-truth I_{t+1} using the standard objective of the chosen conditional generative model. During inference, PhysOCT produces $(\tilde{I}_{t\to t+1}, \text{fluid}_{t+1}, \text{layer}_{t+1}^{\text{pred}})$, which are fed into G_θ to generate $\hat{I}_{t+1}$.

3 Experiments

3.1 Datasets, Metrics, and Implementation Details

Datasets. We evaluate PhysOCT on three longitudinal OCT datasets: a private CNV cohort (CNV disease, Optovue; 39 patients, 49 paired longitudinal 3D cubes at $400 \times 640 \times 400$ resolution), OLIVES TREX-DME (DME disease) [18], and MARIO Task 1 (AMD disease) [19] (both Heidelberg Spectralis). When manual annotations are unavailable, lesion and layer labels are obtained using pretrained U-Net models [20, 21, 25] with manual correction. For each cohort,

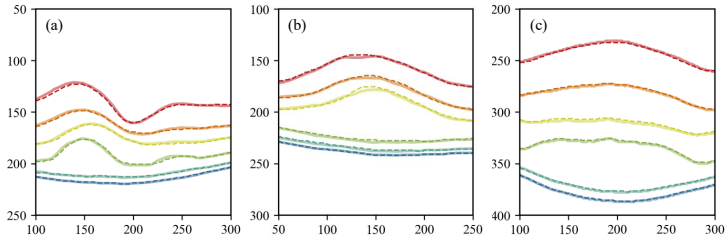

Fig. 2. Overlay of predicted vs. ground-truth retinal layers at next-step on AMD (a), DME (b), and CNV (c) datasets (GT: solid; Pred: dashed).

paired B-scans with substantial lesion change are selected for force predictor training/evaluation (860/1079/936 pairs for AMD/DME/CNV); the generative model is further trained on the full training-split B-scans. All datasets are split 8:2 at the patient level.

Implementation. All experiments are conducted on a single NVIDIA RTX 3090. Diffusion-based generators use the default β-schedule with 1000 diffusion steps and the standard noise-prediction MSE objective. Images are resized to 480×256 (CNV) or 256×256 (DME/AMD). For LDM variants, a VQGAN maps images to latents (120×64×4 for CNV; 64×64×4 for DME/AMD). Physical parameters (Young's modulus, E) are initialized following ocular elasticity-related references used in prior work [22].

Metrics. For deformation prediction, we report Rel-L_2 (Relative L2 Error) and MAE (Mean Absolute Error) on next-step point locations. For synthesis quality, we report PSNR/SSIM and FID.

3.2 PhysOCT Effectiveness for One-Step Deformation Prediction

Qualitative study. Fig. 2 visualizes representative examples by overlaying predicted and ground-truth next-step layer structures. PhysOCT closely follows the target layer trajectories, capturing both global geometry and local variations, and maintaining anatomically plausible layer ordering.

Comparative study. Table 1 compares PhysOCT with two baselines: *Direct regression*, which predicts next-step point locations without force-driven propagation, and a *PIE-NeRF solver* [16] variant that replaces our solver at test time under the same GMLS discretization and energy setting. PhysOCT consistently achieves lower errors across cohorts while remaining efficient per step.

Ablation and parameter sensitivity. We ablate two design choices on DME: (i) predicting point-wise forces followed by GMLS assembly (Eq. 1) and (ii) constructing an SPD global operator (Eq. 4) for stable solves. Table 2 shows that

Table 1. One-step point prediction across cohorts (mean±std). Lower is better for Rel-L_2/MAE. Time includes the runtime of GMLS matrix assembly and network inference.

Method	AMD		CNV		DME		Time (ms)
	Rel-L_2 (%)↓	MAE (px)↓	Rel-L_2 (%)↓	MAE (px)↓	Rel-L_2 (%)↓	MAE (px)↓	
PhysOCT	$\mathbf{0.45 \pm 0.17}$	$\mathbf{0.91 \pm 0.35}$	$\mathbf{0.42 \pm 0.16}$	$\mathbf{0.97 \pm 0.31}$	$\mathbf{0.44 \pm 0.18}$	$\mathbf{0.75 \pm 0.31}$	34 ± 7
Direct regression	3.56 ± 0.42	4.51 ± 1.81	2.22 ± 0.27	3.20 ± 3.34	3.37 ± 0.27	4.39 ± 3.78	$\mathbf{5 \pm 2}$
PIE-NeRF solver	0.49 ± 0.40	1.02 ± 0.88	0.46 ± 0.36	1.09 ± 0.69	0.48 ± 0.19	0.97 ± 0.33	30 ± 5

Table 2. Ablation on DME. "NaN rate" is the percentage of test samples whose predicted next-step points contain NaNs. $\|f\|_{\text{mean}}$ denotes the mean magnitude of the predicted forces. "100×/0.01× parameters" denote scaling of Young's modulus E.

Method	Rel-L_2 (%)↓	NaN rate (%)↓	$\|f\|_{\text{mean}}$
PhysOCT	0.44 ± 0.18	0	0.69 ± 0.21
Kernel-level force regression	86 ± 81	0	–
w/o SPD construction	–	100	–
100× parameters	0.59 ± 0.21	0	3.76 ± 1.84
0.01× parameters	0.56 ± 1.20	0	0.05 ± 0.03

regressing kernel-level generalized forces substantially degrades accuracy, consistent with ill-conditioning induced by coupled DoFs. Removing SPD construction leads to frequent numerical failures. We define a failure as the presence of NaNs in predicted next-step point locations and report the NaN rate.

We also observe that the elastic coefficient (Young's modulus, E) primarily rescales the solver's stiffness, which in turn determines the magnitude of the inferred generalized external forces and the overall gradient scale during end-to-end training. Since supervision is on next-step positions, an improper E can mis-scale stiffness and force magnitudes, resulting in poorly conditioned gradient scales (exploding or vanishing) and unstable optimization. In practice, we therefore adopt physically plausible and numerically stable parameter scales (e.g., $E = 332$ kPa).

3.3 Generative Performance as a Plug-in Conditioner

We evaluate whether PhysOCT improves conditional OCT B-scan synthesis in fidelity, perceptual quality, and structural consistency. We compare DDPM [23] and LDM [24] baselines with their PhysOCT-augmented counterparts. We also report SegGuidedDiff [10] as a strong segmentation-guided baseline. Quantitative results are in Table 3; representative visual comparisons are shown in Fig. 3.

Quantitative results. PhysOCT consistently improves both DDPM and LDM across CNV/AMD/DME. In all settings, adding PhysOCT increases PSNR/SSIM and reduces FID, indicating improved pixel-level fidelity and closer alignment to the real-image distribution. Importantly, this gain comes with negligible runtime overhead: the additional PhysOCT inference takes only ∼34 ms per sample, which is about a 5.0% increase relative to the ∼0.6 s sampling time (DDIM,

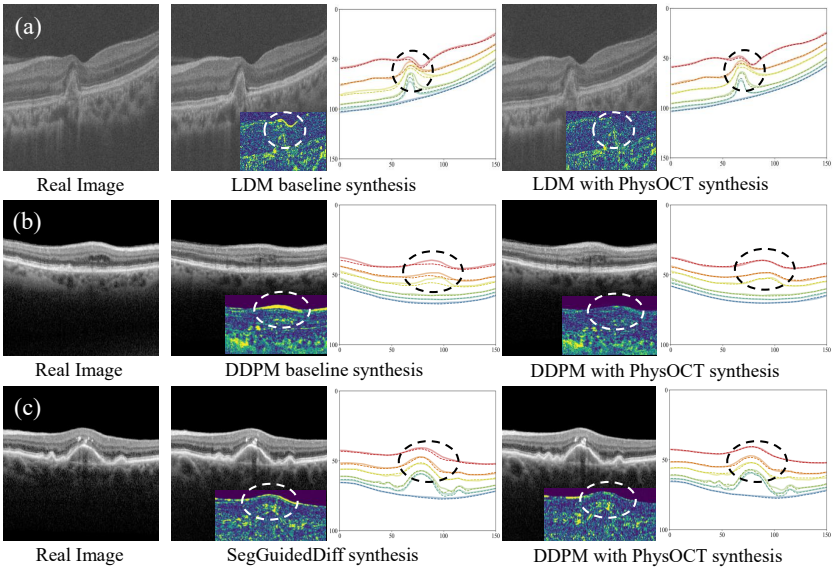

Fig. 3. Qualitative comparison of conditional synthesis. Columns 3 and 5 overlay retinal layer annotations extracted from generated images with ground-truth layers; the bottom-right insets of Columns 2 and 4 show per-pixel difference heatmaps to real targets. PhysOCT reduces boundary-aligned errors under lesion-driven deformation. (a) LDM on CNV. (b) DDPM on AMD. (c) SegGuidedDiff vs. PhysOCT-augmented on DME.

50 steps), yet yields a clear improvement in synthesis realism and lesion–layer anatomical consistency.

Qualitative results. Fig. 3 highlights the structural impact of PhysOCT near lesion-affected regions. Baseline diffusion models often introduce layer deviations that appear as strong responses in difference maps along layer boundaries. Conditioning on the physics-predicted layer prior and deformation-aligned sketch reduces these artifacts, yielding anatomically more consistent layer configurations.

4 Conclusion

We presented PhysOCT, a physics-guided plug-in framework that predicts lesion-induced forces and advances retinal geometry via a stable differentiable mesh-free solver, producing a next-step layer prior that guides conditional diffusion generation for anatomically consistent OCT synthesis. Experiments across multiple OCT cohorts show improved deformation prediction and consistent gains in synthesis quality. The current 2D formulation can extend naturally to 3D via the same force prediction and GMLS solver, with a 2.5D formulation offering a computationally feasible intermediate step left to future work.

Table 3. Conditional generation performance (mean±std). Higher is better for PSNR/SSIM; lower is better for FID.

Method	AMD			CNV			DME		
	PSNR↑	SSIM↑	FID↓	PSNR↑	SSIM↑	FID↓	PSNR↑	SSIM↑	FID↓
DDPM	21.1 ± 2.6	0.57 ± 0.09	53.9	27.9 ± 1.4	0.55 ± 0.03	56.2	21.2 ± 1.3	0.55 ± 0.08	54.4
DDPM+PhysOCT	23.5 ± 2.8	0.62 ± 0.10	**45.5**	$\mathbf{29.7 \pm 4.1}$	$\mathbf{0.58 \pm 0.03}$	**50.9**	$\mathbf{23.9 \pm 1.1}$	0.59 ± 0.08	**46.7**
LDM	21.4 ± 2.8	0.61 ± 0.10	60.4	23.9 ± 1.2	0.48 ± 0.02	72.4	21.3 ± 1.5	0.60 ± 0.07	73
LDM+PhysOCT	$\mathbf{23.8 \pm 2.7}$	$\mathbf{0.65 \pm 0.09}$	48.4	26.4 ± 1.1	0.52 ± 0.02	55.4	23.7 ± 1.3	$\mathbf{0.65 \pm 0.07}$	55.3
SegGuidedDiff	19.5 ± 4.0	0.47 ± 0.07	80.4	19.1 ± 6.4	0.45 ± 0.23	87.8	19.3 ± 4.2	0.46 ± 0.06	83.3

Acknowledgments. This work was supported in part by the Major Research Plan of the National Natural Science Foundation of China (92370109), National Natural Science Foundation of China (62172223), and Natural Science Foundation of Jiangsu Province (BK20252033).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Schmitt, J.M.: Optical coherence tomography (OCT): a review. IEEE Journal of Selected Topics in Quantum Electronics **5**(4), 1205–1215 (2002)
2. Zhang, L. et al.: Quantifying disrupted outer retinal-subretinal layer in SD-OCT images in choroidal neovascularization. Investigative Ophthalmology & Visual Science **55**(4), 2329–2335 (2014)
3. Brandão Fantozzi, C. et al.: A Comparative Study Between Clinical Optical Coherence Tomography (OCT) Analysis and Artificial Intelligence-Based Quantitative Evaluation in the Diagnosis of Diabetic Macular Edema. Vision **9**(3), 75 (2025)
4. Moradi, M. et al.: Deep ensemble learning for automated non-advanced AMD classification using optimized retinal layer segmentation and SD-OCT scans. Computers in Biology and Medicine **154**, 106512 (2023)
5. Goodfellow, I.J. et al.: Generative adversarial nets. In: Advances in Neural Information Processing Systems, vol. 27 (2014)
6. Sohl-Dickstein, J. et al.: Deep unsupervised learning using nonequilibrium thermodynamics. In: International Conference on Machine Learning (ICML) (2015)
7. Liu, X., Gong, C., Liu, Q.: Flow straight and fast: Learning to generate and transfer data with rectified flow. In: International Conference on Learning Representations (ICLR) (2023)
8. Chambon, P. et al.: RoentGen: Vision-language foundation model for chest X-ray generation. arXiv preprint arXiv:2211.12737 (2022)
9. Zhan, C. et al.: MedM2G: Unifying medical multi-modal generation via cross-guided diffusion with visual invariant. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
10. Konz, N. et al.: Anatomically-controllable medical image generation with segmentation-guided diffusion models. In: Medical Image Computing and Computer Assisted Intervention – MICCAI (2024)
11. Li, S. et al.: RetiDiff: Diffusion-based synthesis of retinal OCT images for enhanced segmentation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI (2025)

12. Wu, Y. et al.: Retinal OCT synthesis with denoising diffusion probabilistic models for layer segmentation. arXiv preprint arXiv:2311.05479 (2023)
13. Zienkiewicz, O.C., Taylor, R.L.: The Finite Element Method, 5th edn. Butterworth-Heinemann, Oxford (2000)
14. Faure, F. et al.: SOFA: A multi-model framework for interactive physical simulation. In: Soft Tissue Biomechanical Modeling for Computer Assisted Surgery. Springer (2012)
15. Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. **378**, 686–707 (2019)
16. Feng, Y. et al.: PIE-NeRF: Physics-based interactive elastodynamics with NeRF. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
17. Li, X. et al.: PAC-NeRF: Physics augmented continuum neural radiance fields for geometry-agnostic system identification. In: International Conference on Learning Representations (ICLR) (2023)
18. Prabhushankar, M., Philbrick, K., Burlina, P.: OLIVES Dataset: Ophthalmic Labels for Investigating Visual Eye Semantics. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 35, pp. 9201–9216 (2022)
19. Huber, M., Binder, P., Frohmann, M.: Monitoring Age-Related Macular Degeneration Progression in Optical Coherence Tomography (MARIO), Task 1–MICCAI Challenge 2024. In: International Challenge on Device-Independent Diabetic Macular Edema Onset Prediction, pp. 154–162. Springer Nature Switzerland (2024)
20. Xie, B. et al.: OCT Image Generation for Subretinal Fluid Segmentation with Finite Element Structure Modeling. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1–5. IEEE (2024)
21. Huang, K. et al.: Diverse Data Generation for Retinal Layer Segmentation With Potential Structure Modeling. IEEE Transactions on Medical Imaging **43**(10), 3584–3595 (2024)
22. Ciasca, G., Pagliei, V., Minelli, E., Palermo, F., Nardini, M., Pastore, V., Papi, M., Caporossi, A., De Spirito, M., Minnella, A.M.: Nanomechanical mapping helps explain differences in outcomes of eye microsurgery: A comparative study of macular pathologies. PLoS ONE **14**(8), e0220571 (2019)
23. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: Advances in Neural Information Processing Systems, vol. 33, pp. 6840–6851 (2020)
24. Rombach, R. et al.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10684–10695 (2022)
25. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. pp. 234–241 (2015)