HYPERGEN: CONTROLLABLE PATHOLOGICAL ANOMALY SYNTHESIS VIA HYPERBOLIC SEMANTIC MANIFOLD LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Medical anomaly synthesis has emerged as a promising strategy for alleviating the scarcity of annotated pathological data in clinical image analysis. Recent diffusion-based methods can generate anatomically realistic lesions, yet their latent representations remain largely unstructured, making it difficult to systematically control pathological characteristics and generate diverse abnormalities. To address this challenge, we propose HyperGen, a novel framework that reformulates anomaly synthesis as geometry-guided semantic navigation on a learned hyperbolic manifold. HyperGen first learns a lesion-centric hyperbolic embedding, where the radial coordinate is aligned with Hounsfield Unit (HU) density while the angular coordinates capture complementary morphological and textural variations. By manipulating the manifold in polar coordinates and projecting the navigated embedding back to the Euclidean space, HyperGen provides semantic conditioning for a diffusion model to generate diverse and anatomically consistent pathological images. Extensive experiments on lung CT demonstrate that HyperGen produces more realistic and structurally coherent lesions than strong diffusion-based baselines. Furthermore, the synthesized anomalies consistently improve downstream lesion segmentation, validating the effectiveness of geometry-guided semantic manifold learning for controllable medical anomaly generation.

1 INTRODUCTION

Medical image analysis relies heavily on diverse pathological data, yet abnormal findings are often rare, heterogeneous, and expensive to annotate Tajbakhsh et al. (2021); Chu et al. (2024); Narayanan & Sundaresan (2025). This data scarcity limits the generalization of downstream models and has motivated growing interest in pathological anomaly synthesis, which aims to enrich pathological distributions by generating realistic abnormalities from healthy anatomy Schäfer et al. (2024).

Recent diffusion models have achieved remarkable success in medical image synthesis, producing anatomically realistic and visually coherent lesions. However, existing methods offer limited control over pathological semantics, as anomaly generation is typically driven by masks, prompts, or latent perturbations within largely unstructured latent spaces Bhosale et al. (2025); Qiu et al. (2025); Chang et al. (2025). Consequently, latent manipulations often lead to unpredictable semantic variations, making it difficult to generate specific lesion characteristics or systematically explore diverse pathological patterns. Realistic synthesis alone is therefore insufficient; clinically useful anomaly generation also requires semantic interpretability and controllability

We argue that this limitation stems from the representation of lesion semantics. Most existing synthesis methods organize latent features in Euclidean spaceTehraniNasab et al. (2025); Lei et al. (2025), where semantic variation is implicitly assumed to be relatively flat and uniformly distributed. However, pathological abnormalities rarely evolve in such a manner. Clinically, lesions evolve along multiple factors, including attenuation, morphology, and internal texture Zwanenburg et al. (2020). These factors are correlated yet not uniformly distributed, suggesting that anomaly representations should admit meaningful directions along which semantic attributes can be manipulated independently. Such a manifold is difficult to represent in Euclidean space. In contrast, hyperbolic geometry

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.

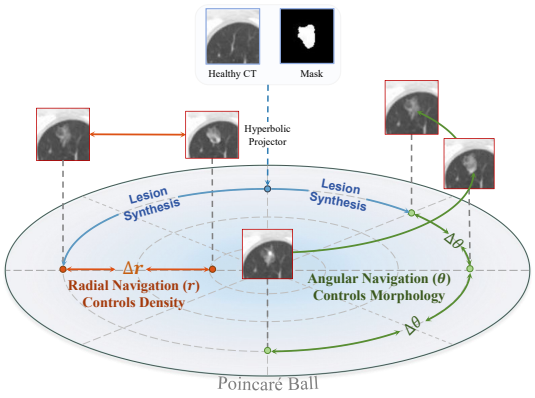


Figure 1: Hyperbolic-space view of unsupervised controllable anomaly generation. By navigating a hierarchically organized hyperbolic feature space through the radial coordinate r and angular coordinate θ, the model enables diverse, controllable, and semantically meaningful anomaly synthesis without requiring explicit anomaly labels.

naturally accommodates hierarchical and non-uniform semantic organization through its negatively curved metric, enabling compact representation of diverse pathological patterns while preserving their latent relationships Desai et al. (2023); Li et al. (2023); Fu et al. (2024); Li et al. (2025).

To this end, we propose **HyperGen**, a hyperbolic manifold learning framework for controllable pathological anomaly synthesis. Given an image and a target insertion region, HyperGen first extracts local anatomical features from the masked region. These local features are projected onto a learned pathological semantic manifold, where geometric navigation transforms them into representations corresponding to diverse abnormalities before being decoded by the diffusion model (Figure 1).

Specifically, the radial coordinate is aligned with lesion density, while the angular component captures morphology and texture variation. The navigated embedding is subsequently projected back to the Euclidean space and injected as semantic conditioning into a diffusion model, enabling anatomically consistent image synthesis while providing explicit control over lesion characteristics.

Our main contributions are summarized as follows:

- We formulate medical anomaly synthesis as semantic manifold navigation. We learn a lesion-centric hyperbolic representation that organizes pathological semantics into an interpretable latent geometry, where radial and angular coordinates capture complementary lesion characteristics.
- We propose a geometry-guided conditioning framework for diffusion models. By navigating the learned manifold before diffusion, HyperGen generates diverse yet anatomically consistent anomalies while enabling controllable manipulation of lesion attributes.
- Extensive experiments on multiple medical anomaly datasets demonstrate that the proposed method can generate diverse lesions, which consistently improves the performance on downstream tasks.

2 Related Work

Diffusion Models for Medical Image Synthesis Diffusion models have emerged as a powerful framework for image generation. Denoising Diffusion Probabilistic Models (DDPM) Ho et al. (2020) formulate generation as a gradual reverse denoising process from Gaussian noise, providing a stable alternative to adversarial training. Denoising Diffusion Implicit Models (DDIM) Song et al. (2020) further improve sampling efficiency by introducing a non-Markovian formulation that enables faster deterministic inference. Latent Diffusion Models (LDM) Rombach et al. (2022) reduce the computational cost of diffusion by performing generation in a compressed latent space, making high-resolution synthesis more practical.In medical imaging, these ideas have been extended to

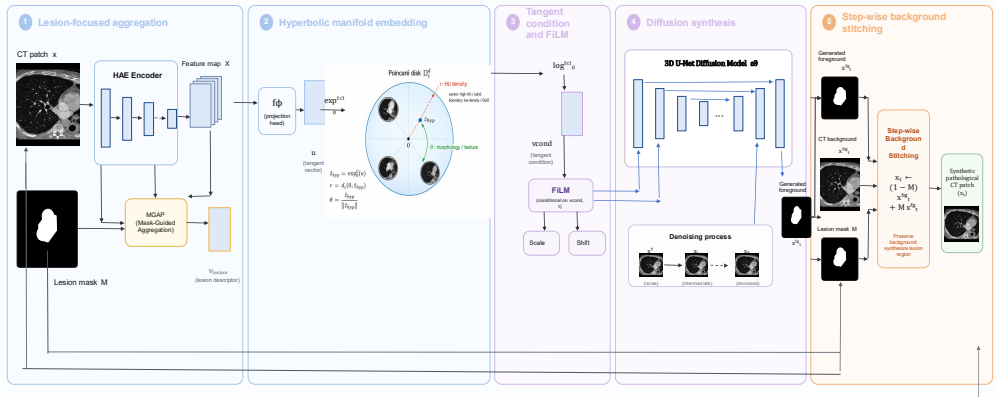


Figure 2: Overview of the proposed HyperGen framework. The model extracts lesion-focused features, maps them into a hyperbolic Poincaré ball for structured semantic organization, and uses the resulting manifold coordinates to condition a diffusion-based synthesis process.

volumetric and modality-specific synthesis tasks. Medical Diffusion Khader et al. (2023) adapts diffusion models to 3D medical image generation and demonstrates their applicability to volumetric data. Brain Imaging Generation with Latent Diffusion Models Pinaya et al. (2022) shows that latent diffusion can synthesize high-quality brain MRI while improving efficiency compared with pixel-space generation. Conditional Diffusion Models for Semantic 3D Brain MRI Synthesis Dorjsembe et al. (2024) introduce semantic conditioning to improve anatomical controllability in 3D synthesis. For lesion-specific generation, LeFusion Zhang et al. (2025) proposes a lesion-focused diffusion framework that decouples lesion and background generation, leading to improved background preservation in lung nodule synthesis. Despite these advances, most existing methods are still built in Euclidean latent spaces, which may be less suitable for organizing hierarchical pathological variations.

Controllable Medical Image Generation and Editing Controllable generation aims to guide synthesis with anatomical, spatial, or semantic constraints. RePaint Lugmayr et al. (2022) formulates image editing as a diffusion-based inpainting problem and shows that pretrained unconditional diffusion models can be repurposed for masked content generation under highly flexible mask shapes. Segmentation-guided diffusion methods Konz et al. (2024) use anatomical masks as spatial conditions to constrain synthesis and preserve structural consistency. MedEdit Alaya et al. (2024) studies counterfactual medical image editing on brain MRI, aiming to induce pathology while maintaining the identity and integrity of the original scan.

Hyperbolic Representation Learning Hyperbolic geometry provides an effective framework for representing hierarchical structures. Poincaré embeddings Nickel & Kiela (2017) show that hyperbolic space can represent tree-like relations with substantially lower distortion than Euclidean embeddings. Hyperbolic neural networks Ganea et al. (2018) extend this idea to deep learning by defining neural operations directly in hyperbolic space, enabling end-to-end representation learning with non-Euclidean geometry.More recent work has explored hyperbolic latent spaces for generative modeling. HypDAE Li et al. (2025) combines hyperbolic representations with diffusion autoencoders to support hierarchical image generation and editing, showing the potential of hyperbolic geometry for structuring generative latent spaces. However, HypDAE is developed for whole-image modeling in natural-image settings rather than lesion-centric medical synthesis.

3 METHODOLOGY

HyperGen performs lesion synthesis through conditional generation in a lesion-centric hyperbolic latent manifold. The central motivation is that pathological abnormalities exhibit structured and non-uniform semantic variation, such as continuous changes in radiological density and heterogeneous changes in morphology and texture, which are difficult to organize and manipulate in a conven-

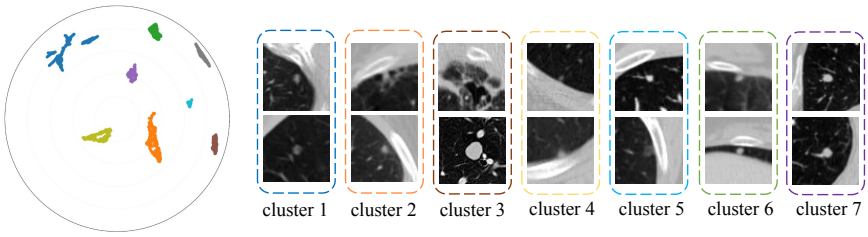


Figure 3: UMAP visualization of the learned hyperbolic manifold embeddings within the Poincaré ball. The embeddings are colored by Hounsfield Unit (HU) density, which serves as the imaging manifestation of the underlying radiological density. This visualization demonstrates that the hyperbolic space effectively captures the continuous spectrum of lesion attenuation. The accompanying gallery (right) shows representative samples from various clusters, illustrating the diverse morphological features and HU profiles learned by HyperGen.

tional Euclidean latent space. Rather than treating lesion synthesis as unconstrained appearance generation, we formulate it as geometry-guided semantic organization and controllable navigation in hyperbolic space. As illustrated in Fig. 2, the proposed framework consists of three stages: (1) lesion-focused feature extraction that maps an input patch to a compact pathology-aware Euclidean representation, (2) projection and embedding into a Poincaré ball for structured semantic organization, and (3) manifold-conditioned diffusion for anatomically consistent lesion synthesis.

3.1 Lesion-Focused Feature Aggregation

In volumetric medical images, lesions typically occupy only a small spatial fraction of the full input, while most voxels correspond to normal anatomical background. Direct global pooling over the entire feature map may therefore dilute pathology-specific information and produce representations dominated by healthy structures. To obtain a compact descriptor that emphasizes lesion semantics, we perform mask-guided feature aggregation.

Let $\mathbf{X} \in \mathbb{R}^{C\times D'\times H'\times W'}$ denote the feature map extracted by the encoder, and let $\mathbf{M} \in \{0,1\}^{D'\times H'\times W'}$ be the binary lesion mask resized to the feature resolution. We define the foreground index set as $\Omega = \{i \mid \mathbf{M}_i = 1\}$. The lesion-focused pooled feature $\mathbf{v}_{\text{lesion}} \in \mathbb{R}^C$ is computed by

$$\mathbf{v}_{\text{lesion}} = \frac{1}{|\Omega| + \epsilon} \sum_{i\in\Omega} \mathbf{X}_i, \tag{1}$$

where $\epsilon = 10^{-6}$ prevents division by zero. This operation restricts aggregation to lesion voxels, suppresses interference from surrounding background anatomy, and provides a compact pathology-aware descriptor for subsequent manifold organization and controllable synthesis. Its computational complexity scales with the lesion volume as $\mathcal{O}(|\Omega|)$.

Compared with standard global average pooling, the resulting representation is more pathology-aware and provides a more suitable input for subsequent manifold organization.

3.2 Hyperbolic Manifold Embedding

To organize lesion semantics in a latent space that supports interpretable control, we embed the lesion descriptor into a d-dimensional Poincaré ball

$$\mathbb{D}_c^d = \{\mathbf{z} \in \mathbb{R}^d : c\|\mathbf{z}\|^2 < 1\}, \tag{2}$$

where $c > 0$ denotes the curvature parameter. In our implementation, c is learnable and initialized as $c = 1.0$, allowing the manifold geometry to adapt to the data distribution.

Since the pooled feature $\mathbf{v}_{\text{lesion}} \in \mathbb{R}^C$ does not necessarily match the manifold dimension d, we first project it to the tangent space at the origin through a learnable projection head:

$$\mathbf{u} = f_\phi(\mathbf{v}_{\text{lesion}}), \qquad \mathbf{u} \in \mathbb{R}^d, \tag{3}$$

where $f_\phi(\cdot)$ is implemented as a linear layer or lightweight MLP. We then treat $\mathbf{u}$ as a tangent vector in $T_{\mathbf{0}}\mathbb{D}_c^d$ and map it into the hyperbolic manifold using the exponential map:

$$\mathbf{z}_{\text{hyp}} = \exp_{\mathbf{0}}^c(\mathbf{u}) = \frac{\tanh(\sqrt{c}\|\mathbf{u}\|)}{\sqrt{c}\|\mathbf{u}\|}\mathbf{u}. \tag{4}$$

In practice, we add a small constant $\epsilon = 10^{-8}$ to the denominator to avoid numerical instability when $\|\mathbf{u}\| \approx 0$.

The Poincaré ball is equipped with metric

$$g_{\mathbf{z}} = \left(\frac{2}{1 - c\|\mathbf{z}\|^2}\right)^2 g^E, \tag{5}$$

which induces a geometry in which geodesic distance from the origin captures radial organization, while direction captures residual semantic variation. Accordingly, the embedding $\mathbf{z}_{\text{hyp}}$ can be decomposed into a geodesic radius and a directional component:

$$r = d_c(\mathbf{0}, \mathbf{z}_{\text{hyp}}), \qquad \theta = \frac{\mathbf{z}_{\text{hyp}}}{\|\mathbf{z}_{\text{hyp}}\|}. \tag{6}$$

Here, r denotes the geodesic distance from the origin, and θ represents the direction on the unit sphere. Figure 3 visualizes the learned hyperbolic manifold, with embeddings colored by HU density. The embeddings exhibit a structured semantic organization, where lesions with similar densities cluster together while distinct densities are well separated. This confirms that the proposed radial alignment loss effectively aligns the manifold geometry with radiological density. This structured hyperbolic representation serves as the semantic condition for diffusion synthesis.

3.3 Hyperbolic-Conditioned Diffusion

To inject semantically organized lesion information into the synthesis model, we project the hyperbolic embedding back to the tangent space at the origin via the logarithmic map:

$$\mathbf{v}_{\text{cond}} = \log_{\mathbf{0}}^c(\mathbf{z}_{\text{hyp}}) = \text{arctanh}(\sqrt{c}\|\mathbf{z}_{\text{hyp}}\|)\,\frac{\mathbf{z}_{\text{hyp}}}{\sqrt{c}\|\mathbf{z}_{\text{hyp}}\|}. \tag{7}$$

For numerical stability, the norm term is clipped to be smaller than $1/\sqrt{c}$ in implementation.

The condition vector $\mathbf{v}_{\text{cond}}$ is injected into the diffusion UNet through Feature-wise Linear Modulation (FiLM) Perez et al. (2018). Specifically, we use lightweight MLPs to predict affine parameters from $\mathbf{v}_{\text{cond}}$ and the diffusion timestep t, and apply them to intermediate feature activations:

$$\mathbf{h}_{\text{mod}} = \gamma(\mathbf{v}_{\text{cond}}, t) \odot \mathbf{h} + \beta(\mathbf{v}_{\text{cond}}, t), \tag{8}$$

where $\gamma(\cdot)$ and $\beta(\cdot)$ denote FiLM scale and bias functions. This modulation allows the denoising trajectory to adapt to lesion semantics organized in the hyperbolic manifold.

To preserve patient-specific anatomy, we enforce spatial consistency through mask-based stitching at each reverse diffusion step rather than only at the final output. Let $\mathbf{x}_t^{\text{fg}}$ denote the denoised foreground estimate and let $\mathbf{x}_t^{\text{bg}} \sim q(\mathbf{x}_t \mid \mathbf{x}_{\text{bg}})$ denote the noised healthy background. We update the intermediate sample by

$$\mathbf{x}_t \leftarrow (1 - \mathbf{M}) \odot \mathbf{x}_t^{\text{bg}} + \mathbf{M} \odot \mathbf{x}_t^{\text{fg}}. \tag{9}$$

This operation preserves the anatomical background while confining the synthesis process to the lesion region. Because the denoising network still operates on the full patch, the generated lesion can remain consistent with surrounding tissues and boundaries. Compared with post-hoc lesion pasting, this step-wise masking strategy allows the reverse process to adapt lesion appearance jointly with local context, which helps maintain smoother lesion-background transitions.

Inference-Time Manifold Navigation. The learned manifold supports interpretable latent manipulation during inference. Given an encoded lesion representation $\mathbf{z}_{\text{hyp}}$, we decompose it into (r, θ) and modify the radial coordinate while preserving direction:

$$\hat{r} = \alpha \cdot r, \tag{10}$$

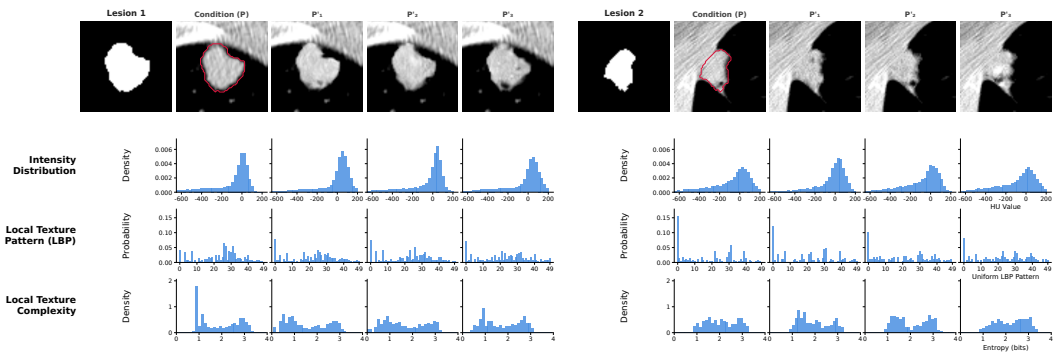


Figure 4: HyperGen generates diverse images from the same conditioning lesion by sampling different latent variables, yet it maintains high consistency in key lesion characteristics. Comparisons of voxel intensity, local texture (LBP), and local entropy distributions confirm that while stochastic sampling introduces realistic visual variations, the underlying clinical features and heterogeneity of the original lesion are effectively preserved.

where α is a user-controlled scaling factor. The manipulated representation is then mapped back to the Poincaré ball as

$$\hat{\mathbf{z}}_{\text{hyp}} = \frac{1}{\sqrt{c}} \tanh\left(\frac{\sqrt{c}\hat{r}}{2}\right)\theta. \tag{11}$$

Finally, $\hat{\mathbf{z}}_{\text{hyp}}$ is projected via Eq. 7 and used to condition the diffusion model. Since the radial coordinate is aligned with lesion density, moving toward the origin corresponds to higher normalized attenuation, while larger radii correspond to lower-density appearances.

In principle, directional manipulation can also be performed by modifying or interpolating θ while preserving r, enabling exploration of morphology-related variation. In this work, we mainly focus on radial navigation because of its direct radiological interpretability.

Algorithm 1 Two-Stage Training of HyperGen

Require: Dataset $\mathcal{D} = \{(\mathbf{x}, \mathbf{x}^{\text{HU}}, \mathbf{M})\}$, HU window $[H_{\min}, H_{\max}]$, weights λ_1, λ_2
Ensure: Trained manifold encoder and diffusion network

Stage I: Hyperbolic Embedding Learning
while not converged **do**
  Sample $(\mathbf{x}, \mathbf{x}^{\text{HU}}, \mathbf{M}) \sim \mathcal{D}$
  $\mathbf{X} \leftarrow E_\psi(\mathbf{x})$
  $\mathbf{v}_{\text{lesion}} \leftarrow \text{MaskPool}(\mathbf{X}, \mathbf{M})$
  $\mathbf{z}_{\text{hyp}} \leftarrow \exp_0^c\big(f_\phi(\mathbf{v}_{\text{lesion}})\big)$
  $\hat{\mathbf{x}} \leftarrow G_\omega(\mathbf{z}_{\text{hyp}})$
  Compute masked reconstruction loss $\mathcal{L}_{\text{recon}}$
  $H \leftarrow \text{MaskMean}(\mathbf{x}^{\text{HU}}, \mathbf{M})$
  $\bar{H} \leftarrow \dfrac{\text{clip}(H, H_{\min}, H_{\max}) - H_{\min}}{H_{\max} - H_{\min}}$
  $r \leftarrow \dfrac{2}{\sqrt{c}} \operatorname{arctanh}\big(\sqrt{c}\|\mathbf{z}_{\text{hyp}}\|_2\big)$
  $\mathcal{L}_{\text{radius}} \leftarrow \|r - (1 - \bar{H})\|_2^2$
  $\mathcal{L}_{\text{HAE}} \leftarrow \lambda_1 \mathcal{L}_{\text{recon}} + \lambda_2 \mathcal{L}_{\text{radius}}$
  Update ψ, ϕ, ω, c using $\nabla \mathcal{L}_{\text{HAE}}$
end while
Freeze $\{E_\psi, f_\phi, c\}$; discard G_ω

Stage II: Diffusion Learning
while not converged **do**
  Sample $(\mathbf{x}_0, \mathbf{M}) \sim \mathcal{D}$
  $\mathbf{v}_{\text{lesion}} \leftarrow \text{MaskPool}(E_\psi(\mathbf{x}_0), \mathbf{M})$
  $\mathbf{z}_{\text{hyp}} \leftarrow \exp_0^c\big(f_\phi(\mathbf{v}_{\text{lesion}})\big)$
  $\mathbf{v}_{\text{cond}} \leftarrow \log_0^c(\mathbf{z}_{\text{hyp}})$
  Sample $t \sim \mathcal{U}\{1, \ldots, T\}$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
  $\mathbf{x}_t \leftarrow \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t}\boldsymbol{\epsilon}$
  $\hat{\boldsymbol{\epsilon}} \leftarrow \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{v}_{\text{cond}}, t)$
  $\mathcal{L}_{\text{diff}} \leftarrow \dfrac{\sum_i \mathbf{M}_i \|\boldsymbol{\epsilon}_i - \hat{\boldsymbol{\epsilon}}_i\|_2^2}{\sum_i \mathbf{M}_i + \varepsilon}$
  Update θ using $\nabla \mathcal{L}_{\text{diff}}$
end while
return trained HyperGen

By normalizing with the lesion volume, the loss remains stable across lesions of different sizes. Although the error is emphasized within the foreground region, the denoising network still receives the full image context, which helps maintain anatomical compatibility between synthesized lesions and surrounding tissues.

4 EXPERIMENTS

In this section, we provide a comprehensive evaluation of HyperGen. We compare our framework against recent state-of-the-art methods in terms of both generative visual fidelity and downstream segmentation utility, followed by an in-depth analysis of the learned hyperbolic manifold.

4.1 EXPERIMENTAL SETUP

Dataset and Preprocessing. All experiments are conducted on the public LIDC-IDRI dataset. For generative training, we extract $16 \times 64 \times 64$ volumetric patches centered on the annotated nodules. To standardize the physical density, all volumes are clamped to the standard lung window $[-1000, 400]$ HU and linearly normalized to $[-1, 1]$.

Downstream Task Protocol. To evaluate the practical utility of the synthesized samples, we adopt the nnU-Net framework for downstream segmentation. We consider augmentation schemes based on real pathological cases (P) and synthetic cases ($P + P'/P + N'/P + N''$). Specifically, P denotes the original real pathological cases. $P + P'$ and $P + N'$ denote synthetic pathological cases generated from pathological and normal volumes, respectively, using standard synthesis strategies and then combined with real cases. We further introduce $P + N''$, namely geometry-guided pseudo-anomalies generated by navigating the learned hyperbolic manifold through (r, θ). Unlike standard pseudo-anomaly generation, this protocol enables controlled variation in radiological density and morphology, directly reflecting the design of the proposed lesion-centric manifold.

Evaluation Metrics. Generative fidelity is measured using Fréchet Inception Distance (FID), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR). Downstream segmentation performance is evaluated using the Dice Similarity Coefficient (Dice) and Normalized Surface Distance (NSD).

Table 1: Quantitative comparison of generative quality on LIDC-IDRI. HyperGen shows improved structural alignment and lower feature distribution divergence.

Method	**FID** $\downarrow$	**SSIM** $\uparrow$	**PSNR (dB)** $\uparrow$
LDM	58.42	0.8923	32.14
RePaint	52.19	0.9156	35.81
LeFusion	48.75	0.9320	38.45
HyperGen (Ours)	**36.12**	**0.9512**	**41.22**

4.2 GENERATIVE QUALITY EVALUATION

To assess the fidelity of the synthesized pathological textures, we compare HyperGen with standard Euclidean diffusion models, including LDM, RePaint, and the lesion-focused LeFusion baseline. As reported in Table 1, HyperGen achieves competitive performance across all image quality metrics.

The reduction in FID suggests that our hyperbolic latent space may better capture the distribution of lung nodules compared to Euclidean approaches. The improvements in SSIM and PSNR are consistent with the hypothesis that spatial masking helps preserve surrounding anatomical structures.

4.3 DOWNSTREAM CLINICAL UTILITY

Beyond visual fidelity, we evaluate whether HyperGen can synthesize anatomically plausible data to improve clinical model training. Table 2 compares the performance of the nnU-Net segmentor trained under various synthetic augmentation strategies.

Performance Gains. HyperGen significantly enhances the baseline Dice coefficient from 78.26% to 83.02% (+4.76%) and NSD from 88.90 to 94.10 (+5.20) through the manifold-guided navigation strategy (N''). This substantial improvement validates the efficacy of hyperbolic (r, θ) coordinate control. Unlike standard pseudo-anomaly methods (N') that rely on fixed templates, the N'' protocol facilitates the synthesis of lesions with precisely controlled radiological density (radius r) and

Table 2: Downstream segmentation performance on LIDC-IDRI under two augmentation routes. P' and N' denote lesions synthesized on pathological and normal cases, respectively. The real-only baseline is repeated as a shared reference. HyperGen + Nav. uses condition-guided navigation during generation. Bold and underlined values indicate the best and second-best results in each setting, respectively.

Method	**Pathological-Case Reconstruction** $P + P'$		**Normal-Case Generation** $P + N'$	
	Dice	**NSD**	**Dice**	**NSD**
Real only (P)	78.26	88.90	78.26	88.90
Hand-Crafted Marino et al. (2023)	76.80	87.94	75.10	85.50
Cond-Diffusion Pinaya et al. (2022); Cao et al. (2022)	77.05	87.69	76.62	86.44
Cond-Diffusion (L) Li et al. (2024)	76.66	87.20	76.71	86.83
RePaint Lugmayr et al. (2022)	77.57	88.07	–	–
Copy-Paste	–	–	77.29	87.60
LeFusion Zhang et al. (2025)	78.77	89.25	77.67	87.94
LeFusion-H Zhang et al. (2025)	<u>80.62</u>	<u>90.90</u>	80.19	89.75
LeFusion-H+DiffMask Zhang et al. (2025)	–	–	<u>82.66</u>	<u>92.49</u>
HyperGen	79.36	90.35	82.64	91.65
HyperGen + Nav.	**80.76**	**91.93**	**83.02**	**94.10**

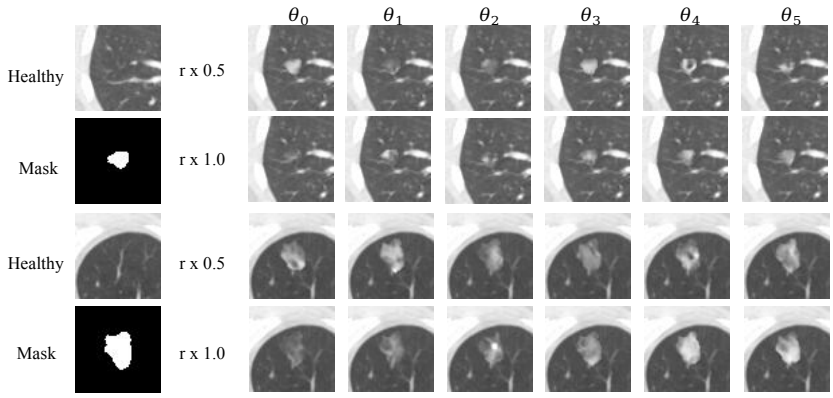


Figure 5: Visualization of manifold interpretability via polar coordinate navigation. (Left) Radial navigation (r) enables continuous control over radiological density (represented by HU values), where moving towards the origin ($r \to 0$) yields more solid structures. (Right) Angular navigation (θ) modulates morphological features such as spiculation and lobulation while preserving the target radiological density.

morphology (angle θ) within healthy anatomical contexts. Such targeted data enrichment provides the segmentor with a more diverse and clinically representative training set, leading to improved generalization.

4.4 Cross-Dataset Evaluation on Emidec

Cross-dataset results are reported in Table 3. HyperGen achieves the highest PSNR on Emidec-MI and Emidec-Avg., together with the best SSIM across all three subsets. The improvement is most pronounced on Emidec-MI, where HyperGen outperforms LeFusion-S and RePaint by 7.50 dB and 13.34 dB, respectively.

On Emidec-PMO, LeFusion-S obtains a marginally higher PSNR (27.71 vs. 27.08 dB), whereas HyperGen achieves the highest SSIM (89.83%), indicating stronger structural consistency. Overall,

Table 3: Comparative results on the Emidec datasets using PSNR and SSIM metrics. HyperGen demonstrates superior generative fidelity across all evaluation subsets. Bold values indicate the best performance per column.

Method	Emidec-MI		Emidec-PMO		Emidec-Avg.	
	PSNR↑	SSIM↑ (%)	PSNR↑	SSIM↑ (%)	PSNR↑	SSIM↑ (%)
Hand-Crafted (2023)	9.39	8.30	7.63	9.15	8.51	8.70
Cond-Diffusion (2022)	13.25	46.92	8.00	9.23	10.62	28.07
RePaint (2022)	19.81	80.68	15.23	70.27	17.52	75.47
LeFusion (2025)	25.65	91.78	**27.71**	89.42	26.68	90.60
HyperGen (Ours)	**33.15**	**92.42**	27.08	**89.83**	**31.12**	**91.27**

these results support the transferability of the learned hyperbolic conditioning representation across different pathological distributions.

4.5 Pathological Condition Preservation

The stability of the learned pathological condition is further examined under stochastic diffusion sampling. Given a fixed conditioning lesion P, HyperGen produces independent reconstructions P'_1, P'_2, and P'_3 using different noise realizations.

As shown in Fig. 4, the reconstructed samples remain closely aligned with P in voxel-intensity, local binary pattern, and local-entropy distributions. This agreement demonstrates that lesion attenuation, local texture, and intralesional heterogeneity are consistently preserved despite variations in the denoising trajectory. The favorable P' reconstruction results on Emidec further indicate that this condition-preserving behavior remains stable across datasets.

4.6 Hyperbolic Manifold Semantic Navigation

Beyond high-quality synthesis, HyperGen learns a semantically interpretable hyperbolic manifold that supports controllable lesion generation. We analyze the learned manifold from both global and local perspectives.

Radial Navigation (r). As shown in Fig. 5, varying the geodesic radius primarily controls lesion attenuation. Moving toward the origin generates denser, solid nodules, whereas increasing r progressively produces lower-density, ground-glass-like lesions. This behavior is consistent with the proposed radial alignment loss, demonstrating that the radial coordinate provides an interpretable control axis for radiological density.

Angular Navigation (θ). With the radial coordinate fixed, varying the angular direction primarily alters lesion morphology and texture, producing diverse boundary irregularities, lobulation, and spiculation while preserving a similar density level. Together, the radial and angular coordinates form an interpretable semantic coordinate system, enabling independent control of lesion density and morphology for geometry-guided anomaly synthesis.

5 Conclusion

We presented HyperGen, a framework that reformulates pathological anomaly synthesis as geometry-guided semantic navigation on a learned hyperbolic manifold. By combining interpretable manifold learning with diffusion-based generation, HyperGen enables controllable synthesis of diverse and anatomically consistent pathological images. Experimental results demonstrate both superior generation quality and consistent improvements in downstream segmentation, validating the effectiveness of the proposed representation. More broadly, this work suggests that structured semantic manifolds offer a principled foundation for controllable medical image generation and may inspire future research on geometry-aware representation learning for medical imaging.

REFERENCES

Malek Ben Alaya, Daniel M Lang, Benedikt Wiestler, Julia A Schnabel, and Cosmin I Bercea. Mededit: Counterfactual diffusion-based image editing on brain mri. In *International Workshop on Simulation and Synthesis in Medical Imaging*, pp. 167–176. Springer, 2024.

Mahesh Bhosale, Abdul Wasi, Yuanhao Zhai, Yunjie Tian, Samuel Border, Nan Xi, Pinaki Sarder, Junsong Yuan, David Doermann, and Xuan Gong. Pathdiff: Histopathology image synthesis with unpaired text and mask conditions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 22415–22424, 2025.

Sheng Cao et al. Conditional diffusion models for medical image synthesis. *Medical Image Analysis*, 79:102456, 2022.

Heng Chang, Yu Shang, Haifeng Wang, Yuxia Liang, Haoyu Wang, Fan Wang, Chen Niu, and Chunfeng Lian. Controllable flow matching for 3d contrast-enhanced brain mri synthesis from non-contrast scans. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 119–128. Springer, 2025.

Yuetan Chu, Changchun Yang, Gongning Luo, Zhaowen Qiu, and Xin Gao. Anatomic-constrained medical image synthesis via physiological density sampling. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 69–79. Springer, 2024.

Karan Desai, Maximilian Nickel, Tanmay Rajpurohit, Justin Johnson, and Shanmukha Ramakrishna Vedantam. Hyperbolic image-text representations. In *International Conference on Machine Learning*, pp. 7694–7731. PMLR, 2023.

Zolnamar Dorjsembe, Hsing-Kuo Pao, Sodtavilan Odonchimed, and Furen Xiao. Conditional diffusion models for semantic 3d brain mri synthesis. *IEEE Journal of Biomedical and Health Informatics*, 28(7):4084–4093, 2024.

Xingcheng Fu, Yisen Gao, Yuecen Wei, Qingyun Sun, Hao Peng, Jianxin Li, and Xianxian Li. Hyperbolic geometric latent diffusion model for graph generation. *arXiv preprint arXiv:2405.03188*, 2024.

Octavian Ganea, Gary Bécigneul, and Thomas Hofmann. Hyperbolic neural networks. *Advances in neural information processing systems*, 31, 2018.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851, 2020.

Firas Khader, Gustav Müller-Franzes, Soroosh Tayebi Arasteh, Tianyu Han, Christoph Haarburger, Maximilian Schulze-Hagen, Philipp Schad, Sandy Engelhardt, Bettina Baeßler, Sebastian Foersch, et al. Denoising diffusion probabilistic models for 3d medical image generation. *Scientific reports*, 13(1):7303, 2023.

Nicholas Konz, Yuwen Chen, Haoyu Dong, and Maciej A Mazurowski. Anatomically-controllable medical image generation with segmentation-guided diffusion models. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 88–98. Springer, 2024.

Wenhui Lei, Hengrui Tian, Linrui Dai, Hanyu Chen, and Xiaofan Zhang. Lesiondiffusion: Towards text-controlled general lesion synthesis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 327–336. Springer, 2025.

Lingxiao Li, Yi Zhang, and Shuhui Wang. The euclidean space is evil: Hyperbolic attribute editing for few-shot image generation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 22714–22724, 2023.

Lingxiao Li, Kaixuan Fan, Boqing Gong, and Xiangyu Yue. Hypdae: Hyperbolic diffusion autoencoders for hierarchical few-shot image generation. In *IEEE/CVF International Conference on Computer Vision*, 2025.

Xiaoxiao Li et al. Large-scale medical image synthesis with conditional diffusion models. *IEEE Transactions on Medical Imaging*, 43(1234–1245), 2024.

Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, and Luc Van Gool. Repaint: Inpainting using denoising diffusion probabilistic models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11461–11471, 2022.

Joseph Marino et al. Synthetic data augmentation for lung nodule classification using generative models. *IEEE Transactions on Medical Imaging*, 42(8):2345–2356, 2023.

Ramanujam Narayanan and Vaanathi Sundaresan. Medlessynth-ld: Lesion synthesis using physics-based noise models for robust lesion segmentation in low-data medical imaging regimes. *Pattern recognition letters*, 188:155–163, 2025.

Maximillian Nickel and Douwe Kiela. Poincaré embeddings for learning hierarchical representations. *Advances in neural information processing systems*, 30, 2017.

Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. Film: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

Walter HL Pinaya, Petru-Daniel Tudosiu, Jessica Dafflon, Pedro F Da Costa, Virginia Fernandez, Parashkev Nachev, Sebastien Ourselin, and M Jorge Cardoso. Brain imaging generation with latent diffusion models. *Medical Image Analysis*, 81:102547, 2022. MICCAI 2022 DGM4MICCAI Workshop.

Kunpeng Qiu, Zhiying Zhou, and Yongxin Guo. Adaptively distilled controlnet: Accelerated training and superior sampling for medical image synthesis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 55–65. Springer, 2025.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10684–10695, 2022.

Raphael Schäfer, Till Nicke, Henning Höfener, Annkristin Lange, Dorit Merhof, Friedrich Feuerhake, Volkmar Schulz, Johannes Lotz, and Fabian Kiessling. Overcoming data scarcity in biomedical imaging with a foundational multi-task model. *Nature Computational Science*, 4(7):495–509, 2024.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502*, 2020.

Nima Tajbakhsh, Holger Roth, Demetri Terzopoulos, and Jianming Liang. Guest editorial annotation-efficient deep learning: the holy grail of medical imaging. *IEEE transactions on medical imaging*, 40(10):2526–2533, 2021.

Zahra TehraniNasab, Amar Kumar, and Tal Arbel. Language-guided trajectory traversal in disentangled stable diffusion latent space for factorized medical image generation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 4846–4851, 2025.

Hantao Zhang, Yuhe Liu, Jiancheng Yang, Shouhong Wan, Xinyuan Wang, Wei Peng, and Pascal Fua. Lefusion: Controllable pathology synthesis via lesion-focused diffusion models. In *International Conference on Learning Representations*, volume 2025, pp. 13232–13253, 2025.

Alex Zwanenburg, Martin Vallières, Mahmoud A Abdalah, Hugo JWL Aerts, Vincent Andrearczyk, Aditya Apte, Saeed Ashrafinia, Spyridon Bakas, Roelof J Beukinga, Ronald Boellaard, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*, 295(2):328–338, 2020.