RESHAPE: Representation Learning for the Explainability of Shapes

Rami Al-Belmpeisi[1,2] (ralbe@dtu.dk), Kristoffer Knutsen Wickstrøm[3], Josefine Vilsbøll Sundgaard[1,2], Peter Hjørringgaard Larsen[2], Anders Bjorholm Dahl[1,†], Robert Jenssen[3,4,5,†], and Michael Kampffmeyer[3,4,†]

[1] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark
[2] Novo Nordisk A/S, Bagsværd, Denmark
[3] Department of Physics and Technology, UiT The Arctic University of Norway, Tromsø, Norway
[4] Norwegian Computing Center, Oslo, Norway
[5] Pioneer Centre for AI, University of Copenhagen, Copenhagen, Denmark

[†]These authors contributed equally to the supervision of this work.

Abstract. Deep learning methods for 3D shape generation and mesh-based reconstruction have advanced rapidly; however, the relation between geometry and learned latent representations remains poorly understood, particularly for shapes. Existing explainable AI methods for 3D shapes predominantly rely on gradient-based techniques, are architecture-specific, or depend on end-to-end differentiability. These limitations challenge their implementation in non-differentiable pipelines of complex generative models. We propose RESHAPE, a model-agnostic framework for latent representation explainability that does not require access to gradients or internal model structure. RESHAPE estimates the geometric importance of individual latent code components by applying stochastic masking to the latent codes, reconstructing the perturbed shapes, and statistically quantifying the geometric deviation of the resulting reconstructions from the original. We demonstrate the properties of RESHAPE on a controlled synthetic dataset and validate its practical relevance on medical shape modeling tasks. Our framework enables (i) highlighting latent dimensions with strong geometric influence within the latent spaces of models, (ii) generating latent saliency maps, and (iii) separating primary shape-defining from stochastic latent dimensions.

Keywords: Representation explainability · Shape reconstruction · Geometric deep learning

1 Introduction

In recent years, deep learning research for 3D shape modeling has explored alternatives to traditional voxel-based approaches [2], including implicit representations [21], point clouds [24], polygonal meshes [3], and implicit surfaces [26].

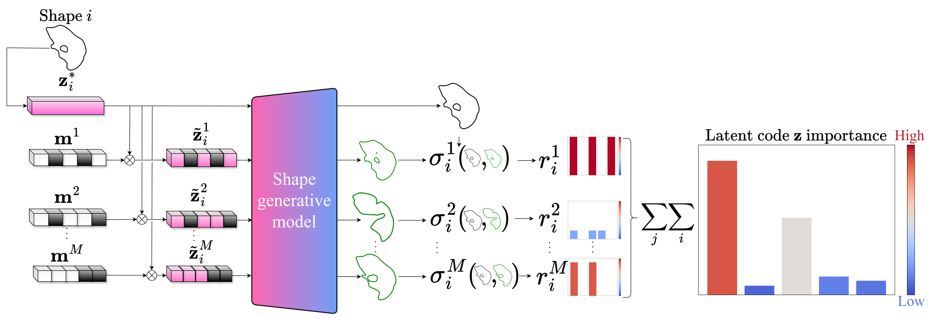

Fig. 1. RESHAPE's framework schematic. Masked latent codes $\tilde{\mathbf{z}}_i^j$ are obtained by perturbing optimized codes $\mathbf{z}_i^*$ for shape i with M stochastic masks m^j $(j = 1, \ldots, M)$. RESHAPE assigns importance scores $r_i(k)$ based on geometric similarity σ between masked reconstructions and original shapes, which are used to identify the latent dimensions k that control shape geometry.

While these approaches offer superior memory efficiency and topological flexibility, their latent spaces are still poorly understood. In this context, we want to answer the question: what is the relation between latent vectors and shapes, i.e., which dimensions affect what parts of the shape, and to what extent?

Answering these questions requires obtaining a better understanding of the geometric abstractions within the latent spaces of shape-generative models. This understanding can find several applications in medical tasks, e.g., identifying latent factors associated with organ enlargement, deformation, or disease progression.

This is paramount in several applications, such as surgical planning, to separate latent factors that drive anatomy and encode nuisance variation, but also more generally for auditing a reconstruction model, which is crucial for organ remodeling and identifying atypical anatomies. Obtaining this understanding requires models that are not only efficient but also simple, therefore easy to interpret. This is particularly true, especially in a medical setting, where computational resources are limited, favoring models with smaller latent spaces over foundational models.

In deep learning-based medical research, meshes are often used to model organ shapes due to their inherent connectivity. However, end-to-end differentiability is compromised in most mesh generation pipelines because they rely on processing steps, such as Marching Cubes or mesh decimation, that break the computational graph. Existing work has not directly quantified the geometric effect of individual latent dimensions in non-differentiable shape-generation pipelines. To address this gap, we propose representation learning for the explainability of shapes (RESHAPE), a framework for shape generation that bridges latent spaces and shape geometry. RESHAPE quantifies the importance of latent dimensions in terms of reconstructed geometry. This procedure does not require gradient flow through the renderer, is model-agnostic, and requires only access to latent codes and the corresponding reconstructed shapes. Our validation experiments of RESHAPE on medical and synthetic meshes demonstrate our framework's ability to (i) identify high-influence latent dimensions within the latent spaces of the models, (ii) rank latent components to assess the monotonic order of geometric influence, and (iii) separate primary shape-defining latent dimensions from low-saliency, noisy, or stochastic dimensions. RESHAPE provides both instance-level and dataset-level explanations of latent representations (see Fig. 1). The implementation of RESHAPE is available at `https://github.com/RamiAlBel/reshape`.

2 Related work

The explainability of representations in non-voxel-based deep learning models for shape reconstruction remains an under-explored research area, a fact reflected in related limitations that hinder progress in this direction [23]. However, a significant amount of effort has been dedicated to similar directions, such as exploring shape representations for 3D objects, ranging from voxels [6], point clouds [16,12], meshes [14,4], implicit neural representations such as signed [18,19] or unsigned distance functions [20], to occupancy networks [22], and radiance fields [13]. The choice of representation used depends on several factors, including shape topology and memory scalability.

Despite increasing attention to representation learning in recent years [9], the central focus is on disentangled latent spaces, where each latent dimension encodes only one humanly understandable concept or factor of variation [5]. Other models explore the mapping from 3D point clouds to latent space using Generative Adversarial Networks [10], but not the reverse.

Explainability involving shapes uses gradient-based methods for identifying significant shape components that drive downstream task decisions, e.g., classification [11] or 3D model generation from single-view images [27]. Some approaches combine both disciplines [7] by providing a disentangled latent space and interpretable latent components that capture topology and shape, achieving impressive results across diverse topologies, yet they do not explicitly quantify per-dimension geometric influence of the latent spaces employed.

Overcoming the aforementioned limitations is not easy, which helps explain why the interpretability of shape representations has stalled for so long. This is further hindered by the lack of work on explaining latent representations for shapes, which hinders the development of such methods and limits the availability of comprehensive baselines.

We address these research gaps by providing a representation-level explainability method relevant to shape modeling. Our occlusion-based approach draws on prior work [25,17,15] to provide usable tools in the absence of an end-to-end differentiable framework. We recommend readers familiarize themselves with our used backbones, DeepSDF [18] and Deep Active Latent Surfaces [8], to better understand the structure of the latent spaces we explore on this work.

3 Methods

3.1 RESHAPE

RESHAPE's objective is to estimate both dataset-level latent code importances $r(k)$ and instance-specific importances $r_i(k)$ to explain shapes, through the per-shape optimized latent codes $\{\mathbf{z}_i^*\}_{i=1}^{N_{\text{test}}}$. To achieve this, we use an occlusion-based approach that stochastically masks latent components to assess their importance by measuring the degradation in the reconstructed geometry.

Let $\mathbf{m}^j \sim \text{Bernoulli}(P)^d$ (for $j = 1, \ldots, M$) be M distinct random d-dimensional masks, where P is the probability of keeping a latent dimension. An element-wise multiplication yields the masked latent codes as $\tilde{\mathbf{z}}_i^j = \mathbf{z}_i^* \otimes \mathbf{m}^j$. Passing these through the decoder for each of the N_{test} test samples yields $N_{\text{test}} \times M$ reconstructed meshes: $\mathcal{R}_i^j = f(\tilde{\mathbf{z}}_i^j)$, obtained through a model f, such as DALS, or DeepSDF.

The Chamfer Distance ($\mathcal{D}_{\text{CD}}$) is used to assess the reconstruction error against the unmasked reconstruction $\mathcal{R}_i$:

$$d_i^j = \mathcal{D}_{\text{CD}}(\mathcal{R}_i, \mathcal{R}_i^j) = \frac{1}{|\mathcal{R}_i|} \sum_{x \in \mathcal{R}_i} \min_{y \in \mathcal{R}_i^j} \|x - y\|_2^2 + \frac{1}{|\mathcal{R}_i^j|} \sum_{y \in \mathcal{R}_i^j} \min_{x \in \mathcal{R}_i} \|y - x\|_2^2. \quad (1)$$

Converting these to normalized similarities $\sigma_i^j \in [0, 1]$ is achieved by applying a per-sample min-max inversion: $\sigma_i^j = 1 - (d_i^j / \max_j d_i^j)$. This results in similar shape pairs having similarity values near 1. While $\mathcal{D}_{\text{CD}}$ is mainly used for defining similarity, we also test against Hausdorff distance (HD95), as it better captures worst-case surface deviations. In RESHAPE, the instance-specific importance, $r_i(k)$, of latent dimension k for shape i is defined as the difference in expected similarity when the dimension is retained versus when it is occluded:

$$r_i(k) = \mathbb{E}_M\left[\sigma_i^j \mid m^j(k) = 1\right] - \mathbb{E}_M\left[\sigma_i^j \mid m^j(k) = 0\right]. \quad (2)$$

Approximating through Monte Carlo integration simplifies this to:

$$r_i(k) \approx \frac{1}{MP(1-P)} \sum_{j=1}^{M} \sigma_i^j (m^j(k) - P). \quad (3)$$

Obtaining a dataset-level importance requires aggregating $r_i(k)$ across test shapes: $r(k) = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} r_i(k)$. Sorting dimensions by decreasing $r(k)$ yields a ranked set of latent factors based on their geometric influence on shapes.

3.2 Data

A series of experiments on synthetic shapes with ground-truth importance verifies the validity of RESHAPE, while experiments on medical shapes demonstrate the method's usefulness for real-world applications (see Fig. 2).

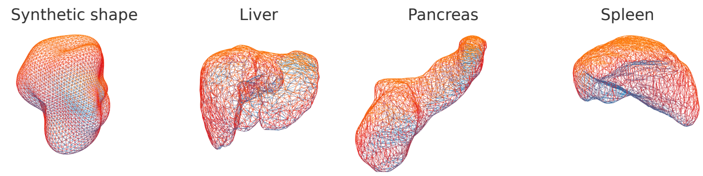

Fig. 2. An example per shape category used in this study, with meshes coloured by surface normals.

Synthetic data: The synthetic dataset is generated using spherical harmonic spectral coefficients with the distance to the surface R being a function of spherical angles θ, ϕ modeled by $R(\theta,\phi) = \sum_{l=0}^{L_{\max}} \sum_{m=-l}^{l} c_{l,m} Y_l^m(\theta,\phi)$, where $Y_l^m(\theta,\phi)$ are the spherical harmonics base of degree l, and order m, $c_{l,m}$ are the spectral coefficients, and $L_{\max}$ is the highest l limit. To obtain smooth shapes, we impose a decaying trend of $c_{l,m}$ with l, and to produce a diverse dataset, we sample $c_{l,m}$ values randomly through $c_{l,m} = \frac{\mathcal{N}(0,1)}{(l+1)^\alpha}$, where $\mathcal{N}(0,1)$ denotes the normal distribution and exponent $\alpha > 0$ controls the energy decay of high-frequency surface oscillations. In this dataset, the set of spectral coefficients $\{c_{l,m}\}$ comprises the latent code, which due to the bandwidth limit $l \leq L_{\max}$ consists of $\sum_{l=0}^{L_{\max}}(2l+1) = (L_{\max}+1)^2$ components.

Medical data: We extract 3D surface meshes from medical segmentation masks from the Medical Decathlon Dataset [1] for the spleen (41 meshes) and liver (126 meshes), both abdominal organs with topologically simple anatomies. To assess performance in more complex topologies, we also include the pancreas (277 meshes), due to its elongated and twisted geometry (see Fig. 1).

3.3 Experiments

Synthetic shape experiments: Our model's ability to rank latent code importance is assessed on synthetic shape data using $P = 0.5$, $M = 1000$, $\alpha = 1$, and $L_{\max} = 6$, which results in $(L_{\max}+1)^2 = 49$ spectral coefficients.RESHAPE should identify low-frequency coefficients as important to validate the pipeline's sensitivity. We also test saliency discrimination between salient and non-salient features by expanding the spectral latent codes $c_{l,m}$ with 49 stochastic components sampled from a Gaussian prior $\mathcal{N}(0,1)$, which do not affect the shape generation. This allows direct verification of ranking correctness, not just consistency.

Medical shape experiments: DeepSDF and DALS models are trained to reconstruct the geometry of abdominal organs, using $P = 0.9$, and $M = 1000/200$ for Chamfer-/HD95-based similarity, respectively. This generates one model per combination of medical dataset and model, which is subsequently processed with the RESHAPE framework to estimate latent-code importance rankings (r_k). The *faithfulness* of our predictions is assessed by measuring the reconstruction error ($\mathcal{D}_{\mathrm{CD}}$), where the top-$k$ dimensions of the latent code $\mathbf{z}_i$ are iteratively replaced by the dataset mean using Area Under the Curve (AUC) metrics. RESHAPE is evaluated against random and variance-based baselines for latent dimension removal. Kendall's τ is used to assess the monotonic correlation of the rankings, while Precision@5 and Normalized Discounted Cumulative Gain (NDGC@5) are used to measure top-k retrieval performance and position-aware ranking quality, along with Rank Accuracy and Mean Absolute Error (MAE) to quantify the exact rank matches and the average displacement of prediction ranks. As no ground-truth latent importance exists, evaluation relies on indirect consistency measures. Deletion AUCs are reported over test shapes with 95% confidence intervals, while RESHAPE's performance is compared to baselines using a one-sided exact sign test across the six organ–model configurations.

4 Results and discussion

A quantitative summary of the synthetic data results is included in Table 1, reflecting perfect monotonic agreement (Spearman's $\rho = 1.0$, Kendall's $\tau = 1.0$) and top-5 accuracy (Precision@5 = 1.0). Individual and ranking metrics underscore RESHAPE's accuracy in the presence of noise ($NDCG$@5 of 0.99 ± 0.01 and 0.97 ± 0.04), further demonstrating the framework's high fidelity and robustness. RESHAPE's aggregated explanation across synthetic shape is shown in Fig. 3, demonstrating the distribution of representation importance across l and m, where a zero-to-low importance is assigned to stochastic latent code components and a monotonic decay trend is observed for the importance of spectral coefficients of increasing order l (see Fig. 3). The medical results of RESHAPE are shown in Fig. 4. These results qualitatively show RESHAPE's sensitivity to changes in the identification of high-influence latent dimensions through the near-zero variance in the deletion curves and their associated variance bands. RESHAPE displays an overall higher deletion AUC compared to the random baseline (one-sided exact signed test, $p = 0.016$), with a median relative gain of 32% (10–43%). RESHAPE comparisons with the variance baseline reveal superior performance in four settings, and similar performance in two ($p = 0.063$). These results underline RESHAPE's competitiveness in identifying important latent dimensions. RESHAPE's importance suggests that latent importance is not solely explained by variance; while this can occur, it is not always the case. DeepSDF displays high sensitivity in its latent space, which is reflected in the curves' noisier appearance and greater overall variance (see Fig. 4). Such noise features are not artifacts of RESHAPE, as it is model-agnostic, but a depiction of the underlying modeling for shape reconstruction. DeepSDF's increased sensitivity is attributed to its SDF-based modeling, as the mesh surface is implicitly extracted from sampled signed distance values.RESHAPE importance is recomputed using HD95, which focuses on worst surface errors. Yet, the derived importance rankings correlate with the Chamfer-derived importances, as

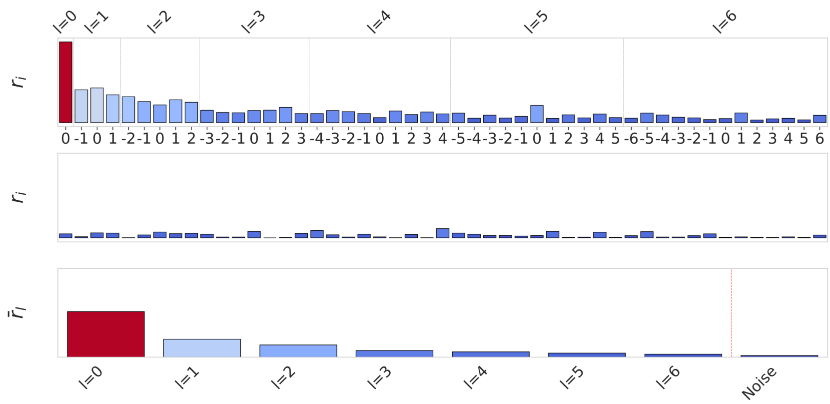

Fig. 3. Importance scores are aggregated across masks and test shapes for the spectral (top) and stochastic (middle) components. The bottom panel displays aggregated importance scores across spectral features of the same l and across stochastic components, highlighting the relative importance identified in the latent dimensions.

Table 1. Evaluation metrics are computed at two distinct evaluation levels: (1) *Aggregated (l-order)* capturing ordering quality across harmonic orders (l=0-6), by aggregating across shapes reflecting global trends. (2) *Individual* reflecting per-shape l-order ranking metrics, obtained by aggregating through 100 synthetic shapes, and reporting the mean ± standard deviation, to assess variance and consistency in identified ranking.

Level	**Localisation metric**	$c_{l,m}$	$c_{l,m} + \mathcal{N}(0,1)$
Aggregated	Spearman (ρ)	1.0	1.0
	Kendall (τ)	1.0	1.0
	Precision@5	1.0	1.0
Individual	Spearman (ρ)	0.90 ± 0.11	0.83 ± 0.21
	Kendall (τ)	0.82 ± 0.15	0.75 ± 0.23
	Precision@5	0.90 ± 0.11	0.88 ± 0.12
	NDCG@5	0.99 ± 0.01	0.97 ± 0.04
	Rank Accuracy	0.79 ± 0.14	0.76 ± 0.15
	Rank MAE	0.51 ± 0.39	0.65 ± 0.51

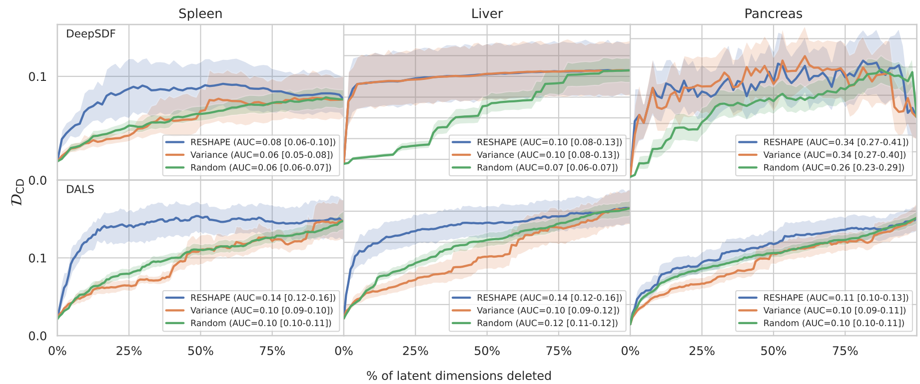

Fig. 4. Deletion experiment results for three medical geometries: the spleen, liver, and pancreas. The top block shows DeepSDF results, while the bottom block shows DALS results. In all panels, RESHAPE is displayed in blue, the random choice of importance ordering in green, and the variance-based estimation of important dimensions in orange.

revealed by Spearman $\rho \geq 0.97$ across all organs, underscoring RESHAPE's robustness.

While the choice of M and P is application-dependent, over-masking (as indicated by low P values) the latent code of DeepSDF affects reconstructed shape geometry and can even lead to shape fragmentation. High P values estimate poorly the second term of Eq. 2, requiring larger M to achieve the same level of uncertainty. We generally advise starting with larger values of P and decreasing P while monitoring convergence.

RESHAPE correctly identifies and ranks features in the presence of stochastic noise, enabling the analysis of latent spaces. This opens up applications in several medical settings, including the study of chronic diseases in relation to induced organ distortions. Other applications include anomaly detection, as RESHAPE importances provide a weighting scheme for latent dimensions. RESHAPE's potential for network pruning is evident in our deletion experiments, which provide dimensionality-reduction tools that respect the reconstructed geometry.

In compact, simple models like DeepSDF and DALS, computational resources are not a bottleneck, underscoring the ease of opening the black box of any shape-generative model. RESHAPE's transparent, model-agnostic attributions come at the cost of multiple forward passes for an accurate Monte Carlo approximation. Still, it is effective in compact latent spaces, while RESHAPE's computational cost is embarrassingly parallel, and it can also be balanced against attribution precision by varying the number of masks to match a desired throughput in the clinic, with processing time on the order of one second for $M = 200$. Its reliance on masking parameters and chosen geometric similarity measure reflects flexible design choices that can be tuned to specific tasks. A promising direction is to expand this framework to learned mask generators, further reducing computational cost [17] and enabling interpretability in larger models.

5 Conclusions

We introduced RESHAPE, a model-agnostic framework for explaining latent representations in 3D shape generation and reconstruction. By stochastically masking latent codes and measuring geometric deviation from the original reconstruction, RESHAPE estimates the instance- and dataset-level importance of latent dimensions without gradients or access to the model's internals. On synthetic shapes, it recovers the expected ordering of low- to high-frequency factors and suppresses noise dimensions. On medical organ meshes with DeepSDF and DALS, it yields faithful importance rankings, consistently outperforming random baselines in all settings ($p = 0.016$) and performing competitively with variance-based baselines, while exposing latent sensitivity differences across models. Its high throughput and performance highlight its clinical utility and potential impact in organ remodeling, thereby enabling post-hoc interpretability across existing pipelines utilizing shapes without requiring modifications or differentiability.

Acknowledgments. This work was supported by the Pioneer Centre for AI, DNRF grant number P1 and by the Research Council of Norway (NFR), through its Centre for Research-based Innovation (Visual Intelligence grant number 309439), IKTPLUSS (grant 303514) and FRIPRO (grants 315029 and 360068). This work was financially supported by the Novo Nordisk A/S NovoSTAR program.

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

References

1. Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., Ronneberger, O., Summers, R.M., et al.: The medical segmentation decathlon. Nature communications **13**(1), 4128 (2022)
2. Caytuiro, N., Sipiran, I.: 3d shape generation: A survey. arXiv preprint arXiv:2506.22678 (2025)
3. Hanocka, R., Hertz, A., Fish, N., Giryes, R., Fleishman, S., Cohen-Or, D.: Meshcnn: a network with an edge. ACM Transactions on Graphics (ToG) **38**(4), 1–12 (2019)
4. He, X., Zou, Z.X., Chen, C.H., Guo, Y.C., Liang, D., Yuan, C., Ouyang, W., Cao, Y.P., Li, Y.: Sparseflex: High-resolution and arbitrary-topology 3d shape modeling. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 14822–14833 (2025)
5. Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., Mohamed, S., Lerchner, A.: beta-vae: Learning basic visual concepts with a constrained variational framework. In: International conference on learning representations (2017)
6. Huang, Y., Zou, S., Liu, X., Xu, K.: Part-aware shape generation with latent 3d diffusion of neural voxel fields. IEEE Transactions on Visualization and Computer Graphics (2025)
7. Hui, K.H., Li, R., Hu, J., Fu, C.W.: Neural template: Topology-aware reconstruction and disentangled generation of 3d meshes. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 18572–18582 (2022)

8. Jensen, P.M., Wickramasinghe, U., Dahl, A., Fua, P., Dahl, V.A.: Deep active latent surfaces for medical geometries. In: Lutchyn, T., Ramírez Rivera, A., Ricaud, B. (eds.) Proceedings of the 6th Northern Lights Deep Learning Conference (NLDL). Proceedings of Machine Learning Research, vol. 265, pp. 120–132. PMLR (07–09 Jan 2025), `https://proceedings.mlr.press/v265/jensen25a.html`
9. Kim, B., Wattenberg, M., Gilmer, J., Cai, C., Wexler, J., Viegas, F., et al.: Interpretability beyond feature attribution: Quantitative testing with concept activation vectors (tcav). In: International conference on machine learning. pp. 2668–2677. PMLR (2018)
10. Kim, J., Hua, B.S., Nguyen, T., Yeung, S.K.: Pointinverter: Point cloud reconstruction and editing via a generative model with shape priors. In: Proceedings of the ieee/cvf winter conference on applications of computer vision. pp. 592–601 (2023)
11. Kim, S., Chae, D.K.: Exmeshcnn: An explainable convolutional neural network architecture for 3d shape analysis. In: Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining. pp. 795–803 (2022)
12. Lan, Y., Zhou, S., Lyuδ, Z., Hong, F., Yangγ, S., Daiδ, B., Pan, X., Loy, C.C.: Gaussiananything: Interactive point cloud latent diffusion for 3d generation. arXiv preprint arXiv:2411.08033 (2025)
13. Li, Z., Wang, D., Chen, K., Lv, Z., Nguyen-Phuoc, T., Lee, M., Huang, J.B., Xiao, L., Zhu, Y., Marshall, C.S., et al.: Lirm: Large inverse rendering model for progressive reconstruction of shape, materials and view-dependent radiance fields. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 505–517 (2025)
14. Li, Z., Wang, Y., Zheng, H., Luo, Y., Wen, B.: Sparc3d: Sparse representation and construction for high-resolution 3d shapes modeling. arXiv preprint arXiv:2505.14521 (2025)
15. Lin, C., Chen, H., Kim, C., Lee, S.I.: Contrastive corpus attribution for explaining representations. arXiv preprint arXiv:2210.00107 (2022)
16. Mo, S.: Scaling diffusion mamba with bidirectional ssms for efficient 3d shape generation. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 19475–19483 (2025)
17. Møller, B.L., Igel, C., Wickstrøm, K.K., Sporring, J., Jenssen, R., Ibragimov, B.: Finding nem-u: Explaining unsupervised representation learning through neural network generated explanation masks. In: Forty-first International Conference on Machine Learning (2024)
18. Park, J.J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S.: Deepsdf: Learning continuous signed distance functions for shape representation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 165–174 (2019)
19. Qi, H., Zhao, C., Salzmann, M., Mathis, A.: Hoisdf: Constraining 3d hand-object pose estimation with global signed distance fields. In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 10392–10402. IEEE (2024)
20. Sørensen, K., Camara, O., De Backer, O., Kofoed, K.F., Paulsen, R.R.: Nudf: neural unsigned distance fields for high resolution 3d medical image segmentation. In: 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2022)
21. Sun, J.M., Wu, T., Gao, L.: Recent advances in implicit representation-based 3d shape generation. Visual Intelligence **2**(1), 9 (2024)

22. Tan, X., Wu, W., Zhang, Z., Fan, C., Peng, Y., Zhang, Z., Xie, Y., Ma, L.: Geocc: Geometrically enhanced 3d occupancy network with implicit-explicit depth fusion and contextual self-supervision. IEEE Transactions on Intelligent Transportation Systems (2025)
23. Vinodkumar, P.K., Karabulut, D., Avots, E., Ozcinar, C., Anbarjafari, G.: Deep learning for 3d reconstruction, augmentation, and registration: A review paper. Entropy **26**(3), 235 (2024)
24. Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M.: Dynamic graph cnn for learning on point clouds. ACM Transactions on Graphics (tog) **38**(5), 1–12 (2019)
25. Wickstrøm, K.K., Trosten, D.J., Løkse, S., Boubekki, A., Mikalsen, K.Ø., Kampffmeyer, M.C., Jenssen, R.: Relax: Representation learning explainability. International Journal of Computer Vision **131**(6), 1584–1610 (2023)
26. Yariv, L., Gu, J., Kasten, Y., Lipman, Y.: Volume rendering of neural implicit surfaces. Advances in neural information processing systems **34**, 4805–4815 (2021)
27. Zhang, B., Yang, T., Li, Y., Zhang, L., Zhao, X.: Compress3d: a compressed latent space for 3d generation from a single image. In: European Conference on Computer Vision. pp. 276–292. Springer (2024)