Supplementary Materials

1 Appendix A: Evaluation of the improvement on image reconstruction task

	StyleGAN		MedGen3D		V2UM2I		ACIS	
	FID	LPIPS	FID	LPIPS	FID	LPIPS	FID	LPIPS
Tlobe	86.98	0.274	41.94	**0.325**	43.63	0.309	**34.94**	0.327
PENet	95.32	0.282	49.32	0.317	45.32	0.322	**38.36**	**0.327**

Table 1. Evaluation of synthesized images. "FID"=Frechet inception distance, "LPIPS"=learned perceptual image patch similarity. A lower value of FID means better image fidelity, and a higher value of LPIPS means higher diversity in synthesized images. The synthesis method yielding the best performance is denoted in **bold** font.

Frechet Inception Distance (FID) [1] and Learned Perceptual Image Patch Similarity (LPIPS) [2] were employed to assess the fidelity and diversity of synthesized images. Lower FID scores indicate higher fidelity, while higher LPIPS scores denote greater diversity. Results were compared across different synthesis methods using two datasets, each comprising 16 training samples. As presented in Table 1, ACIS achieved the best FID values across both datasets, indicating that images synthesized via ACIS most accurately captured the underlying data distributions. The LPIPS scores attained by ACIS were comparable to or modestly below those of MedGen3D, yet significantly surpassed those of StyleGAN, demonstrating ACIS's efficacy in synthesizing images with diverse characteristics.

2 Appendix B: Implementation for Technical Details

Due to the length limitation of the main text, we cannot provide a more detailed description of the concepts and methods we introduced. Here, we try to clarify some necessary concepts.

- **Registration Uncertainty**: In the context of the registration, two primary categories of uncertainty can be identified: transformation uncertainty and appearance uncertainty [3]. Transformation uncertainty pertains to the local ambiguity inherent in the spatial transformation (i.e., the deformation). This type of uncertainty can be utilized for uncertainty-weighted registration, surgical treatment planning, or directly visualized for qualitative assessments [4]. On the other hand, appearance uncertainty relates to the variability in the volumes of the registered organs or annotations. These two uncertainty can be formulated as follows:

$$\hat{\Sigma}^2_{trans} = \frac{1}{T}\sum_{t=1}^{T}\left(\phi_t - {\varphi_t}^{-1}\right)^2 + \left({\phi_t}^{-1} - \varphi_t\right)^2, \tag{1}$$

$$\hat{\Sigma}^2_{app} = \frac{1}{T}\sum_{t=1}^{T}\Big((x \circ \phi_t - \hat{x})^2 + (\hat{x} \circ \varphi_t - x)^2\Big). \tag{2}$$

The notations remain the same as those in the main text. In our studies, we leverage appearance uncertainty to refine anatomical details between natural and generated images as well as annotations.

– **Voronoi Tessellation**: A Voronoi tessellation is a partitioning of a plane into regions based on the distance to a specified set of points $P = \{p_k\}$. Each region, called a Voronoi cell, corresponds to one of the points and consists of all the locations closer to that point than to any other. The Voronoi cell is defined as [5]

$$C(p_k) = \{x_i \in x | \forall q \in P \ d\,(x, q) > d(x, p_k)\}\,. \tag{3}$$

Voronoi cells do not intersect each other in the whole space, and form a tessellation of the whole space, namely the Voronoi Tessellation. There are many applications of Voronoi Tessellation, including geography and urban planning, computer graphics, robotics, and meteorology.

References

1. Soloveitchik M, Diskin T, Morin E, et al. Conditional frechet inception distance[J]. arXiv preprint arXiv:2103.11521, 2021.
2. Zhang R, Isola P, Efros A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 586-595.
3. Luo J, Sedghi A, Popuri K, et al. On the applicability of registration uncertainty[C]//Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22. Springer International Publishing, 2019: 410-419.
4. Chen J, Frey E C, He Y, et al. Transmorph: Transformer for unsupervised medical image registration[J]. Medical image analysis, 2022, 82: 102615.
5. Polianskii V, Marchetti G L, Kravberg A, et al. Voronoi density estimator for high-dimensional data: Computation, compactification and convergence[C]//Uncertainty in Artificial Intelligence. PMLR, 2022: 1644-1653.