

KANResDiff: Learning Local Residual Diffusion via Kolmogorov-Arnold Network for Ambiguous Medical Image Segmentation

Fanding Li¹, Chenglin Wang¹, Xiangyu Li^{1✉}, Xingyu Qiu¹, Xinghua Ma¹, Xiangming Yin¹, Haiyang Li¹, Suyu Dong², Wei Wang³, Kuanquan Wang¹, Gongning Luo¹, and Shuo Li⁴

¹ Faculty of Computing, Harbin Institute of Technology, Harbin, China

² College of Computer and Control Engineering, Northeast Forestry University, Harbin, China

³ Faculty of Computing, Harbin Institute of Technology, Shenzhen, China

⁴ Department of Computer and Data Science and Department of Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio 44106, United States
lixiangyu@hit.edu.cn

Abstract. Ambiguous medical image segmentation aims to provide a series of diverse but plausible segmentation hypotheses. However, existing methods introduce stochasticity in a fixed and pre-defined manner, failing to form a progressive semantic modeling process. To address these challenges, we propose KANResDiff to learn local residual diffusion with Kolmogorov-Arnold Network, thereby assigning distinct roles across stages for ambiguity modeling. Specifically, we propose Independent Time Encoding that offers spline-based time embeddings instead of linear ones from MLPs, which enhances the independence across inference stages and assigns progressive semantic roles to different stages. We propose Residual Schrödinger Bridge that injects deterministic residual prior with learnable weights by constructing local Schrödinger Bridge instead of following manually settings, achieving a flexible deterministic-stochastic interaction and stage-aware ambiguity modeling thanks to local optimal diffusion path. Extensive experimental results on two public datasets demonstrate that KANResDiff achieves SOTA performance on GED and HM-IoU, with maximum improvements of 16.8% and 7.7%, respectively, while maintaining competitive performance on the MDM metric. Source code is available at <https://github.com/PerceptionComputingLab/KANResDiff>.

Keywords: Ambiguous Medical Image Segmentation · Diffusion Model · Kolmogorov-Arnold Network · Schrödinger Bridge.

1 Introduction

Ambiguous medical image segmentation (AMIS), aiming to generate a series of diverse yet anatomically consistent predictions that reflect the underlying annotation distribution [13, 18], relies on the explicit introduction of stochasticity

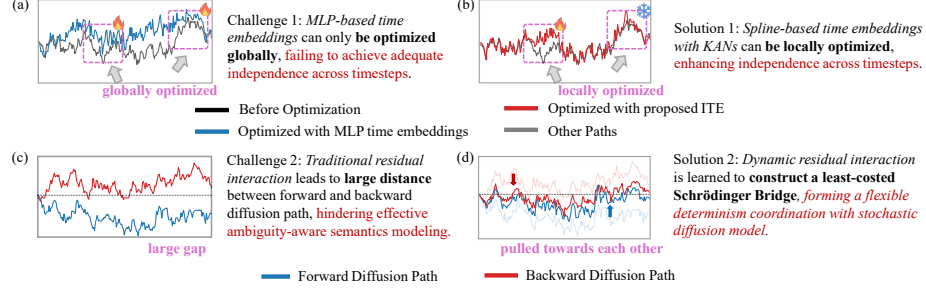


Fig. 1. (a) Traditional MLP-based time embeddings can only be optimized globally. (b) Our proposed KANResDiff enhances independence across timesteps via spline-based time embeddings, which can be locally optimized. (c) Traditional residual diffusion leads to large gap between forward and backward diffusion path, hindering effective semantics modeling. (d) Our proposed KANResDiff constructs local Schrödinger Bridge with dynamic residual components to form a least-costed diffusion path, where forward and backward diffusion paths are close to each other.

during inference to capture the inherent variability in clinical annotations. By exploring multiple anatomically consistent solutions, stochastic inference provides a more reliable and comprehensive basis for clinical decision-making, improving robustness against ambiguous boundaries and inter-observer variability [10,16].

However, existing AMIS methods introduce stochasticity in a fixed and inflexible manner during inference, leading to improper stochastic intensity, where randomness is either over-injected or under-explored. Existing methods can be broadly categorized into three groups: cVAE-based models [10,2], logit distribution-based models [14,21], and diffusion models [13,19]. In the former two categories, ambiguity is introduced only at the final prediction stage, whereas in the latter, stochasticity is incorporated throughout the entire inference trajectory. Consequently, existing methods are constrained by fixed stochastic formulations, failing to establish an AMIS-specific progressive semantic modeling process.

Residual-based diffusion [7,17] has demonstrated potential in alleviating dynamic stochasticity injection thanks to the mechanism that explicitly involves deterministic predictors into diffusion formulation. Residual-based diffusion enables a unified inference mechanism in which the relative roles of determinism and stochasticity can be explicitly controlled, which is widely used in multiple areas such as classification, regression and image super-resolution.

However, directly applying Residual-based diffusion for dynamic and progressive stochasticity injection for AMIS remains challenging due to: **(1) Insufficient independence among inference states hinders the ability to represent heterogeneous ambiguity and undermine effective ambiguity-aware inference in AMIS.** (Fig. 1(a)) Residual-based diffusion typically represents inference stages using time embeddings parameterized by MLPs, which induces strong parameter coupling and limited effective independence across inference steps. While such a design is generally sufficient for holistic genera-

tive tasks such as classification [7], regression [7], and image super-resolution [17], it becomes insufficient for AMIS, where ambiguity must be progressively constructed along inference, requiring greater independence across intermediate stages. **(2) Stage-agnostic residual interaction constrains the flexibility of deterministic–stochastic coordination, limiting effective ambiguity-aware inference in AMIS. (Fig. 1(c))** Residual-based diffusion typically introduces deterministic predictors through a fixed residual formulation, where the relative contribution between deterministic predictors and stochasticity in diffusion is predefined across inference stages. However, inference stages play different roles in ambiguity modeling in AMIS: early stages emphasize structural consistency and later stages express stochasticity to capture diverse plausible segmentations, requiring dynamic and stage-specific residual injection.

To address the aforementioned challenges, we propose KANResDiff to learn local residual diffusion via Kolmogorov-Arnold Networks (KANs) for a flexible deterministic–stochastic coordination with both better independence across inference stages and dynamic, stage-specific residual injection. Specifically, (1) We propose **Independent Time Encoding**, a time reparameterization with localized function representations to reduce cross-stage interference during diffusion inference, enhancing **local independence across inference stages**. Specifically, ITE models the temporal variable with locally supported basis functions, such that optimizations at one inference stage minimally affect others, thereby enhancing step-wise independence and assigning progressive ambiguity modeling across timesteps. (Fig. 1(b)) (2) We propose **Residual Schrödinger Bridge (RSB)**, an AMIS-specific residual interaction strategy by constructing Schrödinger Bridge at each inference stages, assigning **local residual weights**. Specifically, RSB assigns step-specific residual weights to diffusion formulations by constructing Schrödinger Bridge across timesteps, ensuring an optimal stochastic inference by introducing dynamic residual deterministic components, thereby achieving a flexible determinism injection for diffusion in AMIS. (Fig. 1(d))

In summary, our main contributions are as follows:

- We propose KANResDiff to learn local residual diffusion, the first paradigm that enables dynamic and progressive stochastic modeling for ambiguous medical image segmentation, thereby establishing a principled semantic modeling trajectory for ambiguity representation.
- The proposed ITE introduces a novel time-embedding reparameterization strategy based on locally supported basis functions, which significantly enhances the independence across different timesteps and facilitates a progressive ambiguity modeling process.
- The proposed RSB introduces a novel residual guidance mechanism by constructing a local Schrödinger Bridge, which dynamically regulates the contribution of residual deterministic components, thereby enabling flexible and adaptive determinism injection into the diffusion process for AMIS.
- Experimental results on the LIDC dataset and ISIC subset demonstrate that KANResDiff outperforms existing methods across evaluation metrics.

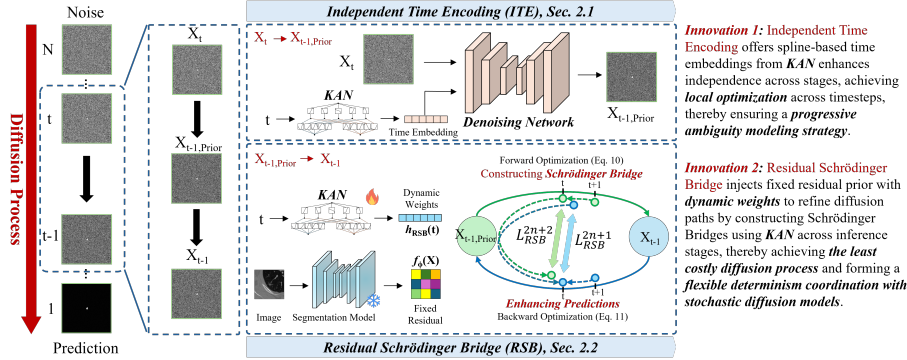


Fig. 2. The proposed KANResDiff achieves superior performance via a flexible deterministic–stochastic interaction in diffusion models by enhancing independence across inference steps (ITE, Sec. 2.1) and assigning dynamic residual weights to construct local Schrödinger Bridges (RSB, Sec. 2.2).

2 Methods

The proposed KANResDiff (Fig. 2) achieves a flexible deterministic–stochastic interaction in diffusion models via Kolmogorov–Arnold Networks from two aspects: Firstly, a time embedding strategy with locally supported basis functions enhances independence across timesteps, assigning a progressive ambiguity modeling role to each timestep (*ITE, Sec. 2.1*). Secondly, local Schrödinger Bridges are constructed via introducing deterministic prior residual with dynamic stage-aware weights to diffusion formulations, achieving a flexible determinism injection to stochasticity of diffusion models, thereby forming a flexible deterministic–stochastic coordination in AMIS (*RSB, Sec. 2.2*).

2.1 Independent Time Encoding Constructs a Progressive Ambiguity Modeling Procedure.

The proposed Independent Time Encoding (ITE) reparameterizes time embeddings with locally supported basis functions, enhancing local optimization independence across timesteps, thereby ensuring a progressive ambiguity modeling strategy along inference stages.

Diffusion models approximate T consecutive denoising networks with a single network taking time embeddings from MLPs as input due to excessive memory. Gradients from MLP-based time embeddings at timestep t_1 is:

$$\frac{\partial L(t_1)}{\partial W} = \frac{\partial L}{\partial g(t_1)} \cdot \frac{\partial g(t_1)}{\partial W} \neq 0. \quad (1)$$

since parameter set W is shared across all timesteps with MLP-based time embeddings, the embedding of any timestep $t_2 \neq t_1$ is optimized as eq. 2. Consequently, all time embeddings are modified with optimization at timestep t_1 ,

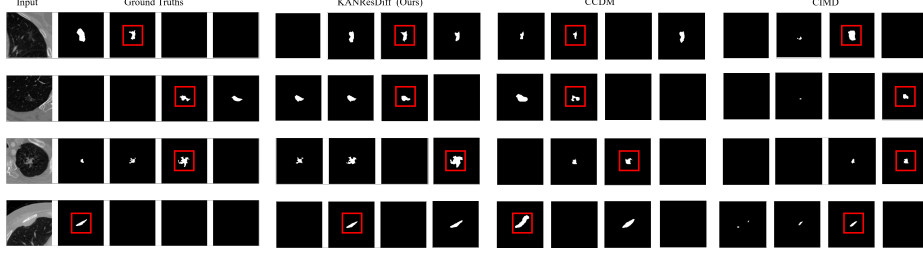


Fig. 3. Comparative qualitative results on LIDC dataset among ground truths, two advanced methods and the proposed KANResDiff demonstrate both better alignment with ground truths and higher per-sample accuracy.

causing inadequate independence across timesteps for AMIS:

$$\frac{\partial L(t_2)}{\partial g(t_1)} = \frac{\partial L(t_2)}{\partial W} \cdot \frac{\partial W}{\partial g(t_1)} \neq 0. \quad (2)$$

In our ITE, time embeddings are formulated with B-splines as eq. 3 for independent optimization during training, enabling a progressive semantic accumulation for plausible ambiguity modeling:

$$g_{spline}(t) = \sum_i c_i \times B_{i,p}(t); B_{i,p}(t) = 0, \text{ when } t \notin [t_i, t_{i+p+1}) \quad (3)$$

where $B_{i,p}(\cdot)$ denotes the i^{th} B-spline basis function with degree p and c_i are learnable coefficients. Therefore, for the optimization at $t_3 \in [t_i, t_{i+p+1})$, the gradient of loss function $L(t_3)$ is calculated as follows:

$$\frac{\partial L(t_3)}{\partial c_i} = \frac{\partial L}{\partial g(t_3)} \cdot B_{i,p}(t_3) \quad (4)$$

Consequently, if t_4 does not fall within the support interval of the B-spline basis functions for t_3 , i.e., $t_4 \notin [t_i, t_{i+p+1})$, then

$$\frac{\partial L(t_4)}{\partial c_i} = \frac{\partial L}{\partial g(t_3)} \cdot B_{i,p}(t_4) = 0 \quad (5)$$

achieving local independent optimization with spline-based parameterizations, thereby ensuring a progressive ambiguity modeling strategy.

For implementation, we utilize cubic B-spline (i.e. $p = 3$) that provides a balanced trade-off between smooth temporal evolution and local parameter updates. Kolmogorov-Arnold Network provides a principled instantiation of spline-based time encoding via its resistance to forgetting and continuous output, making it a suitable and necessary choice for implementation of ITE.

Table 1. Quantitative results on LIDC dataset show the superior performance of KANResDiff. **Bold** represents the best per column.

Methods	GED ₁₆ ↓	GED ₃₂ ↓	GED ₁₀₀ ↓	HM-IoU ₃₂ ↑	MDM ₃₂ ↑
Prob. Unet[10]	0.310±0.010	0.303±0.010	0.252±0.004	0.548±0.000	0.681±0.020
MoSE[6]	0.218±0.003	0.195±0.002	0.189±0.002	0.624±0.004	0.767±0.004
P ² SAM[12]	0.208±0.000	0.206±0.000	0.206±0.000	0.627±0.000	0.939±0.000
CIMD[16]	0.234±0.005	0.218±0.005	0.210±0.005	0.592±0.002	0.915±0.004
AB[4]	0.213±0.001	0.196±0.002	0.193±0.002	0.619±0.001	0.792±0.002
CCDM[19]	0.212±0.001	0.194±0.001	0.183±0.002	0.631±0.002	0.790±0.003
SSB[1]	-	0.245±0.003	0.208±0.001	-	0.942±0.001
ContourMS[18]	0.231±0.003	0.212±0.002	0.203±0.003	0.651±0.005	-
KANResDiff(Ours)	0.187±0.003	0.181±0.002	0.172±0.001	0.701±0.001	0.930±0.002

2.2 Residual Schrödinger Bridge Forms a Flexible Determinism Coordination with Stochastic Diffusion Models.

The proposed Residual Schrödinger Bridge (RSB) assigns dynamic weights instead of manually predefined weights to residual components by constructing local Schrödinger Bridges across inference stages, thereby achieving the least costly diffusion process and forming a flexible determinism coordination with stochastic diffusion models. Deterministic residual prior is significant in ambiguous medical image as a constraint of overall plausibility. Traditional residual diffusion models introduce residual components into inference via fixed formulation as eq. 6 without maintaining the true anatomy of diffusion paths in latent space, failing to preserve the semantic plausibility modeling [13,15].

$$x_t = \sqrt{(\bar{\alpha}_t)} \times x_0 + (1 - \sqrt{(\bar{\alpha}_t)}) \times f_\phi(X) + \sqrt{1 - \bar{\alpha}_t} \times \epsilon \quad (6)$$

To address the limitations, we optimize the diffusion path for AMIS by constructing Schrödinger Bridges across inference stages and regarding residual components as additional refinement to original stochastic diffusion path. Specifically, ambiguity-aware inference procedure of the proposed RSB is:

$$x_t = \sum_{i=1}^n \frac{\sqrt{\bar{\alpha}_t}}{n} \times x_{0i} + h_{\text{RSB}}(t) \times f_\phi(X) + \sqrt{1 - \bar{\alpha}_t} \times \epsilon \quad (7)$$

where $h_{\text{RSB}}(t)$ is the dynamic weight at timestep t , serving both the refinement of original diffusion path to construct a Schrödinger Bridge and a flexible determinism injection maintaining overall plausibility. n denotes the number of plausible predictions, making the proposed RSB an AMIS-effective approach.

To form a Schrödinger Bridge, the inference procedure in eq. 7 should follow SB-formed forward and backward PDEs as eq. 8 and 9, where the residual term $h_{\text{RSB}}(t) \times f_\phi(X)$ serves as the control drift for minimum cost optimization:

$$dX_t = \left(\sqrt{(\bar{\alpha}_t)} X_t + h_{\text{RSB}}(t) f_\phi(X) + \beta(t) \nabla_x \log \Psi(X_t, t) \right) dt + \sqrt{\beta(t)} dW_t \quad (8)$$

Table 2. Quantitative results on ISIC3 subset dataset show the superiority of KAN-ResDiff. **Bold** represents the best per column.

Methods	GED ₁₆ ↓	GED ₃₂ ↓	GED ₁₀₀ ↓	HM-IoU ₃₂ ↑	MDM ₃₂ ↑
Prob. Unet[10]	0.202±0.003	0.187±0.003	0.171±0.002	0.697±0.005	0.927±0.003
SSN[14]	0.197±0.007	0.181±0.004	0.167±0.002	0.700±0.004	0.939±0.001
c-Prob. Unet[20]	0.208±0.004	0.202±0.005	0.179±0.002	0.719±0.004	0.925±0.004
c-SSN[20]	0.200±0.007	0.195±0.007	0.177±0.001	0.725±0.004	0.931±0.003
MoSE[6]	0.304±0.003	0.275±0.004	0.232±0.001	0.673±0.001	-
ContourMS[18]	0.223±0.003	0.202±0.002	0.174±0.001	0.764±0.005	-
KANResDiff(Ours)	0.160±0.003	0.151±0.002	0.139±0.003	0.773±0.002	0.942±0.003

$$dX_t = \left(\sqrt{(\bar{\alpha}_t)} X_t + h_{\text{RSB}}(t) f_{\phi}(X) - \beta(t) \nabla_x \log \Psi(X_t, t) \right) dt + \sqrt{\beta(t)} dW_t. \quad (9)$$

For implementation, we perform refinement based on a pretrained DDPM with fixed noise schedule. The forward and backward procedures of SB are iteratively optimized to be as similar as possible for each timestep t . Specifically, the optimization and training objectives of two procedures are as follows:

Forward Procedure: For the forward procedure, we aim to achieve a Schrödinger Bridge diffusion path. Therefore, $h_{\text{RSB}}(t)$ is optimized to matching the control drift to the SB correction term:

$$\mathcal{L}_{\text{RSB}}^{(2n+1)} = \mathbb{E}_{t, x_t} \left[\|u_{\phi}(x_t, t) - \beta(t) \nabla_x \log \Psi(x_t, t)\|^2 \right]. \quad (10)$$

Backward Procedure: For the backward procedure, we aim to update the predictions with the latest diffusion path. Hence, the magnitude of the control drift is regularized to ensure minimal deviation from the pretrained diffusion:

$$\mathcal{L}_{\text{RSB}}^{(2n+2)} = \mathbb{E}_{t, x_t} \left[\|\nabla_x \log p_t^{\text{SB}}(x_t) - \nabla_x \log p_t^{\text{data}}(x_t)\|^2 \right]. \quad (11)$$

3 Experiments

3.1 Experimental Setup.

Datasets. Two public datasets are selected for evaluation: LIDC-IDRI [8] and ISIC3 subset [5, 20]. LIDC dataset consists of lung CT scans with four corresponding segmentation labels, preprocessed following [10, 11]. ISIC3 subset from [20] provides dermoscopic images of skin lesions with three corresponding labels.

Implementation Details. All training and inference procedures are conducted on a single RTX 4090 GPU. We set $T=1000$ and a batchsize of 8 for all experiments. Models are optimized using an Adam optimizer [9] with a learning rate of 10^{-4} . ITE and RSB are both trained for 600 epochs.

Evaluation Metrics. Generalized Energy Distance (GED) [3], Hungarian Matching IoU (HM-IoU) [6], and Maximum Dice Matching (MDM) [16] are adopted for evaluation. These metrics assess distributional alignment, overall matching accuracy, and maximum individual performance. Subscript n denotes the use of n samples. All results are reported as mean $\pm$ standard deviation over five runs.

Table 3. Ablation study on LIDC dataset demonstrates effectiveness of innovative components. **Bold** represents the best per column.

Methods	ITE	RSB	GED ₁₀₀ ↓	HM-IoU ₃₂ ↑	MDM ₃₂ ↑
Baseline (ResDiff)	×	×	0.187 $\pm$ 0.002	0.667 $\pm$ 0.002	0.903 $\pm$ 0.003
ResDiff with ITE	✓	×	0.181 $\pm$ 0.003	0.683 $\pm$ 0.001	0.908 $\pm$ 0.002
ResDiff with RSB	×	✓	0.174 $\pm$ 0.004	0.688 $\pm$ 0.004	0.916 $\pm$ 0.002
KANResDiff	✓	✓	0.172$\pm$0.001	0.701$\pm$0.001	0.930$\pm$0.002

3.2 Experimental Results.

Quantitative Evaluations. Table 1 and 2 report the quantitative comparisons among KANResDiff and advanced methods on LIDC dataset and ISIC subset. KANResDiff consistently achieves the best performance by reducing GED₁₆, GED₃₂, and GED₁₀₀ by a maximum of 18.8%, 16.6%, and 16.8%, respectively, indicating that the proposed dynamic and progressive stochastic modeling enables more effective ambiguity representation under all sampling strategies. A maximum of 7.7% improvement in HM-IoU₃₂ further demonstrates the superior fidelity of the generated samples from KANResDiff. Meanwhile, KANResDiff maintains MDM performance comparable to supervised sample-level segmentation methods, demonstrating competitive individual accuracy.

Qualitative Evaluations. Fig. 3 illustrates the segmentation visualization comparison across ground truths, the proposed KANResDiff and advanced methods. Predictions from KANResDiff better meet the pattern with ground truths compared with other methods, further validating its superiority.

Ablation Study. Table 3 shows the ablation study on LIDC dataset evaluating four model variants: ResDiff baseline, ResDiff with ITE, ResDiff with RSB and KANResDiff. Introducing ITE and RSB both lead to consistent improvements across all metrics, indicating the effectiveness of innovative components. Performance is further boosted with both components are jointly applied, which further verifies that ITE and RSB are complementary and that their integration effectively improves both segmentation accuracy and diversity.

4 Conclusion

In this work, we proposed KANResDiff for learning local residual diffusion, promoting a flexible stochastic intensity instead of predefined injection for ambiguous medical image segmentation. Specifically, KANResDiff offers a progressive, dedicated ambiguity-aware semantic modeling process by assigning independent and local residual-based diffusion formulation across stages. The proposed Independent Time Encoding offers spline-based time embeddings that can be locally optimized, enhancing independence across stages. The proposed Residual Schrödinger Bridge introduces deterministic residual via learnable weights by constructing local optimal diffusion path with Schrödinger Bridge, forming a

flexible determinism coordination with stochasticity in diffusion models. Experimental results on two public datasets demonstrate our KANResDiff outperforms existing methods. Through multifaceted analysis and comprehensive evaluation, our KANResDiff offers a promising approach for ambiguous medical image segmentation that eliminates excessive or inadequate stochasticity.

Acknowledgments. This work was supported by the National Natural Science Foundation of China under Grant No. 62501195; the Shenzhen Medical Research Fund under Grant No. C2501016; the National Natural Science Foundation of China under Grant No. 82527807; the Science and Technology Innovation Committee of Shenzhen Municipality under Grant No. JCYJ20250604145426036; the Key Research & Development Program of Heilongjiang Province under Grant Nos. 2024ZX12C23 and 2023ZX01A08; and the Natural Science Foundation of Heilongjiang Province under Grant No. LH2024F019.

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

References

1. Baru, L.B., Dadi, K., Chakraborti, T., Bapi, R.S.: Ambiguous medical image segmentation using diffusion schrödinger bridge. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 25–35. Springer (2025)
2. Baumgartner, C.F., Tezcan, K.C., Chaitanya, K., Hötter, A.M., Muehlematter, U.J., Schawkat, K., Becker, A.S., Donati, O., Konukoglu, E.: Phiseg: Capturing uncertainty in medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 119–127. Springer (2019)
3. Bellemare, M.G., Danihelka, I., Dabney, W., Mohamed, S., Lakshminarayanan, B., Hoyer, S., Munos, R.: The cramer distance as a solution to biased wasserstein gradients. arXiv preprint arXiv:1705.10743 (2017)
4. Chen, T., Zhang, R., Hinton, G.: Analog bits: Generating discrete data using diffusion models with self-conditioning. In: The Eleventh International Conference on Learning Representations (2023)
5. Codella, N., Rotemberg, V., Tschandl, P., Celebi, M.E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., et al.: Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). arXiv preprint arXiv:1902.03368 (2019)
6. Gao, Z., Chen, Y., Zhang, C., He, X.: Modeling multimodal aleatoric uncertainty in segmentation with mixture of stochastic experts. In: The Eleventh International Conference on Learning Representations (2023)
7. Han, X., Zheng, H., Zhou, M.: Card: Classification and regression diffusion models. *Advances in Neural Information Processing Systems* **35**, 18100–18115 (2022)
8. Kalpathy-Cramer, J., Zhao, B., Goldgof, D., Gu, Y., Wang, X., Yang, H., Tan, Y., Gillies, R., Napel, S.: A comparison of lung nodule segmentation algorithms: methods and results from a multi-institutional study. *Journal of digital imaging* **29**, 476–487 (2016)

9. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
10. Kohl, S., Romera-Paredes, B., Meyer, C., De Fauw, J., Ledsam, J.R., Maier-Hein, K., Eslami, S., Jimenez Rezende, D., Ronneberger, O.: A probabilistic u-net for segmentation of ambiguous images. *Advances in neural information processing systems* **31** (2018)
11. Kohl, S.A., Romera-Paredes, B., Maier-Hein, K.H., Rezende, D.J., Eslami, S., Kohli, P., Zisserman, A., Ronneberger, O.: A hierarchical probabilistic u-net for modeling multi-scale ambiguities. arXiv preprint arXiv:1905.13077 (2019)
12. Li, C., Lin, Z., Liu, H., Liu, Y., Huang, Y., Ding, X., Tu, X., Yuan, Y., et al.: P²sam: Probabilistically prompted sams are efficient segmentator for ambiguous medical images. In: *ACM Multimedia* (2024)
13. Li, F., Li, X., Su, X., Qiu, X., Dong, S., Wang, W., Wang, K., Luo, G., Li, S.: Ambiguity-aware truncated flow matching for ambiguous medical image segmentation. arXiv preprint arXiv:2511.06857 (2025)
14. Monteiro, M., Le Folgoc, L., Coelho de Castro, D., Pawlowski, N., Marques, B., Kamnitsas, K., Van der Wilk, M., Glocker, B.: Stochastic segmentation networks: Modelling spatially correlated aleatoric uncertainty. *Advances in neural information processing systems* **33**, 12756–12767 (2020)
15. Qiu, X., Yang, M., Ma, X., Li, F., Liang, D., Luo, G., Wang, W., Wang, K., Li, S.: Finding local diffusion schrodinger bridge using kolmogorov-arnold network. In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. pp. 23227–23236 (2025)
16. Rahman, A., Valanarasu, J.M.J., Hacıhaliloglu, I., Patel, V.M.: Ambiguous medical image segmentation using diffusion models. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 11536–11546 (2023)
17. Shang, S., Shan, Z., Liu, G., Wang, L., Wang, X., Zhang, Z., Zhang, J.: Resdiff: Combining cnn and diffusion model for image super-resolution. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 38, pp. 8975–8983 (2024)
18. Wang, J., Zhou, C., Huang, Y.: Contour-aware multi-expert model for ambiguous medical image segmentation. *IEEE Transactions on Medical Imaging* (2025)
19. Zbinden, L., Doorenbos, L., Pissas, T., Huber, A.T., Sznitman, R., Márquez-Neila, P.: Stochastic segmentation with conditional categorical diffusion models. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 1119–1129 (2023)
20. Zepf, K.M., Petersen, E.W., Frelsen, J., Feragen, A.: That label's got style: Handling label style bias for uncertain image segmentation. In: *Eleventh International Conference on Learning Representations* (2023)
21. Zhang, W., Zhang, X., Huang, S., Lu, Y., Wang, K.: Pixelseg: Pixel-by-pixel stochastic semantic segmentation for ambiguous medical images. In: *Proceedings of the 30th ACM International Conference on Multimedia*. pp. 4742–4750 (2022)