
Классические методы как регуляризатор при генерации диффузионных моделей

A Preprint

Войт Руслан Александрович
факультет ВМК
МГУ им. М.В. Ломоносова
Москва

2025

Abstract

В работе исследуется применение классических методов компьютерного зрения для регуляризации процесса генерации в диффузионных моделях. Предлагается использовать структурные дескрипторы, такие как диаграмма Вороного, медиальная ось и контурный анализ, для наложения ограничений на процесс денойзинга. Это позволяет генерировать изображения с улучшенной структурной целостностью, топологической точностью и четкими границами объектов. Эксперименты показывают, что предложенный подход эффективно подавляет артефакты генерации и повышает семантическую согласованность выходных данных, особенно в задачах, требующих точного воспроизведения геометрии сцены.

Keywords Диффузионные модели · Регуляризация · Медиальная ось · Диаграмма Вороного

1 Introduction

Современные диффузионные модели демонстрируют выдающиеся результаты в генерации изображений, однако сталкиваются с проблемами структурной целостности и топологической точности генерируемых данных. Такие артефакты, как нарушение геометрии объектов, разрывы контуров и семантическая несогласованность, ограничивают применение этих моделей в задачах, требующих точного воспроизведения структуры сцены. Необходимость регуляризации генеративного процесса становится особенно актуальной в контексте проблемы запоминания обучающих данных и нарушения нормальности обратного процесса Baptista et al. [2025], Falck et al. [2025].

Существующие подходы к регуляризации диффузионных моделей включают различные методы структурной стилизации. В работах Chen et al. [2024], Pobitzer et al. [2024] предлагается использование контурных представлений для сохранения анатомической структуры при междоменном переводе изображений. Методы скелетной стилизации Ju et al. [2023], Zhu et al. [2023] демонстрируют эффективность структурного контроля при генерации сложных объектов. Перспективным направлением является использование топологических инвариантов, как в работе Gupta et al. [2024], где персистентная гомология используется для контроля чисел Бетти. Классические методы компьютерного зрения, такие как морфологическая диффузия Angulo [2011] и анизотропная регуляризация Pace et al. [2013], также находят применение в современных генеративных моделях.

Несмотря на разнообразие существующих подходов, большинство методов не обеспечивают комплексного учета структурных характеристик различного уровня. Использование отдельных дескрипторов (контуров, скелетов) ограничивает возможность контроля глобальной топологии сцены. Кроме того, многие решения требуют значительной модификации архитектуры модели или сложных процедур обучения, что снижает их практическую применимость. Частотный анализ Falck et al. [2025] показывает, что стандартные

процессы диффузии неравномерно зашумляют различные частотные компоненты, что приводит к потере структурных деталей.

В данной работе предлагается унифицированный подход к регуляризации диффузионных моделей на основе классических структурных дескрипторов компьютерного зрения. Мы используем диаграммы Вороного, медиальные оси и контурный анализ для создания многоуровневых ограничений на процесс денойзинга. Предлагаемый метод позволяет непосредственно встраивать структурные априорные предположения в генеративный процесс без существенного изменения архитектуры модели. Для интеграции классических дескрипторов разработаны специализированные функции потерь и механизмы внимания, обеспечивающие согласованность на различных уровнях абстракции.

Экспериментальные результаты демонстрируют, что предложенный подход значительно улучшает структурную целостность и топологическую точность генерируемых изображений по сравнению с современными методами. Был разработан унифицированный фреймворк для интеграции классических методов компьютерного зрения в современные диффузионные модели, что открывает новые возможности для точного контроля геометрии генерируемых сцен. Предложенное решение эффективно подавляет артефакты генерации и повышает семантическую согласованность, особенно в задачах, требующих точного воспроизведения сложных структур.

2 Headings: first level

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula. See Section 2.

2.1 Headings: second level

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

$$\xi_{ij}(t) = P(x_t = i, x_{t+1} = j | y, v, w; \theta) = \frac{\alpha_i(t) a_{ij}^{w_t} \beta_j(t+1) b_j^{v_{t+1}}(y_{t+1})}{\sum_{i=1}^N \sum_{j=1}^N \alpha_i(t) a_{ij}^{w_t} \beta_j(t+1) b_j^{v_{t+1}}(y_{t+1})} \quad (1)$$

2.1.1 Headings: third level

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Paragraph Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.



Рис. 1: Sample figure caption.

3 Examples of citations, figures, tables, references

3.1 Citations

Citations use **natbib**. The documentation may be found at

<http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf>

Here is an example usage of the two main commands (**citet** and **citep**): Some people thought a thing [??] but other people thought something else [?]. Many people have speculated that if we knew exactly why ? thought this...

3.2 Figures

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetur odio sem sed wisi. See Figure 1. Here is how you add footnotes.¹ Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetur eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

3.3 Tables

See awesome Table 1.

The documentation for **booktabs** ('Publication quality tables in LaTeX') is available from:

<https://www.ctan.org/pkg/booktabs>

¹Sample of the first footnote.

Таблица 1: Sample table title

| Part | | |
|----------|-----------------|------------------------|
| Name | Description | Size (μm) |
| Dendrite | Input terminal | ~ 100 |
| Axon | Output terminal | ~ 10 |
| Soma | Cell body | up to 10^6 |

3.4 Lists

- Lorem ipsum dolor sit amet
- consectetur adipiscing elit.
- Aliquam dignissim blandit est, in dictum tortor gravida eget. In ac rutrum magna.

Список литературы

- Ricardo Baptista, Agnimitra Dasgupta, Nikola B. Kovachki, Assad Oberai, and Andrew M. Stuart. Memorization and regularization in generative diffusion models, 2025. URL <https://arxiv.org/abs/2501.15785>.
- Fabian Falck, Teodora Pandeava, Kiarash Zahirnia, Rachel Lawrence, Richard Turner, Edward Meeds, Javier Zazo, and Sushrut Karmalkar. A fourier space perspective on diffusion models, 2025. URL <https://arxiv.org/abs/2505.11278>.
- Yuwen Chen, Nicholas Konz, Hanxue Gu, Haoyu Dong, Yaqian Chen, Lin Li, Jisoo Lee, and Maciej A. Mazurowski. Contourdiff: Unpaired image-to-image translation with structural consistency for medical imaging, 2024. URL <https://arxiv.org/abs/2403.10786>.
- Markus Pobitzer, Filip Janicki, Mattia Rigotti, and Cristiano Malossi. Outline-guided object inpainting with diffusion models, 2024. URL <https://arxiv.org/abs/2402.16421>.
- Xuan Ju, Ailing Zeng, Chenchen Zhao, Jianan Wang, Lei Zhang, and Qiang Xu. Humansd: A native skeleton-guided diffusion model for human image generation, 2023. URL <https://arxiv.org/abs/2304.04269>.
- Fangqi Zhu, Lin Zhang, Jun Gao, Bing Qin, Ruifeng Xu, and Haiqin Yang. A diffusion model for event skeleton generation, 2023. URL <https://arxiv.org/abs/2305.17458>.
- Saumya Gupta, Dimitris Samaras, and Chao Chen. Topodiffusionnet: A topology-aware diffusion model, 2024.
- Jesús Angulo. Generalised morphological image diffusion. *Nonlinear Analysis: Theory, Methods Applications*, 134, 06 2011. doi:10.1016/j.na.2015.12.015.
- Danielle F Pace, Stephen R Aylward, and Marc Niethammer. A locally adaptive regularization based on anisotropic diffusion for deformable image registration of sliding organs. *IEEE Trans. Med. Imaging*, 32(11):2114–2126, November 2013.