

Diffusion Model的知识整理

论文:

Diffusion Survey:

Diffusion Theory:

Diffusion Application:

Diffusion Compression:

Diffusion Accelerate:

资料:

总览:

论文:

Diffusion Survey:

- Cao H, Tan C, Gao Z, et al. A survey on generative diffusion model[J]. arXiv preprint arXiv:2209.02646, 2022.<https://arxiv.org/abs/2209.02646>
- Kazerouni A, Aghdam E K, Heidari M, et al. Diffusion models for medical image analysis: A comprehensive survey[J]. arXiv preprint arXiv:2211.07804, 2022.<https://arxiv.org/abs/2211.07804>

Diffusion Theory:

- Mittal G, Engel J, Hawthorne C, et al. Symbolic music generation with diffusion models[J]. arXiv preprint arXiv:2103.16091, 2021.<https://arxiv.org/abs/2103.16091>
- Lee S, Kim B, Ye J C. Minimizing Trajectory Curvature of ODE-based Generative Models[J]. arXiv preprint arXiv:2301.12003, 2023.<https://arxiv.org/abs/2301.12003>
- Lee H, Lu J, Tan Y. Convergence of score-based generative modeling for general data distributions[C]//International Conference on Algorithmic Learning Theory. PMLR, 2023: 946–985.<https://proceedings.mlr.press/v201/lee23a.html>
- Chen S, Chewi S, Li J, et al. Sampling is as easy as learning the score: theory for diffusion models with minimal data assumptions[J]. arXiv preprint arXiv:2209.11215, 2022.<https://arxiv.org/abs/2209.11215>
- Lipman Y, Chen R T Q, Ben-Hamu H, et al. Flow matching for generative modeling[J]. arXiv preprint

arXiv:2210.02747, 2022.<https://arxiv.org/abs/2210.02747>

- Liu X, Gong C, Liu Q. Flow straight and fast: Learning to generate and transfer data with rectified flow[J]. arXiv preprint arXiv:2209.03003, 2022.<https://arxiv.org/abs/2209.03003>
- Song J, Meng C, Ermon S. Denoising diffusion implicit models[J]. arXiv preprint arXiv:2010.02502, 2020.<https://arxiv.org/abs/2010.02502>
- Bao F, Li C, Sun J, et al. Estimating the optimal covariance with imperfect mean in diffusion probabilistic models[J]. arXiv preprint arXiv:2206.07309, 2022.<https://arxiv.org/abs/2206.07309>
- Bao F, Li C, Zhu J, et al. Analytic-dpm: an analytic estimate of the optimal reverse variance in diffusion probabilistic models[J]. arXiv preprint arXiv:2201.06503, 2022.<https://arxiv.org/abs/2201.06503>
- Kingma D, Salimans T, Poole B, et al. Variational diffusion models[J]. Advances in neural information processing systems, 2021, 34: 21696–21707.<https://proceedings.neurips.cc/paper/2021/hash/b578f2a52a0229873fefc2a4b06377fa-Abstract.html>
- Sohl-Dickstein J, Weiss E, Maheswaranathan N, et al. Deep unsupervised learning using nonequilibrium thermodynamics[C]//International Conference on Machine Learning. PMLR, 2015: 2256–2265.<http://proceedings.mlr.press/v37/sohl-dickstein15.html>
- Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models[J]. Advances in Neural Information Processing Systems, 2020, 33: 6840–6851.<https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html>
- Song Y, Sohl-Dickstein J, Kingma D P, et al. Score-based generative modeling through stochastic differential equations[J]. arXiv preprint arXiv:2011.13456, 2020.<https://arxiv.org/abs/2011.13456>
- Song Y, Ermon S. Generative modeling by estimating gradients of the data distribution[J]. Advances in neural information processing systems, 2019, 32.<https://proceedings.neurips.cc/paper/2019/hash/3001ef257407d5a371a96dcd947c7d93-Abstract.html>

Diffusion Application:

- Chung H, Ryu D, McCann M T, et al. Solving 3D Inverse Problems using Pre-trained 2D Diffusion Models[J]. arXiv preprint arXiv:2211.10655, 2022.<https://arxiv.org/abs/2211.10655>
- Ajay A, Du Y, Gupta A, et al. Is Conditional Generative Modeling all you need for Decision-Making?[J]. arXiv preprint arXiv:2211.15657, 2022.<https://arxiv.org/abs/2211.15657>
- Ho J, Salimans T. Classifier-free diffusion guidance[J]. arXiv preprint arXiv:2207.12598, 2022.<https://arxiv.org/abs/2207.12598>

- Poole B, Jain A, Barron J T, et al. Dreamfusion: Text-to-3d using 2d diffusion[J]. arXiv preprint arXiv:2209.14988, 2022.<https://arxiv.org/abs/2209.14988>
- Hertz A, Mokady R, Tenenbaum J, et al. Prompt-to-prompt image editing with cross attention control[J]. arXiv preprint arXiv:2208.01626, 2022.<https://arxiv.org/abs/2208.01626>
- Couairon G, Verbeek J, Schwenk H, et al. Diffedit: Diffusion-based semantic image editing with mask guidance[J]. arXiv preprint arXiv:2210.11427, 2022.<https://arxiv.org/abs/2210.11427>
- Dhariwal P, Nichol A. Diffusion models beat gans on image synthesis[J]. Advances in Neural Information Processing Systems, 2021, 34: 8780–8794.<https://proceedings.neurips.cc/paper/2021/hash/49ad23d1ec9fa4bd8d77d02681df5cfa-Abstract.html>
- Liu X, Park D H, Azadi S, et al. More control for free! image synthesis with semantic diffusion guidance[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2023: 289–299.https://openaccess.thecvf.com/content/WACV2023/html/Liu_More_Control_for_Free_Image_Synthesis_With_Semantic_Diffusion_Guidance_WACV_2023_paper.html
- Rasul K, Seward C, Schuster I, et al. Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting[C]//International Conference on Machine Learning. PMLR, 2021: 8857–8868.<http://proceedings.mlr.press/v139/rasul21a.html>
- Nie W, Guo B, Huang Y, et al. Diffusion models for adversarial purification[J]. arXiv preprint arXiv:2205.07460, 2022.<https://arxiv.org/abs/2205.07460>

Diffusion Compression:

- Meng C, Gao R, Kingma D P, et al. On distillation of guided diffusion models[J]. arXiv preprint arXiv:2210.03142, 2022.<https://arxiv.org/abs/2210.03142>
- Yang X, Zhou D, Feng J, et al. Diffusion Probabilistic Model Made Slim[J]. arXiv preprint arXiv:2211.17106, 2022.<https://arxiv.org/abs/2211.17106>
- Theis L, Salimans T, Hoffman M D, et al. Lossy compression with gaussian diffusion[J]. arXiv preprint arXiv:2206.08889, 2022.<https://arxiv.org/abs/2206.08889>
- Luhman E, Luhman T. Knowledge distillation in iterative generative models for improved sampling speed[J]. arXiv preprint arXiv:2101.02388, 2021.<https://arxiv.org/abs/2101.02388>
- Salimans T, Ho J. Progressive distillation for fast sampling of diffusion models[J]. arXiv preprint arXiv:2202.00512, 2022.<https://arxiv.org/abs/2202.00512>


Diffusion Accelerate:

- Deng Y, Kojima N, Rush A M. Markup-to-Image Diffusion Models with Scheduled Sampling[J]. arXiv preprint arXiv:2210.05147, 2022.<https://arxiv.org/abs/2210.05147>
- Luzi L, Siahkoohi A, Mayer P M, et al. Boomerang: Local sampling on image manifolds using diffusion models[J]. arXiv preprint arXiv:2210.12100, 2022.<https://arxiv.org/abs/2210.12100>
- Widadwongsa S, Suwajanakorn S. Accelerating Guided Diffusion Sampling with Splitting Numerical Methods[J]. arXiv preprint arXiv:2301.11558, 2023.<https://arxiv.org/abs/2301.11558>
- Zhang Q, Chen Y. Fast sampling of diffusion models with exponential integrator[J]. arXiv preprint arXiv:2204.13902, 2022.Fast Sampling of Diffusion Models with Exponential Integrator
- Tachibana H, Go M, Inahara M, et al. Quasi-Taylor Samplers for Diffusion Generative Models based on Ideal Derivatives[J].<https://openreview.net/forum?id=7ks5PS09q1>
- Kim B, Ye J C. Denoising MCMC for Accelerating Diffusion-Based Generative Models[J]. arXiv preprint arXiv:2209.14593, 2022.<https://arxiv.org/abs/2209.14593>
- Zheng H, He P, Chen W, et al. Truncated diffusion probabilistic models and diffusion-based adversarial auto-encoders[J]. arXiv preprint arXiv:2202.09671, 2022.<https://arxiv.org/abs/2202.09671>
- Lin X, Jwalapuram P, Joty S. Dynamic Scheduled Sampling with Imitation Loss for Neural Text Generation[J]. arXiv preprint arXiv:2301.13753, 2023.<https://arxiv.org/abs/2301.13753>
- Lu C, Zhou Y, Bao F, et al. Dpm-solver++: Fast solver for guided sampling of diffusion probabilistic models[J]. arXiv preprint arXiv:2211.01095, 2022.<https://arxiv.org/abs/2211.01095>
- Lu C, Zhou Y, Bao F, et al. Dpm-solver: A fast ode solver for diffusion probabilistic model sampling in around 10 steps[J]. arXiv preprint arXiv:2206.00927, 2022.<https://arxiv.org/abs/2206.00927>

资料：

方法	简单描述	链接	我的见解
----	------	----	------

DDPM	DDPM的视频讲解，入门版	 Diffusion Model: 比“GAN”还要牛逼的图像生成模型！公式推导+论文精读，迪哥打你从零详解扩散模型！_哔哩哔哩_bilibili	简单来说，就是不断采样噪声，让图片加上噪声，然后这些噪声是标签，他们可以被UNet还原，UNet的输入时当前时刻t以及t时刻的图像，UNet输出的噪声均值去对齐这些噪声，计算损失进行训练。推理阶段需要一步步还原。
NCSE, DDPM, SDE	医学图像综述	 Diffusion Models for Medical Image Analysis: A Comprehensive Survey	这里提出了Diffusion Model可以分为三类，分别是： <ul style="list-style-type: none"> • denoising diffusion probabilistic models (DDPMs) • noise-conditioned score networks (NCSNs) • stochastic differential equations (SDEs) 后续是一张大表，在医学图像已经应用的例子
DDPM	DDPM的原始论文 (2020)	 Denoising Diffusion Probabilistic Models	从信息熵角度去解释，以及其中的变分界是其理论来源，然而个人理解他之所以work并不在于此，而是在于模型仅仅预测了每个时间段的布朗运动，相较于NCSE是一种两步走策略，更好控制
DDPM	DDPM的推导	 生成扩散模型漫谈 (三) : DDPM = 贝叶斯 + 去噪 - 科学空间 Scientific Spaces	后面的思考很值得人深思，为什么不直接预测 x_0 ? 因为DDPM每一步都会通过 x_t 来预测 x_0 ，但问题是这是一种数值模拟方法。以及样本满足特殊的分布会发生什么? 比如样本群体满足狄拉克分布或自身满足高斯分布。

SDE	Song的论文, SDE提出者	 Score-Based Generative Modeling through Stochastic Differential Equations	提出了SDE方法, 目前还没看完
NCSE	2019年的论文, NCSE提出者	 Generative Modeling by Estimating Gradients of the Data Distribution	早期版本, 不能work, 仅仅使用一次前向传播, 并没有反向扩散进行noise的对齐, 他通过对齐 $\nabla_x \log_{p_\sigma}(\hat{x} x)$ 来进行, 其中 x 是训练样本, \hat{x} 是生成样本, 而 p_σ 是一个固定的概率分布, 为高斯分布。这个通常称为score, 但如果从数学角度理解, 他应该是衡量梯度的大小和方向两个因素。
ODE/SDE	ODE和SDE理论	 无标题	非常长的一本书, 这本书也是陆橙学长推荐的, 内容我写在了目录下的diffusion model的SDE学习笔记中。
SDE	Diffusion Models和SDE的联系以及Score Models和SDE的联系	 Score-based generative models	SDE是包含DDPM和NCSE的, SDE是一种概括。
MCMC	由于在看SDE的论文时遇到, 因此学习MCMC采样。	 MCMC随机采样	MCMC采样是蒙特卡罗方法和马尔可夫链的结合, 它通过一个随机初始化的分布不断经过迭代来得到平稳目标分布, 并在这期间不断进行采样。

DDPM	为什么DDPM的加速都在做推理的加速？	知diffusion model 目前加速感觉很多是inference加速，有没有训练阶段加速的方法？ - 知乎	<p>原因很简单，因为生成能够把过去的过程整合为一个直接到目标step T，因此对加噪过程完全没必要加速。所以要做加速的是inference过程即由噪声到图片的过程。</p>
SDE	Fokker-Planck方程	知Fokker-Planck方程	<p>统计力学（statistical mechanics）中，Fokker-Planck方程被用来描述粒子在随机力的作用下的速度的概率密度函数（PDF），通过推导，其实能够看出求解的是关于x的概率密度函数（PDF），核心就是构建一个随意的函数F(t)并保证和</p> $\left[-\frac{\partial p(t, x)}{\partial t} - \frac{\partial(\mu p(t, x))}{\partial x} + \frac{\partial^2 \left(\frac{1}{2} \sigma^2 p(t, x) \right)}{\partial x^2} \right]$ <p>同号，从而推导出</p> $\frac{\partial p(t, x)}{\partial t} = -\frac{\partial(\mu p(t, x))}{\partial x} + \frac{\partial^2 \left(\frac{1}{2} \sigma^2 p(t, x) \right)}{\partial x^2}。$ <p>以上只是一维推导，多维结果看原文。这个公式被用在推导SDE为ODE的过程中，非常有用。</p>
SDE	伊藤引理	伊藤引理的推导过程 · Saul's Space	
DDPM	路橙学长在知乎的回答	知diffusion model 最近在图像生成领域大红大紫，如何看待它的风头开始超过GAN？ - 知乎	<p>按照陆橙大佬所说的学习路线一步步学习，是一个不错的主意。</p>

ELBO	关于ELBO->DDPM的学习	 无标题	KL散度推出ELBO，然后ELBO推出DDPM的优化目标。
DDIM	知乎关于DDIM的讲解	 扩散模型之DDIM	DDIM是DDPM的加速版本，即对inference进行加速。他的做法是采用ODE，并且跳步组成马尔科夫链，来进行去噪。
多变量分布	用于DDIM的推导	 无标题	在看
MARL	查看高斯分布的贝叶斯定理	 无标题	公式在2.115，是以高斯分布为条件概率下的线性高斯分布的积分公式推导
欧拉积分	用于DDIM的推导	 两类欧拉积分	
DDPM	DDPM从ELBO推导至MSE损失	 一文解释 Diffusion Model (一) DDPM 理论推导	写的非常好的一篇文章，全文非常详细的推导，比原文更详细，不过是离散形式下的推导
SDE	Score-based SDE 理论推导	 一文解释 Diffusion Model (二) Score-based SDE 理论推导	写的还可以的一篇文章，加深理解了SDE中的几个点
Analytic-DDPM	解析 Analytic-DDPM文章，上篇		<p>推导方差的解析解，需要通过协方差公式：</p> $\text{COV}(X, Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{E[XY] - \mu_X \mu_Y}$ <p>求解。最后得出修正方差的上确界为：</p> $\sigma_t^2 + \left(\sqrt{\frac{\bar{\beta}_n}{\alpha_n}} - \sqrt{\bar{\beta}_{n-1} - \sigma_n^2} \right)^2$ <p>而如果采用蒙特卡洛采样，则修正方差可以被解析为：</p> $\sigma_t^2 + \left(\sqrt{\frac{\bar{\beta}_n}{\alpha_n}} - \sqrt{\bar{\beta}_{n-1} - \sigma_n^2} \right)^2 (1 - \bar{\beta}_n \mathbb{E}_{q_n(x_n)} \frac{\ \nabla_{x_n} \log q_n(x_n)\ ^2}{d})$

Extend- Analytic- c-DPM	解析 Analytic-DPM文章， 下篇	 生成扩散模型漫谈 (八)：最优扩散方差估计 (下) - 科学空间 Scientific Spaces	Hadamard product积就是 \odot ,刚刚知道。这篇文章是前一篇文章的推导，是在更加平凡的场景下求最优方差的解析解，然而他需要两阶段训练，因此需要额外思考从而考虑该方案。
高斯分布的最大似然估计	给出了高斯分布的最大似然估计，相当于一个工具	 高斯分布的最大似然估计	得到的是一个直接套用的公式，可以用于推导
条件生成模型	这个是 diffusion model能够部署的核心关键	 生成扩散模型漫谈 (九)：条件控制生成结果 - 科学空间 Scientific Spaces	真的是膜拜苏神，他的理解令我对diffusion model的理解更为深刻，classifier-guidance引导下的生成相对于普通生成究竟该如何，对应到SDE又是什么？真的是很感谢苏神，给出对应的新反向采样的公式： $p(x_{t-1} x_t, y) \approx \mathcal{N}(x_{t-1}, \sigma_t^2 \gamma \nabla_{x_t} \text{sim}(x_t, y), \sigma_t^2 \mathbf{I})$
DPM Solver++	相较于DPM Solver更有效的高分辨率采样	 DPM-Solver++: Fast Solver for Guided Sampling of Diffusion Probabilistic Models	lucheng的相较于dpm solver的近一步改进，核心是更换了ODE采样公式为如下： $\bar{x}_{t_i} = \frac{\sigma_{t_i}}{\sigma_{t_{i-1}}} \bar{x}_{t_{i-1}} + \sigma_{t_i} \underbrace{\sum_{n=0}^{k-1} x_{\theta}^{(n)}(\bar{x}_{\lambda_{t_{i-1}}}, \lambda_{t_{i-1}})}_{\text{estimated}} \underbrace{\int_{\lambda_{t_{i-1}}}^{\lambda_{t_i}} e^{\lambda(\lambda - \lambda_{t_{i-1}})} \frac{1}{n!} d\lambda}_{\text{analytically computed (Appendix A)}} + \underbrace{\mathcal{O}(h_i^{k+1})}_{\text{omitted}}, \quad (9)$
矩阵求导	矩阵求导基本公式	 矩阵求导公式的数学推导 (矩阵求导——基础篇)	可以忘记的时候看一看

DiffEdit	无需训练的通过prompt编辑图像的算法	 使用HuggingFace实现DiffEdit论文的掩码引导语义图像编辑_deephub的博客-CSDN博客	这是一个博客解读，看这篇博客之前还是要看看原文三步是怎么实现的：1) 加噪，然后通过DDIM条件去噪生成mask；2) 原图加噪，到一个latent encode；3) 2得到的加噪去噪，通过mask引导只改变其中的内容，然后完成编辑。这篇博客给出了代码实现以及自己创建了FastDiffEdit算法加速。
DDPM	如何finetune一个stablediffusion	 如何在Stable Diffusion上Fine Tuning出自己风格的模型 - 腾讯云开发者社区-腾讯云	finetune diffusion model需要多少步，如何的配置，unet怎么样，学习一下。
ODE训练	我认可的观点，diffusion model的训练永远从复杂到简单。	 无标题	ODE直线生成，但论文保证了每一时刻速度相同，也就是变化的程度相同，仔细想想这也太像VAE了，认可ODE的训练，但不认可非得每一时刻速度相同，第二个reflow其实挺有意思的，可以理解为自蒸馏的变种。

总览：

