计 算 机 研 究 与 发 展
Journal of Computer Research and Development
https://crad.ict.ac.cn
63(11): 2963−2975, 2026

Wander: 基于预热感知噪声复用的高效自适应扩散模型推理方法

杨迎春[1,2] 陈静波[1] 彭雪梅[3] 纪 东[4] 刘松冉[1]

[1](东北大学计算机科学与工程学院 沈阳 110819)

[2](沈阳理工大学汽车与交通学院 沈阳 110159)

[3](香港科技大学(广州)信息学院 广州 511453)

[4](东北大学工业智能与系统优化国家级前沿科学中心 沈阳 110819)

(2302620772@stu.sylu.edu.cn)

Wander: Warmup-Aware Noise Reuse Based Efficient Adaptive Diffusion Models Inference Method

Yang Yingchun[1,2], Chen Jingbo[1], Peng Xuemei[3], Ji Dong[4], and Liu Songran[1]

[1] (*School of Computer Science and Engineering*, *Northeastern University*, *Shenyang* 110819)

[2] (*School of Automotive and Transportation*, *Shenyang Ligong University*, *Shenyang* 110159)

[3] (*School of Information*, *the Hong Kong University of Science and Technology* (*Guangzhou*), *Guangzhou* 511453)

[4] (*National Frontiers Science Center for Industrial Intelligence and Systems Optimization*, *Northeastern University*, *Shenyang* 110819)

Abstract The diffusion models have made significant achievements in the field of image generation, but they still face some challenges such as high computational cost, long inference latency, and poor image generation quality during image generation. Based on previous research, we propose an efficient adaptive diffusion model inference method based on the warmup-aware and noise reuse methods, which can optimize the performance of the diffusion models according to input conditions and solve the problems of slow inference speed and poor image generation quality. Specifically, our warmup-aware method can dynamically adjust the number of warmup time steps in the initial stage of the diffusion process according to the semantic complexity of the input text prompts to ensure efficient initialization. Subsequently, our noise reuse method dynamically determines the number of time steps for noise reuse in the iterative refinement stage, and reduces unnecessary computational overhead. And we achieve 2.66 times acceleration on 4 NVIDIA V100 GPUs, and compared with DistriFusion, our inference speed increases by 12.8% and the image generation quality improves by 5.03 times. These improvements indicate that our proposed method has significant effects in balancing inference efficiency and image generation quality, and provides a scalable solution for real-time image generation applications.

Key words diffusion models; warmup-aware; noise reuse; image generation quality; inference efficiency

摘 要 扩散模型在图像生成领域取得了显著成就,但在图像生成时仍然面临计算成本高、推理延迟长和生成质量差的挑战。在前人研究的基础上,提出了一种基于预热感知噪声复用的高效自适应扩散模型推

收稿日期: 2025-04-28; **修回日期**: 2026-06-04

基金项目: 国家自然科学基金项目(62302083); 辽宁省联合自然科学基金项目(2023BSBA-123); 中央高校基本科研业务费专项资金(N25ZJL005); 辽宁省自然科学基金项目(2025-BS-0094)

This work was supported by the National Natural Science Foundation of China (62302083), the Liaoning Provincial United Natural Science Foundation (2023BSBA-123), the Fundamental Research Funds for the Central Universities (N25ZJL005), and the Natural Science Foundation of Liaoning Province (2025-BS-0094).

通信作者: 刘松冉(liusongran@cse.neu.edu.cn)

理方法,用于根据输入条件优化扩散模型的性能,解决了推理速度慢和生成质量差的问题。具体而言,预热感知方法能够根据输入文本提示的语义复杂度动态调整扩散过程初始阶段的预热时间步的数量,以确保高效的初始化。随后,噪声复用方法动态决定在迭代细化阶段噪声复用的时间步数量,减少不必要的计算开销。在 4 个 NVIDIA V100 GPUs 上实现了 2.66 倍的加速,与 DistriFusion 相比,推理速度提高了 12.8%,生成质量提升了 5.03 倍。这些进步表明,在平衡推理效率与生成质量方面具有显著的效果,为实时图像生成应用提供了一种可扩展的解决方法。

关键词 扩散模型;预热感知;噪声复用;图像生成质量;推理效率

中图法分类号 TP391

DOI: 10.7544/issn1000-1239.202550290 **CSTR:** 32373.14.issn1000-1239.202550290

人工智能领域的持续发展不断推动着图像生成[1-2]技术的进步。其中,扩散模型在生成质量和多样性上的突出表现成为研究焦点。以 Stable Diffusion[3-6]、DALL.E[7]通过从文本到图像跨模态映射能力为代表的模型,在创意设计、数字内容生产等领域展现出巨大应用潜力。其核心机理是通过多步迭代去噪过程逐步构建高质量图像,值得关注的是,分布式补丁并行技术[8]通过空间分块计算进一步提升了模型推理效率,成为当前优化扩散模型的重要技术路径。

尽管扩散模型在图像生成方面取得了显著进展,但在实际应用中,通常只使用单个设备执行所有时间步的去噪计算,导致生成单张图像耗时长达数十秒,如图 1(a)所示。其根本矛盾源于 2 个层面:1)计算冗余性。每个时间步需要独立计算噪声,而高频细节优化往往集中在关键时间步。2)资源静态分配。现有分布式推理采用固定预热步数与非预热步数,无法根据输入语义复杂度动态调整资源,造成低复杂度任务的计算浪费和高复杂度任务的细节损失,如图 1(b)所示。

虽然分布式补丁并行方法通过分阶段计算缓解了单设备负载压力,但其固定步长策略还存在一定的缺陷:1)初始化鲁棒性不足。不同文本提示存在语义差异,统一预设的预热步数无法动态调整,会引起潜在空间表征偏差。2)时间步优化僵化。多余的计算开销大多集中在对图像质量贡献度较低的时间步,传统方法无法根据具体情况调整计算资源的使用。这些问题严重制约了扩散模型在实时交互场景中的应用。

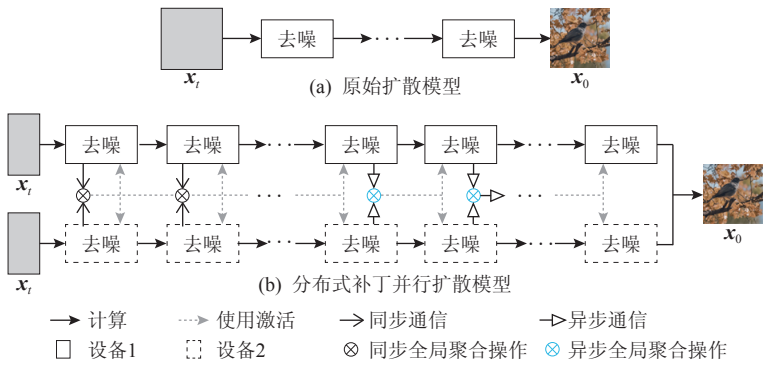


Fig. 1 Running diagram of diffusion model on single and multiple devices

图 1 扩散模型在单设备及多设备运行图

针对上述挑战,本文提出了自适应扩散模型推理方法 Wander,并且建立了动态资源分配机制。1)预热感知方法。我们通过轻量化文本编码器提取输入提示的语义复杂度特征,动态确定分布式补丁并行扩散模型中预热阶段的最佳预热步数,确保不同复杂度的任务获得合适的潜在空间初始化。2)噪声复用方法。我们在非预热阶段选择性跳过低贡献时间步的噪声计算,复用先前时间步的噪声更新潜变量。

相较于传统方法,本文的主要贡献包括以下 3 个方面,生成效果如图 2 所示。

1)提出了一种轻量化的预热感知方法,该方法通过分析输入提示文本的语义特征,动态确定分布式补丁并行扩散模型初始阶段的预热步数,确保模型在初始去噪阶段构建更具鲁棒性的潜在空间表征,

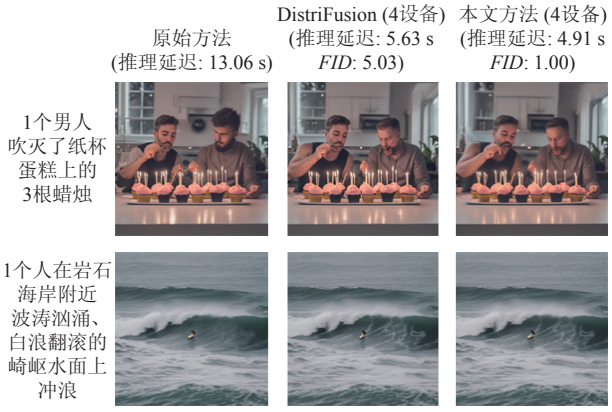


Fig. 2 Comparison of the generation effect between our method and traditional methods

图 2 本文方法与传统方法生成效果对比

提升了模型初始化效果。

2)设计了一种噪声复用方法，该方法通过预测噪声复用步数，在后续去噪阶段可以动态跳过对最终图像质量影响较小的时间步的噪声计算，复用先前时间步的噪声来更新潜变量，显著减少了计算开销并提高了推理效率。

3)在 COCO Captions 数据集上验证了我们方法的有效性，结果表明，我们的方法在推理速度上提高了 12.8%，在图像质量上提升了 80.1%，并对不同语义复杂度任务具有良好的泛化能力。

1 相关工作

扩散模型的基本框架通常包括 2 个过程：正向扩散过程和反向扩散过程。正向扩散过程是逐步将数据添加噪声至高斯分布；反向扩散过程是利用神经网络逐步去噪，恢复目标数据。这种方式虽然为生成任务提供了灵活性和高质量保证，但也因其迭代特性带来了较高的计算成本。

在反向扩散过程中，模型通常通过逐步迭代不断移除噪声，每一步需要对输入特征进行复杂计算。具体来说，对于每个时间步 t，扩散模型通过式(1)计算去噪后的状态。

$$\boldsymbol{x}_{t-1} = Update(\boldsymbol{x}_t, t, \boldsymbol{\epsilon}_t), \boldsymbol{\epsilon}_t = \epsilon_\theta(\boldsymbol{x}_t, t, c)\,, \tag{1}$$

其中 ϵ_θ 表示神经网络预测噪声的过程，$Update$ 为特定采样器的更新函数(DDIM 采样器)，$\boldsymbol{x}_t$ 为当前时间步的潜在特征，t 为当前时间步，c 为提示文本。提升扩散模型的生成质量和推理效率，主要依赖于对反向扩散过程的精细优化。传统方法(DDPM 的线性调度)在去除噪声的过程中缺乏对数据特征的适应性，导致了高频信息过早退化。改进的调度策略(余弦调度、IDDPM 的可学习方差)虽然缓解了这一问题，但往往需要增加额外的计算开销[9]，导致推理效率降低。分类器引导和无分类器引导通过引入条件信息(类别标签或文本描述)提升了模型生成可控性和质量[10-11]。然而，这些方法往往依赖于大规模预训练的模型，这种模型通常需要复杂的联合训练策略，显著增加了内存占用和计算开销。潜在扩散模型(latent diffusion model, LDM)[12] 通过将扩散过程迁移到低维隐空间，大幅降低了计算复杂度。但隐空间的压缩不可避免地引入了信息损失，特别是在细节丰富的图像生成任务中，模型的生成质量对隐空间编码器的敏感性显著增加。

1.1 扩散模型的基本框架与演化

扩散模型的思想源于物理学中的扩散过程，通过模拟图像在噪声干扰下的扩散过程并逆向去噪生成清晰的图像。2015 年 Sohl-Dickstein 等人[13] 在 Deep Unsupervised Learning 中提出了这种生成框架，通过对图像添加噪声并逐步在每个时间步去噪来产生图像。2020 年 Ho 等人[14] 提出的去噪扩散概率模型(denoising diffusion probabilistic model, DDPM)为扩散模型的发展奠定了基础，DDPM 通过在多个时间步中对图像进行去噪操作，生成了高质量的样本，但也出现了计算成本高和推理速度慢的问题。于是，许多研究者在 DDPM 的基础上进行优化来提高扩散模型的生成质量和效率。2020 年，Song 等人[15] 提出的 Score Matching 通过优化模型的目标函数，使得扩散模型在计算复杂度上得到一定程度的改进。2021 年，Nichol 等人[9] 提出的 Improved DDPM 通过改进噪声采样过程，提高了图像的生成质量，特别是在高分辨率图像生成任务中，取得了较好的效果。

1.2 生成质量的改进

生成质量的提升常常依赖于感知损失[16-17] 和对抗训练[18-19] 等技术。2014 年，Goodfellow 等人[20] 提出的 GANs 通过引入对抗损失，增强了模型的生成能力，使图像生成得更加真实。2017 年，Ledig 等人[21] 提出的 SRGAN 使用感知损失优化了高分辨率图像的生成质量。2022 年，Karakas 等人[22] 提出的 StyleGAN2 在图像生成的过程中引入了高层次的感知损失，使图像生成质量得到显著提升，尽管该模型不是扩散模型，但为扩散模型的研究打开了新思路。2017 年，Karras 等人[23] 提出的 Progressive GANs 采用逐渐构建图像的方式，从低分辨率到高分辨率逐渐生成图像，有效提高了图像生成质量并降低了模型训练时的梯

度问题。一些扩散模型的研究者借鉴这种思想，通过在不同尺度上进行生成，逐渐提高了图像的生成质量，Ho 等人[14]提出的 DDPM 就采用了这种思想。

1.3 计算效率的优化

扩散模型虽然在生成质量方面表现出了显著的性能，但这些方法通常面临着推理速度慢、计算资源消耗大的问题。2020 年，Song 等人[24]提出的 DDIM（denoising diffusion implicit model）通过修改扩散模型的采样过程和采用非标准的去噪过程在减少扩散步数与计算量的同时，保持了生成图像的清晰度和细节，该方法为减少扩散时间步提供了一种新思路。

1.4 条件生成

扩散模型在生成过程中通过引入条件信息，可以实现更加定向和多样化的生成任务。Imagen[25]和 DALL.E2[26]等模型通过引入文本描述条件，显著提升了扩散模型的图像生成质量和多样性。在扩散过程中，条件生成的策略通过引入条件向量或条件噪声来控制生成过程。通过这些条件信息，扩散模型生成的图像可以更准确地符合用户需求或应用场景的要求。

2 动　　机

2.1 噪声的相似性

扩散模型的图像生成能力虽然获得了广泛认可，但逐时间步执行的确定性推理范式引入了不可避免的计算开销。实验中我们发现，相邻去噪时间步产生的噪声具有相同或者相似性的特点，$Sim(t)=\frac{1}{4\times128^2}\sum_{c=1}^{4}\sum_{i=1}^{128}\sum_{j=1}^{128}\left|N_t^{(c,i,j)}-N_{t+1}^{(c,i,j)}\right|$，其中 $N\in\mathbb{R}$。如图 3 所示，时间步 6 与时间步 7、时间步 40 与时间步 41 的噪声热力图几乎一致。这一现象表明，模型在连续多个时间步中仅进行了微弱的去噪调整，相邻时间步对生成过程的实际贡献趋于同质化。具体而言，当模型连续执行若干相邻时间步时，潜在特征修正的方向和幅度呈现高度一致性，大量计算资源被重复消耗于噪声的细微调整，而非推动生成过程的有效演进。



(a) 最终生成图像　(b) 噪声差异$|\varepsilon_7-\varepsilon_6|$　(c) 噪声差异$|\varepsilon_{41}-\varepsilon_{40}|$

Fig. 3　The thermal map of noise difference

图 3　噪声差异热力图

2.2 固定预热方法的适应性瓶颈

为了研究固定预热方法的适应性瓶颈，我们用 *FID*（Fréchet inception distance）（详见 4.1.4 节评估指标）来衡量图像质量，并且通过实验发现，如图 4 和图 5 所示，当噪声复用步数过长时，未被补偿的预测误差会随着扩散过程逐渐放大，导致图像质量变差。而随着预热阶段预热步数 T_{warmup} 的增加，模型在初始阶段构建了更具有鲁棒性的潜在空间特征，使这种误差减小，图像质量变好。进一步显示，在不进行噪声复用的情况下，对于整体情况而言，图像质量随着预热步数的增加而升高，其中，预热步数从 4 增加到 30 时，质量改善尤为显著，在 30 步之后，质量改善趋于平缓。这一结果表明，预热步数的增益具有一定限度，过多预热对图像质量的提升有限。

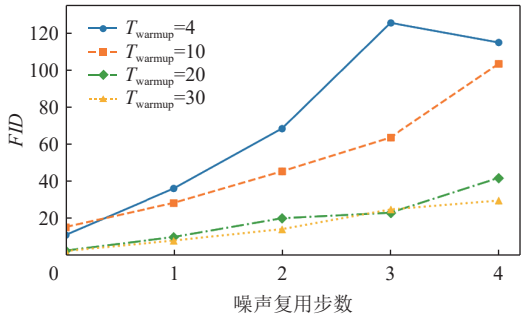


Fig. 4　The mutual influence between the number of warmup steps and the number of noise reuse steps

图 4　预热步数与噪声复用步数的相互影响

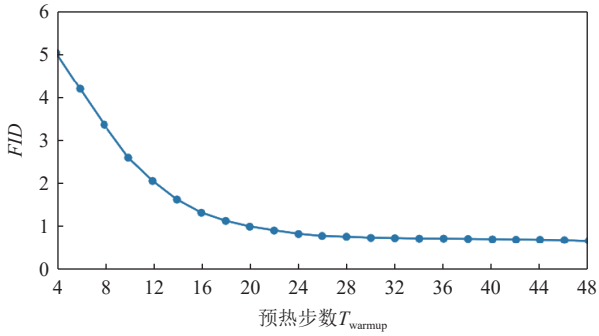
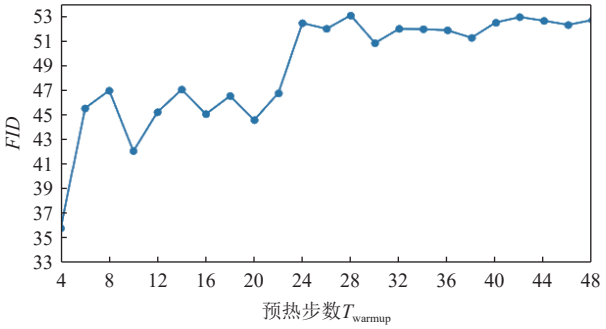


Fig. 5　Image quality trend diagram with increase of the number of warmup steps

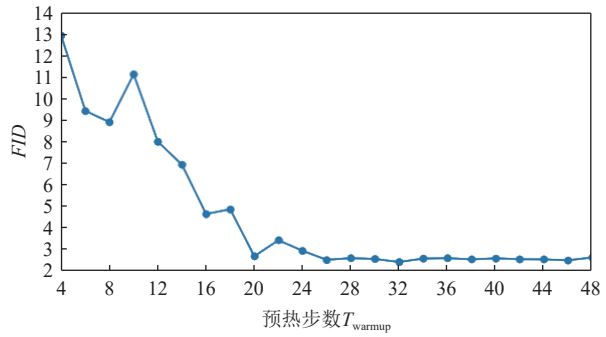
图 5　随着预热步数增加的图像质量趋势图

对于不同复杂度的提示文本，最佳预热步数存在差异。如图 6 所示，对于提示文本语义信息较简单的，最佳预热步数比较少，增加预热步数反而使图像的质量降低，对于提示文本语义信息较复杂的，最佳预热步数比较多，多于或者低于这个最佳预热步

数都会使图像的质量降低。实验表明，预热步数对图像生成质量的影响存在明显的输入相关性，对于简单的提示文本，只需较少的预热步数，就可以使模型在初始阶段达到较好的稳定性，而对于较复杂的提示文本，则需要更多的预热步数以捕捉丰富的语义信息。因此，固定数量的预热步数难以适应多样化的输入需求，既限制了模型在复杂语义场景中的表达能力，又带来了额外的计算开销，降低了整体推理效率，是制约当前扩散模型性能释放的重要瓶颈。



(a) 简单提示文本



(b) 复杂提示文本

Fig. 6 Image quality varies with the number of warmup steps under different semantic complexities

图 6 不同语义复杂度下图像质量随预热步数的变化

2.3 本文的研究动机

基于上述分析，我们认为当前扩散模型在处理动态输入和时间步依赖性方面存在显著局限性。这些局限性主要体现在 2 个方面：一方面是模型在初始阶段的性能瓶颈，即由于缺乏有效的预热策略，导致模型在初始阶段的初始化能力不强；另一方面是现有的方法在时间序列任务中存在明显不足，它们未能充分识别和优化非关键时间步，从而增加了模型计算开销并降低了整体性能。这些问题的根源在于传统模型采用静态的输入处理方式，难以适应动态变化的任务需求，也缺乏对时间步重要性的自适应判断。所以，要让扩散模型提高生成质量和推理效率，关键在于构建一种根据输入条件动态调整推理节奏的方法。

因此，本文设计了 Wander 方法，该方法主要包括预热感知和噪声复用 2 个核心组成部分，使用预训练的语义编码器来分析文本提示的语义复杂度，进而预测最佳预热步数和噪声复用步数，实现推理过程的自适应优化。具体结构将在第 3 节中详细介绍。

3 Wander 的设计

3.1 架构总览与核心理念

图 7 展示了我们方法的整体架构，包含语义驱动的参数预测和扩散控制 2 个核心模块。参数预测模块根据输入的文本提示动态预测最优预热步数 T_{warmup} 和噪声复用步数 k，提高扩散模型的图像生成质量和推理效率。扩散控制模块根据预测参数调整模型的扩散过程，进而实现扩散模型的动态预热和噪声复用。在推理阶段的伪代码如算法 1 所示。其基本流程为：输入文本 c→预测 T_{warmup} 和 k→执行 T_{warmup} 步预热→按 k 间隔复用噪声→输出图像 $\boldsymbol{x}_0$。

算法 1. Wander 推理。

输入：文本提示 c，总时间步 T，扩散模型 $\epsilon_\theta(\cdot)$，DDIM 采样器 $\varphi(\cdot)$，参数预测器 $\omega(\cdot)$；

输出：$\boldsymbol{x}_0$。

① 初始化 $\boldsymbol{x}_T \sim \mathcal{N}(0, I)$;
② $T_{\text{warmup}}, k \leftarrow \omega(c)$;
③ for $t = 0$ to $T-1$ do
④ if $t < T_{\text{warmup}}$ then
⑤ $\boldsymbol{\varepsilon} \leftarrow \epsilon_\theta(\boldsymbol{x}_T, T, c)$;
⑥ else
⑦ if $(t - T_{\text{warmup}})\%(k+1) == 0$ then
⑧ $\boldsymbol{\varepsilon} \leftarrow \epsilon_\theta(\boldsymbol{x}_T, T, c)$;
⑨ $\boldsymbol{\varepsilon}' \leftarrow \boldsymbol{\varepsilon}$;
⑩ else
⑪ $\boldsymbol{\varepsilon} \leftarrow \boldsymbol{\varepsilon}'$;
⑫ end if
⑬ end if
⑭ $\boldsymbol{x}_{T-1} \leftarrow \varphi(\boldsymbol{x}_T, T, \boldsymbol{\varepsilon})$;
⑮ $T \leftarrow T - 1$;
⑯ end for

3.2 语义驱动的预测模块

该模块基于输入文本提示 c 预测 T_{warmup} 和 k。采用 BERT-MLP 架构，其中 BERT 擅长捕捉文本提示的语义细节，MLP 可以高效地将这些特征映射为推理参数，确保对不同输入的适应性。我们使用预训练的 BERT-BASE-UNCASED（12 层 Transformer，768

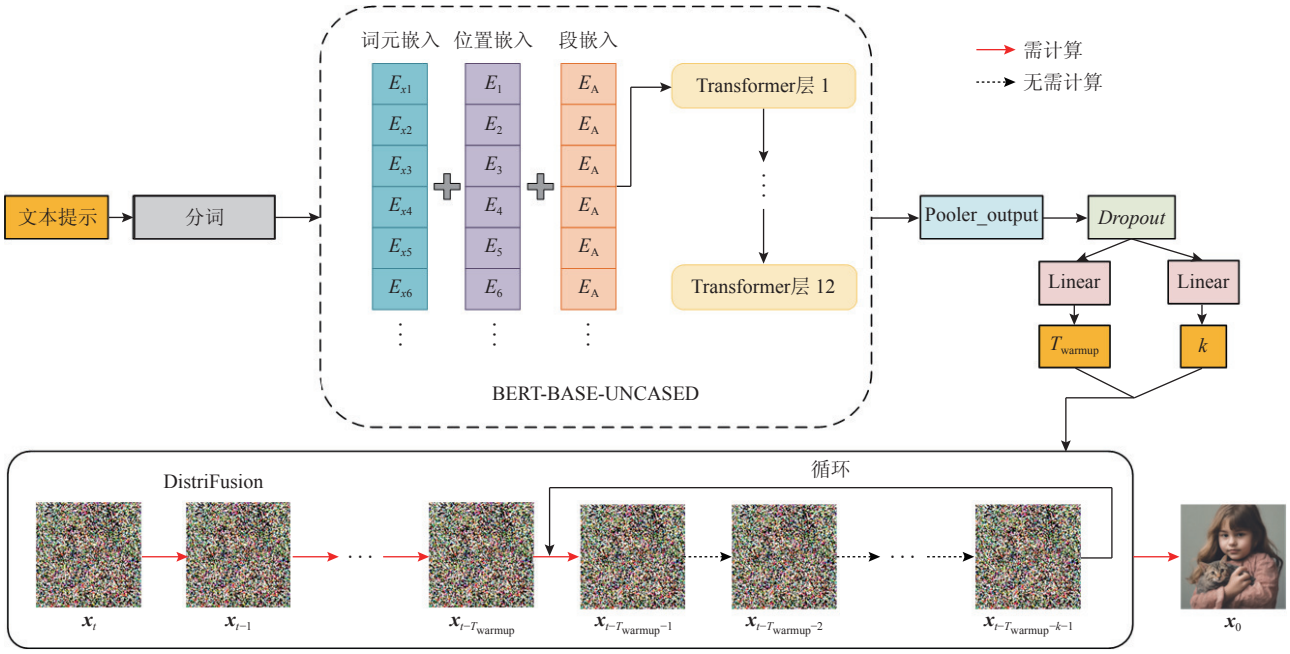


Fig. 7 Overview of Wander

图 7 Wander 概述

维隐藏状态)[27]模型进行语义特征提取,输出池化的[CLS]特征表示:

$$\boldsymbol{h}_{\mathrm{CLS}} = BERT(input_ids, attention_mask), \tag{2}$$

其中$\boldsymbol{h}_{\mathrm{CLS}} \in \mathbb{R}^{768}$为输出[CLS]token的池化表示,*input_ids*为输入文本的token ID序列,*attention_mask*为二进制掩码。然后为了增强模型的泛化能力,对$\boldsymbol{h}_{\mathrm{CLS}}$施加*Dropout*:

$$\boldsymbol{h}'_{\mathrm{CLS}} = Dropout(\boldsymbol{h}_{\mathrm{CLS}}, p = 0.1), \tag{3}$$

其中p为丢弃概率。通过全连接层同时预测2个参数:

$$T_{\mathrm{warmup}} = \boldsymbol{W}_1 \cdot \boldsymbol{h}'_{\mathrm{CLS}} + b_1, \tag{4}$$

$$k = \boldsymbol{W}_2 \cdot \boldsymbol{h}'_{\mathrm{CLS}} + b_2, \tag{5}$$

其中$\boldsymbol{W}_1, \boldsymbol{W}_2 \in \mathbb{R}^{1\times 768}$, $b_1, b_2 \in \mathbb{R}$为线性层参数。

3.3 最优标签生成策略

从COCO数据集[28]的5 000个文本提示中生成监督信号T^*_{warmup}和k^*。首先,为了确定每个文本提示的最佳预热步数T^*_{warmup},我们将预热步数从4逐渐调整到50,步长为2,共测试了24组参数(即4, 6, 8, …, 50),对于每组预热步数,我们生成5 000张图像,总计生成120 000张图像。然后计算每张图像的质量指标(*FID*[29], *LPIPS*[30], *PSNR*[31]),确定了每个图像对应文本提示的最佳预热步数。随后,为了确定每个文本提示的最佳噪声复用步数k^*,我们将噪声复用步数从1调整至3,步长为1,共测试3组参数(即1, 2, 3),每组5 000张图像,总共15 000张图像。通过评估每张图像的质量,选出质量最佳的噪声复用步数作为该文本提示的最佳噪声复用步数。

3.4 训练策略与损失设计

模型采用监督学习训练,同时优化T_{warmup}和k的预测。损失函数为2个任务的加权均方误差。

$$\mathcal{L} = \alpha \times \mathcal{L}_{\mathrm{warmup}} + (1-\alpha) \times \mathcal{L}_{\mathrm{skip}}, \tag{6}$$

$$\mathcal{L}_{\mathrm{warmup}} = E_c\left(\left\|T_{\mathrm{warmup}} - T^*_{\mathrm{warmup}}\right\|_2^2\right), \tag{7}$$

$$\mathcal{L}_{\mathrm{skip}} = E_c\left(\left\|k - k^*\right\|_2^2\right), \tag{8}$$

其中α为加权系数。表1为超参数配置表。

Table 1 Hyperparameter Configuration

表 1 超参数配置

超参数	取值
优化器	Adam
学习率	0.001
批次大小	32
学习率衰减	30轮无改善时衰减0.1倍
Dropout	0.1

3.5 扩散控制模块

推理阶段利用T_{warmup}和k控制扩散过程,整个过程分为2个阶段。在预热阶段,执行T_{warmup}步标准去噪过程,稳定潜在空间表示。在非预热阶段,由于相邻去噪时间步噪声具有相似性的特点,我们复用前次噪声并不会对最终图像质量造成很大影响,只需每$k+1$步更新一次噪声,直至生成$\boldsymbol{x}_0$。假设我们需要

跳过时间步$t-1$到$t-k$的噪声计算，那么我们需要复用时间步t的噪声进行更新。具体地说，对于时间步$t-1$到$t-k$，更新公式可简化为

$$\boldsymbol{x}_{t-i-1}=\boldsymbol{x}_{t-1}-\sum_{j=1}^{i}\epsilon_\theta(\boldsymbol{x}_t,t,c),i=1,2,\cdots,k, \tag{9}$$

即跳过的每个时间步$t-i$潜变量在更新时，统一复用时间步t预测的噪声$\epsilon_\theta(\boldsymbol{x}_t,t,c)$。

4 实　　验

在本节中，我们对提出的预热感知方法和噪声复用方法进行了全面的实验评估。首先介绍实验设置，包括数据集、模型和基线方法；随后，我们报告实验结果，并从生成质量与推理效率2个角度进行分析；最后进行消融实验验证各组件的有效性。

4.1 设　置

4.1.1 数据集

我们在COCO Captions 2014[32]数据集上选取了5 000个文本提示进行实验，这些文本提示覆盖简单、中等和复杂的语义场景。为了确定预热感知方法和噪声复用方法的最佳参数，我们测试了预热步数4~50的24个取值(步长为2)，每组生成5 000张图像，总共生成120 000张图像。测试了噪声复用步数1~3的3个取值，每组生成5 000张图像，总计15 000张图像。所有图像生成都是基于SDXL(Stable Diffusion XL)[33]模型，使用50步DDIM采样器在多GPU环境下完成，总计生成135 000张图像用于参数优化。质量指标(*FID*, *LPIPS*, *PSNR*)用于选出每个文本提示的最佳参数训练映射函数。

4.1.2 模　型

实验基于扩散模型SDXL，我们的方法与Original方法分别在4个与1个NVIDIA V100 GPUs上实现(除非另有说明)。我们使用50步DDIM采样器和4步预热步数作为默认采样方法。

4.1.3 基　线

为了全面评估我们方法的有效性，我们将其与以下3种基线方法进行对比。

1)标准扩散模型。该模型基于DistriFusion在推理过程中使用4个GPUs，采用固定数量的预热步数且在后续时间步中无噪声复用。

2)无预热感知方法。该模型在推理过程中采用固定数量的预热步数，但在后续过程中会进行噪声复用。

3)无噪声复用方法。该模型在推理过程中实现了预热感知方法，根据输入文本语义信息的复杂度动态调整预热步数，但在后续时间步中无噪声复用。

4.1.4 评估指标

为了全面评估生成质量和推理效率，我们使用以下指标评估135 000张图像。

1)使用*FID*衡量生成图像之间的分布差异，值越小表示图像质量越好。

$$FID=\|\boldsymbol{\mu}_{\mathrm{r}}-\boldsymbol{\mu}_{\mathrm{g}}\|_2^2+\mathrm{tr}\left(\boldsymbol{\Sigma}_{\mathrm{r}}+\boldsymbol{\Sigma}_{\mathrm{g}}-2\sqrt{\boldsymbol{\Sigma}_{\mathrm{r}}\boldsymbol{\Sigma}_{\mathrm{g}}}\right)。 \tag{10}$$

2)使用*LPIPS*(learned perceptual image patch similarity)评估生成图像之间的感知相似性，值越小表示生成的图像视觉质量越高。

$$LPIPS(\boldsymbol{x},\boldsymbol{x}')=\sum_l\omega_l\|\phi_l(\boldsymbol{x})-\phi_l(\boldsymbol{x}')\|_2^2。 \tag{11}$$

3)使用*PSNR*(peak signal-to-noise ratio)衡量生成图像中信号强度(*MAX*)与噪声强度(*MSE*)的比值，值越高表示生成的图像质量越好。

$$PSNR=10\lg\left(\frac{MAX^2}{MSE}\right)。 \tag{12}$$

4)使用推理延迟记录生成每张图像所需的时间。

5)使用计算成本乘加操作数(单位为MAC)[34]度量推理效率。

4.2 实验结果与分析

我们基于对生成图像的实验结果确定了每个文本提示的最佳参数。在此基础上，Wander方法的性能通过额外生成的测试图像进行评估。

4.2.1 敏感性分析

我们对预热感知方法和噪声复用方法的生成质量和推理延迟进行了权衡。图8展示了T_{warmup}与k对生成质量和推理延迟的敏感性分析。

如图8所示，我们在整个测试集上分析了T_{warmup}和k对生成质量和推理延迟的影响，也说明了预热感知方法和噪声复用方法对生成质量和推理延迟的权衡关系。

实验结果表明，随着k的增加，推理延迟和生成质量呈下降趋势，但这种趋势随着k的增加会趋于平缓。随着T_{warmup}的增加，图像生成质量显著提高，当T_{warmup}较大时图像质量提升趋于饱和，而推理延迟却受到限制。进一步分析可知，当T_{warmup}较小时，k对生成质量和推理延迟的影响较明显；当T_{warmup}较大时，k对生成质量和推理延迟的影响较微弱。基于这种现象，我们对T_{warmup}和k进行了调优，以实现生成质量和

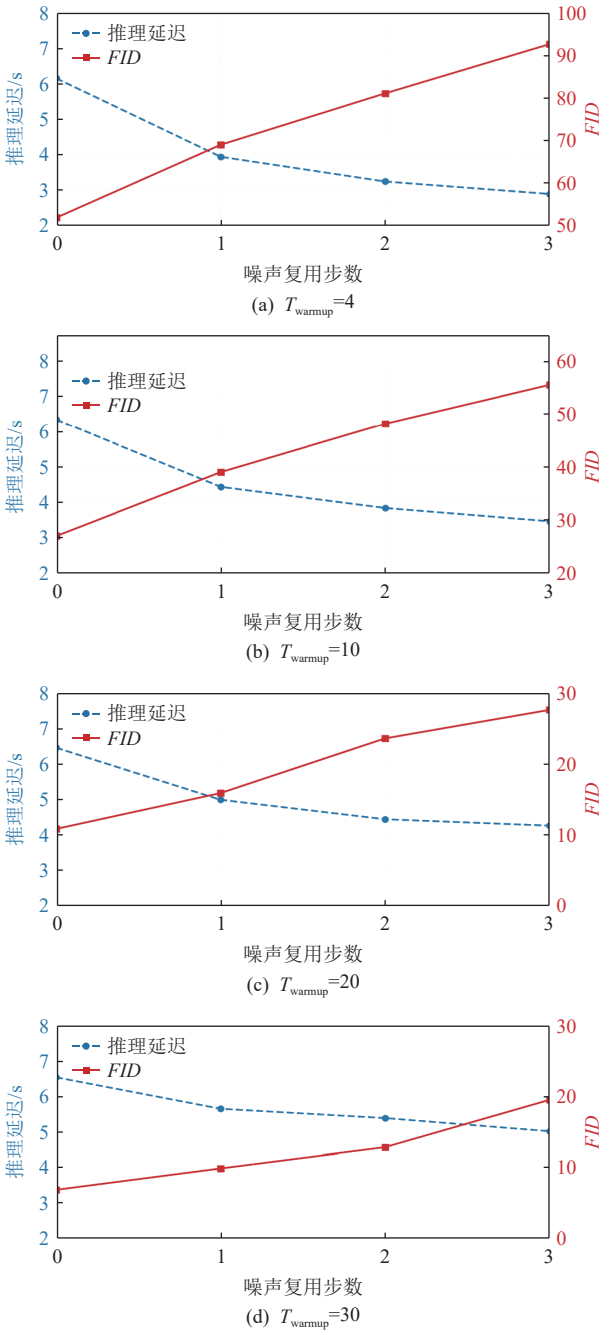


Fig. 8 Sensitivity analysis of generating quality and inference latency with T_{warmup} and k

图 8 T_{warmup}与k对生成质量和推理延迟的敏感性分析

效率之间的权衡。

4.2.2 与现有方法的定量评估

我们在采用标准 DiT 的 Pixart-α(0.6B)模型上进行当前实验, 针对图像生成任务对 DistriFusion 进行了调整, 将其从 U-Net 结构调整为 DiT 结构。为了确保公平性, 在 DistriFusion 和 PipeFusion[35] 方法中, 我们均使用 50 步 DPM-Sover 采样策略和 4 步预热步数。对于 DistriFusion、PipeFusion 和本文方法均使用 4 个 NVIDIA V100 GPUs 进行并行计算, 而对于 Original 只使用 1 个 NVIDIA V100 GPU。实验结果如表 2 所示, 生成效果如图 9 所示。

Table 2 Quantitative Evaluation for Comparison of Our Method and Existing Methods

表 2 本文方法与现有方法对比的定量评估

方法	*FID*	*LPIPS*	*PSNR*/dB	推理延迟/s	加速比
原始方法				16.6	
DistriFusion	55.48	0.159	22.016	6.08	2.73
PipeFusion	49.47	0.137	22.694	7.32	2.26
本文方法	**35.24**	**0.087**	**26.466**	**5.45**	2.99

注：黑体数值表示最优值。

在推理效率方面, 本文方法展现出最优的推理效率, 实现了 5.45 s 的推理延迟, 较 Original 方法提升了 2.99 倍的加速, 优于 DistriFusion 方法 2.73 倍和 PipeFusion 方法 2.26 倍。在生成质量方面, 本文方法显著优于基线方法, 在 *FID*, *LPIPS*, *PSNR* 指标上, 分别较次优方法 PipeFusion 提升了 28.7%, 36.5%, 16.6%, 表明本文方法有效缓解了传统并行方法导致的生成质量退化的问题。实验结果表明, 本文方法在图像生成质量和推理效率方面均取得了最佳性能。相比于 DistriFusion 和 PipeFusion, 本文方法大幅提升了图像生成质量, 展现出更强的应用潜力。

4.2.3 全部方法的效果

为了验证我们所提方法的有效性, 我们提取 COCO 数据集中相同数量不同复杂度的文本提示进行了测试, 并记录了对应的生成质量和推理延迟。结果如表 3 所示。

实验结果表明, 在生成质量方面, 本文方法显著优于标准模型。从简单到复杂提示 *FID* 分别降低 24.28%, 68.89%, 92.19%, 其中复杂提示的 *FID* 从 69.61 降至 5.44, *LPIPS* 降低了 82.76%, *PSNR* 升高了 42.27%, 反映出本文方法在感知相似性和像素级质量上的优势。随着提示复杂度的增加, 质量的提升幅度增大, 表明本文方法对复杂语义的适应性更强。在推理延迟方面, 本文方法的推理延迟平均降低 10.9%, 表明噪声复用方法确实起到了减少计算开销的作用。这些结果表明, 本文方法对复杂语义场景的生成质量提升尤为显著, 并且实现了质量与效率的协同优化。

4.2.4 预热感知方法的效果

为了验证预热感知方法的有效性, 我们提取 COCO 数据集中相同数量不同复杂度的文本提示进行了测试, 并记录了对应的生成质量和推理延迟。



Fig. 9 Quantitative evaluation effectiveness diagram for comparison of our method and existing methods

图 9 本文方法与现有方法对比的定量评估效果图

Table 3 Generation Quality and Inference Latency When Employing All Methods

表 3 使用全部方法的生成质量和推理延迟

提示文本复杂度	*FID*		*LPIPS*		*PSNR*/dB		推理延迟/s	
	Comm.	Ours.	Comm.	Ours.	Comm.	Ours.	Comm.	Ours.
简单	97.17	**73.61**	0.207	**0.170**	22.15	**23.66**	5.614	**4.694**
中等	83.53	**25.99**	0.187	**0.072**	21.70	**26.51**	5.604	**5.194**
复杂	69.61	**5.44**	0.145	**0.025**	23.30	**33.15**	5.606	**5.096**

注：Comm.表示在使用标准扩散模型的条件下生成的效果；Ours.表示在使用全部方法的条件下生成的效果。黑体数值表示最优值。

Table 4 Generation Quality and Inference Latency When Employing Solely Warmup-Aware Method

表 4 仅使用预热感知方法的生成质量和推理延迟

提示文本复杂度	*FID*		*LPIPS*		*PSNR*/dB		推理延迟/s	
	Comm.	D.W.	Comm.	D.W.	Comm.	D.W.	Comm.	D.W.
简单	97.17	**47.04**	0.207	**0.133**	22.15	**24.71**	**5.614**	5.639
中等	83.53	**20.02**	0.187	**0.062**	21.70	**26.70**	**5.604**	5.723
复杂	69.61	**5.15**	0.145	**0.022**	23.30	**33.36**	**5.606**	5.803

注：Comm.表示在使用标准扩散模型的条件下生成的效果；D.W.表示在仅使用预热感知方法的条件下生成的效果。黑体数值表示最优值。

结果如表 4 所示。

结果表明，预热感知方法显著提升了生成质量，在简单、中等、复杂提示下 *FID* 分别降低 51.59%, 76.03%, 92.60%, *LPIPS* 和 *PSNR* 也有明显改进。表明预热感知方法通过自适应调整预热步数，有效提升了初始阶段的稳定性，特别是在复杂语义场景下表现尤为突出。但推理延迟略微增加，平均增幅 2.37%，也表明预热感知方法会带来额外的计算开销。

4.2.5 噪声复用方法的效果

我们进一步评估了噪声复用方法的性能。我们

提取 COCO 数据集中相同数量不同复杂度的文本提示进行了测试,并记录了对应的生成质量和推理延迟。结果如表 5 所示。

Table 5 Generation Quality and Inference Latency When Employing Solely Noise Reuse Method

表 5 仅使用噪声复用方法的生成质量和推理延迟

提示文本复杂度	*FID*		*LPIPS*		*PSNR*		推理延迟/s	
	Comm.	D.S.	Comm.	D.S.	Comm.	D.S.	Comm.	D.S.
简单	97.17	**84.35**	**0.207**	0.222	**22.16**	22.15	5.614	**4.318**
中等	83.53	**79.52**	**0.187**	0.203	**21.70**	21.28	5.604	**4.810**
复杂	69.61	**67.38**	**0.145**	0.151	**23.30**	23.22	5.606	**4.984**

注:Comm.表示在使用标准扩散模型的条件下生成的效果;D.S.表示在仅使用噪声复用方法的条件下生成的效果。黑体数值表示最优值。

结果表明,噪声复用方法显著提升了推理效率,在简单、中等、复杂提示下推理延迟分别降低了 23.09%, 14.17%, 11.09%, 平均降幅 16.11%, 验证了噪声复用方法减少计算开销的效果。但生成质量略微下降, *FID* 降幅仅为 3.20%~13.18%, *LPIPS* 和 *PSNR* 分别增加 4.14%~8.56% 和降低 0.05%~1.94%, 反映出噪声复用方法会导致细节损失。

4.3 消融实验

为了验证预热感知方法和噪声复用方法的独立贡献,我们分别去除预热感知方法或噪声复用方法进行了消融实验,并记录生成质量和效率的变化。生成效果如图 10 所示,实验结果如表 6 所示。

实验结果表明,本文方法显著提升了图像生成质量,相比于标准模型,本文方法表现出了明显优势。在使用全部方法的情况下, *FID* 降低了 80.11%, *LPIPS* 减少了 78.52%, *PSNR* 增加了 40.65%, 反映出本文方法在逼真度、感知相似度和像素级质量上的改进。消融实验进一步揭示了本文方法的作用,去除预热感知方法后, *FID* 和 *LPIPS* 分别为 4.90 和 0.159, *PSNR* 接近标准模型 22.751 dB, 虽优于标准模型但不及本文方法,表明预热感知方法可以根据语义复杂度自适应选择最佳预热步数,获得更一致、更高质量的潜在表示,提升对初始阶段的稳定性。相比之下,去除噪声复用方法后, *FID* 升至 0.63, *LPIPS* 升至 0.019, *PSNR* 升至 33.877 dB, 显然噪声复用对质量影响较小,其作用主要体现在效率优化而非质量提升。

本文方法在推理延迟和计算成本上也实现了显



Fig. 10 Generation effectiveness diagram of ablation study

图 10 消融实验生成效果图

Table 6 Ablation Study Results

表 6 消融实验结果

方法	*FID*	*LPIPS*	*PSNR*/dB	推理延迟/s	计算成本/MAC
标准模型	5.03	0.149	23.008	5.63	338×10^{12}
本文方法	1.00	0.032	32.361	4.91	300×10^{12}
无预热感知	4.90	0.159	22.751	**4.35**	**271×10^{12}**
无噪声复用	**0.63**	**0.019**	**33.877**	5.87	338×10^{12}

注：黑体数值表示最优值。

著提升。推理延迟从标准模型的 5.63 s 降至 4.91 s，计算成本从 338×10^{12} MAC 减至 300×10^{12} MAC，表明本文方法不仅提高了生成质量还提升了推理效率。消融实验显示，噪声复用方法是效率优化的关键，去除该方法后，延迟增至 5.87 s，计算成本恢复至 338×10^{12} MAC，验证了噪声复用方法在低贡献时间步跳过噪声预测，大幅减少计算开销的作用。相反，去除预热感知方法后，延迟降至 4.35 s，计算成本减至 271×10^{12} MAC，但此效率的提升以质量略微下降为代价，突显出预热感知方法在质量与效率权衡中的重要性。

5 结论与讨论

本研究提出了一种基于预热感知噪声复用的高效自适应扩散模型推理方法——Wander，以优化扩散模型在多 GPU 环境下的性能。实验结果表明，所提方法显著提升了图像生成质量和推理效率。预热感知方法通过自适应调整预热步数，根据输入提示文本的语义复杂度灵活适应多样化场景；噪声复用方法则通过跳过部分时间步的噪声计算，复用前一时间步的噪声更新潜变量，有效解决了后续时间步的计算开销问题。评估显示，所提方法相较于基线方法在图像质量上提升了 80.1%，在推理速度上提升了 12.8%，充分验证了方法的有效性和应用潜力。

尽管所提方法在图像生成质量和推理效率上取得了显著改进，但预热感知方法在初始阶段引入了额外的计算负担。因此，在计算资源受限的场景下，所提方法可能难以充分发挥其潜力，参数预测过程对计算资源的需求可能进一步加剧这一问题。未来，在内存不够用的场景中我们会进一步探索方法的可行性。

作者贡献声明：杨迎春开展实验研究，撰写论文并修改论文；陈静波阅读文献并提供研究方法；彭雪梅与刘松冉共同提出了研究思路及工作方案；纪东提供了实验条件并提出论文修改意见；刘松冉提出论文修改意见。

参 考 文 献

[1] Li Zhixin, Wei Haiyang, Zhang Canlong, et al. Research progress on Image captioning[J]. Journal of Computer Research and Development, 2021, 58(9): 1951−1974 (in Chinese)
(李志欣, 魏海洋, 张灿龙, 等. 图像描述生成研究进展[J]. 计算机研究与发展, 2021, 58(9): 1951−1974)

[2] Liu Dawei, Yu Bihui, Shi Jiawei, et al. Method for generating clothing model images based on stable diffusion and adaptive enhancement techniques[J]. Application Research of Computers, 2025, 42(8): 2267−2273 (in Chinese)
(刘大伟, 于碧辉, 石珈维, 等. 基于稳定扩散与自适应增强技术的服装模特图像生成方法[J]. 计算机应用研究, 2025, 42(8): 2267−2273)

[3] Chen Qiang, Zheng Yuhui, Sun Quansen, et al. Patch similarity based anisotropic diffusion for image denoising[J]. Journal of Computer Research and Development, 2010, 47(1): 33−42 (in Chinese)
(陈强, 郑钰辉, 孙权森, 等. 片相似性各项异性扩散图像去噪[J]. 计算机研究与发展, 2010, 47(1): 33−42)

[4] Xing Zhen, Feng Qijun, Chen Haoran, et al. A survey on video diffusion models[J]. ACM Computing Surveys, 2024, 57(2): 1−42

[5] Wang Zelong, Wu Yuhang, Li Jian, et al. A review of diffusion models for inverse problems in image processing[J]. Journal of National University of Defense Technology, 2025, 47(4): 91−110 (in Chinese)
(王泽龙, 吴宇航, 李健, 等. 面向图像处理逆问题的扩散模型研究综述[J]. 国防科技大学学报, 2025, 47(4): 91−110)

[6] Li Xinwei, Su Busheng, Xu Haoran, et al. Diffusion Models: From Principles to Practice[M]. Beijing: Posts & Telecom Press, 2023 (in Chinese)
(李忻玮, 苏步升, 徐浩然, 等. 扩散模型从原理到实战[M]. 北京: 人民邮电出版社, 2023)

[7] Sapkota R, Karkee M. Creating image datasets in agricultural environments using DALL. E: Generative AI-powered large language model[J]. arXiv preprint, arXiv: 2307.08789, 2023

[8] Li Muyang, Cai Tianle, Cao Jiaxin, et al. DistriFusion: Distributed parallel inference for high-resolution diffusion models[C]//Proc of the 41st IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2024: 7183−7193

[9] Nichol A Q, Dhariwal P. Improved denoising diffusion probabilistic models[C]//Proc of the 38th Int Conf on Machine Learning. New York: PMLR, 2021: 8162−8171

[10] Dhariwal P, Nichol A. Diffusion models beat GANs on image synthesis[J]. Advances in Neural Information Processing Systems, 2021, 34: 8780−8794

[11] Ho J, Salimans T. Classifier-free diffusion guidance[J]. arXiv preprint, arXiv: 2207.12598, 2022

[12] Rombach R, Blattmann A, Lorenz D, et al. High-resolution image

synthesis with latent diffusion models[C]//Proc of the 39th IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2022: 10684−10695

[13] Sohl-Dickstein J, Weiss E, Maheswaranathan N, et al. Deep unsupervised learning using nonequilibrium thermodynamics[C]// Proc of the 32nd Int Conf on Machine Learning. New York: PMLR, 2015: 2256−2265

[14] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models[J]. Advances in Neural Information Processing Systems, 2020, 33: 6840−6851

[15] Song Yang, Garg S, Shi Jiaxin, et al. Sliced score matching: A scalable approach to density and score estimation[C]//Proc of the 35th Uncertainty in Artificial Intelligence Conf. New York: PMLR, 2020: 574−584

[16] Sami S M, Hasan M M, Dawson J, et al. HF-Diff: High-frequency perceptual loss and distribution matching for one-step diffusion-based image super-resolution[J]. arXiv preprint, arXiv: 2411.13548, 2024

[17] Zhu Suxia, Wang Jinyin, Sun Guanglu. Perceptual similarity-based multi-objective optimization for stealthy image backdoor attack[J]. Journal of Computer Research and Development, 2024, 61(5): 1182−1192 (in Chinese)
(朱素霞, 王金印, 孙广路. 基于感知相似性的多目标优化隐蔽图像后门攻击[J]. 计算机研究与发展, 2024, 61(5): 1182−1192)

[18] Huang Peixin, Zhao Xiang, Fang Yang, et al. End-to-end knowledge triplet extraction combined with adversarial training[J]. Journal of Computer Research and Development, 2019, 56(12): 2536−2548 (in Chinese)
(黄培馨, 赵翔, 方阳, 等. 融合对抗训练的端到端知识三元组联合抽取[J]. 计算机研究与发展, 2019, 56(12): 2536−2548)

[19] Formento B, Feng Winjie, Foo C S, et al. Semrode: Macro adversarial training to learn representations that are robust to word-level attacks[J]. arXiv preprint, arXiv: 2403.18423, 2024

[20] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. Advances in Neural Information Processing Systems, 2014, 27, 2672−2680

[21] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proc of the 30th IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 105−114

[22] Karakas C E, Dirik A, Yalçınkaya E, et al. Fairstyle: Debiasing StyleGAN2 with style channel manipulations[C]//Proc of the 17th European Conf on Computer Vision. Berlin: Springer, 2022: 570−586

[23] Karras T, Aila T, Laine S, et al. Progressive growing of GANs for improved quality, stability, and variation[J]. arXiv preprint, arXiv: 1710.10196, 2017

[24] Song Jiaming, Meng Chenlin, Ermon S. Denoising diffusion implicit models[J]. arXiv preprint, arXiv: 2010.02502, 2020

[25] Saharia C, Chan W, Saxena S, et al. Photorealistic text-to-image diffusion models with deep language understanding[J]. Advances in Nural Information Processing Systems, 2022, 35: 36479−36494

[26] Ramesh A, Dhariwal P, Nichol A, et al. Hierarchical text-conditional image generation with clip latents[J]. arXiv preprint, arXiv: 2204.06125, 2022

[27] Hyder S B, Tariq N, Moqurrab S A, et al. BERT-based deceptive review detection in social media: Introducing deceptiveBERT[J]. IEEE Transactions on Computational Social Systems, 2024, 11(6): 7234−7243

[28] Jain S, Dash S, Deorari R. Object detection using COCO dataset[C]// Proc of the 2022 Int Conf on Cyber Resilience (ICCR). Piscataway, NJ: IEEE, 2022: 1−4

[29] Jayasumana S, Ramalingam S, Veit A, et al. Rethinking FID: Towards a better evaluation metric for image generation[C]//Proc of the 34th IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2024: 9307−9315

[30] Wu Jiangyan, Zhang Guanghui, Fan Yugang. LM-CycleGAN: Improving underwater image quality through learned perceptual image patch similarity and multi-scale adaptive fusion attention[J]. Sensors, 2024, 24(23): 7425

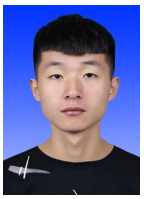
[31] Martini M G. Measuring objective image and video quality: On the relationship between SSIM and PSNR for DCT-based compressed images[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 1−13

[32] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: Common objects in context[C]//Proc of the 13th European Conf on Computer Vision. Berlin: Springer, 2014: 740−755

[33] Gupta Y, Jaddipal V V, Prabhala H, et al. Progressive knowledge distillation of stable diffusion XL using layer level loss[J]. arXiv preprint, arXiv: 2401.02677, 2024

[34] Howard A G, Zhu Minglong, Chen Bo, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint, arXiv: 1704.04861, 2017

[35] Wang Jianan, Fang Jiaru, Li Aoyu, et al. PipeFusion: Displaced patch pipeline parallelism for inference of diffusion transformer models[J]. arXiv preprint, arXiv: 2405.14430, 2024



Yang Yingchun, born in 2001. Master candidate. His main research interest includes diffusion models.
杨迎春, 2001 年生。硕士研究生。主要研究方向为扩散模型。



Chen Jingbo, born in 1997. Master candidate. Her main research interest includes diffusion models.
陈静波, 1997 年生。硕士研究生。主要研究方向为扩散模型。



Peng Xuemei, born in 1995. PhD candidate. Her main research interests include machine learning systems and model inference acceleration.
彭雪梅, 1995 年生。博士研究生。主要研究方向为机器学习系统、模型推理加速。



Ji Dong, born in 1989. PhD. His main research interests include image understanding and real-time systems.

纪　东, 1989 年生。博士。主要研究方向为图像理解、实时系统。



Liu Songran, born in 1990. PhD. Member of CCF. His main research interests include intelligent robot system, edge computing, and system software.

刘松冉, 1990 年生。博士。CCF 会员。主要研究方向为智能机器人系统、边缘计算、系统软件。

《计算机研究与发展》投稿指南

征稿范围

计算机体系结构；计算机网络与通信；网络与信息安全；人工智能；软件技术；并行与分布式计算；图形图像、信息检索等应用技术；其他计算机相关领域。

征稿类型

1）学术论文：有创新学术见解的研究成果的完整论述，对该学术领域的发展有积极意义，或者支持我国学术生态和产业生态的重要成果。

2）综述：对新兴的、活跃的学术研究领域或技术开发领域的评述。

3）快报短文：对目前国内外计算机领域新思想、新观点、新技术的解读文章。

征稿要求

1）本刊只接收中文稿，不受理英文稿.学术论文建议不超过 15 页，综述不超过 20 页，短文不超过 4 页。

2）要求来稿未在正式出版物上发表过，不存在一稿多投问题。

3）作者在投稿时，需同时向编辑部提交所有作者手写签字确认的“投稿声明”扫描版，如不提供不予受理.对于违反投稿声明相关条款的来稿，本刊将视情节做严肃处理，后果自负。

4）本刊双盲评审，作者投稿时上传系统的电子版中需要隐去作者、单位、基金、作者简介等信息，初审通过后给编辑部邮寄的纸质版本需要包含前述署名信息。改后复审的稿件在提交修改版本以及修改说明时，同样需符合双盲评审要求。

5）稿件评审通过后，作者需提交单位审核盖章的“不涉密证明”和所有作者签字确认的“著作权转让声明”。

6）投稿时间在 6 个月以上的，若没有收到编辑部发出的审理结果通知，在书面通知编辑部并收到编辑部确认回复之后，作者方可自行处理此稿件。

7）无论何种原因撤回投稿，需以书面形式通知编辑部并确保所有作者知情。