

ASD

状态：AAAI2024

单位：西北工业大学+华为诺亚方舟研究所

文章链接：<https://ojs.aaai.org/index.php/AAAI/article/view/28589>

Github 链接：<https://github.com/ProAirVerse/Any-Size-Diffusion>

目录

摘要	1
评价	1
背景	1
贡献	1
实现	2
实验	2

摘要

diffusion 通常使用某一固定尺寸的训练集进行训练，这使得其无法生成质量较高的其他尺寸图像。但直接使用不同尺寸的数据进行训练是不现实的，这需要巨量的训练集以及额外算力开销。本文提出了一种两阶段的方案 ASD 来进行任意尺寸的扩散生成任务。第一阶段为 ARAD，第二阶段为 FSTD。

评价

一篇质量不佳的论文，实现细节不够清晰，且其主要贡献应该在提出了一种取代分块时 overlap 的加速方案。（代码也不容易读）

背景

扩散模型大多只能适用于特定的尺寸。想要训练出一个兼容不同尺寸输入的模型是不现实的。而有时候会想生成特定尺寸的图像。

贡献

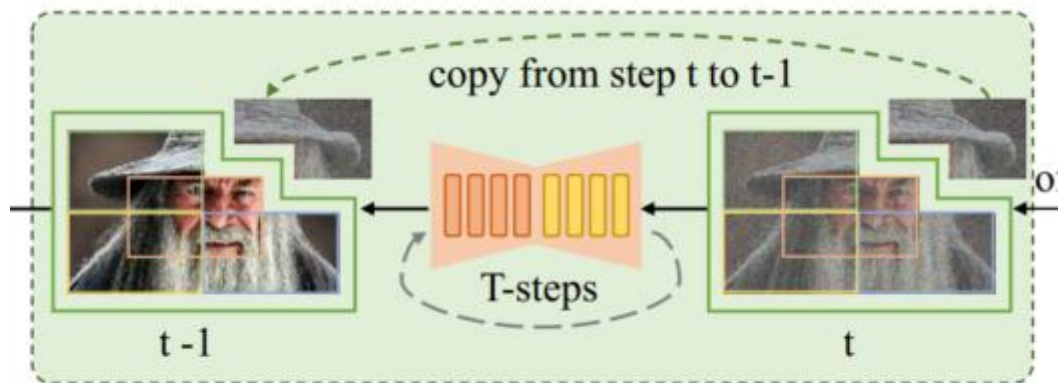
1. 本文提出了一种生成任意尺寸图像的整体模型。
2. 通过实验论证了其可行性。

实现

实现部分主要分为两个阶段：

1. 生成目标比例的较小图像。
2. 将这个小图像放大到目标尺寸。

这一两阶段的生成方案并不十分有创新性，但其主要贡献在于，提出了一种精确切分的方案，来减少了重叠部分所造成的额外开销。

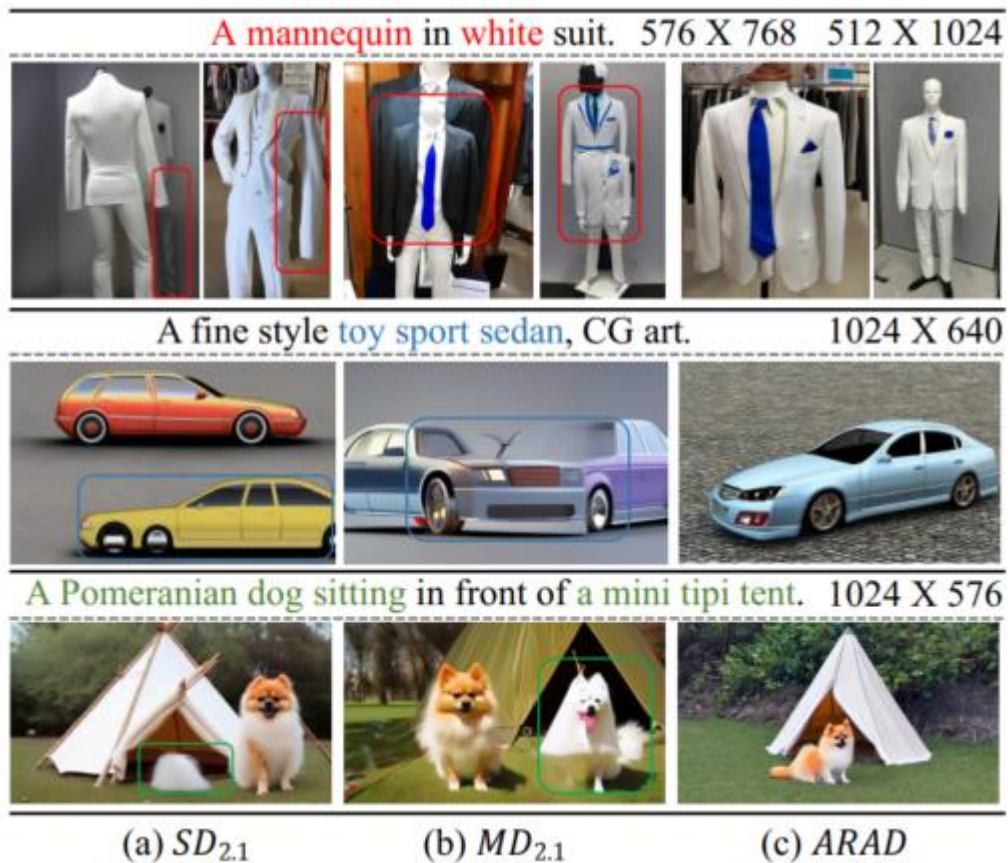


(c) Implicit-Overlap in Tiled Sampling

具体实现方式如图，首先进行不重叠的切分，但为了保证其与周围图像的连贯性，每次去噪时都会偏移一部分，从而对包含周围图像的部分进行去噪，在迭代过程中引入了周围像素的信息，从而避免了明显边界的出现。

实验

基础的定性对比：



以及简单的定量实验：

Method	MA-LAION-COCO			MA-COCO		
	FID ↓	IS ↑	CLIP ↑	FID ↓	IS ↑	CLIP ↑
$SD_{2.1}$	14.32	31.25	31.92	42.50	30.20	31.63
$MD_{2.1}$	14.57	28.95	32.11	43.25	28.92	30.92
ARAD	13.98	34.03	32.60	40.28	29.77	31.87