

# 23.12.11~23.12.17周报

## 一：本周工作

- 读论文《Emu: Enhancing Image Generation Models Using Photogenic Needles in a Haystack》
- Machine Learning课程学习
- games101课程学习

课程学习笔记更新在仓库：<https://github.com/orange-v-soda/Learning-recoder>

## 二：下周计划

- 继续diffusion相关论文的阅读
- 继续Machine Learning课程学习
- 继续games101课程学习

## 三：详细工作



### 3.1 概述

diffusion模型在生成领域得到了极大的应用，但由于数据集过大，预训练模型所输出的结果往往不够满足人类的审美。本文提出了一种高校训练高质量图片的模型，主要基于两个阶段：1.学习基础生成能力的预训练阶段。2.学习生成高质量图片的微调阶段。

本文将文生图模型的质量提高与LLM的微调进行了联系，指出两者所共有的特点：

1. 需要通过微调来提高产出质量
2. 微调通过少量但高质量的训练集来进行
3. 基本保留预训练所学习的知识

latent diffusion模型和pixel diffusion模型有什么区别？

本文贡献：

1. 构建了超过眼下模型的高质量图片生成模型Emu
2. 首次强调了训练集中图片质量相较于数量更加重要
3. 该范式也可以运用于其他模型中

### 3.2 相关工作

**文生图模型相关工作：**目前的问题在于，预训练模型生成的图片质量不够

**文生图模型的微调：**虽然有关文生图的微调有许多尝试工作，但本文是第一篇在大范围多领域提高图片质量的工作。

**LLM的微调工作：**语言模型的微调和生成模型的微调之间存在的关系，启发了本文

### 3.3 实现

**Latent Diffusion模型：**简单介绍了本次使用的Diffusion模型

**预训练模型：**本次预训练模型的训练情况

**如何选择高质量图片：**详细介绍了本文中训练集图片的选择标准

**微调：**非常简短的有关微调过程的介绍

### 3.4 实验

主要与微调之前的预训练模型进行了对比，并与当前的模型SDXL进行对比，体现了本文工作有所突破。

不足之处在于：

1. 训练集需要人工分类，因为要求十分严格

2. 对于预训练模型没能很好学会的物体，并不会有很大的提高
3. 可能会像大多数生成模型一样，有不恰当的结果生成