在本次课程中，我认识到了开源技术在推动算法平权方面的重要作用。此外，在本次课程中，通过实际案例展示，与课堂练习，我认识到了利用API对大模型进行调用与微调的方法。作为论坛题目，在此对我在此次课程中比较感兴趣的LDM进行简要的介绍。

LDM是建立在DDPM基础上的重大改进，主要解决了 DDPM 在高分辨率图像生成上面临的核心瓶颈：‌计算效率‌。它通过将扩散过程转移到‌潜空间‌来实现这一目标。

LDM的核心改进在于通过引入预训练的自编码器，将传统 DDPM 在高维像素空间进行的扩散过程转移到低维潜空间，从而显著提升计算效率。具体而言，LDM 利用编码器将原始图像压缩为低维潜变量（保留语义信息但大幅降低维度），并在该潜空间内执行噪声添加与去除的扩散过程，最后通过解码器将去噪后的潜变量重建为高质量图像。这一设计解决了 DDPM 在高分辨率生成时计算量庞大的问题，使训练和推理速度提升数倍，同时降低内存需求，且生成质量与像素空间扩散相当甚至更优。此外，潜空间的紧凑性增强了多模态条件（如文本）的融合能力，通过交叉注意力机制实现精准的条件控制（如文本到图像生成），其模块化结构（自编码器与扩散模型分离）也提高了灵活性和可扩展性。这一改进使扩散模型得以广泛应用于高分辨率、复杂条件的生成任务（如 Stable Diffusion），成为 AI 生成领域的里程碑技术。‌

LDM的核心优势在于通过潜空间扩散机制实现了高效与高质量的平衡：其利用预训练自编码器将高维像素空间压缩至低维潜空间进行扩散，大幅降低计算复杂度，使训练和推理速度提升数倍，同时减少内存消耗；潜空间保留的语义信息与紧凑表征不仅维持了生成图像的感知质量（甚至优于像素级扩散），还通过交叉注意力等机制强化了多模态条件控制（如文本到图像的精准对齐），成为Stable Diffusion等应用的技术基础；模块化设计（分离自编码器与扩散模型）进一步提升了灵活性和可扩展性，使LDM能够高效支持高分辨率生成与复杂条件任务，推动扩散模型从理论走向大规模实际应用。‌

值得特别留意的是，在与我相关的实际的研究应用中，LDM相较于现有的方法显示出了较大的优势。当前，我正着重于研究AIGC技术在扩充高质量的研究试验数据方面的应用。例如，在某材料的疲劳寿命试验中，试验开展费时费力，试验测得的数据量不足以为建立基于机器学习的增材疲劳寿命预测模型提供足够的支撑。为了解决这一问题，便开展了相应的研究。目前的SOTA是我们以添加了先验物理信息约束的GAN所形成的数据增强框架，如图1所示。数据增强结果如图2所示。此外，我们也开展了利用DDPM进行相应的数据增强的先导研究，结果如图3所示。能够发现，DDPM本身显示出了较好的性能。然而，如何添加边界约束条件却始终是将DDPM进行部署时亟待解决的问题。在LDM中能够以MHA模块得到应用的Prompt工程也许能够为其实际应用提供新的思路。此外，通过利用LDM在低维空间中进行前向传播及反向去噪的过程，能够在低维空间中直观地绘制出相应的扩散路径，进而在此空间中进行额外的约束。这也不失为一种部署DDPM的新思路。



图1 当前基于GAN的数据增强框架



图2 当前基于GAN的数据增强框架所得结果



图3 当前基于DDPM的数据增强框架所得初步结果