

会议纪要

参会人：樊老师，邹老师，徐宇航，董林伟，Teme，王忠豪，史怡晨，郭奕宏

时间：2024.7.29 下午 3: 00

Teme:

- Presented my progress report that includes quantitative and qualitative results of the first version and my attempts to address the limitation observed in the experiments
- Discussed on how to improve the work to make it publishable on top-tier conferences. These includes ⁱusing similar dataset setting as previous works, improve novelty of the work by using own dataset or proposing better modeling ideas

Feedbacks:

- 邹老师
 - Build own dataset and try to solve fine-grained image retouching to have better control on experiments and potentially add novelty
 - Try to add more architectural improvement to the retouch transfer module to improve novelty, if there exist problems that the current module can't address
- 樊老师
 - Retouch transfer should be trained on a separate dataset that is different from the dataset used for training retouch adapter
 - Improve the sample efficiency of previous image retouching methods and test on the same test-set as previous methods
 - Do experiments to analyze this
 - All experiments that will be done next should focus on supporting the two contributions: 1. Few-shot image retouching and retouch transfer modules

Explanation: my presentation using the second slide might have created a confusion that I am doing experiments different from the two main contributions because I didn't clearly explain the motivation. The main reason I did the experiments were to improve the retouch adapters performance as I found that the performance degradation was mainly there rather than retouch transfer.

徐宇航:

介绍了 UHD/HD 领域的大致情况。

介绍了目前对几个 idea 的排布方案。更优化的切分方案并不会特别适应高尺寸图像，而智能切分方案则在大图像上表现更佳。因此，尝试将这两个 idea 分布到两份工作中。

介绍了大致的论文框架，包括文章背景，相关工作发展，目前方案的不足之处，本文的创新点与贡献。

最后介绍了基于文章 SeD 的思路：在训练时加入语义信息以提高判别器 D 的能力，从而进一步提升生成器 G 的效果。并减少预测时所需要的计算开销。

樊老师：

工业界需要简洁有效的方案，但学术界往往更在乎想法即思路的尝试，所以以解决问题为最先导向。下次汇报最好放上更多的实验结果图来说明出现的问题。方向技术方案比较明确了，对比方法，训练测试集需要确定好，方便与实验结果进行对比。

邹老师：

要多读最新发表的文章，对目前的行业进展有更为清晰的了解。可以尝试将这一 idea 用于更多相关领域，如更大尺寸的图像生成，或者短视频生成。

董林伟：

介绍了 SD3 的基础数据。

介绍了几种基于 SD3 的蒸馏方案。

介绍了一些基于目前方案的改进方案。对 VAE 进行额外的训练。

主要有三种方案：OSDiff，GAN，Improved。展示了这几组模型的可视化结果，进行定性对比，以及测试指标，进行了定量对比。

并进行了剪枝训练的数据及结果展示，总结了目前遇到的问题：

1. 生成的结果存在规律网格或竖线噪声
2. 目前网络重建做的多，生成做的少，对于细节的修补能力弱，细节不太行
3. 可能因为像素做 mse 的原因，导致图像过于平滑，看似仍存在一层噪声

对于问题一，猜测可能是 sd3 的原因，发现 sd3 少步生成时，也会出现这种情况。而且，发现可能与输入图像尺寸有关，sd3 可能是多尺度训练，倾向于 1024 数据，我们训练时 512 数据，故训练初期会产生这一现象，且推理如果使用 2048 做输入，也会出现这种现象，将 2048 分片成 512 得到缓解

对于问题二，猜测为单步去噪点 transformer 未充分学好，可能单步生成能力不好，一个观察是，通过这个 Vae 拥有了重建的去模糊能力，而后面的去噪网络的作用可能是对于细节的还原和生成。

樊老师：

后续的汇报希望逻辑更为清晰一些，将自己的贡献更明确的强调清楚。图像对比，数值对比都最好更为清晰一些。目前的结果很难去具体衡量改进的程度。

vae 编码结果和 hr 加噪做正则损失和文章的卖点可能有出入。taesd 工作可以参考一下，但是不建议剪枝 vae 的架构。数据集统一，退化方式统一。图像的展示方式可以用新工具来展示，要展示的很清楚。剪枝效果后续需要数值指标做证明，而且说明计算模型缩小了多少或者浮点计算少了多少。剪枝不能太过于简单粗暴。搞一个退化能力强的数据集来做测试，验证模型的效果。
