

APISR

状态：CVPR 2024

单位：密歇根大学

文章链接： <https://arxiv.org/abs/2403.01598>

Github 链接： <https://github.com/Kiteretsu77/APISR>

目录

摘要	错误!未定义书签。
背景	错误!未定义书签。
贡献	错误!未定义书签。
实现	错误!未定义书签。
实验	错误!未定义书签。

摘要

这是一项针对动漫图像的超分工作，其通过获取动漫视频中的关键帧构建了一个独有数据集，并以此训练模型。此外，其还提出了动漫超分所面对的几大问题：

1. 变形
2. 模糊的线条
3. 额外的色块

并提出了相应的解决方案。

背景

Anime SR（动漫超分）是 SR 的一个子任务，在当下的社会中有越来越大的重要性和经济利益。但过往的工作都尝试将真实图像 SR 中的方案直接搬运到这一领域，而忽略了动漫本身的独特属性。

在本文中，我们考虑了动漫本身的制作流程，并尝试从其中获得灵感，利用其特有的性质来帮助动漫超分。首先，动漫被以草图绘制，并经过上色，增强后放入电脑中，然后利用关键帧形成视频。由于在视频中使用接近的画面是常见现象（为了减少工作量），因此本文放弃使用完整的视频作为训练素材。转而采用从视频中截取图片来构建训练集。

此外，目前对动漫超分的方案普遍存在以下问题：

- 1.无法还原真实动漫中清晰，明确的线条。生成的线条往往变形且模糊。
- 2.存在额外不该出现的色块。

贡献

- 1.提出了一种从动漫视频中提取关键图像的流程，并提供了一个动漫 SR 的数据集。
- 2.构建了一个对动漫图像压缩的模块，用于模仿动漫图像在传播过程中受到的损坏。
- 3.针对性的对两个动漫超分的独特问题进行了解决。
- 4.通过一系列的实验论证了这些方案的可行性。

API 数据集及流程

由于要从视频中获取图像作为数据时希望所选择的图像尽可能的不同，且含有足够多的信息，随机取样或者等距取样显然难以达到期望。因此，本文开发出了一种获取关键图像帧的流程，并以此提取出了一个动漫 SR 数据集——API 数据集。

特殊问题解决方案

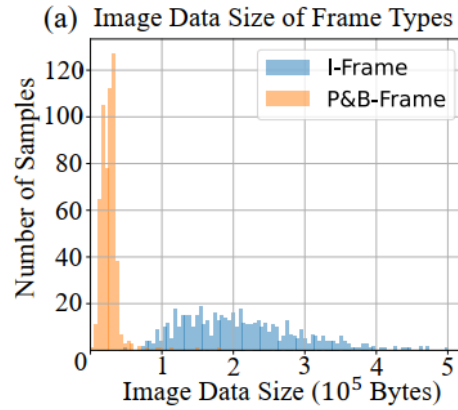
线条不清晰问题：

在训练时加入了一个额外的线条增强构件，以给 GT 图像进行进一步的线条增强，同时以此作为 loss 计算的图像，来引导生成模型生成更具有清晰线条的图像。

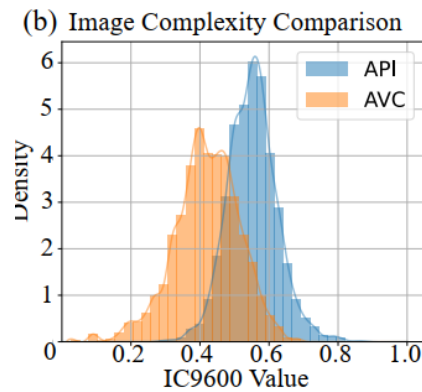
额外色块问题：

本文认为这一问题的出现来自真实图像的训练集，而本文通过额外加入一个感知 loss 来进行解决。

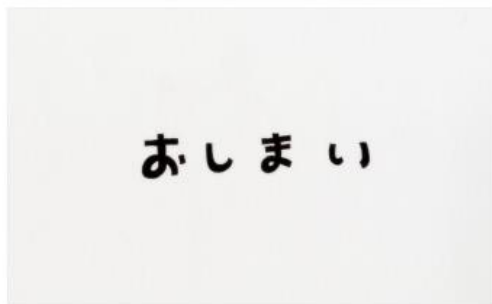
实现



此图代表 I-Frame 所含的数据量要远超过其他的非 I-Frame，即关键帧所含的数据量显著更大，因此要尽可能获取动漫中的关键帧作为数据集。



此图将本文得到的数据集 API 和过往首个动漫图像数据集 AVC 进行对比，可以看出 API 在整体分布上评分更高，因此论证本文的 API 数据集更佳。

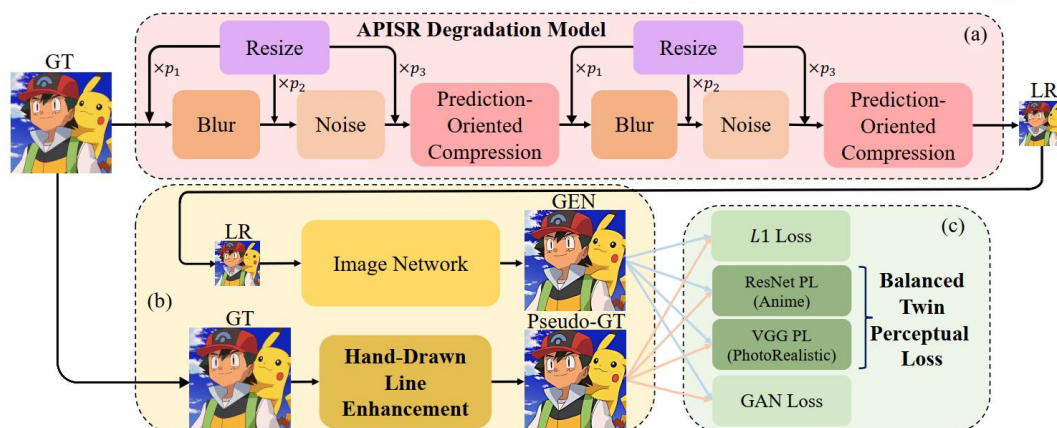


HyperIQA↑: 0.519 Brisque↓: 20.48
Image Complexity (IC9600) ↑: 0.097



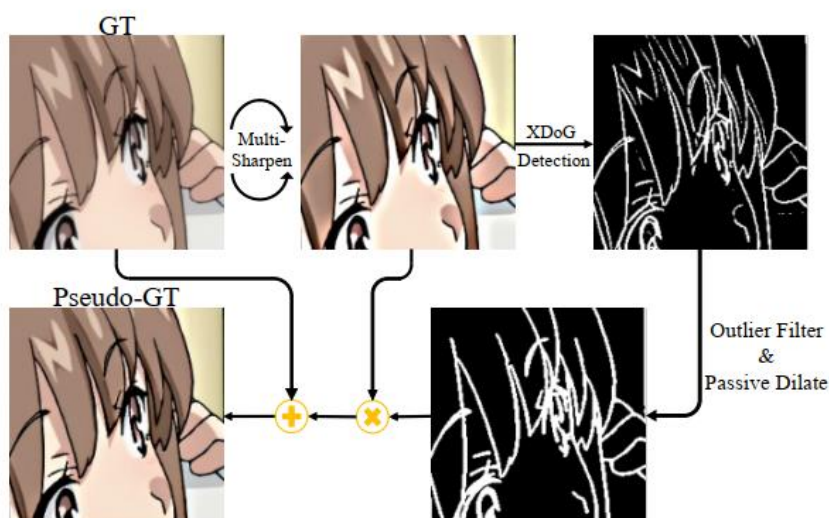
HyperIQA↑: 0.294 Brisque↓: 33.78
Image Complexity (IC9600) ↑: 0.613

更清晰的数据展示，可以看出，右图中的信息量和复杂程度都远远超过左图，因此更适合作为训练集。同时其各项指标中，IC9600 最适合用于评估复杂程度。



本文的整体流程图，结构较为简单，上方红色区域为将 **GT** 生成 **LR** 的缩放过程，其间模仿实际可能在互联网上经历的传播过程，进行了模糊和加噪的不同程度叠加。

下方则将线条增强后的 **GT** 图像作为 **loss** 计算图像，用于引导超分网络生成更佳鲜明的图像。右下则是通过平衡 **Anime** 和真实图像两种感知损失来减少额外的色块产生。



上图为线条增强的方案，简单来说就是提取了 **GT** 图像中的线条部分，并将锐化了的原图的线条部分保留，与原图进行复合得到增强过线条的图像。

实验

Methods	Params ↓	NIQE↓	MANIQA↑	CLIPQA↑
Real-ESRGAN* [54]	16.70	8.281	0.381	-
BSRGAN* [65]	16.70	8.632	0.376	-
RealBasicVSR* [8]	6.30	8.621	0.362	-
AnimeSR [56]	1.50	8.109	0.462	0.539
VQD-SR [46]	1.47	8.202	0.464	0.567
APISR (Ours)	1.03	6.719	0.514	0.711

实验部分较为简单，只是将几种方案与本方案进行了四组参数上的对比。但值得注意的是下三种方案是针对动漫的超分方案，因此参数显著提高不少。

Dataset	NIQE↓	MANIQA↑	CLIPQA↑
AVC-Train [56]	7.681	0.476	0.658
Random Select	8.006	0.446	0.625
I-Frame + IQA Select	7.876	0.493	0.675
I-Frame + ICA Select	6.912	0.499	0.683
I-Frame + ICA Select + 720P Rescale	6.719	0.514	0.711

关键帧提取方案的消融实验，以此得到了最佳的关键帧提取法。

Degradation Model	NIQE↓	MANIQA↑	CLIPQA↑
High-Order [54]	6.667	0.483	0.663
Random Order [65]	6.975	0.491	0.674
Prediction-Oriented Compression	7.133	0.506	0.709
Compression + Shuffled Resize	6.719	0.514	0.711

本实验用于论证本文的退化方案更为优秀。