

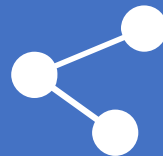


# 基于改进型生成对抗网络的图像 超分辨率重建方法研究

——以 Real-ESRGAN 为核心模型

答辩人：吴抒航、杨宇涵、宋语婷

第三组小组成员：吴抒航、张敬善、杨宇涵、曾丽婷、宋语婷



研究背景与技术原理

1

使用方法与实用技巧

2

Real-ESRGAN

在不同场景中的应用

3

4

AI工具及传统对比

5

与传统图像增强技术比较

6

研究总结与反思

The slide features a light gray background with a pattern of overlapping, semi-transparent squares in various shades of gray and beige. A solid blue horizontal bar spans the width of the slide, containing the title text. On the left side, a dark blue square contains the number '1'.

1

# 研究背景与技术原理



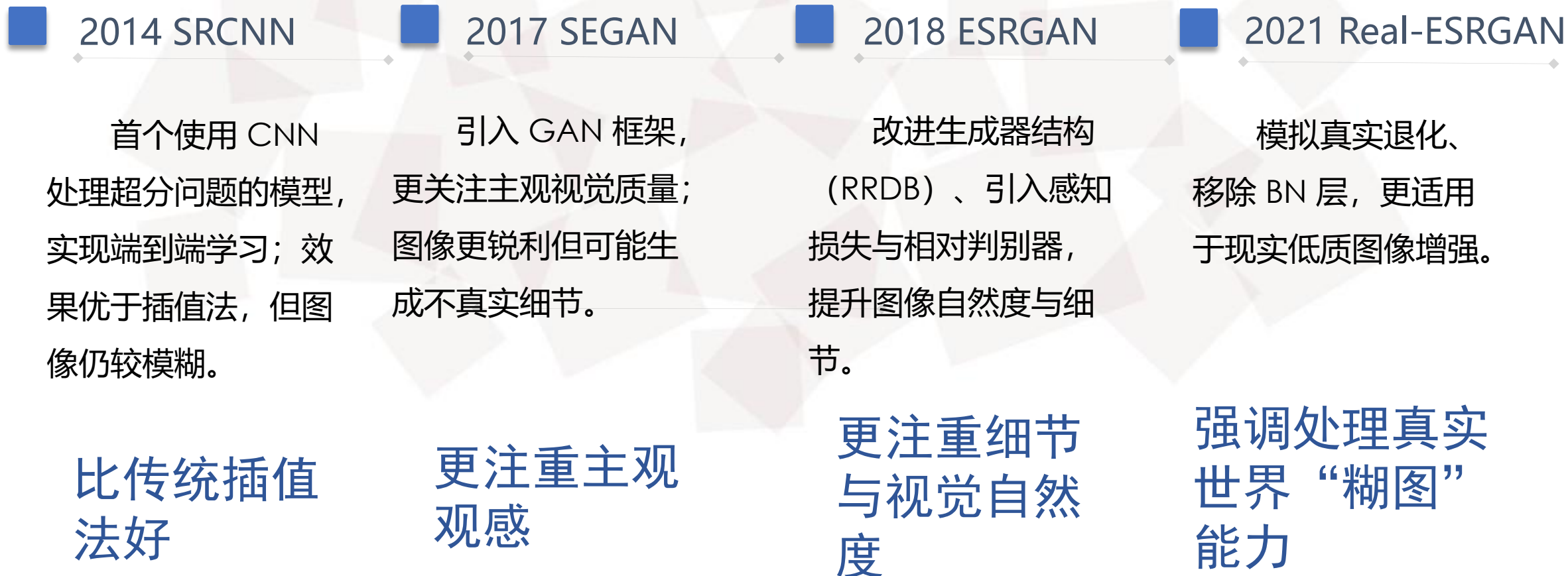
图像超分辨率重建 (Super-Resolution, SR) ,指将低分辨率图像还原为高分辨率图像的技术,常用于提升图像的清晰度和细节质量。

SR从早期的插值和重建方法逐步发展为以深度神经网络为核心的智能化方法。

图像超分方法常分为三类：插值法、重建法与基于学习的方法。

- 传统方法（如插值、重建）简单易用但存在模糊、细节丢失问题，难以满足实际需求。
- 随着深度学习发展，基于神经网络的 SR 方法展现出强大图像建模能力。
- 本研究旨在探究 Real-ESRGAN 模型的原理与性能，并通过实验验证其在多种图像场景中的增强表现。

## 超分辨率模型发展脉络





### 🔧 网络结构组成:

- **生成器 (Generator)** : 基于 RRDB (残差块) 结构, 去除 BN 层, 增强细节还原能力
- **判别器 (Discriminator)** : 采用**相对判别机制**, 更关注哪张图像更接近真实
- **退化建模模块**: 模拟模糊、噪声、压缩等真实图像退化过程
- **损失函数设计**:
  - 像素损失 (Pixel Loss)
  - 感知损失 (Perceptual Loss)
  - 对抗损失 (Adversarial Loss)
  - 全变差损失 (TV Loss)



## 2

# 使用方法与实用技巧

## 使用流程

1

在FFmpeg官网，根据操作系统选择合适的版本，下载并解压 FFmpeg 到本地目录。

2

从 GitHub 项目页获取 Real-ESRGAN 的可执行文件，并解压至同一目录。

3

在同一文件夹中创建 out\_frames 与 tmp\_frames 两个子文件夹。

4

将需要处理的图片或视频放入可执行文件夹，打开命令行中切换至可执行文件目录，执行以下命令：  
**realesrgan-ncnn-vulkan.exe -i input.jpg -o output.png -n realesrgan-x4plus -s 4**。  
程序处理完毕后，输出图像将保存在指定位置。





新标签页 新标签页 新标签页 视频执... x + - □ x

视频执行代码

帧率一定要契合

```
ffmpeg -r 23.98 -i out_frames/frame%08d.jpg -c:v libx264 -r 23.98 -pix_fmt yuv420p output.mp4
```

Shell

### 图片执行代码

```
realesrgan-ncnn-vulkan.exe -i A2.png -o outputA2.png -n realesrgan-x4plus -s 4
```

Shell

个人展示 ? 0 条反向链接 117 个词 641 个字符

realesrgan-ncnn-vulkan-20220424-windows > 在 realesrgan-ncnn 中

新建 排序 查看 详细信息

名称	修改日期	类型	大小
.obsidian	2025/7/3 11:05	文件夹	
models	2025/7/3 11:05	文件夹	
out_frames	2025/7/1 17:43	文件夹	
tmp_frames	2025/7/1 17:43	文件夹	
A2.png	2025/7/1 17:50	PNG 文件	
README_windows.md	2022/4/24 19:17	MD 文件	
realesrgan-ncnn-vulkan.exe	2022/4/24 17:07	应用程序	6.0
vcomp140.dll	2022/1/5 23:34	应用程序扩展	
vcomp140d.dll	2022/1/5 23:34	应用程序扩展	

9 个项目

命令提示符 x + - □ x

Microsoft Windows [版本 10.0.26100.4351]  
(c) Microsoft Corporation. 保留所有权利。

C:\Users\Shu>

2025/06/20 19:43:01



3

## 在不同场景中的应用



## 实验设计与评价指标

- 实验场景： 人物图像、自然场景、动画图像三类典型场景
- 评价方法：
  1. 主观视觉观察（细节、自然感、结构）
  2. 客观指标：
    - NIQE（自然图像质量评估）： 数值越低越自然
    - BRISQUE（盲感知图像质量）： 数值越低视觉质量越好



原图



Real-ESRGAN处理后图像

Real-ESRGAN能**有效**  
**提升面部保真度和边缘纹**  
**理细节上的增强效果**，未  
出现明显过锐或过度平滑  
现象。

模型在某些原图中比  
较模糊的区域，例如眼睛  
部分，**出现了轻微伪影**，  
**对真实感略有影响**。



### 3 自然场景图像增强分析



在自然图像处理方面，Real-ESRGAN有较好表现，整体来看，处理过后的图像细节丰富，船只与建筑轮廓更加清晰锐利，色彩饱和度略有提升，视觉感受更鲜明。



原图



Real-ESRGAN处理后图像

### 3 自然场景图像增强分析



但将图片放大来看，  
墙壁上砖块的纹理出现  
较明显的边缘模糊。树  
叶上由于锐化的原因，  
将部分光影锐化为白边。

总体来说在自然图  
像处理方面整体结构保  
持良好，但在细节表现  
上略有不足。

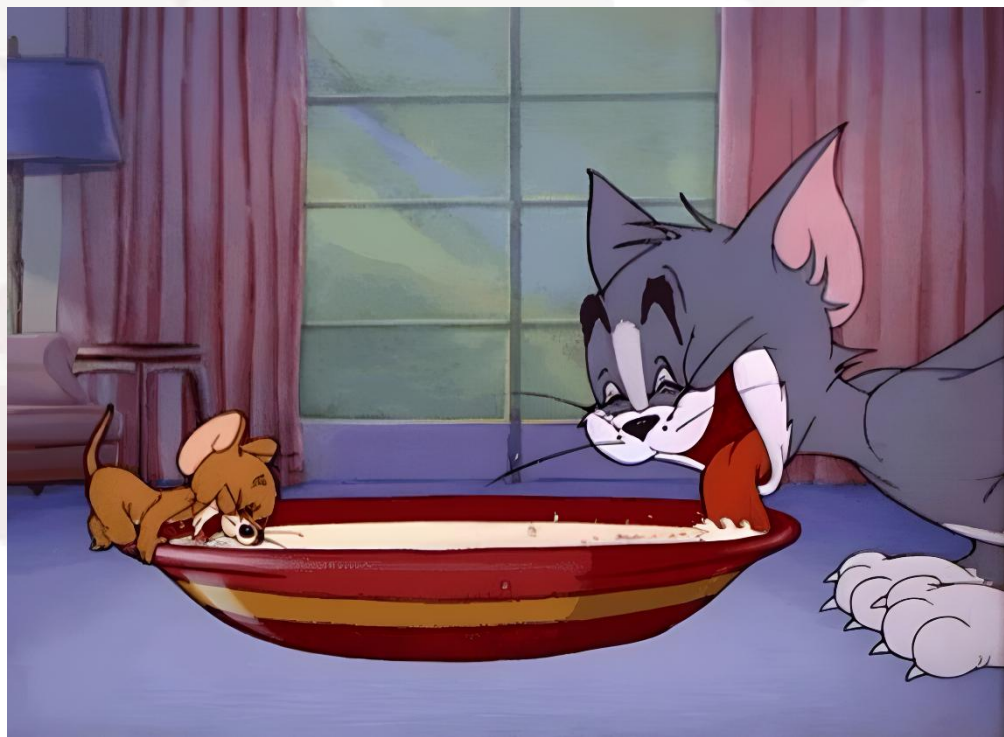






原图

Real ESRGAN在边缘线条重建方面表现突出，线条流畅锐利色彩分区过度自然，画风保持能力表现稳定，增强后的图像整体风格与原图一致，未出现风格大面积突变或“写实化”等失真现象。



Real-ESRGAN处理后图像

尽管从数据来看Real-ESRGAN处理后两个评估值普遍呈现下降趋势，但是主观看来，增强后的图片比原图更加清晰。这种现象恰好体现了Real-ESRGAN优化目标侧重于提升人眼主观感知下的图像清晰度与真实感。

图像 编号↵	类型↵	原图↵ NIQE↵	原图 BRISQUE↵	Real- ESRGAN 增强图像 NIQE↵	Real- ESRGAN 增强图像 BRISQUE↵
A1↵	↵ 人物图像↵	4.32↵	50.24↵	2.44↓↵	26.11↓↵
A2↵		2.79↵	29.63↵	3.46↑↵	38.98↑↵
A3↵		3.29↵	29.54↵	3.11↓↵	11.43↓↵
B1↵	↵ 动画图像↵	2.93↵	33.36↵	3.58↑↵	20.54↓↵
B2↵		3.96↵	34.03↵	6.03↑↵	46.54↑↵
B3↵		5.83↵	34.76↵	5.10↓↵	38.64↑↵
C1↵	↵ 自然场景图像↵	3.29↵	32.32↵	3.9↑↵	22.37↓↵
C2↵		3.85↵	12.24↵	3.33↓↵	28.16↑↵
C3↵		2.32↵	13.10↵	3.10↑↵	31.14↑↵





4

## 同类AI工具对比

为全面评估Real-ESRGAN的图像增强性能，我们选取了三款典型的同类AI图像增强工具进行对比分析，分别为GFPGAN、waifu2x和 Bigjpg。三者和功能定位、适用场景上各具特色，能够与Real-ESRGAN形成具有代表性的对照。

### GFPGAN

专注人脸修复与增强，擅长对人脸区域进行结构还原与自然补全，尤其适合处理面部模糊、面部特征丢失等问题图像。

### waifu2x

面向二次元图像，具有良好的线条保护能力和抗锯齿效果。

### Bigjpg

是一款通用型图像放大工具，适配艺术类插画、绘图等图像处理场景。



原图



GFPGAN



Real-ESRGAN

图 像 编 号	类 型	原图 NIQE	原图 BRISQUE	Real-ESRGAN 增强图像 NIQE	Real-ESRGAN 增强图像 BRISQUE	GFPGAN 增强图像 NIQE	GFPGAN 增强图像 BRISQUE
A1	人物	4.32	50.24	2.44	26.11	3.65	44.13
A2	图像	2.79	29.63	3.46	38.98	3.74	30.78
A3		3.29	29.54	3.11	11.43	3.56	15.62

分析结论

Real-ESRGAN 与GFPGAN的指标起伏基本一致，Real-ESRGAN的指标表现优于GFPGAN。这表明GFPGAN可能在某些方面引入过锐。



原图



Bigjpg



Real-ESRGAN

图像 编号	类型	原图 NIQE	原图 BRISQUE	Real-ESRGAN 增强图像 NIQE	Real-ESRGAN 增强图像 BRISQUE	Bigjpg 增 强图像 NIQE	Bigjpg 增 强图像 BRISQUE
C1	自然	3.29	32.32	3.9	22.37	3.56	15.62
C2	场景	3.85	12.24	3.33	28.16	3.56	15.62
C3	图像	2.32	13.10	3.10	31.14	3.56	15.62

分析结论

对于大部分图像来说Bigjpg 在 NIQE 与 BRISQUE 上的得分均低于 Real-ESRGAN，表现出更优的客观指标。尤其是在大面积纹理区域，Bigjpg 更能保留原始细节。

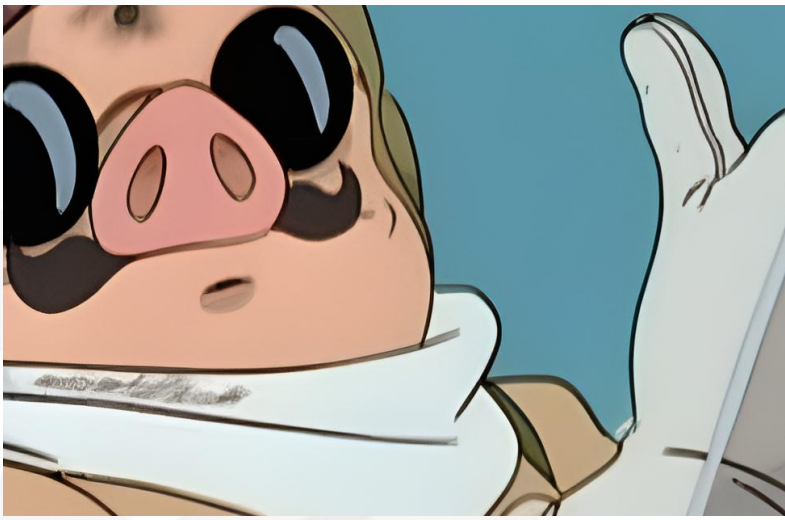




原图



waifu2x



Real-ESRGAN

图像 编号	类型	原图 NIQE	原图 BRISQUE	Real-ESRGAN 增强图像 NIQE	Real-ESRGAN 增强图像 BRISQUE	waifu2x 增强 图像 NIQE	waifu2x 增强 图像 BRISQUE
B1	动画 图像	2.93	33.36	3.58↑	20.54↓	3.65↑	44.13↑
B2		3.96	34.03	6.03↑	46.54↑	3.74↓	30.78↓
B3		5.83	34.76	5.10↓	38.64↑	3.56↓	15.62↓

分析结论

Real-ESRGAN在两个评估指标中上升幅度较大，而waifu2x则是在两个指标中基本下降。



场景	推荐
人像	GFPGAN（追求清晰美化）/Real-ESRGAN(追求自然)
动画	Real-ESRGAN
自然	Bigjpg(追求细节自然)/Real-ESRGAN(追求整体观感和边缘清晰度)



5

# 与传统图像增强技术比较

Real-ESRGAN处理人  
物

双三次处理人物



未处理人物

双线性处理人物

拉普拉斯锐化处理人  
物



Real-ESRGAN处  
理自然场景

双三次处理自然  
场景



未处理自  
然场景

双线性处理自然  
场景

拉普拉斯锐化处理  
自然场景

Real-ESRGAN处理动  
画



双三次处理动  
画



未处理动  
画

双线性处理动  
画

拉普拉斯锐化处理动  
画



从数据来看，Real-ESRGAN、双三次插值、双线性插值的两个评判指标上升下降趋势几乎一致，拉普拉斯锐化因引入噪点使得两个指标普遍上升。

在主观视觉上，Real-ESRGAN的增强效果表现明显优于传统方法处理效果。

类型	图像编号	提取值	原图	拉普拉斯锐化	双三次插值	双线性插值	Real-ESRGAN
老照片修复	A1	NIQE	4.32	7.09 ↑	3.90 ↓	4.18 ↓	2.44 ↓
		BRISQUE	50.24	43.39 ↓	48.16 ↓	47.94 ↓	26.11 ↓
	A2	NIQE	2.79	7.02 ↑	4.10 ↑	3.88 ↑	3.46 ↑
		BRISQUE	29.63	43.46 ↑	35.87 ↑	37.90 ↑	38.98 ↑
	A3	NIQE	3.29	6.15 ↑	2.85 ↓	2.95 ↓	3.11 ↓
		BRISQUE	29.54	43.46 ↑	37.21 ↑	38.06 ↑	11.43 ↓
动漫图像处理	B1	NIQE	2.93	7.91 ↑	3.56 ↑	3.58 ↑	3.58 ↑
		BRISQUE	33.36	43.46 ↑	15.62 ↓	20.54 ↓	20.54 ↓
	B2	NIQE	3.96	6.02 ↑	3.56 ↓	3.88 ↓	6.03 ↑
		BRISQUE	34.03	22.81 ↓	15.62 ↓	37.90 ↑	46.54 ↑
	B3	NIQE	5.83	6.79 ↑	4.10 ↓	3.88 ↓	5.10 ↓
		BRISQUE	34.76	27.23 ↓	35.87 ↑	37.90 ↑	38.64 ↑
自然图像处理	C1	NIQE	3.29	8.59 ↑	2.85 ↓	2.95 ↓	3.91 ↑
		BRISQUE	32.32	43.46 ↑	37.21 ↑	38.06 ↑	22.37 ↓
	C2	NIQE	3.85	13.76 ↑	3.33 ↓	3.56 ↓	3.33 ↓
		BRISQUE	12.24	43.46 ↑	28.16 ↑	15.62 ↑	28.16 ↑
	C3	NIQE	2.32	9.94 ↑	3.10 ↑	3.56 ↑	3.10 ↑
		BRISQUE	13.1	43.46 ↑	31.14 ↑	15.62 ↑	31.14 ↑



## 使用场景总结

1

传统图像增强方法：操作简便、运算速度快、资源占用低，适合对图像质量要求不高或资源受限的场景（如嵌入式设备、实时监控）。

2

深度学习方法（Real-ESRGAN）：部署复杂、计算开销大，但细节恢复和纹理保真效果更优，适用于高画质要求的任务（如图像修复、内容创作、视觉AI）。



# 6

## 研究总结与反思





## 局限性

处理速度相对较慢，不适合实时视频处理任务；对部分模糊区域的还原仍有限；在硬件支持不充分时，程序稳定性和兼容性可能受到影响。

## 改进方向

针对上述问题，未来研究可在以下方向持续深化与优化：

- (1) 轻量化设计：探索更高效的网络结构以适应移动端与实时场景；
- (2) 自适应增强机制：提高模型对不同图像风格与退化程度的适应能力；



## 实验设计反思

### 1. 客观指标选取

- NIQE（自然图像质量评估）：数值越低越自然
- PSNR（峰值信噪比）：数值越高，表示图像越接近原图。
- SSIM（结构相似性指数）：取值范围为 0 ~ 1，越接近 1 表示越相似。
- BRISQUE（盲感知图像质量）：数值越低视觉质量越好

### 2. 实验样本数量有限



# Thanks

答辩人：吴抒航、杨宇涵、宋语婷