

# DENOISING DIFFUSION PROBABILISTIC MODELS FOR ROBUST IMAGE SUPER-RESOLUTION IN THE WILD

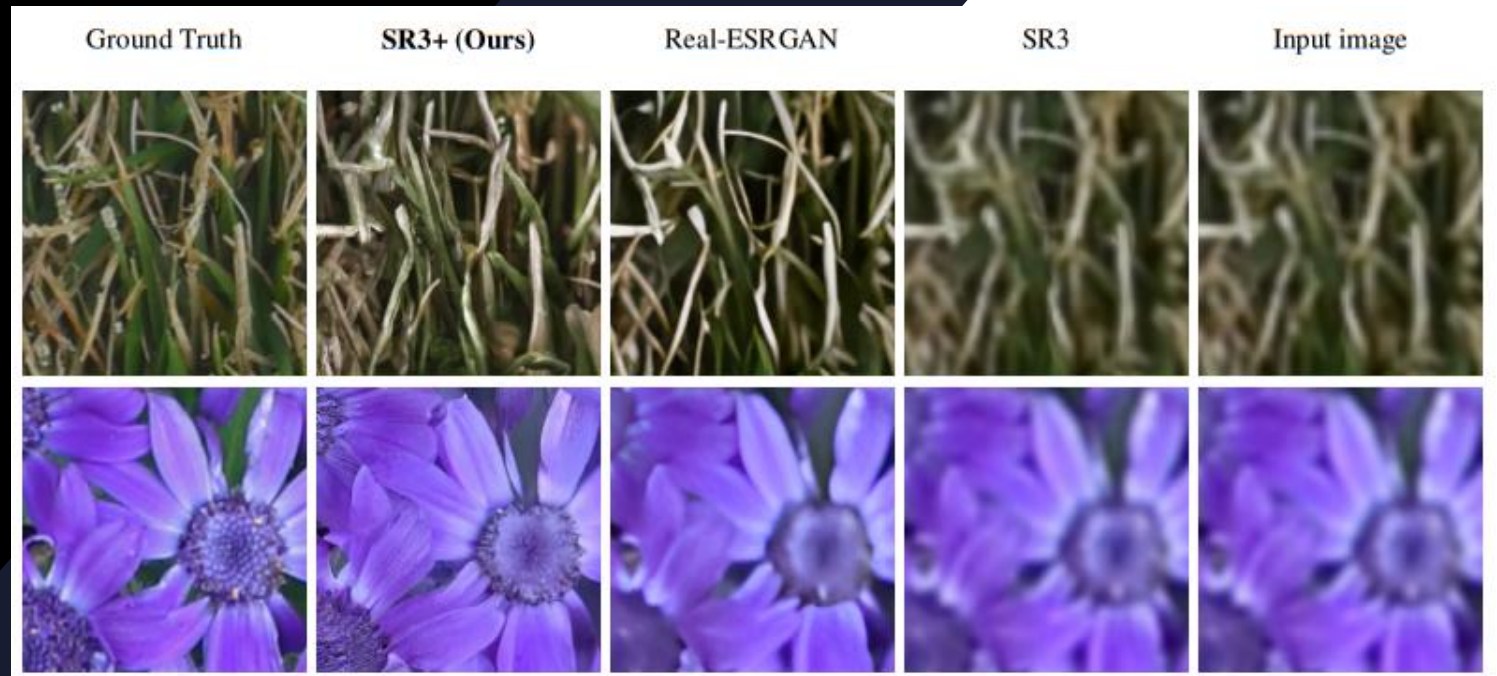
2023-11-29

Presenter: 傅于軒



# Introduction

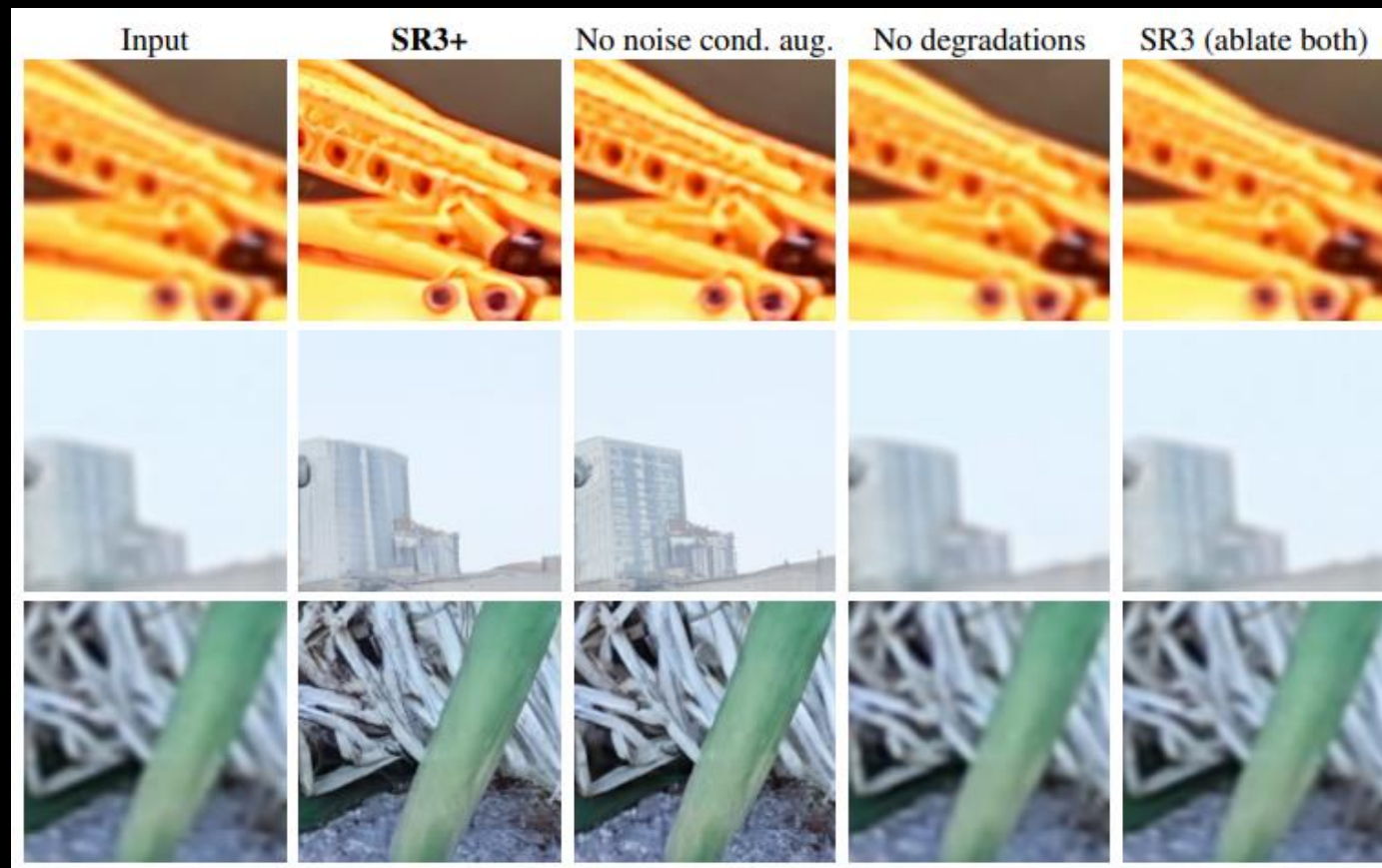
- Blind super-resolution
- SR3 falls short on out-of-distribution (OOD) data
- SR3+ achieve state-of-the-art results on OOD data



# Contributions

1. outperforming SR3 and the previous SOTA on zero-shot RealSR and DRealSR benchmarks
  2. 透過 Ablation study ，證明了參數退化和噪聲調節增強技術的互補優勢
  3. SR3+ 隨著模型大小和資料集的增加，其性能顯著提升
-

# Ablation samples



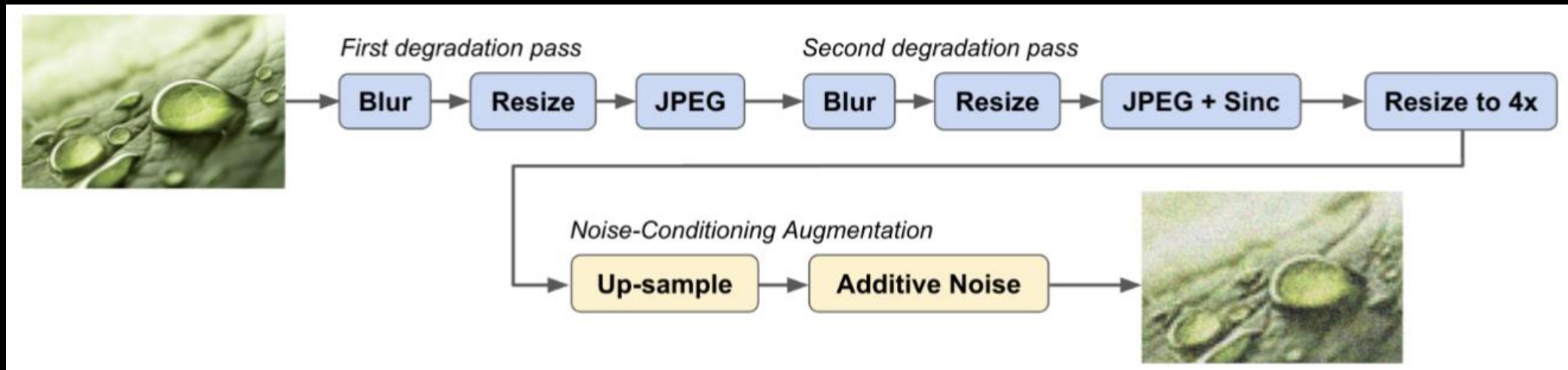
# Related Work

- Wang et al. demonstrates applying this degradation make LR closer wild image
  - Blind super resolution 的 degradation 包含 explicit(顯式) 和 implicit (隱式)
    - *implicit (隱式): need large datasets*
    - *explicit(顯式): 退化在訓練期間直接合併為資料增強*
  - Diffusion model 能較好的捕捉視覺世界的統計數據，比GAN更多樣性
-

# Methodology

- Higher-order degradations
    - Higher-order degradations 的 Self-Supervised Learning 對 HR 圖像進行採樣，並獲得相對應的 LR 輸入
  - Noise Conditioning Augmentation
-

# Degradations & Noise Conditioning



# Higher-order degradations

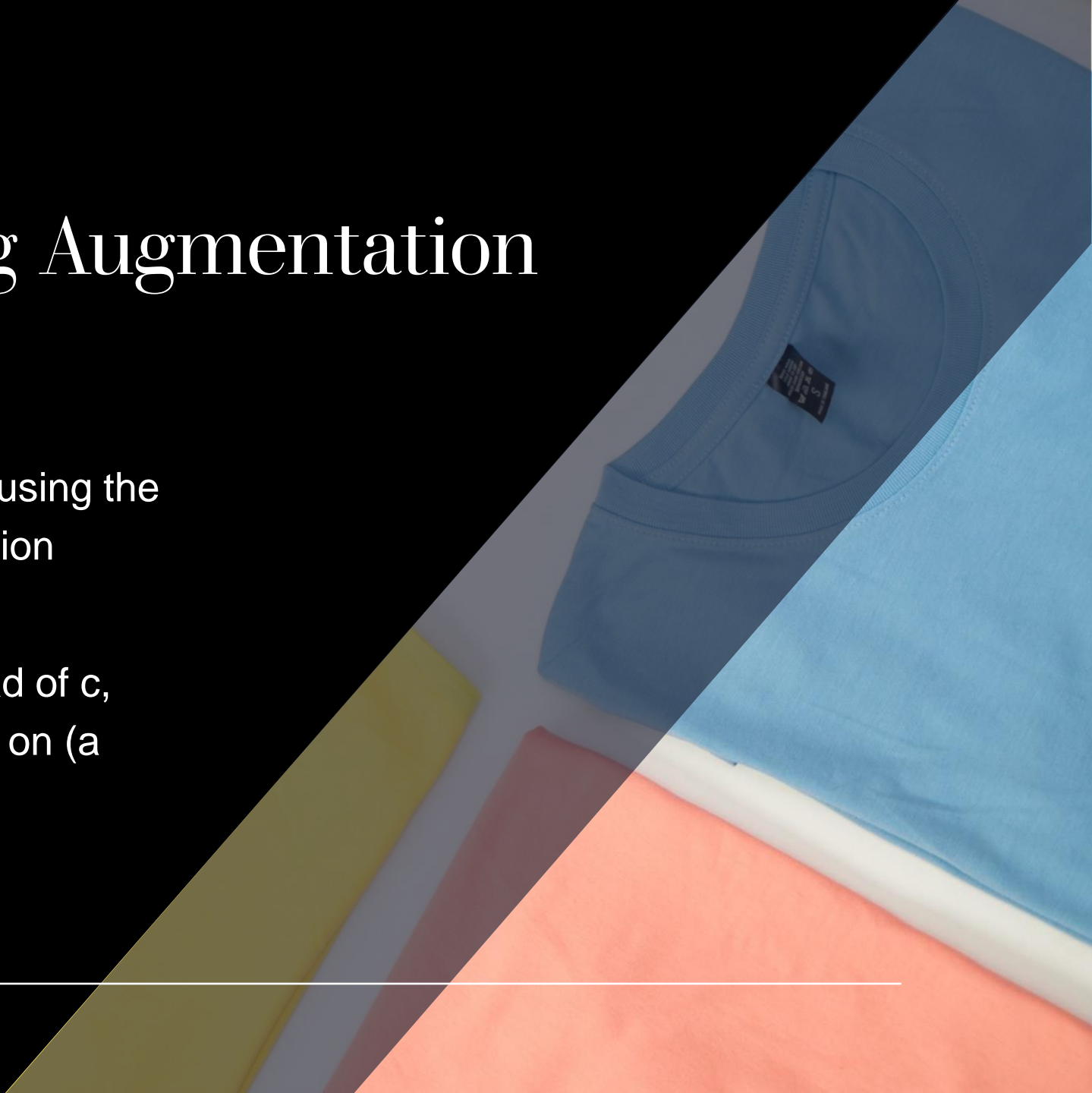
1. Self-supervision 會對 HR 進行 downsample 得到 LR Image
2. Data augmentation 的重複順序，對OOD的 data 泛化有重大影響



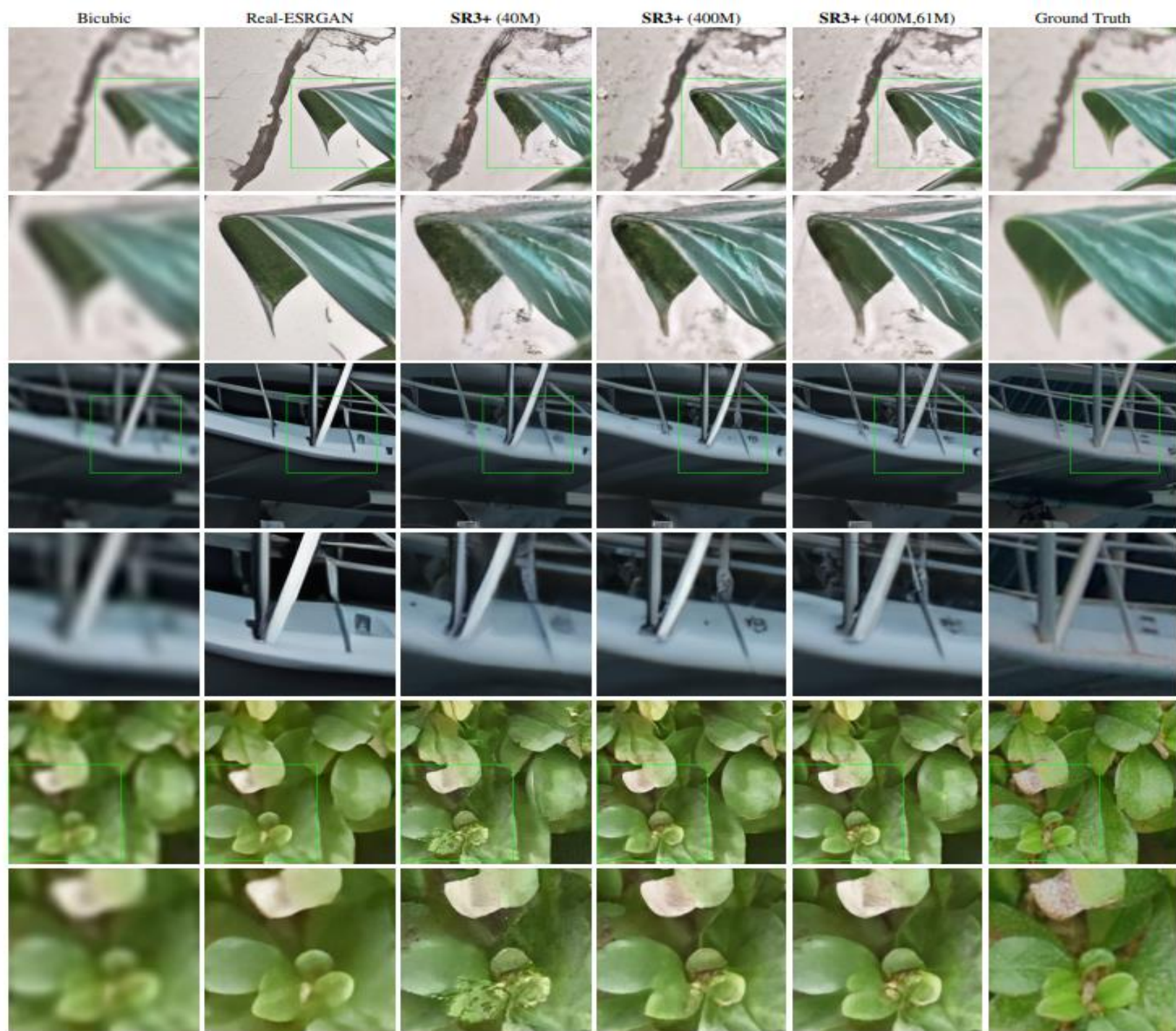


# Noise Conditioning Augmentation

1. Sample  $\tau \sim \text{Uniform}(0, \tau_{\max})$ .
2. Add noise to get  $c_{\tau} \sim q(z_{\tau} | c)$ , reusing the marginal distribution of the diffusion forward process.
3. Condition the model on  $c_{\tau}$  instead of  $c$ , and we also condition the model on (a positional embedding of)  $\tau$



# Comparison



# Experiments

SR Model (Parameter Count, Dataset)	FID(10k) ↓		PSNR ↑		SSIM ↑	
	RealSR	DRealSR	RealSR	DRealSR	RealSR	DRealSR
Real-ESRGAN	34.21	37.22	<b>25.14</b>	<b>25.85</b>	<b>0.7279</b>	<b>0.7808</b>
SR3+ (40M, DF2K + OST)	<u>31.97</u>	40.26	24.84	25.18	0.6827	0.7201
SR3+ (400M, DF2K + OST)	<u>27.34</u>	<u>36.28</u>	23.84	24.36	0.662	0.719
SR3+ (400M, 61M Dataset)	<b>24.32</b>	<b>32.37</b>	24.89	25.74	0.6922	0.7547

# Experiments

SR Model (400M parameters, 61M Dataset)	FID(10k) ↓		PSNR↑		SSIM↑	
	RealSR	DRealSR	RealSR	DRealSR	RealSR	DRealSR
SR3+	<b>24.32</b>	<b>32.37</b>	24.89	25.74	0.6922	0.7547
SR3+ (no noise cond. aug.)	34.20	49.93	22.34	22.28	0.6469	0.6994
SR3+ (no degradations)	36.93	44.18	<u>25.00</u>	<u>26.22</u>	0.6824	<u>0.7687</u>
SR3 (i.e., ablating both)	85.77	93.05	<b>27.89</b>	<b>28.25</b>	<b>0.784</b>	<b>0.83</b>

# Conclusion

- 透過高階降級方案和雜訊調節增強，相較於其他模型，FID分數較高
  - 在大資料集更有效率
  - 在資料充足情況下，能處理有文字的圖像
-

Thank you for listening !

---