Explicit MAP Optimization for Photo-Realistic Super Resolution

Nick Huang Dustin Wu Shijie Mao Rohit Mohanty

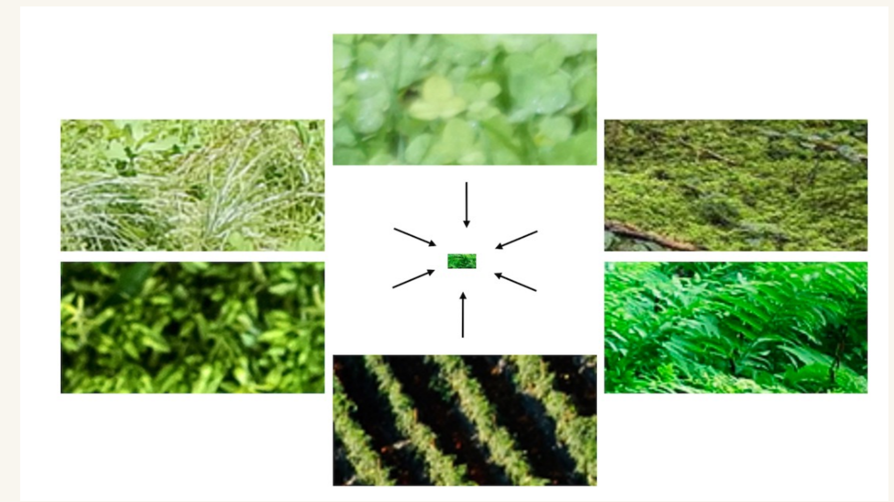
BROWN

Motivation

Traditional super resolution techniques lead to **visually unpleasant** reconstructions, can we do better?

Problem Setting

We attribute the visually unpleasant reconstruction to the over-constrained optimization objective which admits only one of infinitely many possible solutions.



Contribution

1. We identify that the pixelwise loss used in traditional super resolution methods is directly at odds with allowing multiple feasible solutions.
2. We relax the aforementioned objective by explicitly modeling the MAP image restoration objective.
3. We model the regularization term in the MAP framework with the relativistic adversarial loss [1].
4. We adopt noise injection [2, 3] to explore the super resolution space [4].

Method

We optimize the super resolution network G using the following objective:

$$L_G = \alpha L_{\text{Fidelity}} + L^G_{\text{Adversarial}}$$

where in our experiment setting:

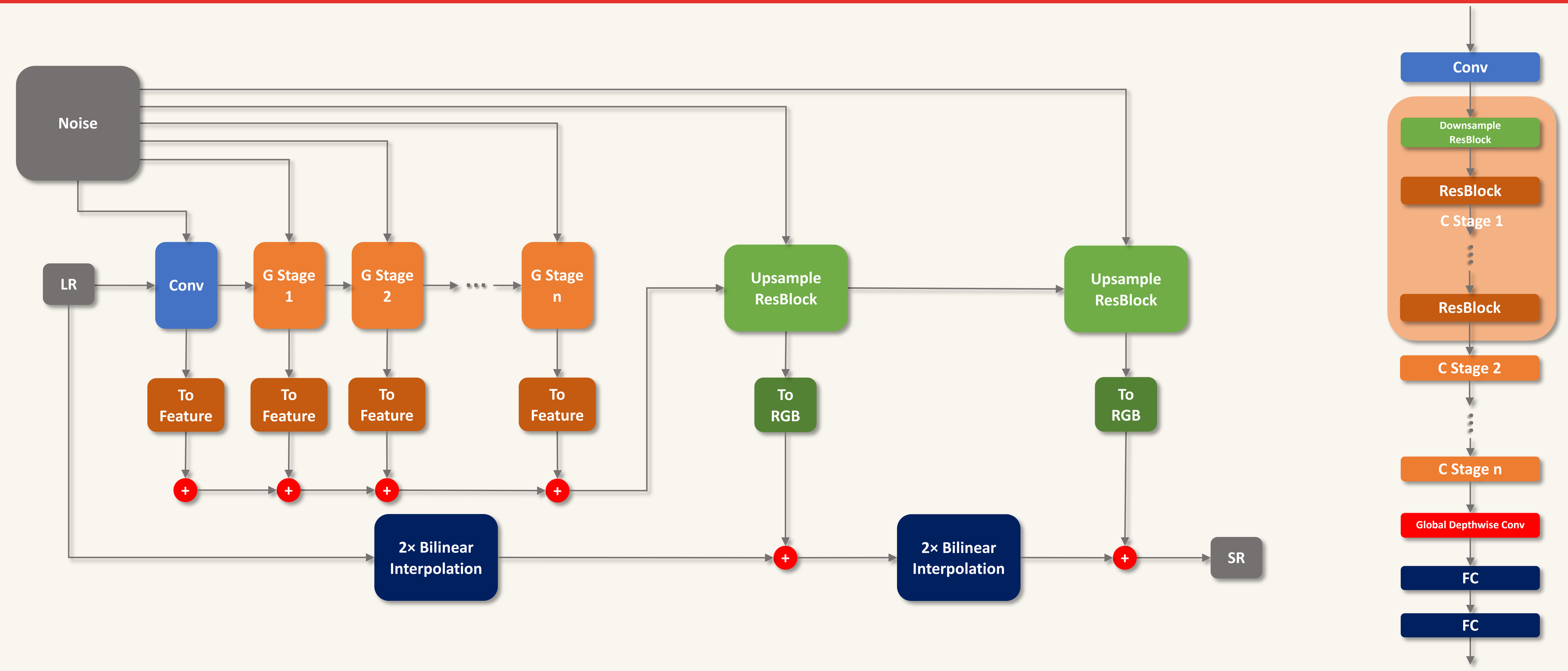
$$\alpha = 10 \quad L_{\text{Fidelity}} = \|y - G(y)_{\downarrow\text{bicubic}}\|_1 \quad L^G_{\text{Adversarial}} = \sup_{G:\mathcal{Y}\to\mathcal{X}} \mathbb{E}_{\substack{x\sim\text{HR}\\ y\sim\text{LR}}}[f(C(G(y)) - C(x))]$$

The critic network C is adversarially trained with R_1 gradient penalty [5]:

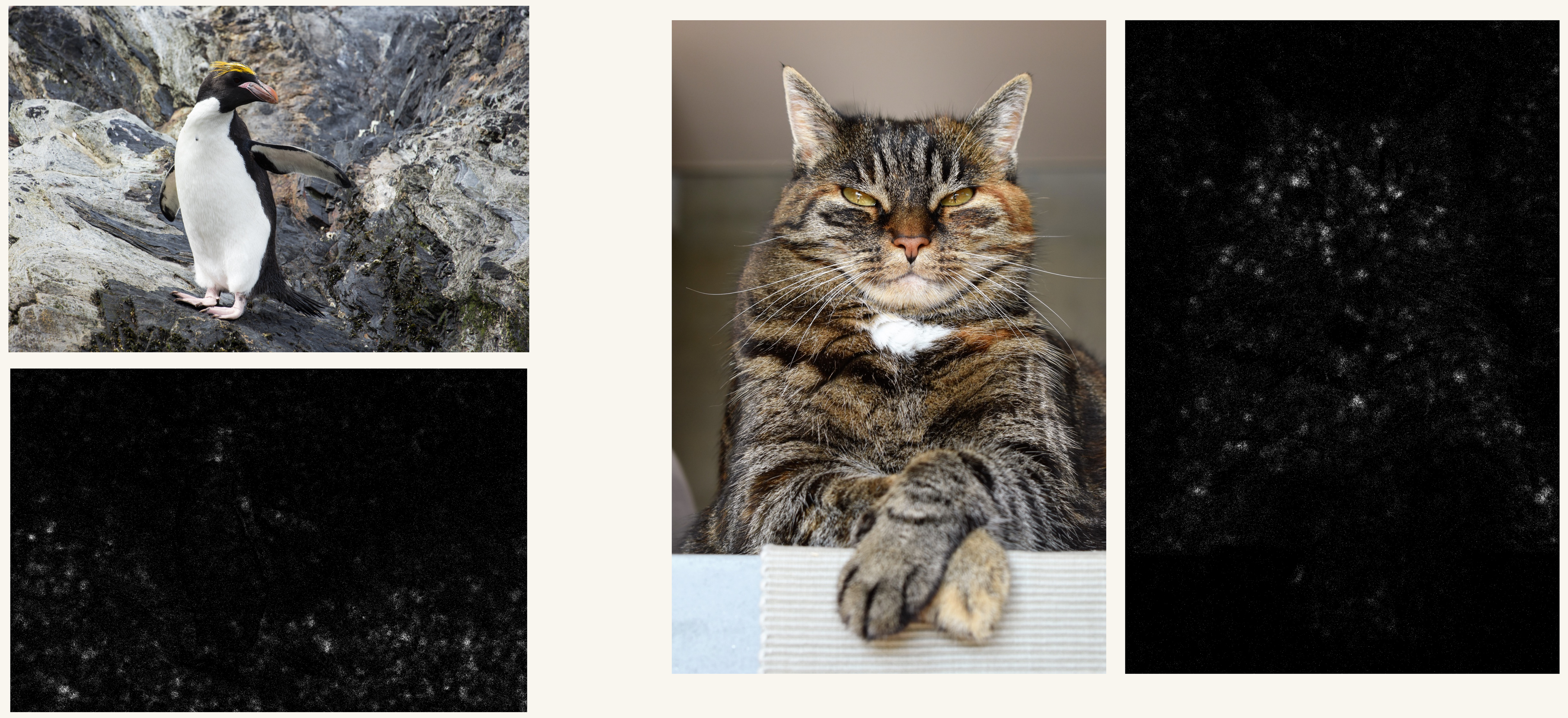
$$L_C = \sup_{C:\mathcal{X}\to\mathbb{R}} \mathbb{E}_{\substack{x\sim\text{HR}\\ y\sim\text{LR}}}[f(C(x) - C(G(y)))] + \frac{\gamma}{2}\mathbb{E}_{x\sim\text{HR}}\left[\|\nabla_x C(x)\|^2\right]$$

It is proved in [6] that $L^G_{\text{Adversarial}}$ and $L^C_{\text{Adversarial}}$ define a divergence, thus our objectives push G to produce results which lie on the natural image manifold.

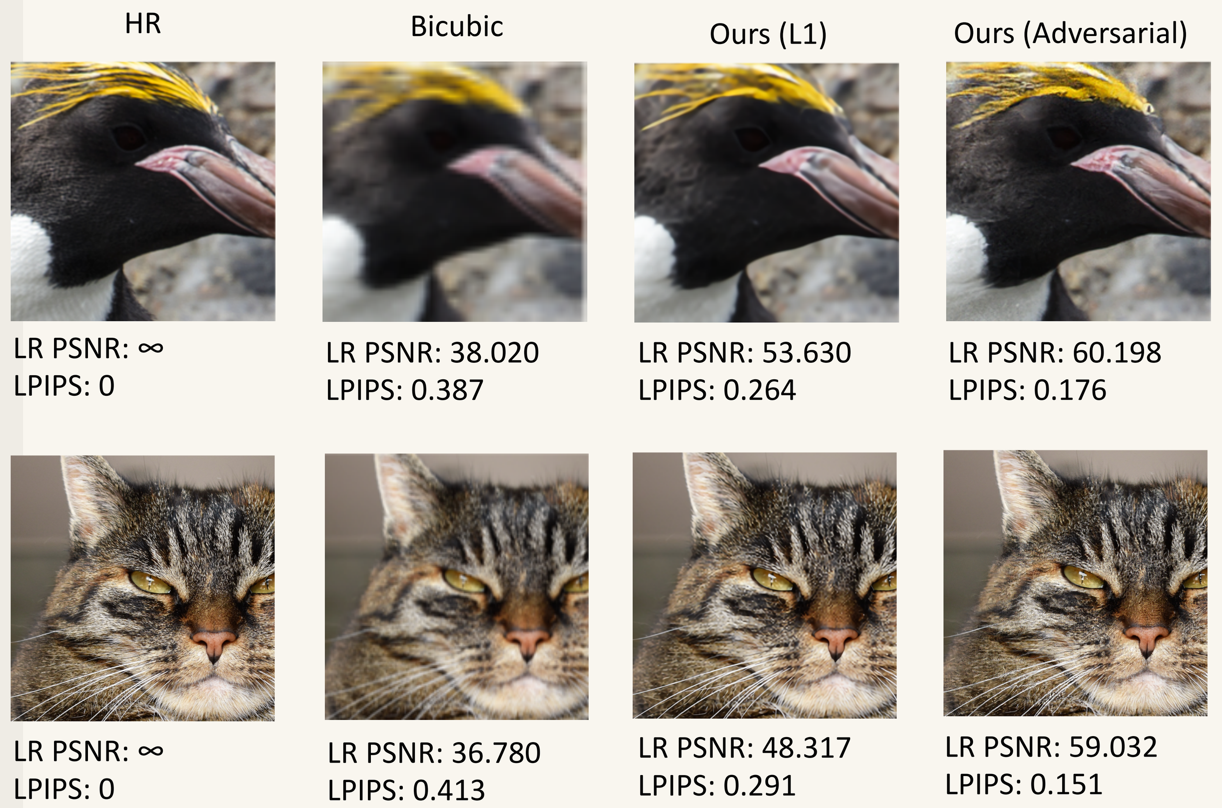
Model Architecture

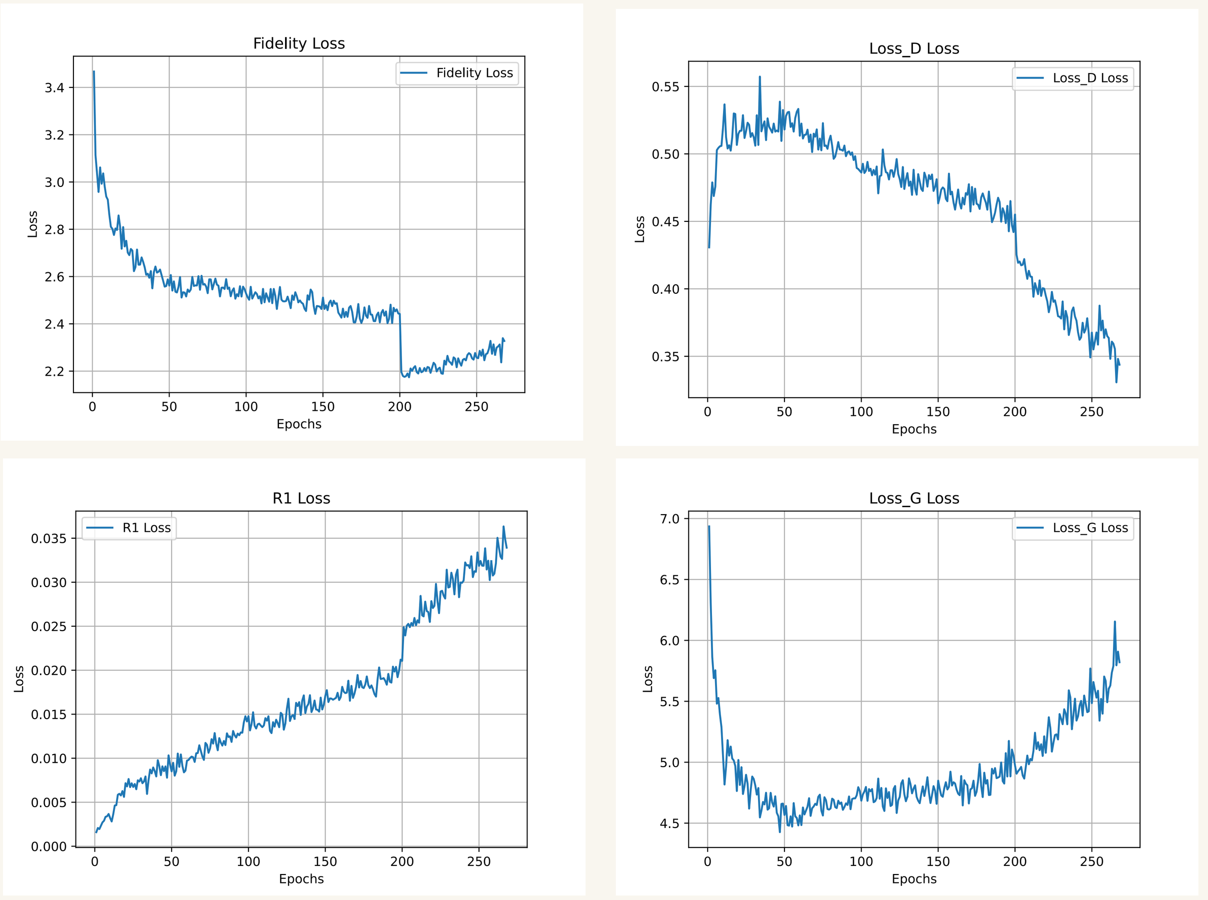


Reconstruction Stochasticity



Experiments





References

[1] Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard gan. In International Conference on Learning Representations, 2018.
[2] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 4401–4410, 2019.
[3] Ruili Feng, Deli Zhao, and Zheng-Jun Zha. Understanding noise injection in gans. In Marina Meila and Tong Zhang, editors, Proceedings of the 38th International Conference on Machine Learning, volume 139 of Proceedings of Machine Learning Research, pages 3284–3293. PMLR, 18–24 Jul 2021.
[4] Andreas Lugmayr, Martin Danelljan, and Radu Timofte. Ntire 2021 learning the super-resolution space challenge. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pages 596–612, June 2021.
[5] Lars Mescheder, Andreas Geiger, and Sebastian Nowozin. Which training methods for gans do actually converge? In International conference on machine learning, pages 3481–3490. PMLR, 2018.
[6] Alexia Jolicoeur-Martineau. On relativistic f-divergences. In International Conference on Machine Learning, pages 4931–4939. PMLR, 2020.

Acknowledgements

We would like to thank Professor Tompkin for the helpful discussion :D