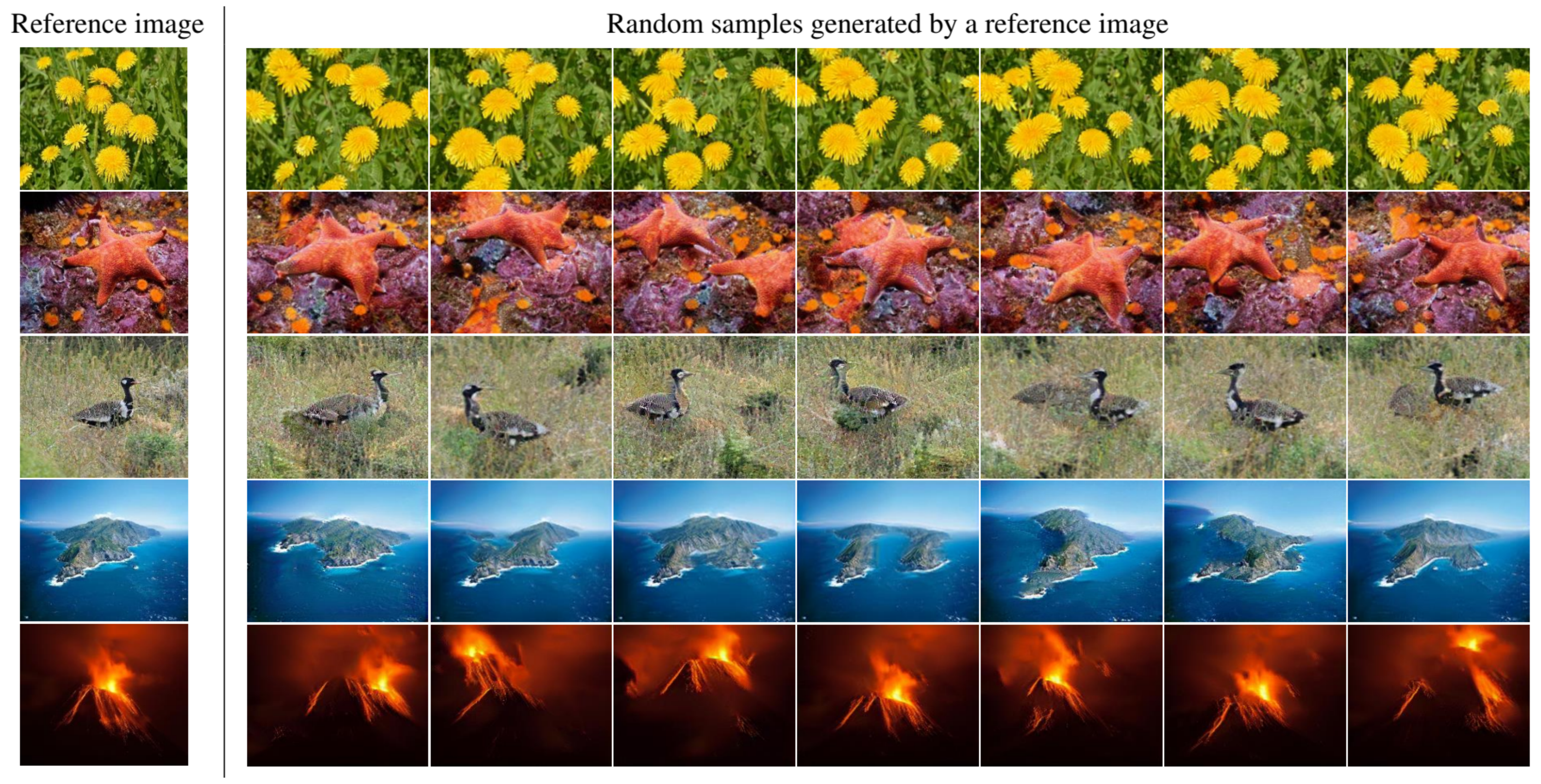
IMAGINE: Image Synthesis by Image-Guided Model Inversion

Pei Wang*, Yijun Li# , Krishna Kumar Singh#, Jingwan Lu#, Nuno Vasconcelos*

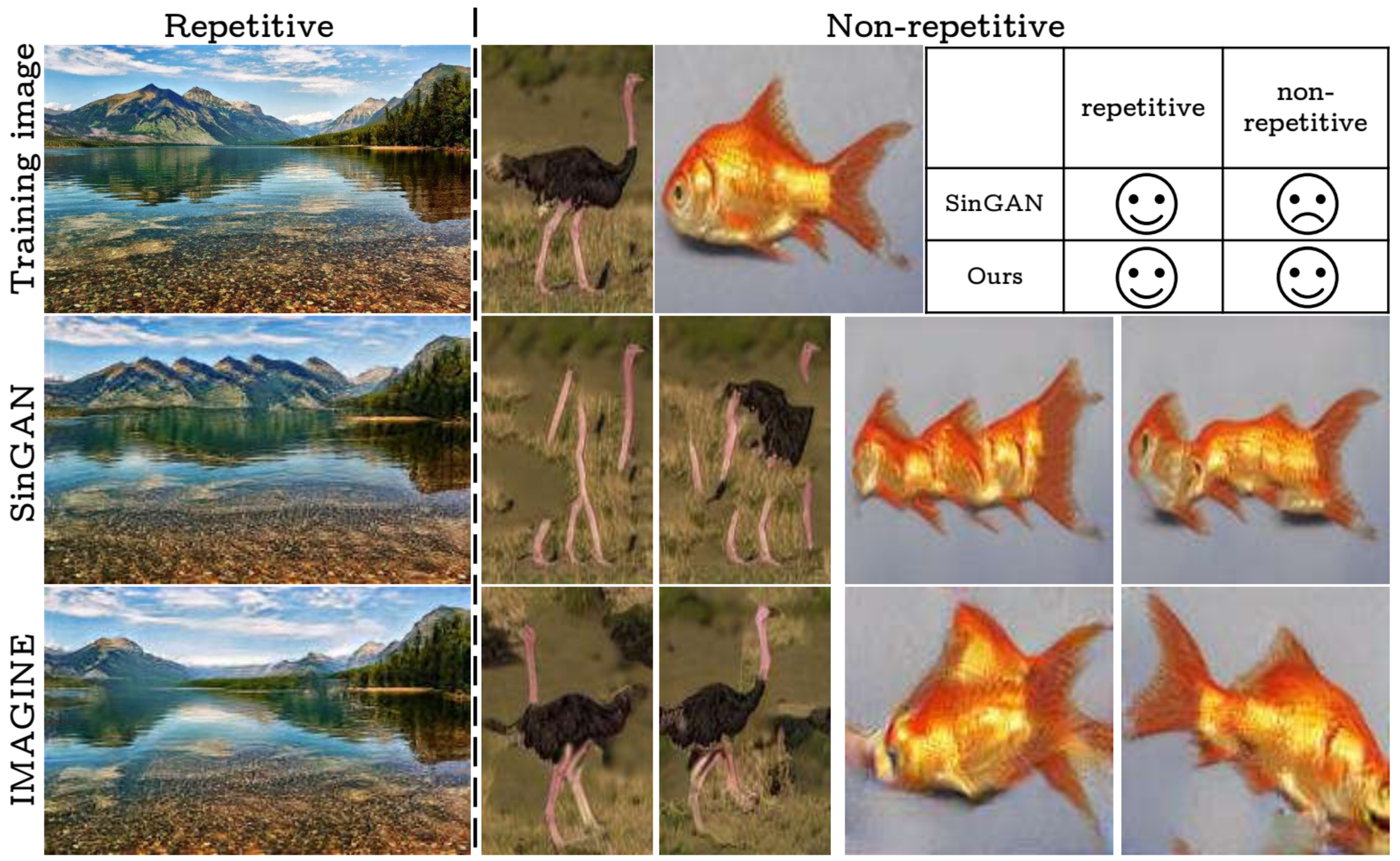
*SVCL, UC San Diego; #Adobe Research

Introduction

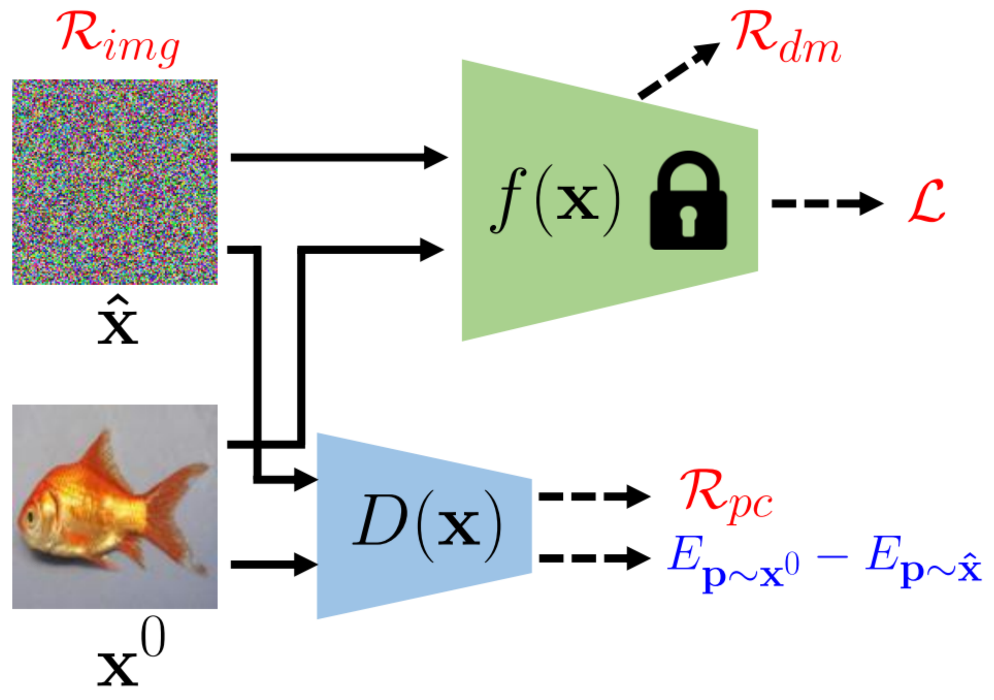
- We consider the problem of generating images from one single reference.



- Previous work mainly use GAN inversion, e.g., DGP or learning individual GANs for individual target images, e.g., SinGAN
  - Training of generators is not easy
  - One generator per reference image
  - Cannot handle non-repetitive images



Proposed Method



- Given a well-trained frozen classifier, an image is optimized from a noise under a couple of constraints.

$$\mathbf{x}^* = \arg\min_{\hat{\mathbf{x}}} \mathcal{L}(f(\hat{\mathbf{x}}), y^*) + \mathcal{R}(\hat{\mathbf{x}})$$

$$\mathcal{R}(\hat{\mathbf{x}}; \mathbf{x}^0, \Phi) = \mathcal{R}_{img}(\hat{\mathbf{x}}) + \lambda \mathcal{R}_{dm}(\hat{\mathbf{x}}; \mathbf{x}^0, \Phi) + \gamma \mathcal{R}_{pc}(\hat{\mathbf{x}})$$

- Basic regularizer

$$\mathcal{R}_{img}(\hat{\mathbf{x}}) = \alpha R_{TV}(\hat{\mathbf{x}}) + \beta ||\hat{\mathbf{x}}||^2$$

- Feature distribution matching regularizer

$$\mathcal{R}_{dm}(\hat{\mathbf{x}}; \mathbf{x}^0, \Phi) = \sum_{l\in\Phi} ||\mu_l(\hat{\mathbf{x}}) - \mu_l(\mathbf{x}^0)||_2 + ||\sigma_l(\hat{\mathbf{x}}) - \sigma_l(\mathbf{x}^0)||_2$$

- Adversarial loss

$$\mathcal{R}_{pc}(\hat{\mathbf{x}}) = -D(\hat{\mathbf{x}})$$

$$D(\hat{\mathbf{x}}) = E_{\mathbf{p}\sim\hat{\mathbf{x}}}[d^*(\mathbf{p})] \qquad d^* = \arg\max_{d} E_{\mathbf{p}\sim\mathbf{x}^0}[d(\mathbf{p})] - E_{\mathbf{p}\sim\hat{\mathbf{x}}}[d(\mathbf{p})]$$

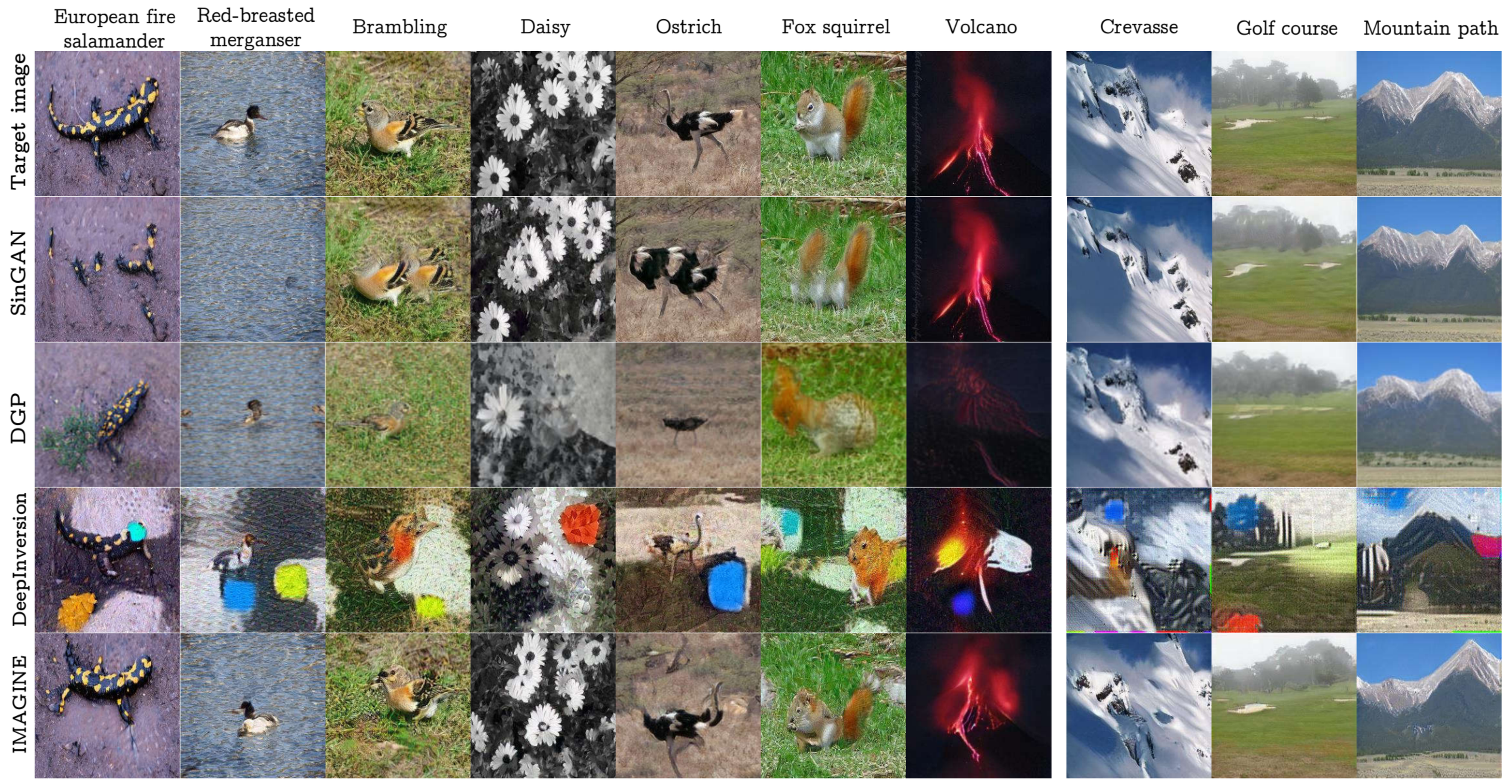
Reference

[1] Xingang Pan, Xiaohang Zhan, Bo Dai, Dahua Lin, Chen Change Loy, and Ping Luo. Exploiting deep generative prior for versatile image restoration and manipulation, ECCV2020.

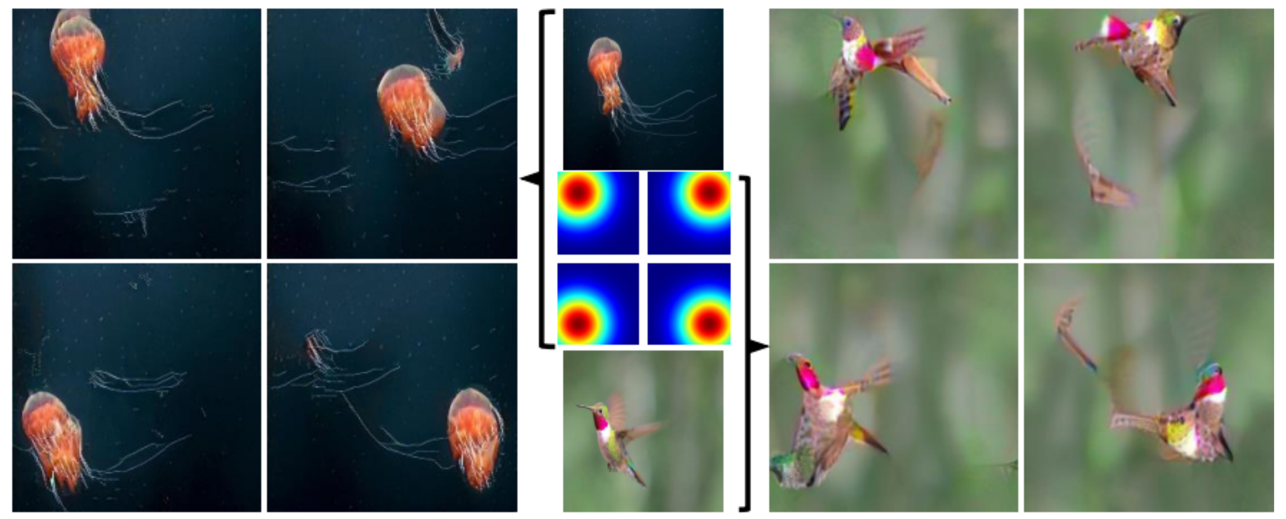
[2] Tamar Rott Shaham, Tali Dekel, and Tomer Michaeli. Singan: Learning a generative model from a single natural image, ICCV2019.

[3] Hongxu Yin, Pavlo Molchanov, Jose M Alvarez, Zhizhong Li, Arun Mallya, Derek Hoiem, Niraj K Jha, and Jan Kautz. Dreaming to distill: Data-free knowledge transfer via deepinversion, CVPR2020.

Image Synthesis



Position Control



Shape Control

Conclusion

- First attempt to utilize model inversion for image-guided synthesis
- Introduce adversarial training to improve the image quality of optimization-based model inversion
- Demonstrate the superiority of IMAGINE over GAN counterparts
- Show multiple multifaceted image control applications