

Supplementary material for It's All About *Your* Sketch: Democratising Sketch Control in Diffusion Models

Subhadeep Koley^{1,2} Ayan Kumar Bhunia¹ Deeptanshu Sekhri¹ Aneeshan Sain¹
Pinaki Nath Chowdhury¹ Tao Xiang^{1,2} Yi-Zhe Song^{1,2}

¹SketchX, CVSSP, University of Surrey, United Kingdom.

²iFlyTek-Surrey Joint Research Centre on Artificial Intelligence.

{s.koley, a.bhunias, d.sekhri, a.sain, p.chowdhury, t.xiang, y.song}@surrey.ac.uk

A. Additional Qualitative Results

Fig. 1 delineates qualitative comparison of our method with pix2pix [1], CycleGAN [6], ControlNet [5], T2I-Adapter [2], and PITI [4]. Whereas, Fig. 2-7 shows additional results generated by our method.

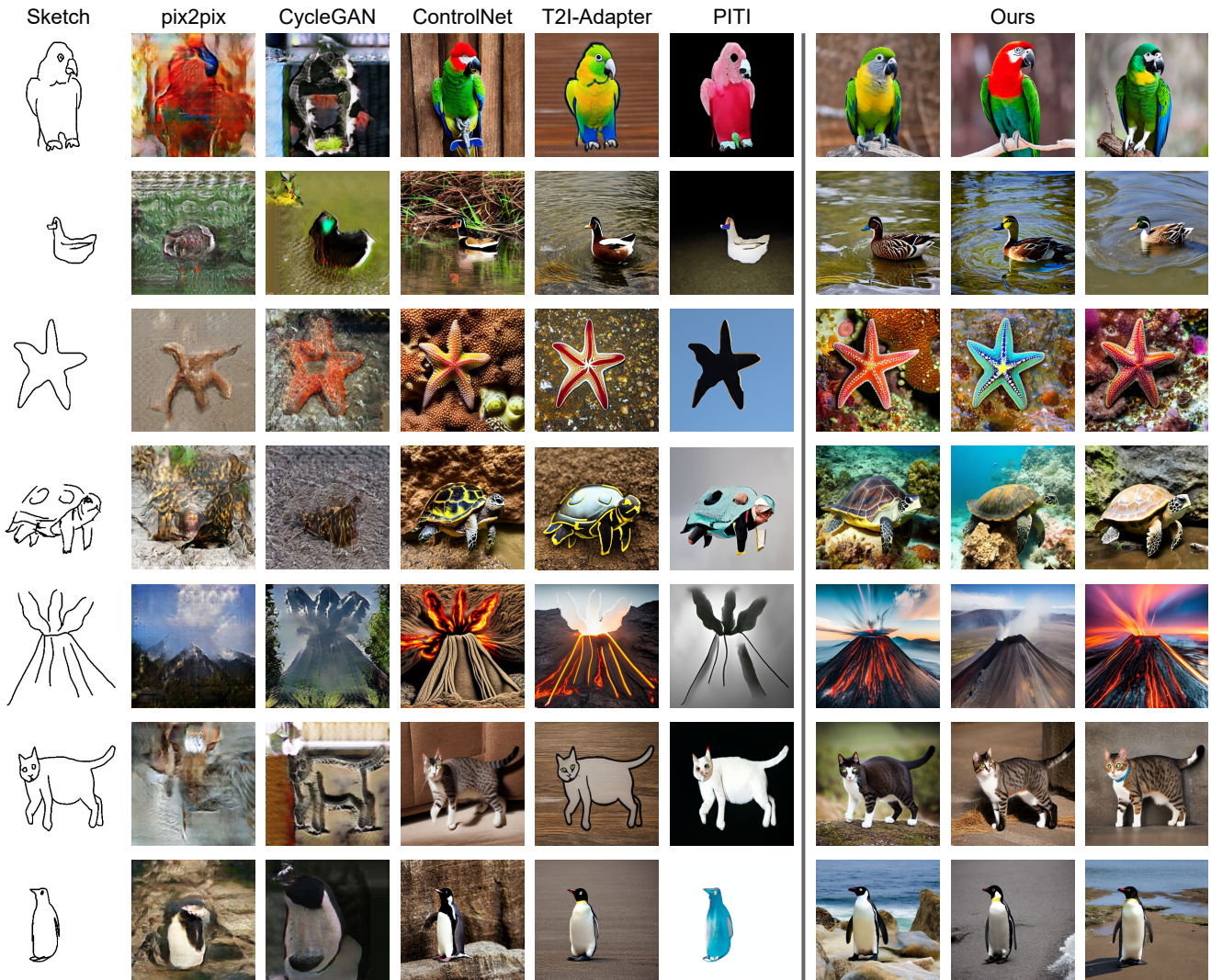


Figure 1. Qualitative comparison with SoTAs. For ControlNet [5], T2I-Adapter [2], and PITI [4], we use the fixed prompt “a photo of [CLASS]”, with [CLASS] replaced with corresponding class-labels of the input sketches. (Best view when zoomed in.)



Figure 2. Images generated by our method.

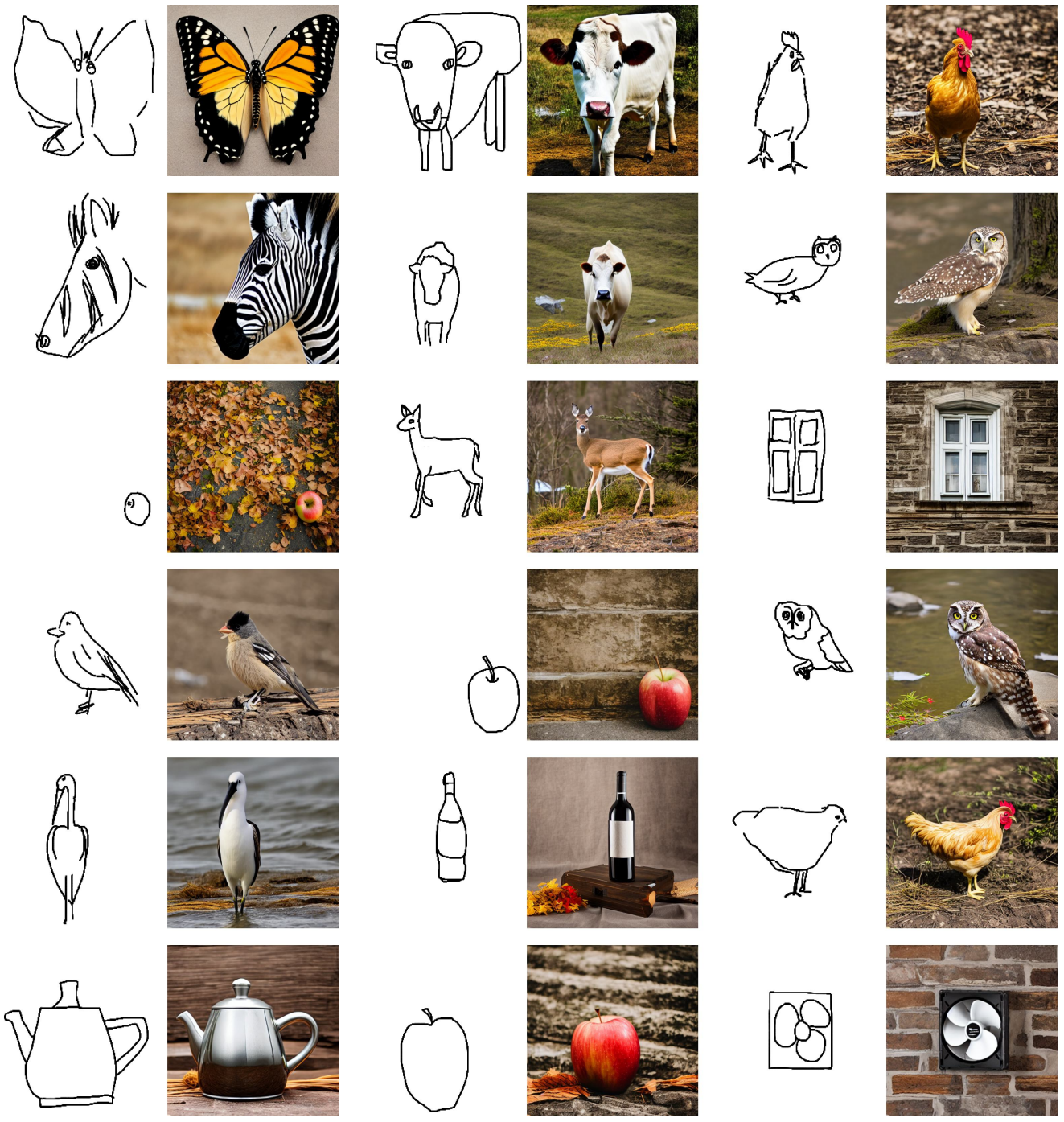


Figure 3. Images generated by our method.



Figure 4. Images generated by our method.



Figure 5. Images generated by our method.

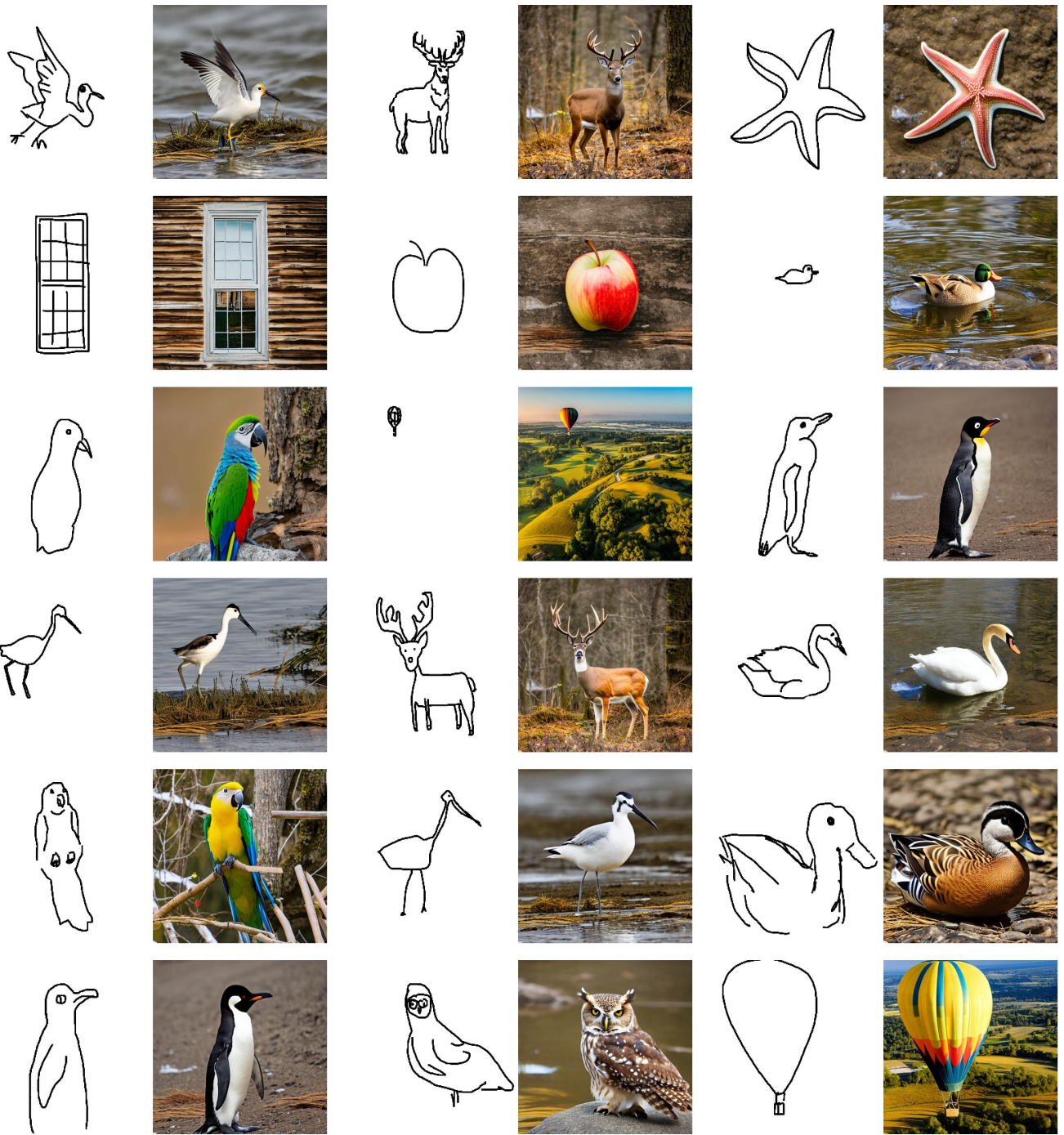


Figure 6. Images generated by our method.



Figure 7. Images generated by our method.

B. Results on Out-of-distribution Sketches

Keeping the pre-trained diffusion model *frozen*, we fully leverage its generalisation potential. We posit that our design enables out-of-distribution generalisation. Fig. 8 shows a few sketches, absent in Sketchy [3].



Figure 8. Results on out-of-distribution sketches.

References

- [1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-Image Translation with Conditional Adversarial Networks. In *CVPR*, 2017. 1
- [2] Chong Mou, Xintao Wang, Liangbin Xie, Jian Zhang, Zhongang Qi, Ying Shan, and Xiao Hu Qie. T2I-Adapter: Learning Adapters to Dig out More Controllable Ability for Text-to-Image Diffusion Models. *arXiv preprint arXiv:2302.08453*, 2023. 1
- [3] Patsorn Sangkloy, Nathan Burnell, Cusuh Ham, and James Hays. The sketchy database: learning to retrieve badly drawn bunnies. *ACM TOG*, 2016. 8
- [4] Tengfei Wang, Ting Zhang, Bo Zhang, Hao Ouyang, Dong Chen, Qifeng Chen, and Fang Wen. Pretraining is All You Need for Image-to-Image Translation. *arXiv preprint arXiv:2205.12952*, 2022. 1
- [5] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding Conditional Control to Text-to-Image Diffusion Models. In *ICCV*, 2023. 1
- [6] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In *ICCV*, 2017. 1