

Supplementary material for *DemoCaricature: Democratising Caricature Generation with a Rough Sketch*

Dar-Yen Chen Ayan Kumar Bhunia Subhadeep Koley Aneeshan Sain
Pinaki Nath Chowdhury Yi-Zhe Song
SketchX, CVSSP, University of Surrey, United Kingdom.

{d.chen, a.bhunia, s.koley, a.sain, p.chowdhury, y.song}@surrey.ac.uk

<https://democaricature.github.io>

A. Details on Random Mask Reconstruction

We present further implementation details on the random mask reconstruction (RMR) as shown in Fig. 1. We create the masked image x_0^m by randomly occluding several patches with different ratios to simulate a caricature having local variation. Besides the random occlusion, we also apply id- and style-specific masks on M to isolate regions of interest when calculating loss using Eq. (5). Specifically, for identity finetuning, M contains a binary mask over the background, making the model capture the distinguishing facial features exclusively. For style reference, M adopts a small value (0.2 in this work) on the face area, subtly nudging the model to infuse the stylistic elements from both the background and the face into caricatures.

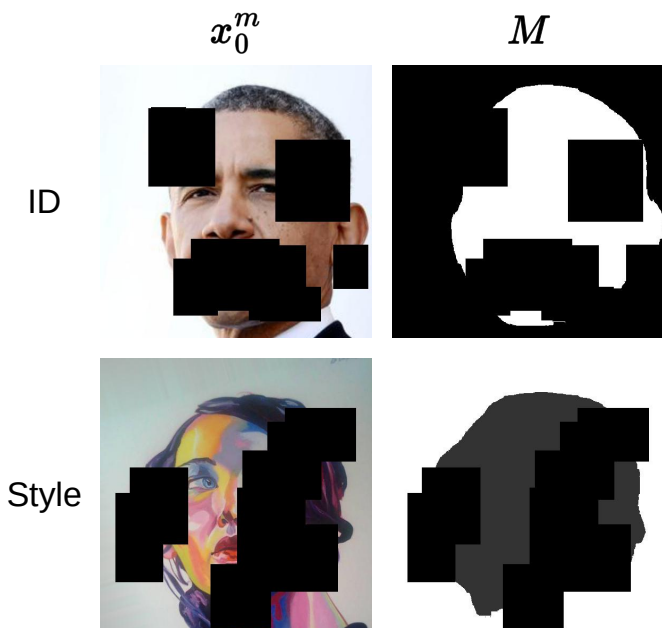


Figure 1. **Random Mask Reconstruction.** x_0^m mimics an image with local variation, a critical feature of caricature. M makes the objective function focus on the region of interest and ignore the occluded area.

B. More Qualitative Results

In this section, we provide additional qualitative results across various fields. Fig. 2 showcases more caricatures of celebrities. Beyond celebrities, we exhibit our method’s capability of learning identity and style from artistic and synthetic¹ portraits in Fig. 3 and Fig. 4 respectively. The results demonstrate the versatility of our democratising caricature generation, allowing users to flexibly and artistically create caricatures with the desired identities and styles.

¹We collect the synthetic id image from Ruiz *et al.* [1] and synthetic style reference from <https://civitai.com>.



Figure 2. Results on celebrities.

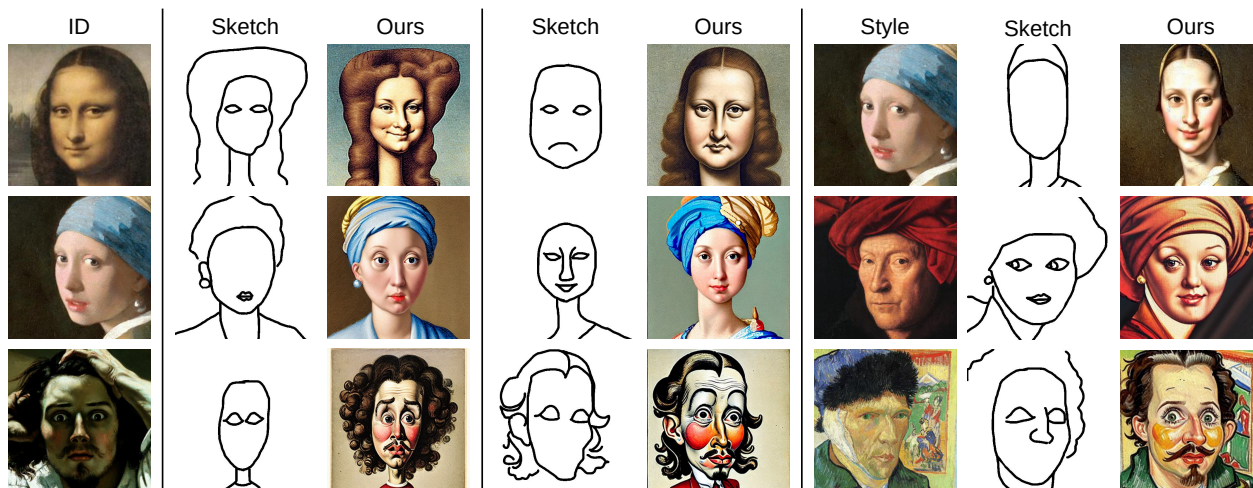


Figure 3. Results on famous artworks.

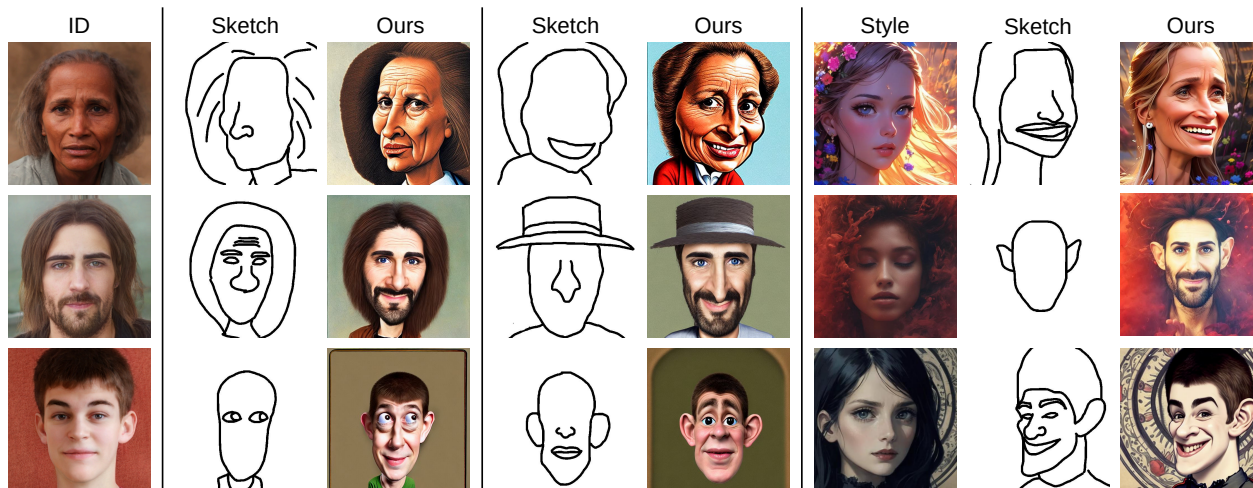


Figure 4. Results on synthetic human faces.

References

- [1] Nataniel Ruiz, Yuanzhen Li, Varun Jampani, Wei Wei, Tingbo Hou, Yael Pritch, Neal Wadhwa, Michael Rubinstein, and Kfir Aberman. HyperDreamBooth: HyperNetworks for Fast Personalization of Text-to-Image Models, 2023. 1