GANs Spatial Control via Inference-Time Adaptive Normalization Supplementary Materials

Karin Jakoel* Liron Efraim* Tamar Rott Shaham
Technion - Israel Institute of Technology

We present additional results for the applications presented in the paper. Figure 1 presents additional results for the task of local generation. Figure 2 shows the effect of choosing different StyleGAN [1] scales to participate in this task. Figures 3 4 present additional results for the tasks of semantic attributes transfer and class hybridization respectively.

References

[1] 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. 1, 4

*These authors contributed equally to this work.

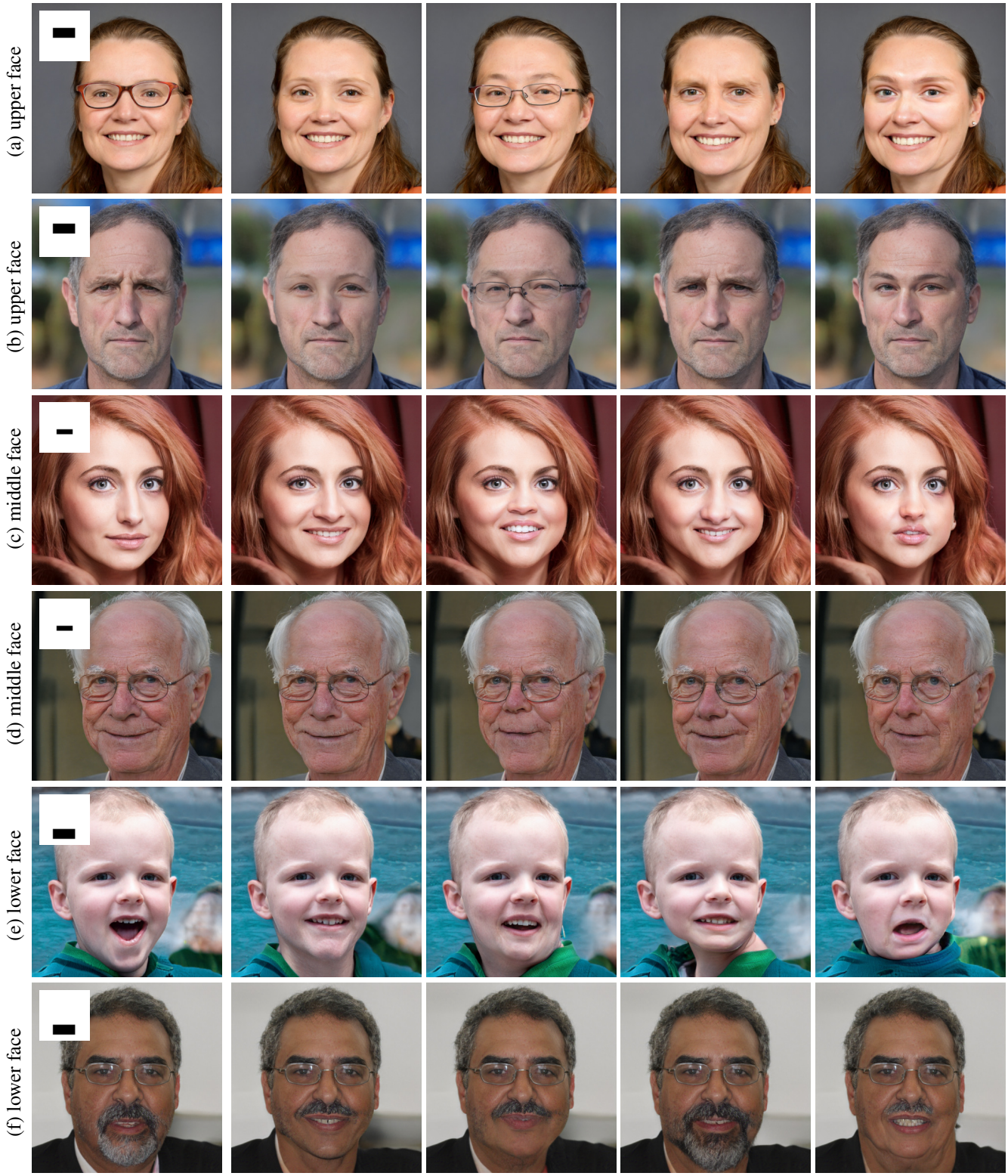

Figure 1: **Local generation- Additional result.** Our approach enables to randomly resample only specific regions of an image. We first generate an initial image (leftmost column) using StyleGAN and then randomly re-generate only a specific part of the face indicated by the mask (left column, upper left corner) using our inference-time adaptive normalization. By doing so we are able to create new image content inside the mask, while keeping the rest of the image intact.

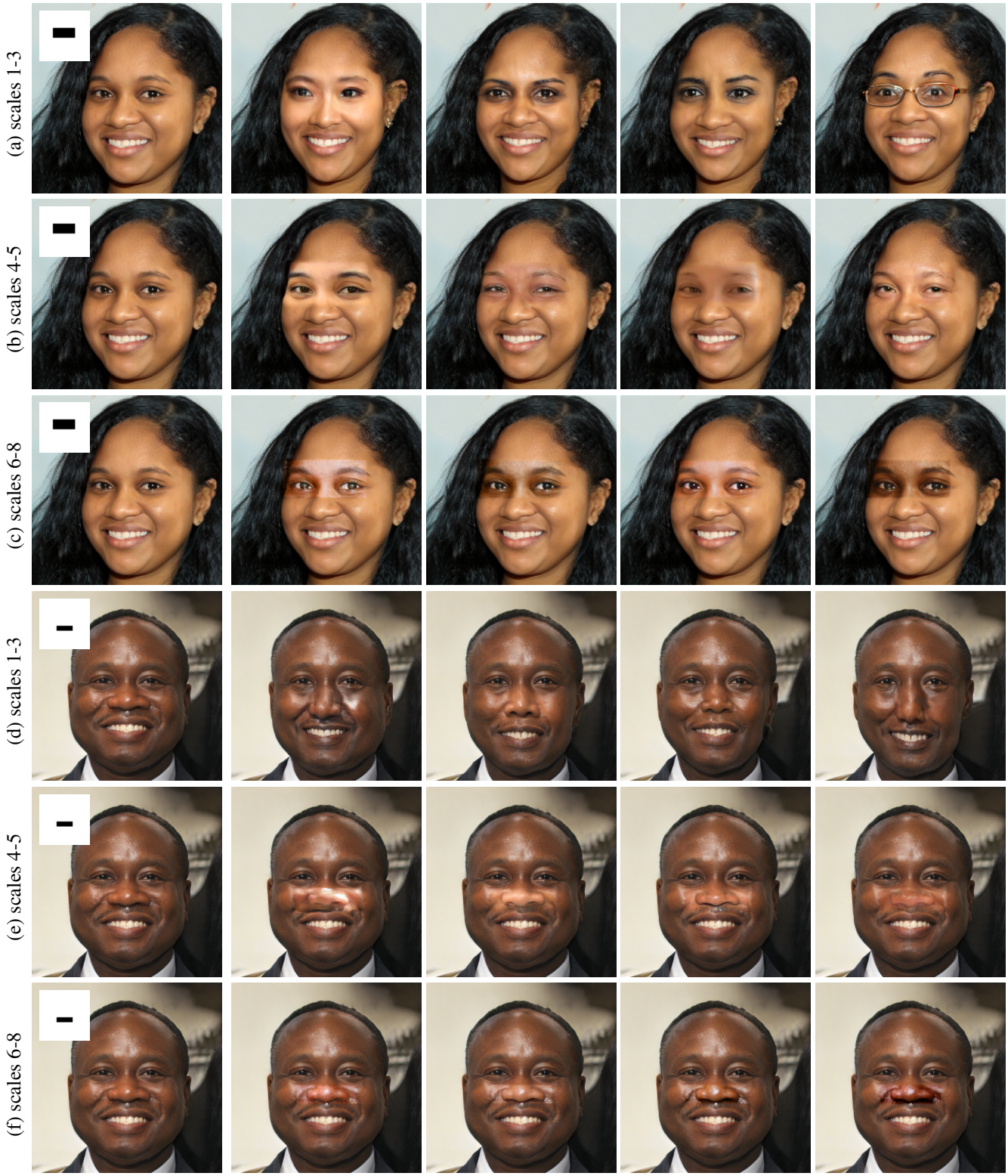

Figure 2: **Local generation- Scale selection effect.** When choosing relatively coarse scales (e.g. scales 1-3) to be adapted with ITAN, we generate new coarse structures that are blended smoothly with the initial generation (leftmost column). However when applying this to finer model scales (e.g. 4-5, 6-8), we generate new finer details like skin tone and lighting are regenerated in a way that does not necessarily fits the original image anymore. We find empirically that choosing scales 1-3 works best in terms of both generating new image content while coherently blending the image. This is also true for the tasks of semantic attribute transfer.

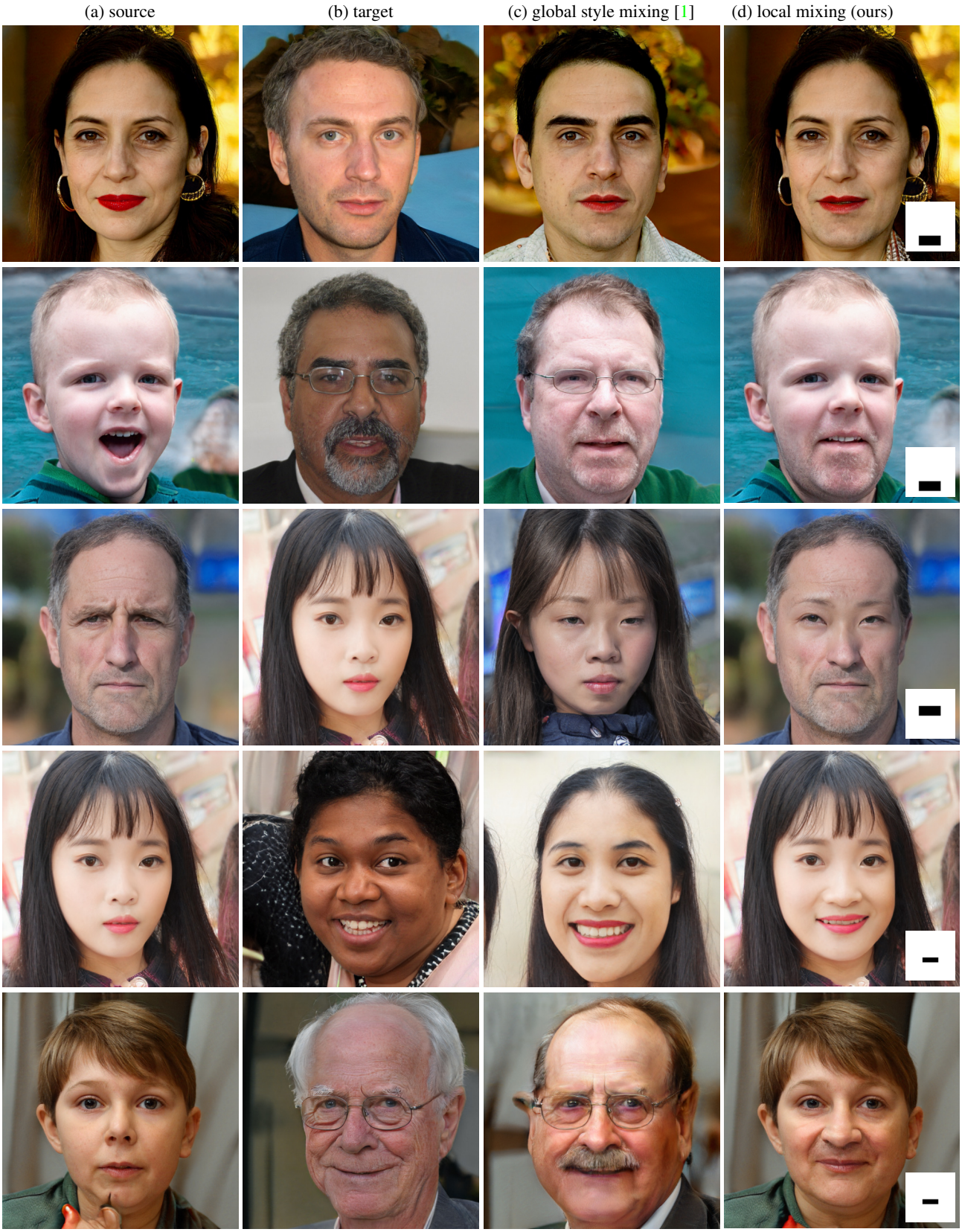

Figure 3: **Semantic attribute transfer- Additional results.** Our test-time adaptive normalization enables to perform local style mixing; The effect is that only local attributes are transferred from the target image to the source image, while keeping the area outside the mask fixed. Note that we manage to transfer relatively coarse structures (*e.g.* lips, eyes and nose shape) while maintaining realistic appearance. This is in contrast to styleGAN global mixing effect that completely changes the structure of the source image to match the target.

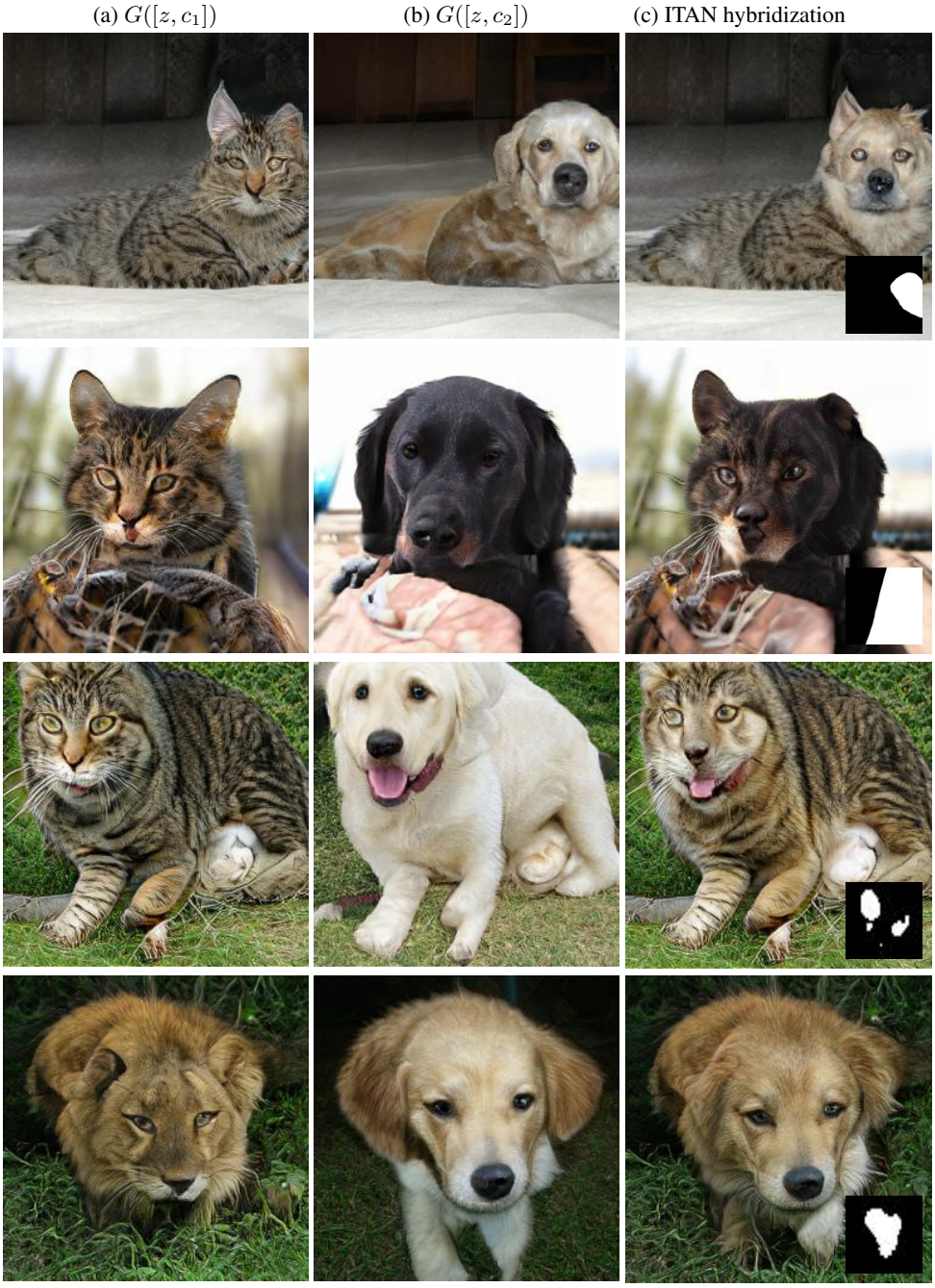

Figure 4: **Class hybridization- Additional result.** We use our inference-time adaptive normalization technique to generate images of hybrid classes, blended spatially according to a mask.