Interpretable-through-prototypes deepfake detection for diffusion models:
Appendix

1. Appendix A tSNE Plots

Figure 1: image-to-image ImageNet Deepfake subset training samples’ 2D tSNE visualization with fine-tuned features

Figure 3: image-to-image ImageNet Deepfake subset training samples’ 2D tSNE visualization with pre-trained features

Figure 2: image-to-image ImageNet Deepfake subset testing samples’ 2D tSNE visualization with fine-tuned features

Figure 4: image-to-image ImageNet Deepfake subset testing samples’ 2D tSNE visualization with pre-trained features
Figure 5: Oxford-IIIT Pet Deepfake dataset testing samples’ 2D tSNE visualization with cross-dataset approach