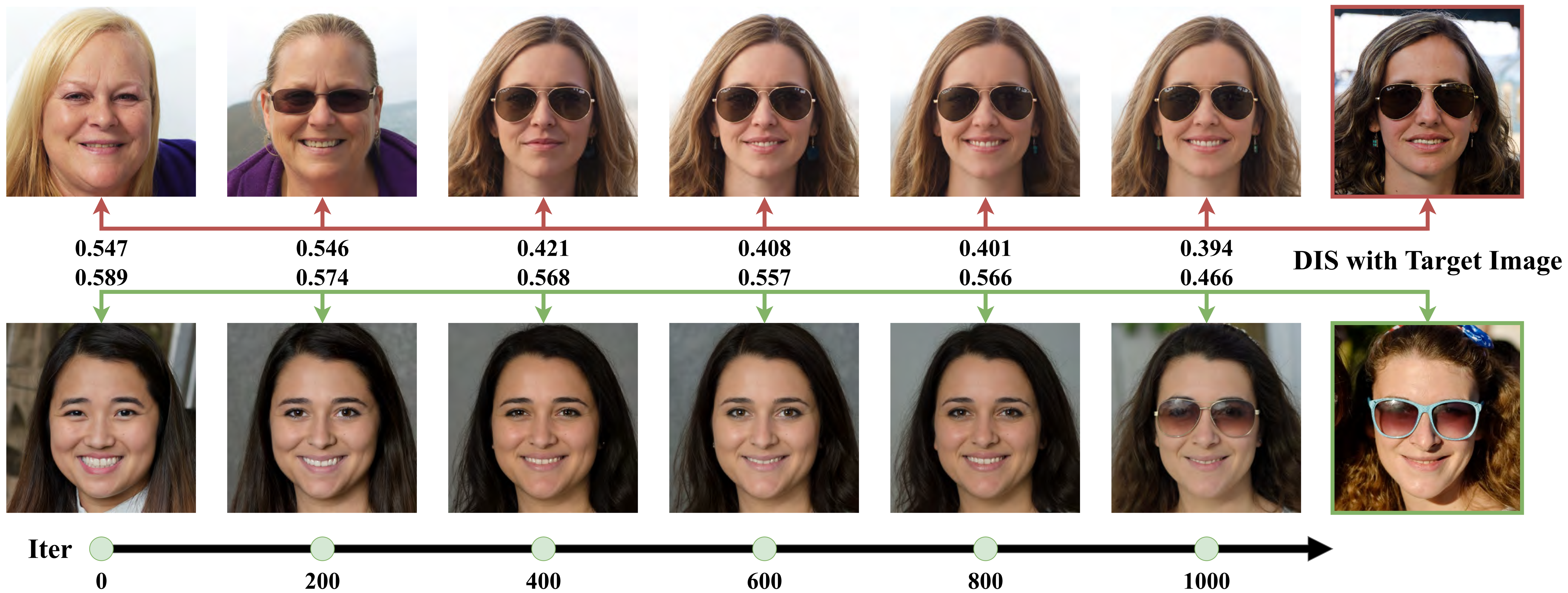
Bridging Data Gaps in Diffusion Models with Adversarial Noise-Based Transfer Learning

Xiyu Wang[1], Baijiong Lin[2], Daochang Liug[1], Ying-Cong Chen[2], Chang Xu[1]

[1]The University of Sydney [2]The Hong Kong University of Science and Technology (Guangzhou)

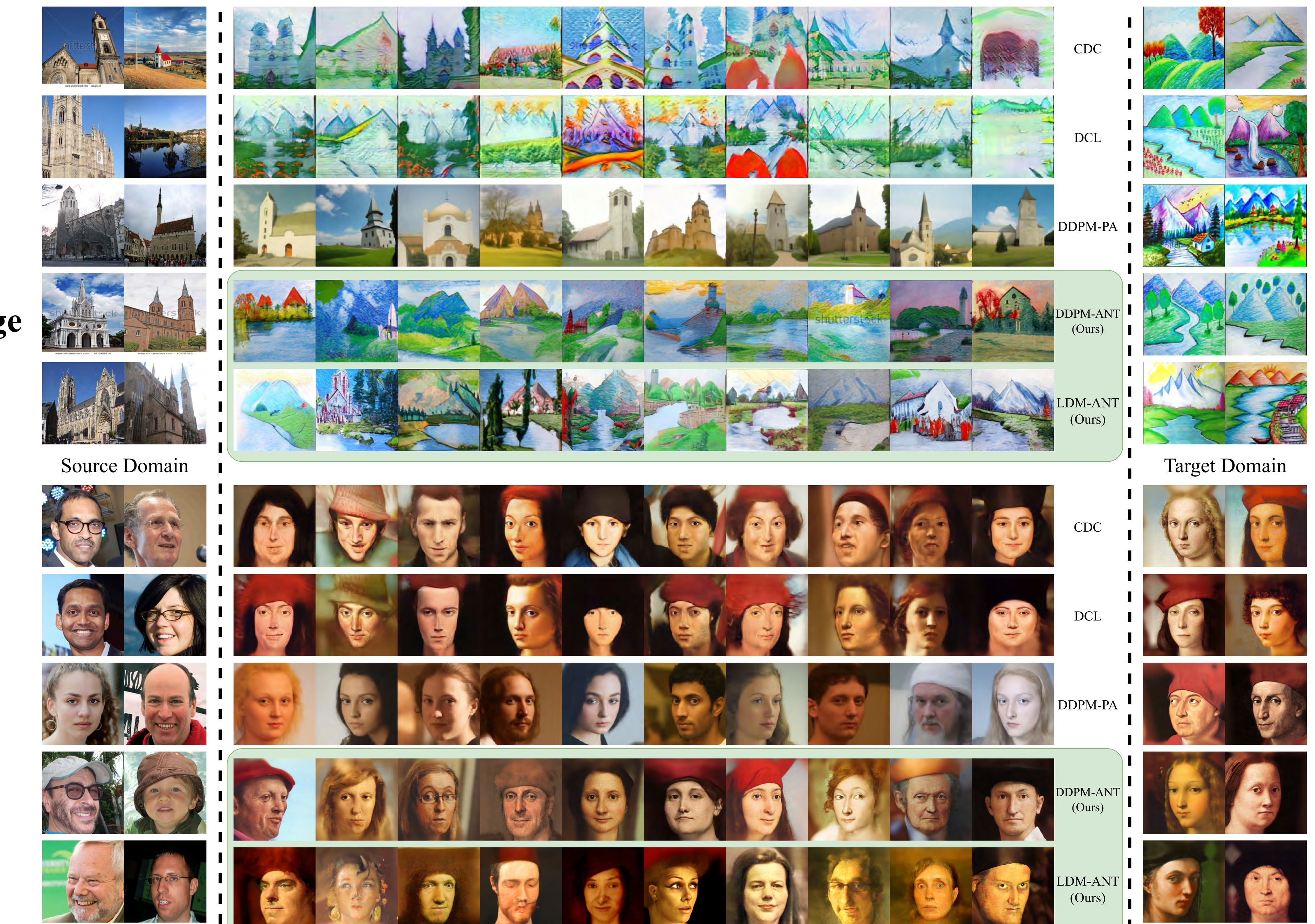
ICML International Conference On Machine Learning

Unbalance Transfer Result



Two sets of images generated from corresponding fixed noise inputs at different stages of fine-tuning DDPM from FFHQ to 10-shot Sunglasses. When the bottom image successfully transfers to the target domain, the top image is already overfitting.

Qualitative Evaluation



The 10-shot image generation samples on LSUN Church → Landscape drawings (top) and FFHQ → Raphael's paintings (bottom). When compared with other GAN-based and DDPM-based methods, our method, ANT, yields high-quality results that more closely resemble images of the target domain style, with less blurring.

Motivation

- **Estimating the transfer direction** for each timesteps of Diffusion Models training to boosting the transfer.
- **Searching the targeted domain related noise** instead of mechanism noise in diffusion models training for less iteration.

Contribution

- **ANT: an efficient fusion method** leverage existing knowledge from the source domain to aid in training the target domain
- Experiments on few-shot image generation tasks affirm the effectiveness of our ANT.

Our Method

- **Similarity-guided training** employs a classifier to estimate the divergence between the source and target domains, leveraging existing knowledge from the source domain to aid in training the target domain.

$$\min_{\theta} \mathbb{E}_{t,x_0,\epsilon}\left[\left\|\epsilon_t - \epsilon_\theta(x_t,t) - \hat{\sigma}_t^2 \gamma \nabla_{x_t} \log p_\phi(y=\mathcal{T}|x_t)\right\|^2\right]$$

- **Adversarial noise selection** dynamically choose the noise according to the input image.

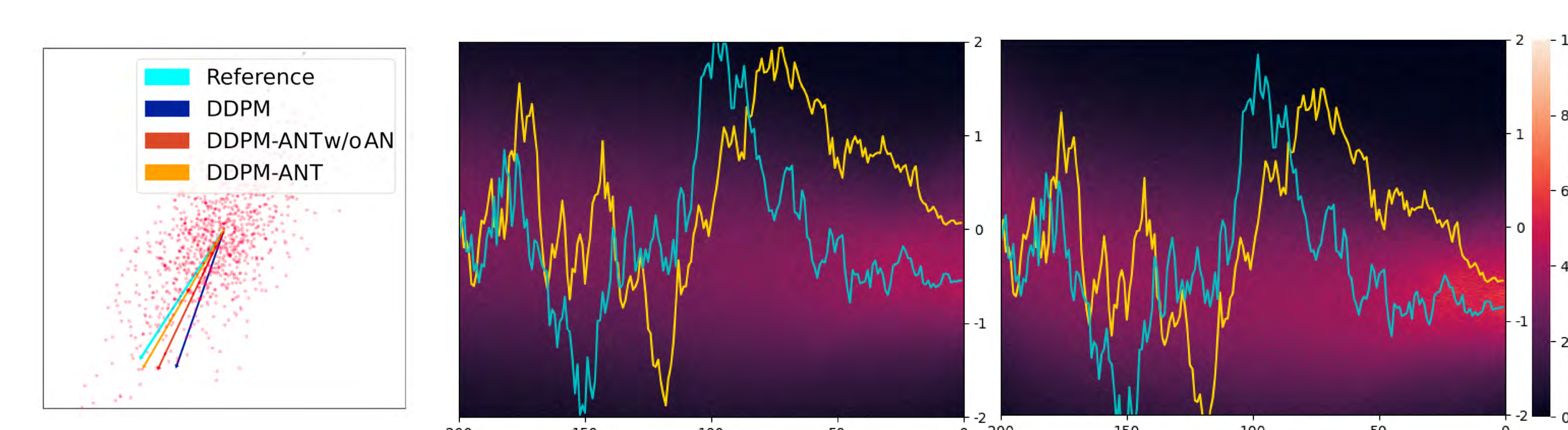
$$\min_{\theta} \max_{\epsilon} \mathbb{E}_{t,x_0}\left[\left\|\epsilon - \epsilon_\theta(x_t,t) - \hat{\sigma}_t^2 \gamma \nabla_{x_t} \log p_\phi(y=\mathcal{T}|x_t)\right\|^2\right]$$

Quantitative Evaluation

Intra-LPIPS results for both DDPM and GAN-based baselines are presented for 10-shot image generation tasks.

Methods	Parameter Rate	FFHQ → Babies	FFHQ → Sunglasses	FFHQ → Raphael's paintings	LSUN Church → Haunted houses	LSUN Church → Landscape drawings
TGAN	100%	0.510±0.026	0.550±0.021	0.533±0.023	0.585±0.007	0.601±0.030
TGAN+ADA	100%	0.546±0.033	0.571±0.034	0.546±0.037	0.615±0.018	0.643±0.060
EWC	100%	0.560±0.019	0.550±0.014	0.541±0.023	0.579±0.035	0.596±0.052
CDC	100%	0.583±0.014	0.581±0.011	0.564±0.010	0.620±0.029	0.674±0.024
DCL	100%	0.579±0.018	0.574±0.007	0.558±0.033	0.616±0.043	0.626±0.021
DDPM-PA	100%	0.599±0.024	0.604±0.014	0.581±0.041	0.628±0.029	0.706±0.030
DDPM-ANT (Ours)	1.3%	0.592±0.016	0.613±0.023	**0.621**±0.068	0.648±0.010	0.723±0.020
LDM-ANT (Ours)	1.6%	**0.601**±0.018	**0.613**±0.011	0.592±0.048	**0.653**±0.010	**0.738**±0.026

Visualization on Toy Data



(a) Gradient of Output Layer (b) Heat-map of DDPM (c) Heat-map of DDPM-ANT