SHANGHAI JIAO TONG UNIVERSITY

Dual-path Image Inpainting with Auxiliary GAN Inversion

Wentao Wang[1], Li Niu[1*], Jianfu Zhang[2], Xue Yang[1], Liqing Zhang[1*]
[1]Department of Computer Science and Engineering, MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University
[2]Tensor Learning Team, RIKEN AIP

CVPR JUNE 19-24 2022 NEW ORLEANS • LOUISIANA

Introduction:

➢ Abstract:

- Feed-Forward inpainting methods: Train a convolutional neural network with adversarial training and inpaint the corrupted image by feed-forward inference.
- GAN inversion inpainting methods: First seek for the closest latent code in the latent space (input to the pretrained generator) for the corrupted image and then invert the latent code back to a complete image using the pretrained generator.
- Our method: We propose a hybrid inpainting framework which applies GAN inversion inpainting to assist feed-forward inpainting.

➢ Our Contributions:

- We propose a hybrid dual-path inpainting framework which assists in feed-forward inpainting with GAN inversion.
- We propose a novel deformable fusion module in the generator to solve the misalignment issue when fusing the features from two paths.
- Extensive experiments prove that our method can produce more semantically reasonable and high-fidelity results than other state-of-the-art methods.

Method:

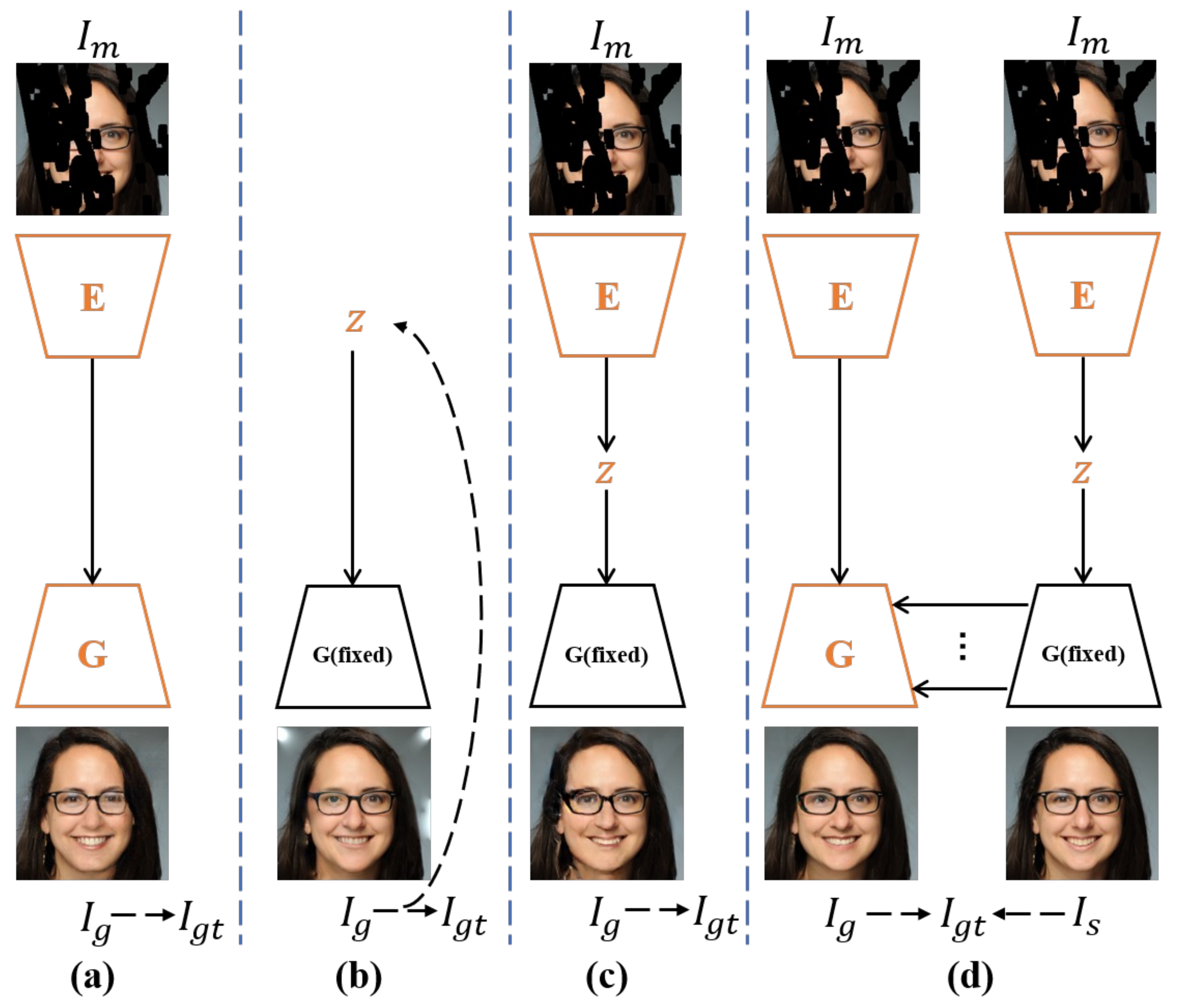

Figure 1. Illustration of four types of inpainting methods. (a) Feed-forward inpainting method. (b) Optimization-based GAN inversion inpainting method (c) Learning-based GAN inversion inpainting method. (d) Ours. The modules marked in orange are optimized during training.

Network Architecture:

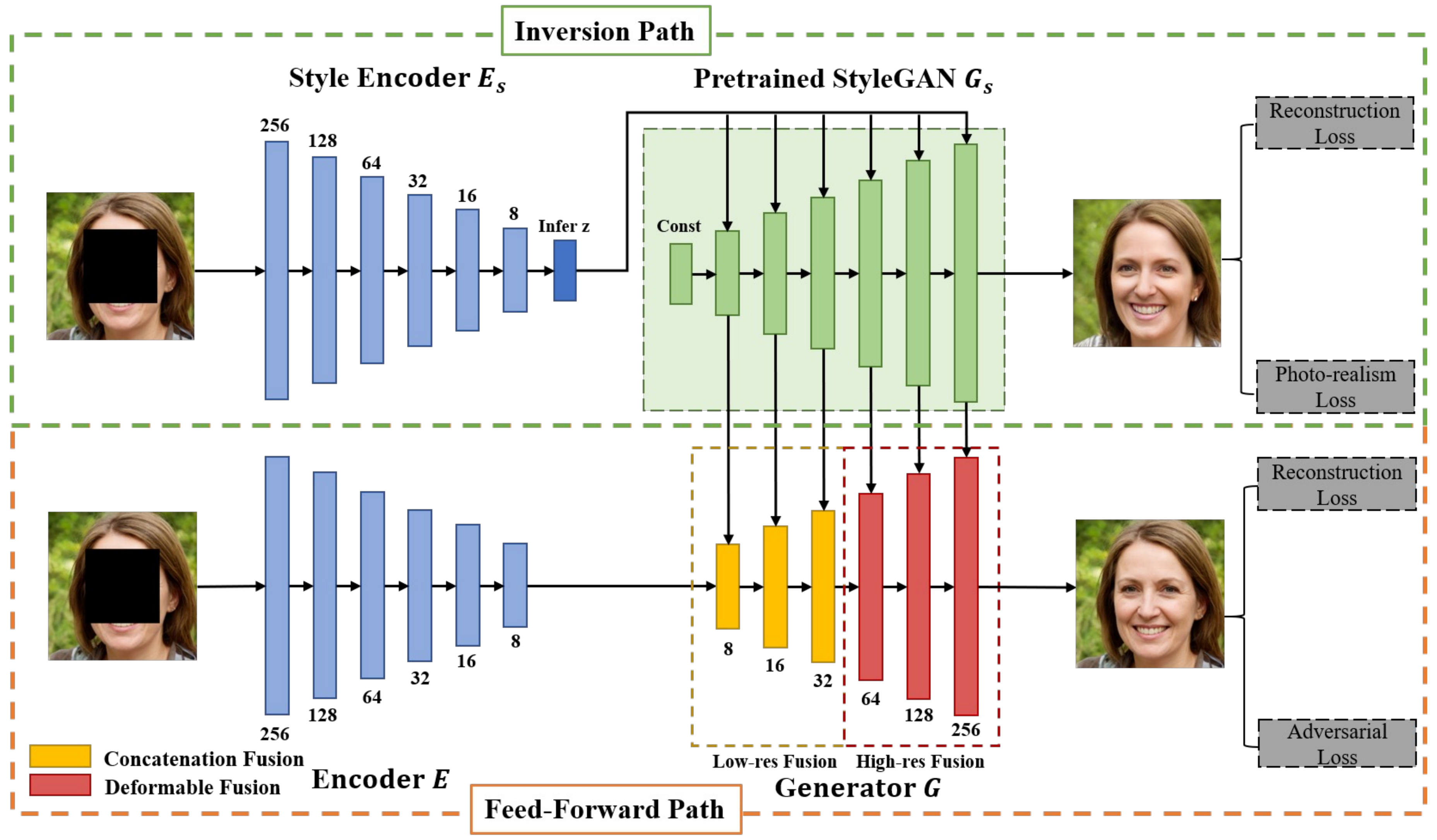

Figure 2. Hybrid Two-path Inpainting Network

➢ Inversion Path

Find the closest latent code of the corrupted image I_m and extract the corresponding semantic intermediate features F_s from the pretrained GAN.

➢ Feed-Forward Path

An auto-encoder inpainting network which fuses the the extracted semantic intermediate features from Inversion Path as extra semantic prior.

➢ Deformable Feature Fusion Module

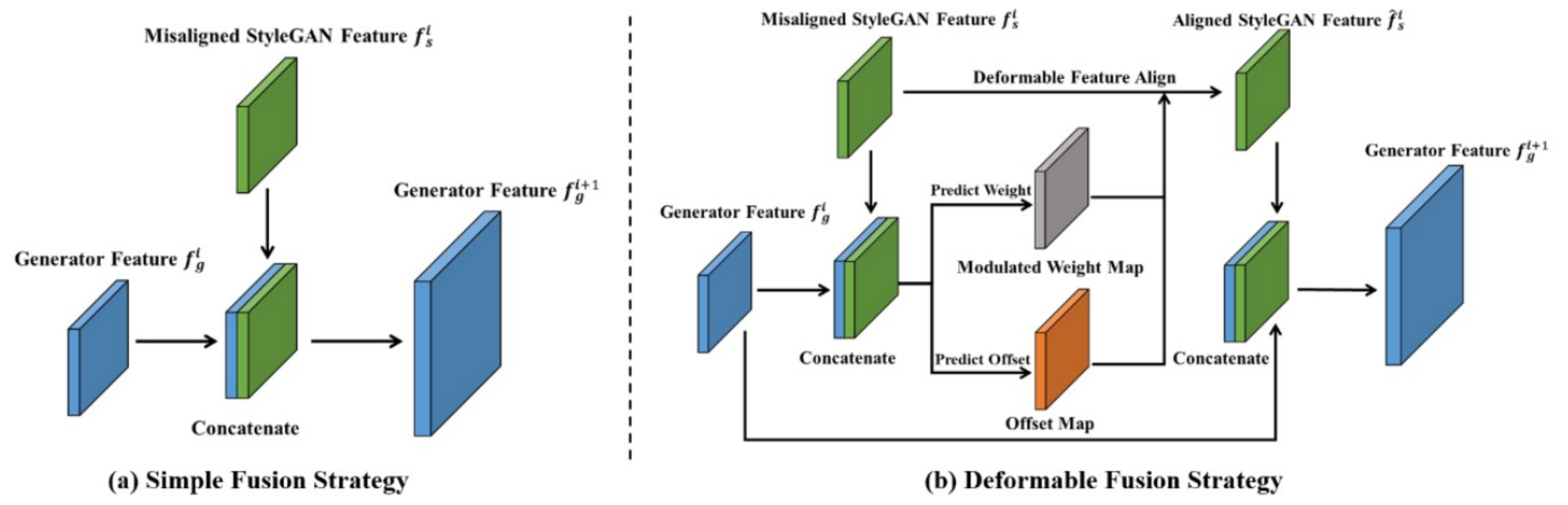

Figure 3. Two fusion strategies. (a) Simple fusion strategy. (b) Our deformable fusion module. In (a), simply concatenate features from two path will cause misalignment. In (b), our deformable fusion module can align the inversion path feature and solve the misalignment.

Experiments:

➢ Quantitative Comparison on FFHQ

	Mask	Yeh *et al.* [25]	Lahiri *et al.* [12]	GC [29]	PICNet [33]	CoModGAN [32]	Ours
ℓ_1 (%)↓	0-10%	0.94	1.27	0.73	0.74	0.64	**0.62**
	10-20%	1.59	1.83	1.23	1.23	1.11	**1.06**
	20-30%	2.53	2.38	1.97	1.95	1.75	**1.41**
	30-40%	3.64	4.05	2.83	2.79	2.61	**2.16**
	40-50%	5.06	5.94	3.90	3.84	3.69	**3.21**
	50-60%	7.73	9.21	5.73	5.76	5.62	**4.59**
	Ave%	3.58	4.18	2.73	2.71	2.54	**2.17**
SSIM↑	0-10%	0.969	0.911	0.974	0.973	0.978	**0.979**
	10-20%	0.932	0.876	0.941	0.939	0.948	**0.951**
	20-30%	0.881	0.827	0.895	0.893	0.905	**0.914**
	30-40%	0.827	0.774	0.845	0.843	0.857	**0.879**
	40-50%	0.767	0.711	0.789	0.785	0.802	**0.827**
	50-60%	0.688	0.621	0.713	0.702	0.727	**0.743**
	Ave%	0.844	0.787	0.859	0.856	0.870	**0.882**
PSNR↑	0-10%	33.576	31.994	35.600	35.726	36.211	**36.342**
	10-20%	28.937	27.311	30.807	31.053	31.209	**31.607**
	20-30%	25.714	24.729	27.467	27.813	27.703	**28.365**
	30-40%	23.281	22.512	25.113	25.474	25.214	**26.251**
	40-50%	23.282	20.201	23.093	23.462	23.069	**24.155**
	50-60%	21.087	17.039	20.625	20.804	20.490	**21.751**
	Ave%	25.152	23.964	27.117	27.388	27.366	**28.078**
FID↓	0-10%	1.83	1.97	1.50	1.57	1.31	**1.20**
	10-20%	3.33	3.85	2.40	2.71	2.14	**2.00**
	20-30%	5.42	6.70	3.93	4.55	3.86	**2.99**
	30-40%	7.92	10.01	6.25	6.90	5.56	**4.13**
	40-50%	11.04	14.42	9.69	10.64	6.25	**5.67**
	50-60%	13.89	21.73	15.91	16.71	9.08	**8.13**
	Ave%	7.24	9.77	6.61	7.18	4.75	**4.02**

➢ Qualitative Comparison