Guided CycleGAN via Semi-Dual Optimal Transport for Photo-Realistic Face Super-resolution

Wenbo Zheng[1,3], Chao Gou [2,✉], Lan Yan[3], Wenwen Zhang[1,3], Fei-Yue Wang [3]

[1] School of Software Engineering, Xi'an Jiaotong University

[2] School of Intelligent Systems Engineering, Sun Yat-sen University, China

[3]The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences

Introduction

Face super-resolution has been studied for decades, and many approaches have been proposed to upsample low-resolution face images using information mined from paired low-resolution (LR) images and high-resolution (HR) images. However, most of this kind of works only simply sharpen the blurry edges in the upsampled face images and typically no photo-realistic face is reconstructed in the final result. In this paper, we present a GAN-based algorithm for face super-resolution which properly synthesizes photo-realistic super-recovered face. To this end, we introduce semi-dual optimal transport to optimize our model such that the distribution of its generated data can match the distribution of a target domain as much as possible. This way would endow our model with learning the mapping of distribution from unpaired LR images and HR images with desired properties. We demonstrate the robustness of our algorithm by testing it on Color FERET database and show that its performance is considerably superior to all state-of-the-art approaches.

Semi-Dual Optimal Transport Guided CycleGAN

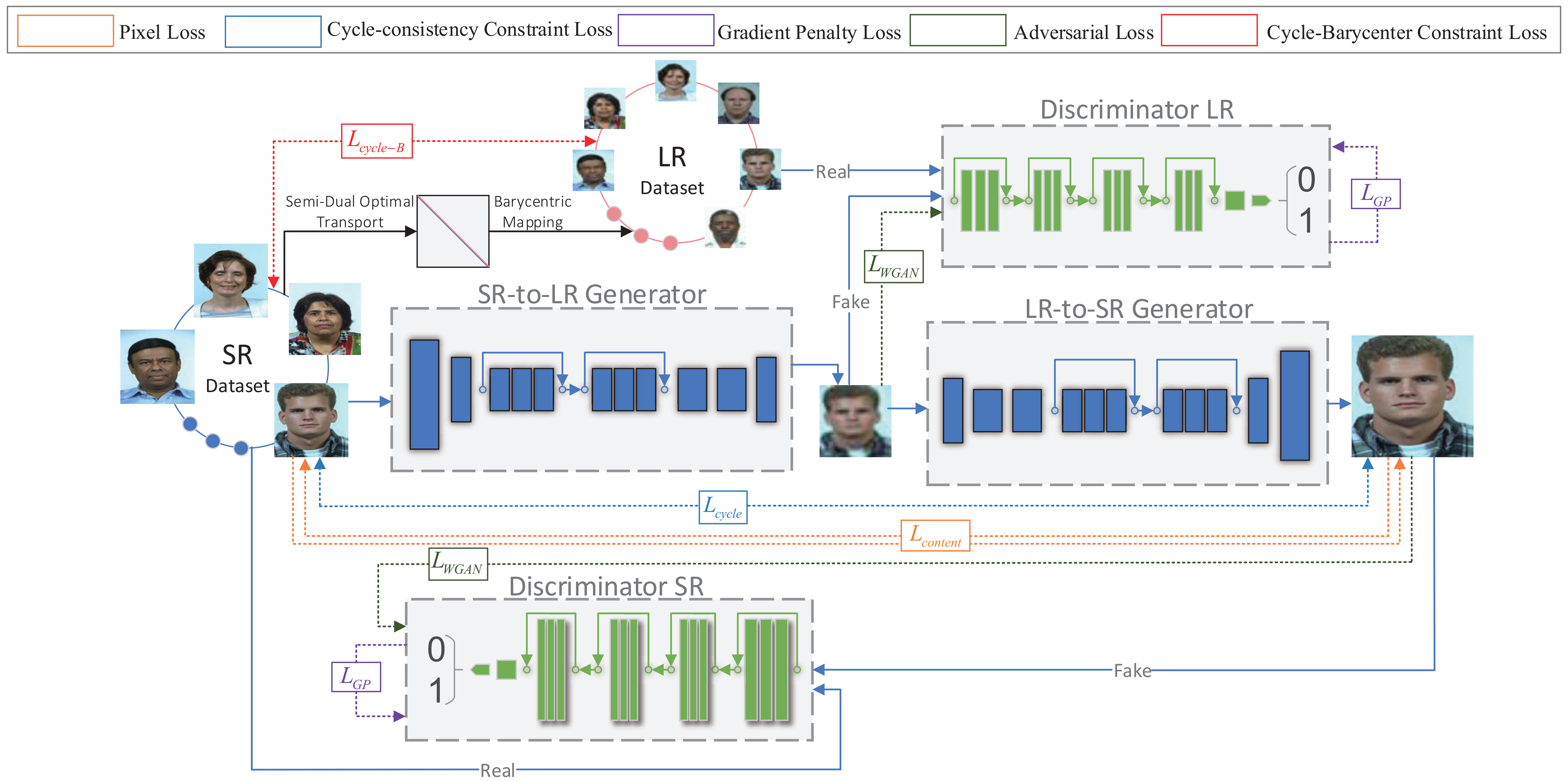

Figure 1: The Flow Chart of Our Network During Training

We use the Wasserstein GAN (WGAN) to build our SDOT-CycleGAN shown in Figure 1. The definition of our perceptual loss function L_{total} is critical for the performance of our SDOT-CycleGAN model.

Experimental Results

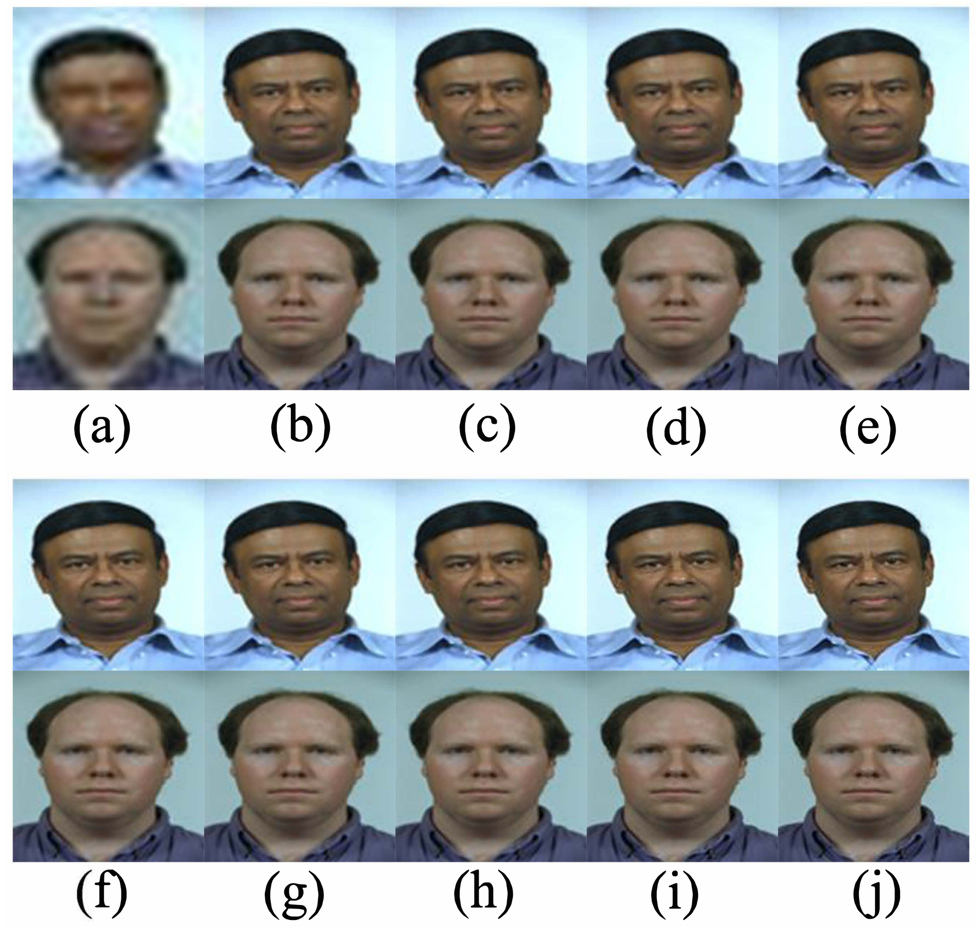

Figure 2: Comparison of Face Image Super-Resolution Methods' Results on The FERET Dataset by Different Methods. (a) LR Input Images. (b) SRCNN (c) VDSR. (d) URDGN. (e) TDAE[?]. (f) SRGAN. (g) FSRNet. (h) ESRGAN. (i) Ours. (j) Original HR Images.

Method \ Measure	PSNR(dB)↑	SSIM↑	IFC↑	FSIM↑	LPIPS↓
SRCNN	33.7108	0.9577	5.5489	0.9696	0.0692
VDSR	35.0336	0.9687	6.4907	0.9792	0.0469
URDGN	35.9018	0.9735	6.9776	0.9832	0.0369
TDAE	36.6110	0.9774	7.4177	0.9864	0.0298
SRGAN	37.4647	0.9807	7.8502	0.9890	0.0254
FSRNet	39.8043	0.9881	9.0108	0.9938	0.0114
ESRGAN	40.7114	0.9903	9.1733	0.9951	0.0084
Ours	**41.8163**	**0.9913**	**9.4333**	**0.9955**	**0.0075**

Table 1: Averaged Evaluations on Comparison Experiment for The FERET Dataset. ↑ Means The Larger The Values, The Better The Performance. ↓ Means The Smaller The Values, The Better The Performance.

Qualitative Evaluation It can be seen from the Figure 2 that, the results of the algorithm can cleanliest reconstruct the details of radius of the corner of the mouth in the face images of two men. In the face images of a man with blue shirt, the hair above the ear is better reconstructed than others. Therefore, our algorithm in this paper has higher overall image reconstruction than other algorithms.

Quantitative Evaluation From Table 1, it can get the following these points: PSNR of our method is 8.1055, 6.7827, 5.9145, 5.2053, 4.3516, 2.0120, and 1.1049 higher than SRCNN, VDSR, URDGN, TDAE, SRGAN, FSRNet, and ESRGAN, respectively. In SSIM, IFC, FSIM and LPIPS evaluation, there are also true of the case. In accordance with these points, our method achieves the best results among all methods. This means our method has the best ability of recovering photo-realistic face and outperforms state-of-the-art methods.

Conclusion and Future Work

In this paper, a novel semi-dual optimal transport guided CycleGAN (SDOT-CycleGAN) is proposed for face super-resolution. The key idea of SDOT-CycleGAN is using the semi-dual optimal transport to guide CycleGAN. This way not only helps to endow CycleGAN with learning a one-to-one mapping between unpair samples with desired properties, but also provides a solution for the photometric recovery in terms of PSNR/SSIM/IFC/FSIM/LPIPS. Extensive experiments show that SDOT-CycleGAN is superior to the state-of-the-arts on unaligned face images, both quantitatively and qualitatively. Future research can be expanded in various aspects, including video super-resolution.