Multi-Laplacian GAN with Edge Enhancement for Face Super Resolution

Shanlei Ko, Bi-Ru Dai

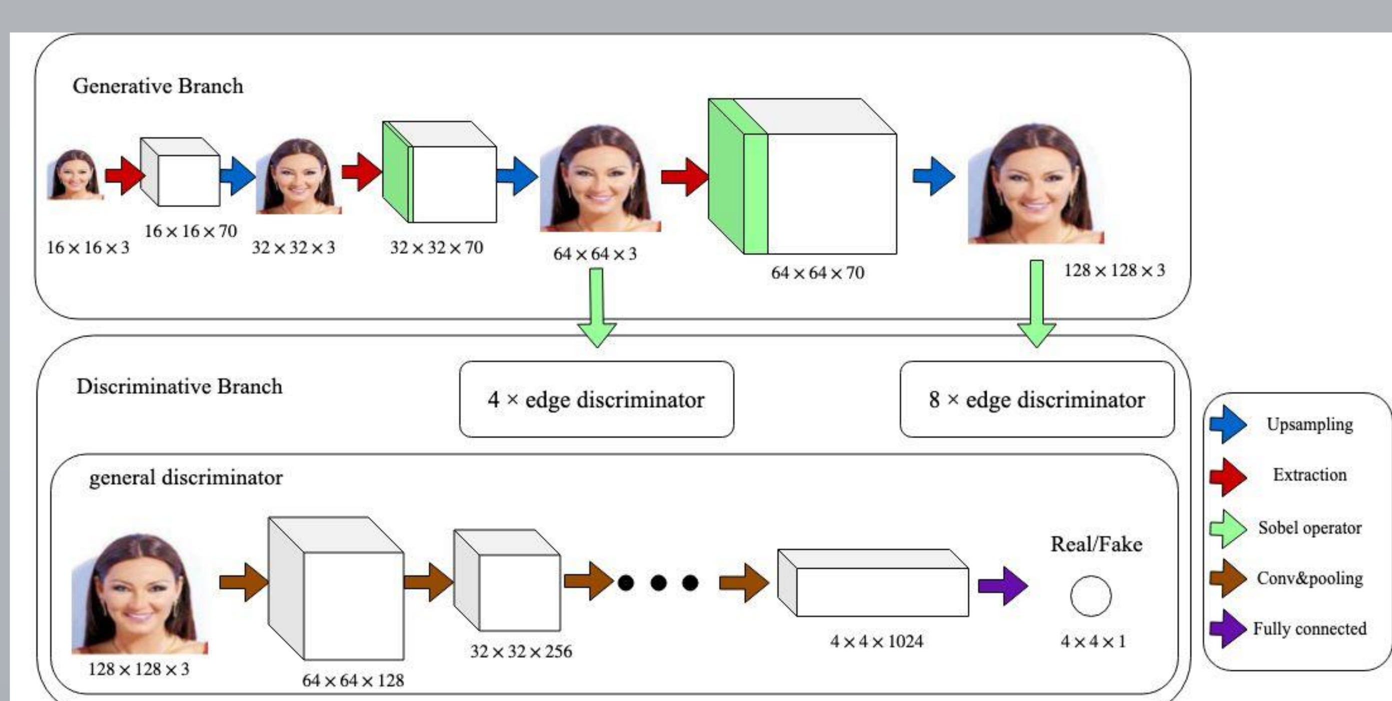
Department of Computer Science and Information Engineering,
National Taiwan University of Science and Technology, Taipei, Taiwan, R. O. C.

Knowledge Discovery and Data Mining Laboratory

Abstract

Face image super-resolution has become a research hotspot in the field of image processing. Nowadays, more and more researches add additional information, such as landmark, identity, to reconstruct high resolution images from low resolution ones, and have a good performance in quantitative terms and perceptual quality. However, these additional information is hard to obtain in many cases. In this work, we focus on reconstructing face images by extracting useful information from face images directly rather than using additional information. By observing edge information in each scale of face images, we propose a method to reconstruct high resolution face images with enhanced edge information. In additional, with the proposed training procedure, our method reconstructs photorealistic images in upscaling factor 8× and outperforms state-of-theart methods both in quantitative terms and perceptual quality.

Methods



The architecture of our edge enhancement upsampling network.

The purpose of this paper is to extract edge information from images directly to reconstruct high resolution face images. To enhance edge information for reconstructing high resolution images effectively, we proposed a GAN architecture which is named as multi-Laplacian GAN with edge enhancement and abbreviated as MLGE. MLGE consists of two branches. The first one is the generative branch which upsamples low resolution images to multiscale (2×, 4×, 8×) ones. The second one is the discriminative branch which consists of a general discriminator and two edge discriminators.

Edge Enhancement

Example of gradients in each channel (R, G, B). Images in the first column are gradients of HR images, and images in the second column are generated gradients by MLGE.

There are two edge enhancement parts in our architecture. First, Sobel filters are used to calculate gradients and concatenate gradients to feature maps when the generator predicts 4×, 8× SR images. As shown in Figure 1, green feature maps in the generative branch are extracted gradients. Second, edge discriminators are designed to determinate whether gradients are similar to those of the original image or not.

Experiment

We use the Celebrity Face Attributes (CelebA) dataset [21] to train our network. We randomly select 24K cropped face images from the CelebA dataset, and then resize them to 128 × 128 pixels as HR. For data augmentation, we randomly transform HR images (translate and rotation) and downsample HR images to 16×16 pixels as their corresponding LR images. We use 22K LR and HR face pairs for training, and 2K LR face images for testing.

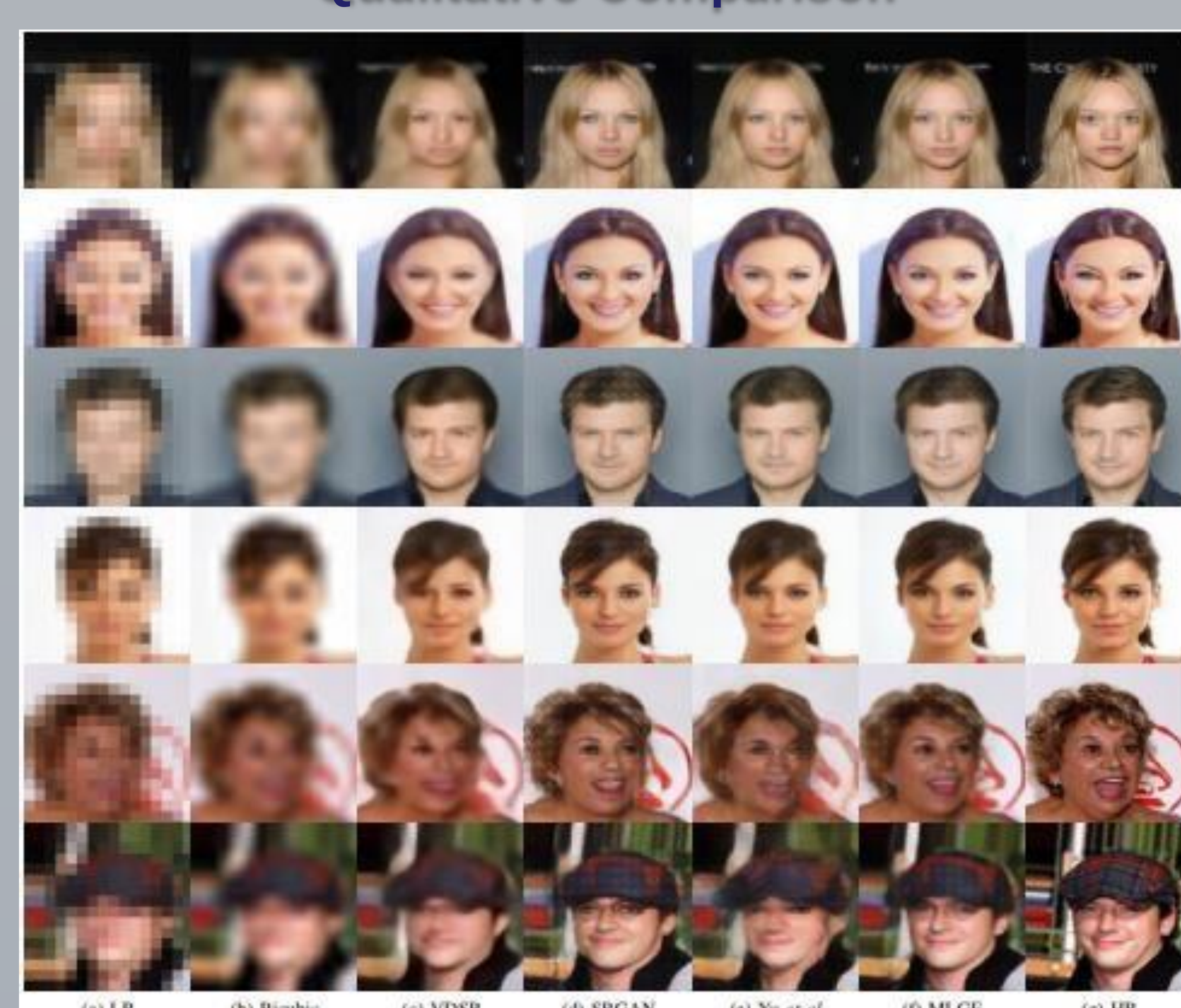
Quantitative Comparison

Method	Bicubic	VDSR	SRGAN	Yu et al	MLGE
PSRN	21.76	23.12	23.76	22.85	**25.07**
SSIM	0.72	0.80	0.78	0.753	**0.83**

Quantitative comparison for PSNR and SSIM

Two popular metrics, including the average Peak Signal to-Noise Ratio (PSNR) and the structural similarity (SSIM) scores are used as our quantitative metrics. Our method outperforms the second best with a large margin of 1.95 dB in PSNR. Based on the design of multiple pixel-wise losses and edge information enhancement, MLGE has the best performance of PSNR and SSIM and is able to reconstruct HR images both in face regions and non-face regions effectively.

Qualitative Comparison



Qualitative comparison for 8× SR on the CelebA dataset.

Our method MLGE generally produces better looking and sharper face images than the state-of-the-art methods. Note that our method MLGE, which only uses 22K training images in this experiment, not only effectively reconstructs face regions but also non-face regions, such as the background region of the fifth row.

CONCLUSION AND FUTURE WORK

We proposed the multi-Laplacian GAN with edge enhancement (MLGE) model to reconstruct 128 × 128 HR face images from 16 × 16 pixel LR face images. MLGE reconstructed HR face images efficiently and no additional prior facial information is required in both training and testing phases.

In future work, we will try to explore better information extracting methods from images and apply extracted information to reconstruct HR images effectively. Moreover, MLGE is possible to be extended to reconstruct all SR tasks.