Super-FAN: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with GANs

Wenjie Niu

June 26,2018

Abstract

As is known in the paper, the two goals are: improving the quality of low resolution facial images and accurately locating the facial landmark. 5 main contributions are made here:

1. *Authors propose Super-FAN.*

2. *They illustrate the benefit of training the two networks jointly.*

3. *They inprove upon the state-of-the-art in face super-resolution by proposing a new residual-based architecture.*

4. *Face super-resolution and alignment get great improvement over state-of-the-art.*

5. *On real-world low-resolution images like ones of Fig. 1, good results is the first time shown.*

1. Introduction

The paper aims at improving the quality and understanding of very low resolution facial images. In terms of quality, the aim is to increase the resolution and recover the detalis of real-world low resolution facial images as Fig. 1. In terms of understanding, they wish to extact mid-and high-level facial information by localizing a set of predefined facial landmarks with semantic meaning.

2. Closely related work

This section reviews related work in image and facial super-resolution, and facial landmark localization.

Image super-resolution. Early attempts on super-resolution using CNNs [3],[5] used standard L_p losses for training which result in blurry super-resolved images. To alleviate this, rather than using an MSE over pixels (between

Table 1. AUC across pose (calculated for a threshold of 10%; see [2]) on our LS3D-W balanced test set. The results, in this case, are indicative of visual quality. See Fig. 2

Method	[0-30]	[30-60]	[60-90]
FAN-bilinear	10.7%	6.9%	2.3%
FAN-SR-ResNet	48.9%	38.9%	21.4%
FAN-SR-GAN	47.1%	36.5%	19.6%
Retrained FAN-bilinear	55.9%	49.2%	37.8%
FAN-Ours-pixel	52.3%	45.3%	28.3%
FAN-Ours-pixel-feature	57.0%	50.2%	34.9%
Super-FAN	67.0%	63.0%	52.5%
FAN-HR images	75.3%	72.7%	68.2%

the super-resolved and the ground truth HR image), the authors of [4] proposed an MSE over feature maps, coined perceptual loss.

Face super-resolution. The recent work of [7] uses a GAN-based approach (like the one of [6] without the perceptual loss) to super-resolve very low-resolution faces.

Face alignment. A recent evaluation of face alignment [2] has shown that when resolution drops down to 30 pixels, the performance drop of a state-of-the-art network trained on standard facial resolution (192×192) for medium and large poses is more than 15% and 30%, respectively.

3. Conclusions

The paper proposed Super-FAN: the very first end-to-end system for integrated facial super-resolution and landmark localization. The method incorporates facial structural information in a newly proposed architecture for superresolution, via integrating a sub-network for face alignment and optimizing a novel heatmap loss.

References

[1] A. Bulat and G. Tzimiropoulos. Super-FAN: Integrated facial landmark localization and super-resolution of real-world low

Figure 1. A few examples of visual results produced by our system on real-world low resolution faces from WiderFace.[1]
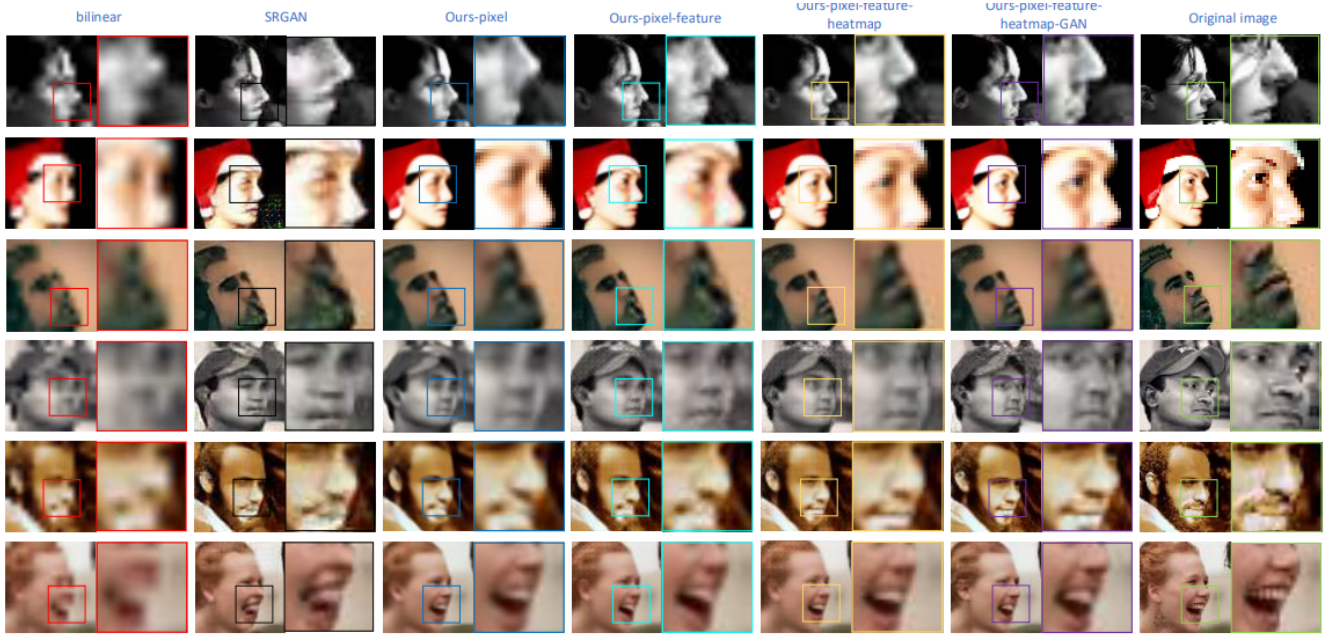

Figure 2. Visual results on LS3D-W.[1]

resolution faces in arbitrary poses with GANs. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2018. 2

[2] A. Bulat, G. Tzimiropoulos, A. Bulat, and G. Tzimiropoulos. How far are we from solving the 2D & 3D face alignment problem? (and a dataset of 230,000 3D facial landmarks). In *IEEE International Conference on Computer Vision*, 2017. 1

[3] C. Dong, C. L. Chen, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295–307, 2016. 1

[4] J. Johnson, A. Alahi, and F. Li. Perceptual losses for real-time style transfer and super-resolution. In *n European Conference on Computer Vision*, 2016. 1

[5] J. Kim, J. K. Lee, and K. M. Lee. Accurate image super-resolution using very deep convolutional networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2015. 1

[6] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. Photo-realistic single image super-resolution using a generative adversarial network. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016. 1

[7] X. Yu and F. Porikli. Ultra-resolving face images by discriminative generative networks. In *European Conference on Computer Vision*, 2016. 1