Finding Tiny Faces in the Wild with Generative Adversarial Network

Xuewen Yang

July 24 2018

Abstract

Face detection techniques have been developed for decades, and one of remaining open challenges is detecting small faces in unconstrained conditions. The reason is that tiny faces are often lacking detailed information and blurring. In this paper,the author proposed an algorithm to direct- ly generate a clear high-resolution face from a blurry small one by adopting a generative adversarial network (GAN). Toward this end, the basic GAN formulation achieves it by super-resolving and refining sequentially (e.g. SR-GAN and cycle-GAN). However,they design a novel network to address the problem of super-resolving and refining jointly.They also introduce new training losses to guide the generator net- work to recover fine details and to promote the discrimina- tor network to distinguish real vs. fake and face vs. non-face simultaneously. Extensive experiments on the challenging dataset WIDER FACE demonstrate the effectiveness of this proposed method in restoring a clear high-resolution face from a blurry small one, and show that the detection perfor- mance outperforms other state-of-the-art methods.

1. Introduction

Face detection is a fundamental and important problem in computer vision, since it is usually a key step towards many subsequent face-related applications, including face parsing, face verification, face tagging and retrieval, etc.Modern face detectors have achieved impressive results on the large and medium faces, however, the performance on small faces is far from satisfactory. The main difficulty for small face detection is that small faces lack sufficient detailed information to distinguish them from the similar background,To detect small faces, [2] directly up-samples images using bi-linear operation and exhaustively searches faces on the up-sampled images. However, this method will increase the computation cost and the inference time will increase significantly too. Moreover, images are often zoomed in with a small upscaling factors in [2], otherwise, artifacts will be generated. Besides,it use the intermediate conv feature maps to represent faces at specific scales, which keeps the balance between the computation burden and the performance. However, the shallow but fine-grained intermediate conv feature maps lack discrimination, which causes many false positive results. More importantly, these methods take no care of other challenges, like blur and illumination.

2. Proposed Method

In this section,the author introduce our proposed method in details. First,they give a brief description on the classical GAN network. Then, the whole architecture of our method is presented. Finally,they introduce each part of the network in details and define the loss functions for training the generator network and discriminator network respectively.

2.1. GAN

GAN [1] learns a generative model via an adversarial process. It trains a generator network and a discriminator network simultaneously. The training process alternately optimizes the generator and discriminator, which compete with each other. The generator is trained for generating the samples to fool the discriminator, and the discriminator is trained to distinguish the real image and the fake image from the generator.

2.2. Network Architecture

Their generator network includes two components and the first sub-network takes the low-resolution images as the inputs and the outputs are the super-resolution images. Since the blurry small faces lack fine details and due to the influence of MSE loss, the generated super-resolution faces are usually blurring. So they design the second subnetwork to refine the super-resolution images from the first sub-network. Furthermore,they add the classification branch to the discriminator network for the purpose of detection, which means their discriminator can classify faces and non-faces as well as distinguish the fake and real images.

2.3. Loss Function

The author adopt the pixel-wise loss and adversarial loss from some state-of-the-art approaches to optimize their generator network. In contrast,the author remove the VGG feature matching loss due to the calculation cost and we introduce the classification loss to drive the generator network to recover fine details from the blurry small faces.

References

[1] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *International Conference on Neural Information Processing Systems*, pages 2672–2680, 2014. 1

[2] X. Xu, D. Sun, J. Pan, Y. Zhang, H. Pfister, and M. H. Yang. Learning to super-resolve blurry face and text images. In *ICCV*, 2017. 1