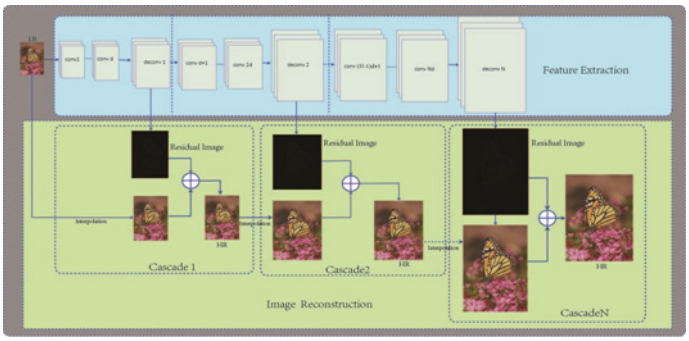
A single super-resolution method via deep cascade network

Wang Fei*, Wang Wei, Qiu Zhiliang

Intelligent equipment laboratory, Xidian University, Xi'an, Shaanxi 710071, China



Deep cascaded network architectures. The light green part is the image reconstruction of each cascade, and the light blue part is the part of the feature extraction for each cascade. This network includes several convolution layers and one transposed convolutions layer (upsampling) in each cascade

Overview: Recovering high resolution (HR) image from its low resolution (LR) image is an important issue in the field of digital image processing and other vision tasks. Recently, Dong et al. found that a convolutional neural network (CNN) can be used to learn end-to-end mapping from LR to HR. The network is expanded into many different forms, using sub-pixel convolutional network, very deep convolutional network, and recursive residual network. Although these models have achieved the desired results, the issues still exist some problems as described as following. First, most methods use up-sampling operators, such as bi-cubic interpolation, to upscale the input image to the bigger size. This pre-processing adds considerable unnecessary computations and often results in visible reconstruction artifacts. To solve this problem, there are several algorithms such as ESPCN using sub-pixels and FSRCNN with transposed convolution. However, the network structures of these methods are extremely too simple to lean complex and detailed mappings. Second, most existing methods use only L2 to optimize the network, which will result in an excessively smooth image less suitable for human vision. Third, those methods cannot reconstruct more than one scale, which means a model is only for one scale, and this will increase the extra-works of training for the other scales, especially for large-scale training.

To address these defects, we propose a deep cascaded network (DCN). DCN is a cascade structure, and it takes an LR image as input and predicts a residual image in each scale. The predicted residual for each scale is used to efficiently reconstruct the HR image through up-sampling and adding operations. We train the DCN with L2 and perceptual loss function to obtain a robust image.

Our approach differs from existing CNN-based methods in the following aspects:

1) Multiple scales with cascade layers. Our network has a cascade structure and generates multiple intermediate SR predictions in feed-forward process. This progressive reconstruction can get more accurate results. Our 4× model can obtain 1.5×, 2×, 2.5×, 3×, 3.5× reconstructed images.

2) Optimize network with L2 and perceptual loss function. Using L2 can get more accurate pixel-level reconstruction and using perceptual loss function may be closer to human vision.

3) Features extraction on LR image. Our method does not require traditional interpolation methods to up-sample images as a pre-processing, thus greatly reducing the computational complexity.

Extensive experiments on several large benchmark datasets show that the proposed approach performs better than existing methods in terms of accuracy and visual improvement.

Citation: Wang F, Wang W, Qiu Z L. A single super-resolution method via deep cascade network[J]. *Opto-Electronic Engineering*, 2018, **45**(7): 170729

* E-mail: 290727048@qq.com