Deep Learning and Computer Vision
P1 : Deep Learning CNN for Super Resolution

Vishisht Sharma

1 Introduction

Single image super-resolution (SISR) is a notoriously difficult ill-posed problem because a particular low-resolution (LR) input can correspond to a crop of potential high-resolution (HR) images, and the HR space (which is typically the natural image space) that we intend to map the LR input to is typically intractable. The two fundamental shortcomings of previous SISR techniques are the ambiguity of the mapping between the LR space and the HR space that we intend to create and the inefficiency of creating a sophisticated high dimensional mapping given a large amount of raw data. Recent DL-based SISR methods have seen notable gains, both numerically and qualitatively, thanks to their great ability to extract useful high-level abstractions that connect the LR and HR domains.

2 What are the main trends on the topic since the publication of the paper discussed?

Since the publication of SRCNN [1] paper which used a Deep convolutional Neural Network, the publications following have addressed problems such as different up-sampling methods such as the one in SRCNN bi-cubic interpolation and Deconvolution were used to improve the performance in super resolution and changing the network design from linear network to residual networks, recursive learning and multi-path learning proved fundamental in improving performance.

3 What are the key ideas of related published works since the original publication?

SRCNN [1] was the pioneering work of using a convolutional neural network (CNN) in image super-resolution reconstruction development [24,48]. The idea of SRCNN was inspired by sparse coding-based super-resolution methods. SRCNN consisted of three main parts, namely patch extraction, non-linear mapping, and image reconstruction. FSRCNN [2] was an improvement to the SRCNN model as it used deconvolutional layer as an upsampling method, it also introduced a shrinking layer and reduced the number of filters used in SRCNN. VDSR [3] was a much deeper network architecture compared to SRCNN with 20 weight layers in contrast to 3 in SRCNN. It also introduced residual learning between the input and output layers of the final mapping layer. DRCN [4] was the first algorithm to apply a recursive method for image super-resolution.It consisted of 3 sub networks namely embedding, inference and reconstruction networks. It also made use of skip connection to address the problem of vanishing gradients with recursive networks. FGLRL [5] used FSRCNN as the base and it consisted of five parts in patch extraction, shrinking, non-linear mapping, expanding and reconstruction. It used PReLU as an activation function and deconvolutional layer as the upsampling method. DBCN [6] was a dual branch-based image super resolution algorithm. the network split into two branches, where one branch adopted a convolutional layer with Leaky ReLU as the activation function, while the other branch adopted dilated convolutional layer with Leaky ReLU as the activation function. The output from each branch would then be fused through the concatenation process before it was upsampled. It also combined both bi-cubic and deconvolutional layers for the recontruction process.

4 What are the main problems solved or improvements over the original work?

FSRCNN introduced deconvolutional layers, shrinking layer and lower number of filters which reduced the amount of parameters which inturn made the network train faster and be more efficient. VSDR had two major benefits over SRCNN.First, the changes in VSDR made the network to converge faster due to the LR image had a high correlation with the HR image.Second, VSDR had better performance than SRCNN.VSDR trained over 93 percent more efficiently than SRCNN. DRCN was designed to overcome the problem of requiring many mapping layers to achieve better performance in SRCNN. Since the recursive method was used, shared weight allowed the network to widen the receptive field without increasing the model capacity, and therefore, fewer resources were required during the training process. FGLRL combined the model architecture from FSRCNN, immediate output design from DRCN and some other modification to the overall architecure which resulted in a significant decrease in training time as FGLRL was training twice as fast as DRCN which was significantly faster than SRCNN. DBCN gave many benefits. First,the dual-branch structure solved the complex network problem that is often observed in chain-way-based networks (Linear layer networks). Second, the adoption of a dilated convolutional filter enhanced image quality during the reconstruction. Third, residual learning gave additional benefit to the model to achieve convergence faster.

5 What are the remaining problems from the published works so far?

(In context of this paper as i haven't discussed any attention based or Generative models) First, the networks discussed so far give equal importance to all spatial locations and channels. In general, giving selective attention to different regions in an image can give much better results. Second, the networks optimize the pixel difference between predicted and output HR images. Although this metric works fine, it is not ideal; humans don't distinguish images by pixel difference, but rather by perceptual quality. Generative models (or GANs) try to optimize the perceptual quality to produce images which are pleasant to the human eye.

6 What is an unsolved problem on the topic most interesting to you to solve and why?

We often see images from the cameras in extreme environments such as in space or small cameras such as the ones in medical surgical applications that they are often blurry and very low resolution. One way of tackling this problem is using better cameras but there is a limit to the amount of light a camera sensor can sense. A much better way in my opinion is to focus on super resolution for on-board computers which will involve writing not only more robust and efficient algorithms but also lighter algorithms so that they can run on slower hardware. I would like to work on algorithms that help these areas of surgical medical application or algorithms for low powered computers that take pictures in space such as the curiosity rover.

References

[1] Image Super-Resolution Using Deep Convolutional Networks, https://arxiv.org/abs/1501.00092

[2] Accelerating the Super-Resolution Convolutional Neural Network, https://arxiv.org/abs/1608.00367

[3] Accurate Image Super-Resolution Using Very Deep Convolutional Networks, https://arxiv.org/abs/1511.04587

[4] Deeply-Recursive Convolutional Network for Image Super-Resolution, https://arxiv.org/abs/1511.04491

[5] A Novel and Effective Image Super-Resolution Reconstruction Technique via Fast Global and Local Residual Learning Model , https://www.mdpi.com/2076-3417/10/5/1856

[6] Single Image Super-Resolution Using Dual-Branch Convolutional Neural Network, https://ieeexplore.ieee.org/document/8588998