IMPROVING GENERATIVE ADVERSARIAL NETWORK FOR SUPER RESOLUTION

Dhaval Parmar

Artificial Intelligence, Indian Institute of Science

ABSTRACT

Generating high-resolution (HR) image from its corresponding low-resolution (LR) input image is referred to image super-resolution (SR). In 2016, Ledig et al. proposed Generative Adversarial Network for Super Resolution (SRGAN)[1] to generate perceptually better HR images. Taking inspiration from their work, I have implemented SRGAN. I have also tried different model architecture and training strategy to get better results.

***Index Terms*—** Super Resolution, GAN

1. INTRODUCTION

Image SR has shown significant performance in video surveillance, remote sensing etc. Recently Deep convolution networks [2] [3] has shown significant performance in image SR but it is not be able to capture finer texture details. While SRGAN is able to generate texture rich realistic images. GAN drives the reconstruction of image towards the natural image manifold producing perceptually more convincing images.

2. TECHNICAL DETAILS

2.1. Architecture Details

Initially, I have considered same architecture as shown in figure 4 of [1]. For generator network 16 residual blocks are considered. I have also tried different architecture for both the model to make Updated SRGAn(USRGAN). I have removed Batch normalization layer to get better texture details and brightness. I have made residual block more deeper by repeating same block 4 times. To effectively utilise features learned in initial layers, I tried to make connections more denser using skip connection in similar fashion as in U-Net [4]. For Discriminator I have considered Global Average Polling on output of final CNN layer to get better performance.

2.2. Loss function

I have considered mean squared error(MSE) loss for pretraining of generator model.For discriminator I have used adversarial loss and for generator perceptual loss which is weighted combination of MSE, adversarial and VGG feature based loss given by equation 2 to 6 in [1]. For USRGAN wights are 0.1, 0.01 and 0.06 respectively.

2.3. Implementation Details

I have considered up-sampling factor of 4. DIV2K bicubic data set having 800 training and 100 validation data is used for experiments. Generator is pretrained for 200 epochs to build SRResNet. I have trained GAN models for 1000 epochs. Randomly cropped image with size of 96 is considered for training GAN. I have used lower learning rate for discriminator and two time scale update rule for efficient training. I have used features at 28^{th} layer of VGG16 model that was before activation function to compute generator loss.

3. RESULTS

	SRResNet	SRGAN	USRGAN
PSNR (DB)	24.0724	22.82	19.54
SSIM	0.7245	0.6557	0.5699

Table 1. Comparison Results on validation data

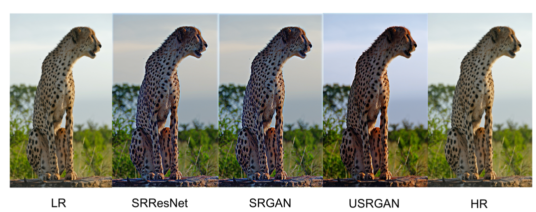

Fig. 1. Loss variation for USRGAN

PSNR and SSIM are reduces for SRGAN and USRGAN but it does not capture perceptual information efficiently. From images we can see that contrast and perceptual quality has improved for USRGAN.

4. CONTRIBUTIONS

I have implemented SRGAN from scratch. Tried modifying generator and discriminator network to improve performance. I have also adopted better training strategy for GAN. Tried different combination for perceptual loss to get better results.

5. RESOURCES

- https://data.vision.ee.ethz.ch/cvl/DIV2K/ (Data Set)
- My Code : https://github.com/DhavalParmar61/SRGAN.git

6. REFERENCES

[1] F. Husz´ar J. Caballero A. Cunningham A. Acosta A. Aitken A. Tejani J. Totz Z. Wang C. Ledig, L. Theis and W Shi, "Photo-realistic single image super-resolution using a generative adversarial network," arXiv:1609.04802v5, 2017.

[2] K. He C. Dong, C. C. Loy and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proceedings Title*. European Conference on Computer Vision (ECCV), 2014, pp. 184–199.

[3] K. He C. Dong, C. C. Loy and X. Tang, "Image super-resolution using deep convolutional networks," in *Proceedings Title*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, vol. 38(2), pp. 295–307.

[4] P. Fischer O. Ronneberger and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," arXiv:1505.04597v1, 2015.