Single Image Super Resolution Using a Generative Adversarial Network

PRATEEK KESHRI
CSE DEPARTMENT
PES UNIVERSITY

Md. SHIRAZ KHAN
CSE DEPARTMENT
PES UNIVERSITY

KUNAL MISRA
CSE DEPARTMENT
PES UNIVERSITY

Dr. S NATARAJAN
PES UNIVERSITY

***Abstract*—In Single Image Super Resolution we have Super-Resolution Generative Adversarial Network (SRGAN) that is capable of generating realistic texture. This is especially useful for the enhancement of an image by working on specific characters of an object, ex face. In this paper, we describe deep learning of understanding the machine processing for upscaling a low image resolution by 25-30PSNR (Peak Signal to Noise Ratio). The underlying model is based on the SRGAN and the enhanced SRGAN with its three key components network architecture, adversarial loss, and perceptual loss.**

***Index Terms*—Image Super Resolution, ESRGAN**

I. Introduction

In a Single Image Super-Resolution, we focus on making the low-resolution image to a higher resolution by upscaling up to 25-30 PSNR (Peak Signal to Noise Ratio). The need for higher resolution is for a better understanding of pattern recognition and analysis of an image.

Various techniques have been put forward to enhance the visual quality such as perceptual loss (optimize resolution avoiding pixel space), the introduction of an adversarial network (encouraging more natural image), semantic image prior (improve texture details that are to be recovered), use of residual blocks (improves structure by providing higher capacity and its easier to upskill). The model is being improved in three aspects as there is a clear existing gap between the lower resolution image and the result acquired using Super Resolution Generative Adversarial Network (SRGAN).

1. Residual-in-Residual Dense Block (RRDB)
   - ➔ Residual scaling is used to remove Batch normalization.
2. Reativistic Average (RaGAN)
   - ➔ Compares the quality of the images.
3. VGG features
   - ➔ Before activation of SRGAN enhances perceptual loss.

In order to reach our main objective of acquiring a higher PSNR (Peak Signal to Noise ratio), a higher frequency, in general, a higher resolution image we propose a network architecture and mention the image perceptual loss. This can be explained by the proposed method giving us a balanced perceptual quality that can be explained by the network interpolation strategy.

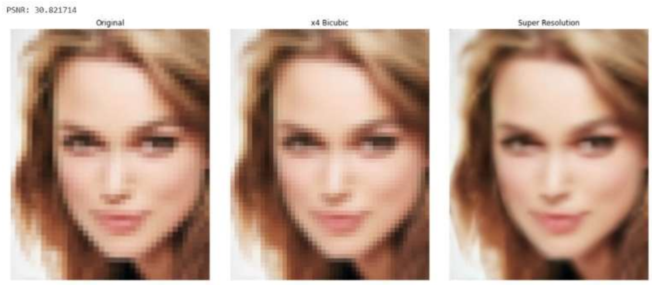


Fig. 1: Super Resolution of a Face

The above fig[1] is the Low-resolution image that we chose to demonstrate upliftment to a higher resolution. An example is a face that has a dimension of 185 x 232 pixels. The peculiarity of the image is concluded by regulating the output image which we gave rise to, to that of the original lower resolution image. the face upliftment could be observed by looking at the upliftment of the nose, mouth, and ear and comparatively a clearer image than the prior image before we processed the image.

II. Related Work

By concentrating on solving Super Resolution scenarios, we use the approach of deep neural network that proposes a Super-resolution neural network to achieve an end-to-end manner of higher performance compared to using preceding regular approaches. We on observing the deeper networks have seen variations of network architectures and few of them were-

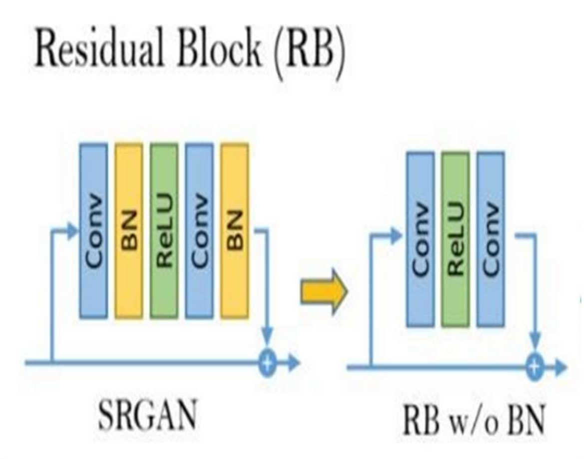
1. Laplacian Pyramid
2. Residual Dense Network
3. Residual Block

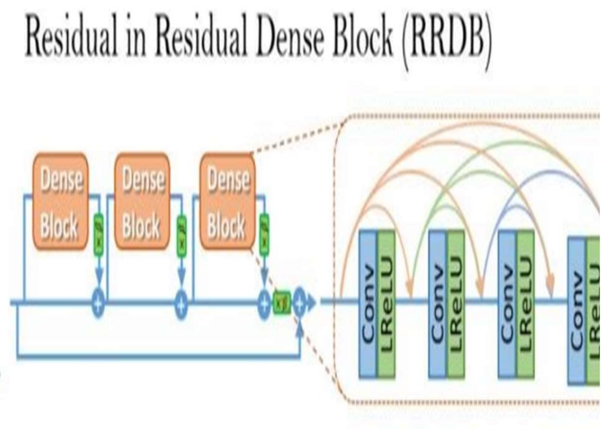
The network architecture mentioned above could be resolved easily by proposing enhanced Deep residual networks (EDSR) which basically functioned in removing unnecessary Batch Normalization by layers by expanding the size of the model to provide a deeper network we use the functionality of residual in a residual dense block which improves the perceptual quality. The models so obtained using SRGAN contain the image's perceptual loss and adversarial loss providing numerous obtained outputs for the natural images

SR algorithms are used for measuring PSNR and SSIN (Structural Similarity Index Measure), providing us with the perceptual index and helping us in understanding the distortion and perceptual quality

III. METHODOLOGY

1) Network Architecture





Network Architecture mainly consists of two modernization of its generator:

1. Dispersing of all BN (Batch Normalization) layers
2. Implementing residual in residual dense block (RRDB) and replacing it with basic blocks.

The benefits so acquired by the removal of batch normalization layers were so that it increased performance and reduced computational complexity and turn out to be useful for deep learning as well. The removal of BN layers was important as they bought limit to the generalization ability and unpleasant artifacts, these artifacts tend to violate the stability of upscaling our SR image, hence the removal of BN layers provides a consistent performance along with memory usage.

The proposal of RRDB took place when it was observed that for more layers and connections the performance could be boosted and hence RRDB being more complex and deeper than basic blocks in SRGAN are more useful

By exploiting even more several techniques we can upscale our deep network by:
-residual scaling (multiplication of constants between zero and one to the main path to increase the stability)
-smaller initialization (Easier to train when the architecture becomes smaller)

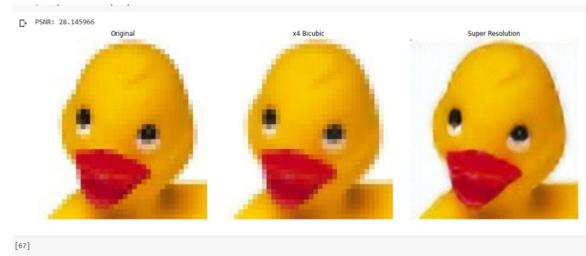
2) Perceptual Loss

The main use of Perceptual Loss is constraining on features before activation. We introduce this method for using features before activation layers as on observation it helped us to overcome limitations.

1) The activated features were very scattered after being introduced to a very deep network. This scattering of activation led to blurry and weak performance.
2) Using of features after the activation layers tend to show inconsistency and ambiguous behavior of the brightness of the super-resolution of the acquired image compared to the ground truth image.

A better perceptual loss for a super-resolution image as we use the fine tune VGG network for recognition. Although on observing perceptual loss it came to our notice that focusing on texture is extremely important for super-resolution images.

Sample outputs



References-

1. Xintao Wang, Ke Yu , Shixiang Wu , Jinjin Gu , Yihao Liu , Chao Dong, Chen Change Loy , Yu Qiao , Xiaoou Tang: "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks": University of Cambridge.

2. Armin Mehri , Parichehr B.Ardakani, Angel D. Sappa"LiNet: A Lightweight Network for Image Super Resolution" Computer Vision Center, Edifici O, Campus UAB

3. Royson Lee, Łukasz Dudziak, Mohamed Abdelfattah, Stylianos. Venieris, Hyeji Kim, Hongkai Wen, and Nicholas D. Lane: "Journey Towards Tiny Perceptual Super-Resolution'. Samsung AI Center, Cambridge.

4. Ben Niu, Weilei Wen, Wenqi Ren, Xiangde Zhang, Lianping Yang, Shuzhen Wang ,Kaihao Zhang Single Image Super-Resolution via a Holistic Attention Network.

5. Zhen Li,Jinglei Yang,Zheng Liu,Xiaomin Yang, Gwanggil Jeon,Wei Wu: "Feedback Network for Image Super-Resolution."

6. Tao Dai, Jianrui Cai, Yongbing Zhang, Shu-Tao Xia, Lei Zhang."Second-order Attention Network for Single Image Super-Resolution"