Perceptual Image Super Resolution Using Deep Learning and Super Resolution Convolution Neural Networks (SRCNN)

Nagaraj P [a][1], Muthamilsudar K [a], Naga Nehanth S [a], Mohammed Shahid R [b] and Sujith Kumar V [b]

[a] *Assistant Professor-CSE ,* [b]*UG Scholar-CSE,*
[a,b] *Kalasalingam Academy of Research and Education, India*

Abstract. The main objective of Perceptual Image Super Resolution is to obtain a high resoluted image from a normal low resolution image. The task is very simple that we just want to make a Low firmness appearance into a extraordinary resolution image. To perform this task we have various methods like Classical Approach in which we try to maximize the mean squared error, evaluate by PSNR(Peak-Signal-to-Noise-Ratio). The first method used to perform this operation was SRCNN (Super Resolution Convolution Neural Network) and these days many of them use DRCN and VDSR which are slightly upgraded methods. Another technique used for the purpose of upscaling to get a high resoluted image from normal little resolution image is the state of art by PSNR. This method was a quite simple one in which we take a low determination image as input and place in a convolution neural network(CNN) and produce a high resolution image as the output. In this technique the edges will be clearly defined, but the whole image will be blurred. This method is unable to produce good-looking textures.

Keywords. Image Resolution,PSNR, SRCNN, Padding, Pooling (Max pooling and Average Pooling, Low Resolution, High Resolution, Convolution Neural Networks).

1. Introduction

In order to know about Deep Learning, one must have idea about Artificial Intelligence and Machine Learning which are the most booming technologies across the world. Deep Learning is a subset of Machine Learning which is again a subset of Artificial Intelligence. Artificial Intelligence is a technique which enables a machine to mimic human behaviour. Machine Learning is technique to achieve Artificial Intelligence through algorithms trained with the data and Deep Learning is type of Machine Learning motivated by the structure of humanoid brain. In terms of humanoid brain, the structure is called Artificial Neural Networks [4]. Artificial Intelligence is accompanied by the study of how human brain thinks and how humanoid brain decide and work

---

[1] Nagaraj P, Assistant Professor-CSE, Kalasalingam Academy of Research and Education, India; Email: nagaraj.p@klu.ac.in

while resolving a problem or task. There are millions of applications of AI, some are speech recognition, Understanding NLP, Image Detection [1-2]. Single Image Super Resolution(SISR) is one of the most complicated and challenging problems which tries to get a extraordinary determination image from a low determination image. In previous methods, we have some disadvantages like the edges of the image can clearly defied but the whole image will be burred, the other one is blurred definition of mapping which we try to develop in between the spaces of great firmness and low firmness. The Single Copy Super Firmness methods which are based on Deep Learning algorithms achieved a lot of improvements both quantitatively and qualitatively. The process of creating a high resolution copy from a low-slung firmnessimage is referred to as SISR (Single Image Super Resolution) [3-5]. Based upon the number of low resolution images given as inputs, the Super resolution can be categorized into two ways. If only one image is given as input, it is called SISR and if the input has multiple images, it is called MISR (Multiple Image Super Resolution). When we compare SISR and MISR, SISR is more famous due its efficiency in upscaling and space definition between the pixels of the image. So most of the people use SISR to convert a LR image into a HR image [3]. By the use of DL techniques in SISR, we can overcome many drawbacks such as estimating above maximum values and also we can estimate the below minimum values. Due to the effects of smoothening, the fault lines cannot be presented well. SPLINE interpolation will not work when sample points are nearer and have high differences [6].

2. Related Work

There are some techniques through which we can perform the task of getting extraordinary determination image from a low determination image. They are:

2.1. Residual Learning

Recent researches state that the Residual Learning technique has increased and improved its performances on SR. This residual learning concept includes the residual blocks [8] for the purpose of image recognition. This concept is mainly used in solving the computer vision problems and also the classification and detection of the image.

3. Proposed Work

In pre-processing, first we have to upscale the image to the desired size by using bi-cubic interpolation technique. Let us consider an interpolated image and denote it as 'Y'. Now our main aim is to improve from Y an image F(Y), that is similar as possible as ground truth Great-Determination Image X. Let us take F(Y) as Y^. Now let us find the error term E, which is mean square error between Y and Y^ [7-9]. In this process the main aim is the estimation of High Resolution(HR) image ISR (Interpolated Great Determination) of a given Low Determination image that is downscaled from the respective High Resolution(HR) image IHR (Interpolated High Resolution). In this architecture the input will be a single image with a bi-cubic interpolated low resolution image that consists of RGB channels of same size as the High Resolution image. There are three layers in the proposed architecture namely convolution layer #1, convolution

layer #2, convolution layer #3 between bi-cubic interpolated LR image as input and provides the output as an image with HR by reducing the PSNR value for image and meet the state of art techniques [6].
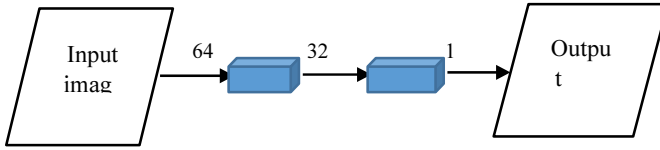

Figure 1. Proposed Architecture

The main objective of this architecture is taking a LR image as input and in the second layer of the architecture, the image will be zoomed in or pinched where it loses its pixel and increases the PSNR due to improper definition of spaces and undefined edges [1-3]. In the above process we find the mean squared error by using the below equation,

$$MSE = \frac{1}{nm}\Sigma_i^n\Sigma_i^m(L_{ii} - 1_{ii})^2 \qquad (1)$$

Now we evaluate the PSNR by using the below equation [2],

$$PSNR = -10 \log 10 \ (MSE) \qquad (2)$$

In many algorithms the main drawback in upscaling is sudden decrease in noise that is PSNR value. There must be a gradual decrease in the PSNR value in order to define the edges clearly without any blurredness in the image. This will be a great advantage in order to get the pinched or zoomed out image without losing its pixels and edges

4. Experiments

4.1. Data Sets used

We have used a trained dataset consisting of large number of images to be taken as inputs which already trained by using different techniques. The our trained images will be taken as inputs and run under the tensor flow backend which passes the image from first convolution layer to the second layer. The operation between these two layers is, the image will be separated into three colours like Red, Green, Blue. In our dataset, the images consists of different such colours like RedGreenBlue(RGB), HSV, CMYK, etc. The purpose of second Convolution layer is to reduce the loss function of the image by using 32 1*1 filters. The key purpose of the third convolution layer uses 1 5*5 filter in order to reduce the signal and noise of the image completely. The output can be seen as three images where the first image is the input image and then the pinched or zoomed in image and the final image will be the output image which has low PSNR [11] values and looks like a high resoluted image.

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Image

4	3	4
2	4	3
2	3	4

Convoluted Feature

From the above figures, we can notice that the dimensions of image be 5x5x1 which has height 5, width 5 and channels are 1. To perform the operation of convolution in the convolution, we multiply the first part of convolution by Kernel (K).

1	0	1
0	1	0
1	0	1

kernel Filter(K)

4.2. Padding

The main aim of the CNN layer is to get high-level geographies like edges and color positioning etc..The results can be obtained by applying the techniques of padding called Valid padding and Same Padding [9-11].In case of Same padding, we convert a 5x5x1 image into 6x6x1 image and perform 3x3x1 kernel to it which results in 5x5x1 matrix image. This process is known as same padding.In case of Valid padding, we can perform the same operation without padding in which the result would be the 3x3x1 kernel. This process is called Valid Padding.

4.3. Pooling

As similar to the convolution, the pooling layer is for decreasing the spatial size. By the use of pooling, we can reduce the computational control that is required to progression of the data [10]. There are twofold kinds of pooling techniques. They are Max pooling and Average Pooling. In case of max pooling, we return the extreme value from the portion covered by the kernel where as in Average pooling, the average of the values of portion covered by kernel will be returned.

4.4. Max Pooling

Let us assume an image of matrices of 4x4x1 as follows:

8	12	0	2
20	12	30	0
34	112	13	25
100	70	5	37

In max pooling, we return the maximum values from the portion covered by the kernel K [10]. Hence the outcome of the above image using max pooling is as follows,

20	30
112	37

Max Pooling

4.5. Average Pooling

In the process of the average pooling, the average value of the portion covered by the kernel will be returned as the outcome. Hence the output would be as follows.

13	8
79	20

Average Pooling

The process done in average pooling is, calculation of the average values. They are:
From the first portion covered by kernel -----→ 8+12+20+12/4= 13
The above calculation is the average of the first portion covered by kernel
K→ 0+2+30+0/4 = 8
The above calculation is the average of the second portion covered by kernel
K →34+112+100+70/4 = 79
The above calculation is the average of the third portion covered by kernel
K→ 13+25+5+37/4 = 20

5. Results

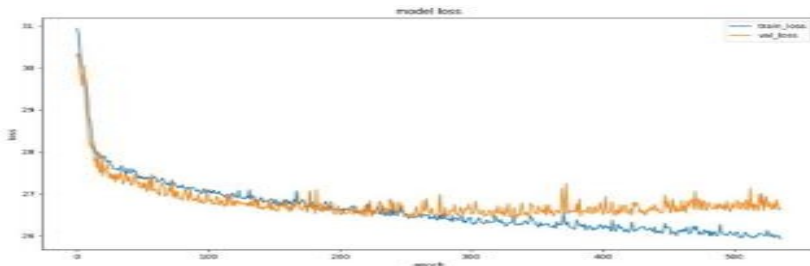Normal Image   Zoomed Image   Resolution Image

Figure 2. Model Loss Plot

MSE	1.254
PSNR	31.567

5.1. Demonstration

In this process, we obtain LR images by down sampling the image using bilinear kernel using a down sampling factor [6]. In the above architecture, the first convolution layer operates given input image by using 64 9x9 filters.The second layer is a pooling layer, where we perform Max pooling and Average Pooling using Kernel K portions [6]. Finally, the output image will be same size as the given input image. The size of the output will not differ with the given Low Resoluted input image.

6. Conclusion

Introduce a combination of loss functions for single image super resolution which will help in gaining better super resolution perceptually. Perceptual loss can be created by using intermediate VGG architecture layers. State of art in both quantitative and qualitative benchmark has been reached. Super resolution can be used for gaining better results for satellite images and MRI scans in medical field.

References

[1] Johnson, J., Alahi, A., & Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. 2016 October. In European conference on computer vision Springer, Cham. p. 694-711.

[2] V.D.Ambeth Kumar, Ramakrishnan, S.Malathi and V.D.Ashok Kumar Performance Improvement using an automation system for recognition of multiple parametric features based on human footprint. 2016 October. Kuwait Journal of Science, 42(1), p.109-132.

[3] Dong, C., Loy, C. C., He, K., & Tang, X. Image super-resolution using deep convolutional networks. 2015 IEEE transactions on pattern analysis and machine intelligence, 38(20), p.295-307.

[4] Choi, J. H., Kim, J. H., Cheon, M., & Lee, J. S. Deep learning-based image super-resolution considering quantitative and perceptual quality 2019. Neuro computing. Dai, D., Wang, Y., Chen, Y., & Van Gool, L. Is image super-resolution helpful for other vision tasks?. 2016 March. In 2016 IEEE Winter Conference on Applications of Computer Vision (WACV) IEEE.p. 1-9.

[5] Dong, C., Loy, C. C., He, K., & Tang, X. Learning a deep convolutional network for image super-resolution 2014 September. In European conference on computer vision Springer, Cham. p. 184-199.

[6] Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J. H., & Liao, Q. Deep learning for single image super-resolution: A brief review. 2019 IEEE Transactions on Multimedia.

[7] Yan, B., Bare, B., Ma, C., Li, K., & Tan, W. Deep Objective Quality Assessment Driven Single Image Super- Resolution.2019. IEEE Transactions on Multimedia

[8] Niu, Z. H., Liu, L. F., Zhang, K. J., Dong, J. F., Yang, Y. B., & Mao, X. J. Single Image Super-Resolution via Perceptual Loss Guided by Denoising Auto-Encoder 2019 July. In Pacific Rim International Conference on Artificial Intelligence Springer, Cham. p. 126-136.

[9] Yang, G., Cao, Y., Xing, X., & Wei, M. Perceptual Loss Based Super-Resolution Reconstruction from Single Magnetic Resonance Imaging 2019 July. In International Conference on Artificial Intelligence and Security Springer, Cham. p. 411-424.

[10] Pandey, R. K., & Ramakrishnan, A. G. Improving the Perceptual Quality of Document Images Using Deep Neural Network 2019 July. In International Symposium on Neural Networks Springer, Cham. p. 448-459.

[11] Ambeth Kumar S. Malathi R. Venkatesan K Ramalakshmi, Weiping Ding, Abhishek Kumar "Exploration of an innovative geometric parameter based on performance enhancement for foot print recognition", Journal of Intelligent and Fuzzy System, vol. 38, no. 2, pp. 2181-2196, 2020.