IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-7, July 2018 E-ISSN: 2347-2693

Image Super-Resolution Using Deep Learning Technique

Nisha Singh[1*], Myna A. N[2]

[1*]Dept. of ISE, Ramaiah Institute of Technology, Bangalore, India
[2] Dept. of ISE, Ramaiah Institute of Technology, Bangalore, India

Corresponding Author: nishamili8@gmail.com

Available online at: www.ijcseonline.org

Accepted: 13/Jul/2018, Published: 31/July/2018

***Abstract*—** With recent advancement in deep learning areas, computer vision research has changed from hard coded features to end-to-end trained deep neural network. Super-resolution is one of such areas which is influenced by deep learning advancement. Super-resolution is the technique for reconstructing high-resolution images from a given set of images. It is very important to acquire better quality images in satellite images, medical images and surveillance monitors where analysis of low quality images is extremely difficult. In this paper a novel approach to solve the problem of super-resolution image is presented. Proposed method trained the network using feedforward convolutional neural network and combined with perceptual loss function which measure the semantic differences between images and helps in reducing the computational complexity of overall super-resolution images. The proposed method also uses the adversarial network which helps in achieving the finer details in images.

***Keywords*—** Super-Resolution, Convolutional Neural network, Sub-Pixel Convolutional Layer, Perceptual Loss

I. INTRODUCTION

Digital imaging system requires better image quality of high-resolution. The recovery of high resolution images is one of the main concerns of digital image processing. It has many application areas where we need high-resolution images like medical imaging, satellite imaging, face recognition, surveillance monitors. Some surveillance system devices do not produce high-resolution images. Super resolution images are used in image processing because low-resolution images raise fundamental issues while examining the relevant information. Due to the blurred and noisy version of images analysis become extremely difficult. Many methods predict that multiple images of same scenes are available in low-resolution but having different perspective. Super-resolution image techniques learn implicit redundancy which recovers missing high-resolution information from low-resolution instance that is present in natural data. The reconstruction of super-resolution image is gaining the attention in research areas because it helps in overcoming the resolution limitation of image system.

Deep learning method solves the problem of super-resolution by providing better quality of images. Recently several methods of deep neural network have achieved great success but there are some drawbacks in terms computational complexity and performance.

Among them super-resolution methods based on Super-resolution convolutional neural network(SRCNN) is faster but it uses bicubic interpolation to upscale resolution of images which increases the computational complexity. To overcome these problems the proposed approach uses feed-forward convolutional neural network to achieve super-resolution image and perceptual loss function which measures the differences of high level feature extracted from the network instead of differences among the pixels of image. It uses sub-pixel convolutional layer where upscaling operation is performed in feature map extracted from low-resolution images and learn upscaling filter to upscale the high-resolution output. Convolutional neural network produces the reconstructed image but it losses some finer texture details. Proposed method also uses the adversarial network to solve this problem depending on deep network with skip connection and discriminator which generates high-quality images and make it hard to distinguish the output image with original high-resolution image.

This research paper is organised in the following manner. Section I contains the introduction of super-resolution images, Section II contains related work of proposed topic, Section III describe the proposed method, Section IV describe the experimental analysis, Section V describe the result achieved by proposed method, Section VI describe the conclusion with future works.

II. RELATED WORK

The main goal of super-resolution method is to achieve high-resolution images from low-resolution input images. Recent popular super-resolution methods have been classified into interpolation based methods, edge based methods, statistical based methods and patch based methods[1],[2].These methods are carefully investigated.

A. Interpolation based super-resolution methods

Interpolation based methods includes bicubic and nearest neighbour. While these filtering approaches can be be fast, they oversimplify the single image super-resolution problem[3]. These methods provide high-resolution image from the low-resolution image which is based on sampling theory. These methods perform a simple weighted sum operation at the interpolated image grid to estimate pixel values and assign the values to output coordinates. Such approaches are not effective as they blur the high-frequency information. Another interpolation technique is nearest neighbor[4]. It is one of the simplest interpolation techniques among all. In this method output pixel value after interpolation is assigned to nearest sample point of input image and used for image scaling. Although nearest neighbour method is effective but quality of image is not good.

B. Edge based super-resolution methods

Edge based methods produces high-resolution images by applying prior knowledge on the upsampled image, while performing the super-resolution process the main task is to preserve the edges information. The performance of such methods depends on prior information and compatibility of given image. According to use of Non Sub-sample Contourlet Transform based Learning approach (NSCT) is to increase the size of images[5]. This helps to decompose a low resolution image directionally as per the directional information it constitutes. Thus, information at minute orientations as well as hard and curved edges are conserved.

C. Statistical methods based super-resolution methods

Statistical methods are used to estimate high-resolution details from the large training set, In this method the information is learned from training data[6]. Based on the similarities between high-resolution and low-resolution images training set this method is effective to provide missing details. Single based super-resolution method is also based on sparse signal representation and considered two dictionaries patches one for low resolution and another for high resolution[7]. The appropriate patch base is selected by computing sparse representation which represents the patch of the given low resolution image. Sparse model builds the dictionary based on mathematical model of the data. But main drawback of using this method is model itself is restrictive, its flexibility and performance is low.

D. Patch based super-resolution methods

Patch based method takes the high-resolution and low-resolution image pairs and takes each patch which means one chunk of image at a time to learn the mapping functions[8]. Among all the methods generally patch based methods perform better[9]. Recent methods of super-resolution provide better performance. It uses three convolution layers and bicubic interpolation to learn the mapping among low-resolution and high-resolution images. It upscales the low-resolution images before feeding to network and uses the loss function based on pixels to measure the differences of reconstructed image from original image but it increases the computational complexity.

III. METHODOLOGY

The proposed method uses feed-forward convolutional neural network and adversarial training approach to achieve super-resolution images.

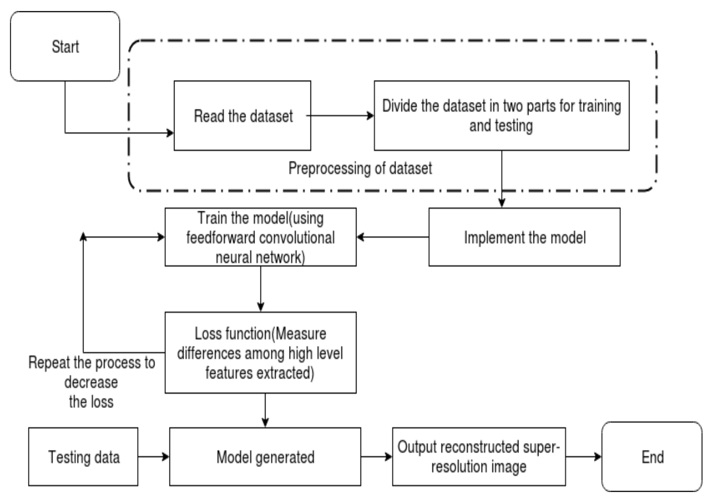


Figure 1. System Overview of proposed model

Proposed method estimates high-resolution image(I_{HR}) from the given input low-resolution image(I_{LR}) . I_{LR} is the low-resolution version of high resolution image I_{HR} and it is downsampled from original image high-resolution images. Only high-resolution images have been used for training. During training gaussian filter operation is used to obtain low-resolution images then downsampling operation is performed on resultant image with upscaling ratio r. Both low-resolution image and high-resolution image are have color channel C and size $W * H * C$ and $rW * rH * C$ respectively.

This method varies from other techniques as we do not use interpolated low-resolution image before feeding to the network. In the proposed network during the first layer convolutional operation is applied directly to the low-resolution image for feature map extraction. Then sub-pixel convolution layer is applied on next layer that upscales the feature map to produce reconstructed high-resolution image.

The network consists of L layers, the first layer L_1 is represented as follows:

$$L_{1,}(I_{LR},W_1,b_1) = \varphi W_1 * I_{LR} + b_1 \quad (1)$$
$$L_{m-1}(I_{LR},W_{m-1},b_{m-1}) = \varphi(W_{m-1} * L_{m-1}(I_{LR}) + b_{m-2}) \quad (2)$$

Where W_m and b_m are learnable network parameters. W_m is the weight of size $n_{m-1} * n_m * k_m * k_m$ where n_m is the number of features at layer and k_m is the filter size at each layer. φ is a activation function which is applied element-wise.

1. *Sub-Pixel Convolutional Layer*

The upscaling process takes place with filter weight and a convolution with a stride of 1/r for activating different parts of filters during the convolution operation [10]. Filter that falls between the pixels cannot be used for computation. The number of activation patterns is r^2. The activation patterns are activated by applying filter to an input image which depend upon different sub-pixel location during convolution operation and it is representing the modules of output pixel as scale factor r.

2. *Loss Function*

Loss function is basically used to guide the training process in neural network. Pixel wise loss function cannot be used as large scale factor low-resolution image tends to be blurry and lack of high frequency details. Proposed method uses perceptual loss function which can measure high level features extracted from images [11]. It makes use of loss network of deep convolutional neural network which is based on VGG-16 loss network.

2.1 Loss in reconstructed features

Instead of matching the pixel of reconstructed images $I^{'}_{HR} = L_1(I_{LR})$ to original images. Proposed method calculates feature representation. Let φ_i is the activation of ith layer of network where i is a convolutional layer $\varphi_i(I_{LR})$ and is a feature map extracted which is having shape of $C_i * H_i * W_i$. Loss in reconstructed features can be calculated as Euclidean distance between represented features:

$$L^{\varphi}\ (I^{'}_{HR,}I_{HR}) = 1/(C_i * H_i * W_i)\sqrt{\varphi}(I'_{HR}) - \varphi(I_{HR}) \quad (3)$$

To minimize the feature reconstruction loss network aims to produce reconstructed images identical to target images in terms of features. During the reconstruction of features when layer increases the contents of images and structures are conserved but as we increase layer shapes, colors, texture details are not conserved. So use of feature reconstruction loss for training the network allows the reconstructed images to be perceptually similar to original images.

2.2 Adversarial loss

Proposed method computes adversarial loss of the network. To generate the images with high perceptuality, GAN (generative adversarial network) based network is optimized and it uses ResNet architecture [12]. The main reason of this technique is that, it allows one to train the generator model that is able to fool the discriminator by producing super-resolution images which is indistinguishable from high resolution images. Discriminator is trained to differentiate super-resolved reconstructed images from original images [13]. To improve perceptual quality of upscaled image the adversarial loss helps to generate upscaled image that is similar to natural image.

$$y_{adversarial} = \sum_{n-1}^{N} - \log D(G(I_{LR})) \quad (4)$$

Where $D(G(I_{LR}))$ is the probability that super-resolved image is natural high-resolution image.

IV. EXPERIMENT

The super-resolution method is implemented using python. All the function of the proposed method is written in python using theano and lasagne deep learning libraries with additional libraries numpy and scipy.

A. Datasets

Training datasets are available at [14] consists of 20,813 images. Set2, Set14 and BSD100 are used as test images. During training first image are converted to yCbCr format and luminance channel in yCbCr is used. The training set is used to build the model and adjust the weights on the neural network. Performance of super-resolution method can be improved using large training set. If the performance is not good process has to be repeated.

B. Training details

Model is trained using the training images. During the training phase sub-images are formed into 16r*16r pixels which are extracted from original images. During the training process low-resolution samples are formed after blurring the images with gaussian filter of sigma 1.0. Model is trained to generate super-resolution image with scale factor 2 and 4 by minimizing the loss occurred while reconstructed features at layer relu2_2 of VGG-16 loss network. Network learns some phase filtered at each layer. During the first layer filtered learned the following features like gaussian, texture detector, edge detector and shape etc. Every further layer extracts the feature map with more details. We train the model with batch size of 4 for every 10 epoch. Optimization of network for low-resolution image converges faster than high-resolution image. For better training rate adam optimizer has been used to provide better learning rate of $1 * 10^{-4}$ without dropout. All the layers are having same learning rates. Training stops after 50 epochs.

C. Testing

Proposed model reconstructs super-resolution image for any random size of image. Testing images in model are used to compare the performance of existing methods as compared to proposed method. Set2, Set14 and BSD100 are used as test images. Downsampling of feature map reduces the size but upsampling gains the original size of images back which is responsible for good testing efficiency. Here input to the system is low-resolution image and it produces output as reconstructed high-resolution image.

V. RESULTS AND EVALUATION

Proposed method has been tested on images taken from set2, set14 and BSD100. Proposed method is compared with Bicubic and SRCNN(Super-Resolution Convolutional neural network) methods. Here first we are comparing the original images with Bicubic and SRCNN methods and results are tabulated in the Table1 and Table 2. Then we are comparing original image with our proposed method and results are tabulated in Table1 and Table2. From these tabulated results, we found that proposed method generate high quality of images as compared to Bicubic and SRCNN.

We are using PSNR and SSIM as the performance metrics to evaluate the visual quality of super-resolved image. Result is shown in Table 1 and Table 2.

$$\text{PSNR} = 10\log_{10} (S^2/(1/C_i * H_i * W_i\sqrt{\varphi(I'_{HR}) - \varphi(I_{HR})} \quad (5)$$

Where s is maximum possible value that exist in original image , maximum possible value is 255.

$$\text{SSIM} = \frac{(2\mu_{I'_{HR}}\mu_{I_{HR}} + C_1)\,(2\sigma_{I_{HR}}\sigma_{I'_{HR}} + C_2)}{\left(\mu^2_{I'_{HR}}\mu^2_{I_{HR}} + C_1\right)(\sigma^2_{I'_{HR}}\,\sigma^2_{I_{HR}} + C_2)} \quad (6)$$

Where $\mu_{I'_{HR}}, \mu_{I_{HR}}, \sigma_{I_{HR}}\sigma_{I'_{HR}}$ and $\sigma_{I'_{HR}I_{HR}}$ are the local means, standard deviations, and cross-covariance for images.

Original Method Bicubic SRCNN Proposed

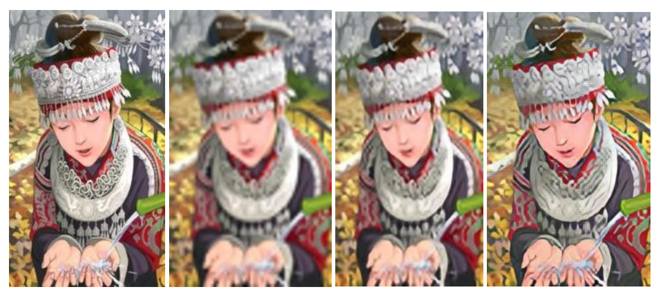
Figure 2. Result of super-resolution reconstructed image(Test image1)

Original Method Bicubic SRCNN Proposed

Figure 3. Result of super-resolution reconstructed image(Test image 2)

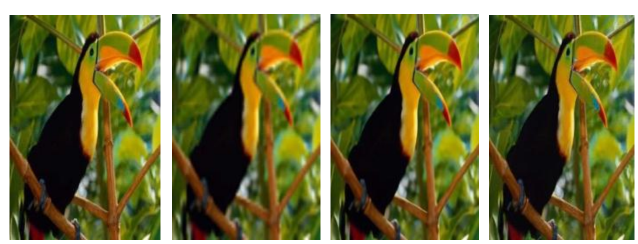
Original Method Bicubic SRCNN Proposed

Figure 4. Result of super-resolution reconstructed image(Test image3)



Original Method Bicubic SRCNN Proposed

Figure 5. Result of super-resolution reconstructed image(Test image4)



Original Method Bicubic SRCNN Proposed

Figure 6. Result of super-resolution reconstructed image (Test image5)

Results obtained after testing the model on trained model. Original image is high-resolution image and other three images are reconstructed images. In overall, Output image indicate that proposed model gives better result than other mentioned method both visually and mathematically. Proposed method has constructed high resolution images with good perceptual quality, pixelwise loss based method have generated blurry and over-smooth images.

A. Performance Evaluation

PSNR and SSIM metrics are used to evaluate the quality of super-resolution method. We have tested our model on scale factor 2x and 4x. We evaluated the model and report PSNR and SSIM values.

Table 1. Comparison of PSNR on test datasets

Dataset	Scale Factor	Bicubic	SRCNN	Proposed Method
Test image1	2	28.07	29.56	31.23
Test image2	2	30.82	32.83	33.47
Test image3	4	26.11	29.53	30.59
Test image4	4	25.78	26.89	29.99
Test image5	4	28.07	29.27	30.23

Table 2. Comparison of SSIM on test datasets

Dataset	Scale Factor	Bicubic	SRCNN	Proposed Method
Test image1	2	0.9612	0.9725	0.9843
Test image2	2	0.9770	0.9833	0.9854
Test image3	4	0.9573	0.9787	0.9821
Test image4	4	0.9475	0.9623	0.9756
Test image5	4	0.9579	0.9675	0.9699

We compare the performance of proposed approach to bicubic and SRCNN methods. Proposed model is trained for feature reconstruction having high PSNR and SSIM values as compared to other methods and performance is better as it is reconstructing sharp edges with fine details and also generating high quality images as compared to other methods.

VI. CONCLUSION AND FUTURE WORK

In this paper, super-resolution method is presented using feed-forward convolutional combined with adversarial training and perceptual loss to achieve better results. To reduce computational complexity a sub-pixel convolutional neural network is used where upscaling operation is perform in feature map extracted from low-resolution image with little computation. Evaluations performed on dataset with upscaling factor 2 and 4 and achieve high quality and performance as compare to other methods.

In future work perceptual loss functions can be explored for other image transformation tasks. One can achieve better network structure by adding more number of filters. One can go for more optimized network. We can also explore different loss networks trained on different tasks.

REFERENCES

[1] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks" IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2):295–307, 2016.
[2] Yang, C.Y, Huang,J.B., Yang, M.H,"Exploiting self-similarities for single frame super-resolution", In: IEEE Asian Conference on Computer Vision, pp.497–510 (2010).
[3] D. Glasner, S. Bagon and M. Irani,"Super-Resolution from a Single Image",Proc. Int. Conf. Computer Vision, Kyoto, Japan,2009.
[4] Shreyas Fadnavis Int, Image Interpolation Techniques in Digital Image Processing:"An Overview, Journal of Engineering Research and Applications" ISSN: 2248-9622, Vol. 4, Issue 10(Part -1), pp.70-73, October 2014.
[5] Amisha J Shah, Suryakant B.Gupta and Rujul Makwana, "Single Image Super-Resolution via Non Sub-sample Contourlet Transform based Learning and a Gabor Prior", International Journal of Computer Applications (0975–8887)Volume 64–No.18, February 2013.
[6] M. Elad. "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing", Springer Publishing Company, Incorporated, 1st edition, 2010.
[7] Yang C.Y, Ma C, Yang M.H,"Single-image super resolution: a benchmark", Springer, Computer Vision (ECCV) pages 372-386,2014.
[8] S. Schulter, C. Leister, and H. Bischof,"Fast and accurate image upscaling with super-resolution forest" IEEE, Conference on computer vision and pattern recognition, pages 3791-3799,2015.
[9] N.S.Lele,"Image Classification Using Convolutional Neural Network", International Journal of Scientific Research in Computer Science and Engineering, Vol.6, Issue.3, pp.22-26 , 2018.
[10] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network", In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1874–1883, 2016.
[11] Johnson, Justin, Alexandre Alahi, and Li Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution", European Conference on Computer Vision. Springer International Publishing, 2016.
[12] He, K., Zhang, X., Ren, S., Sun, J. "Deep residual learning for image recognition" arXiv preprint arXiv: 1512.03385 (2015).
[13] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunning-ham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network", arXiv.org, Sept. 2016.
[14] Lin, T.Y, Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L: "Microsoft coco: Common objects in context. In: Computer Vision– ECCV 2014", Springer (2014) 740–755.

Authors Profile

Nisha Singh, She received B.E degree in Computer Technology from Nagpur University. She is pursuing M.Tech in Software Engineering at Ramaiah Institute of Technology, Bangalore. Her research interest is in Image Processing and Machine Learning.

Myna *A.N* is a M.Sc (Engg) degree holder and is serving as an assistant officer in Information science and Engineering department of Ramaiah Institute of Technology, Bangalore. Her areas of interest includes Image Processing,Machnie Learning, Soft Computing.