Image Super-Resolution using Shallow Neural Network

Khagendra Kumar[1], Balakrishna Mandal[1], Yash Tejwani[1], D. Punitha[2]

[1]Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani, Chennai, India

[2]Assistant Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani, Chennai, India

Abstract

The Super Resolute (SR) model enhances the quality of a low-resolution image to a high resolution. The quality of the output image will be significantly improved and contain more details. The existing system uses deep neural network to enhance the image quality by applying the little filters many times in deep network which help get more contextual information over the low resolute image. The proposed model uses CNN (Convolutional Neural Network) to extract features from the input image and Residual patch data will be calculated. The proposed architecture uses shallow Neural Network which significantly enhances the computational time. It will extracts more information using residual patch map data to restore lost image details. The proposed model is better than the existing CNN models in terms of accuracy and computational time and it is significantly faster than existing models in terms of batch processing.

Index Terms - **Super-Resolution, Shallow Neural Network, Convolutional Neural Network**

1. Introduction

Producing a high-resolution or high quality image from a low-resolution or low-quality image is termed as Super-resolution (SR). Super-resolution are different kind of techniques that generates a high-resolution (HR) image from various observed low-resolution (LR) images, which enhances the high-frequency components and removes the degradation which is done by the low resolution camera. The single image interpolation approach is a related strategy with Super-resolution. The single image interpolation technique can also be used to increase the size of image. The quality of single image interpolation is limited as there is no extra information provided, because of the improperly posed problem, and the lost features components cannot be recovered.

A Super-resolution for single images is used widely in applications for computer vision which include security and surveillance imaging in which more details of an image are required.

C. Dong [5] was the first one who introduced a convolutional neural network (CNN) into image super-resolution and also explained that CNN can be used in order to learn mapping from Low Resolution to High Resolution. Their method is termed as SRCNN.

In this model, we proposed a new model which solve the following issues.

Convergence our model suggests a way to speed up training. The existing model is taking too much time for training because of a deep neural network and generating a low resolution image to high resolution image.

Batch processing the existing model is not capable of processing more than one image at a single time. Our model is taking a group of images and processing all the images in a comparatively faster time than existing models.

2. Related Work

So many works have been done in super-resolution using a deep neural network. These deep network has achieved so many advantages over conventional methods in computer vision.

The SRCNN model proposed by Dong at el[6] having three layers that are patch extraction/representation, non- linear mapping and reconstruction.

The model proposed by Jiwon Kim [1] construct a deep neural network for super-resolution which consists of 20 CNN layers but with high layers, the convergence rate also increases.

The residual dense network Yulun Zhang[2] which has four parts that is residual dense blocks (RDBs), up-sampling net (UPNet), dense feature fusion (DFF), shallow feature extraction net (SFEnet).

Along with this model VDSR model[3] is very efficient in generating HR images from LR images and also capable of providing highly accurate construction. It is inspired by VGG-net and it made efficient by residual learning.

Another model which uses GANs for image super-resolution is known as SRGAN[3]. It uses the deep residual network as a generator to generate HR images from downscaled input. Another model which uses Conditional GAN along with SRGANs.

Saifuddin Hitawla[3] proposed a VDSR-ResNeXt model which is inspired by ResNeXt architecture and then applied the idea of multi-branch design to VDSR. The model is made up of series of block sharing the same hyper-parameters (filter size, number of filters, number of layers).

The above mentions DL based SR network has achieved a significantly great improvement in super-resolution as compared to conventional SR techniques, but of all them has looses some useful information and features from the low-resolution input image.

3. Proposed Method

For Super Resolute (SR) reconstruction, the proposed network uses the convolutional neural network. The CNN extracts the features from input low resolute image. The proposed architecture uses the shallow neural network and passes the extracted features from CNN for Residual Patch map computation.

For a color image, in preprocessing it changes the input image from RGB to YCbCr color space. This part will isolates the luminance and chrominance of the image. The luminance of the picture will be put away in the Y-Channel and the Cb will hold chrominance estimation of blue and Cr hold the chrominance estimation of red. The Y channel holds the luminance component which is the brightness of the image and this is more sensitive to the human eye and need to be more accurate than other channels.

In order to get a low-resolution image, the input image is transformed by image augmentation in which by applying flipping, cropping and resizing the image by 256x256.

Here, the model uses the bilinear interpolation technique to upscale the image factor of 3, i.e. 3 times upscaling the factor.

For initializing the weight and values the model uses the orthogonal seed generation. It is a square matrix whose rows and columns are orthogonal unit vectors.

$$A A^T = I,$$

3.1. Shallow Neural Network

A shallow network is a neural network, which has only one hidden layer compared to less than a deep neural network. As the shallow neural network is wide which is simply memorize the train data and due to which the computation power is significantly faster. So as by increasing the depth of the model, there will be an increase in the complexity of the computation. In order to decrease the complexity of computation shallow network is used.
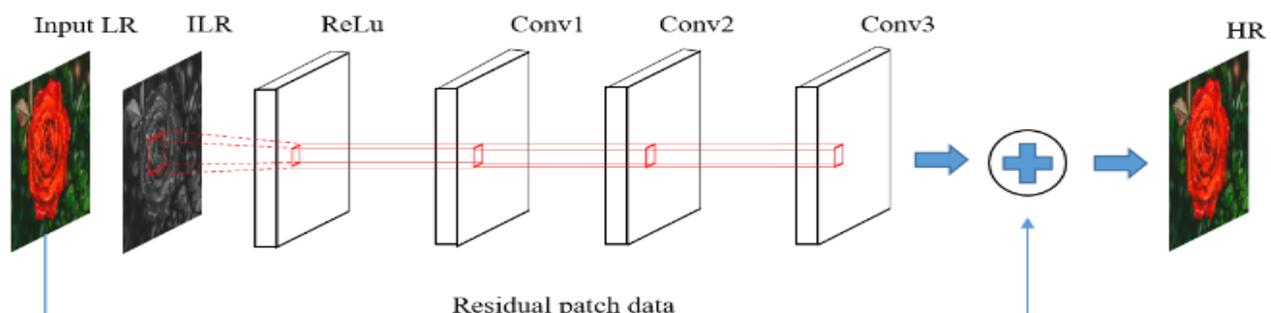

Figure: 1

Figure 1: The architecture has only 3 layers of Convolutional neural network followed by a ReLU layer. The low resolution normal RGB image is converted into YCbCr format and then passed into CNN layers and converted into HR image. Also, the network passes the residual patch data from the original image to give the desired output. The filters for each convolutional layers are in decreasing form with 64, 32 and 9 respectively for each layer.

3.2 ReLU (Rectified Linear Unit)

The ReLU is an activation function which is used in the convolutional neural network. Here, an output is zero when an input is negative and output is equals to input when an input is positive or zero.

$$f(x) = \max(x, 0)$$

ReLU activations are the simplest non- linear activation function which can be used in the neural network easily.

3.3. Training

The proposed model uses BCDS300 dataset where dataset consists of 200 images for training model and 100 images for testing model. For the training process, the model uses mean square error (MSE) for error computation.

3.3.1 Residual Patch Data

In residual patch data, the targeted image is compared with the upscale low-resolution image, the Y luminance channel of both, i.e. target image and upscale low resolute image are being compared. The residual image r = y – x is get predicted from the training data. The residual image helps to reconstruct the high-resolution image by adding the calculated residual image to the upscale low resolute image.

3.3.2 Mean Square Error (MSE)

Mean square error is a function which finds the average of squared of difference. Mean Square Error is also known as risk function. Mean Square Error is always positive.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y_i} - Y_i)^2$$

4. Understanding Properties

In this proposed system, there are two properties which help the model to perform better than the existing systems. First, this system uses the shallow neural network that helps the system to perform faster while training and also increase the computation time which is significantly faster than the existing one.

Second, in the system, the data loss is significantly lesser. The size of the feature map does not get reduced by the time when Convolutional operation is used, unlike the existing system. As in the model shallow neural network consisting of only 3 layers combined with input and the output layer due to which it reduces the data loss while training.

5. Experimental Results

The proposed model provides more detailed information with more accuracy and computation time is significantly faster than the existing models in terms of batch processing. With the upscaling factor of 3, the model is able to generate the output image of the PSNR value equivalent to 26.5.
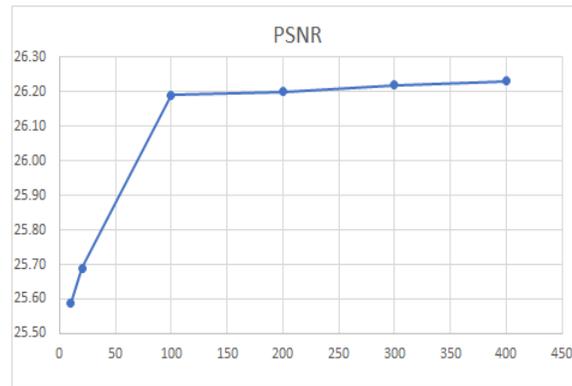
Figure: 2

Figure 2: The graph shows the PSNR value of the proposed system with number of Epoch. With increasing number of Epoch the PSNR value is increasing.

5.1. Comparison

In the existing system for reconstructing the super-resolution image very deep neural network. The existing system follows deeper the network better will the result. As depth increases the contextual information also increases. But due to increasing the depth of the network layer they acquiring a lot of memory. Also, a very deep neural network is difficult because of the slow convergence rate.

So here in the proposed model, the Shallow Neural Network is used which gives the more detailed image as output without using a deeper network and also the computation is significantly faster than the existing methods. Unlike other models in the Shallow network, the size of the feature map will not be affected while a convolutional operation is applied.
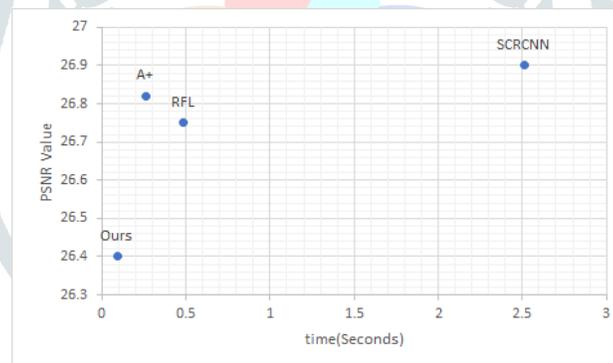

Figure: 3

Figure 3: This system provides improved computational time in compare with related models (SRCNN, RFL, A+).

6. Conclusion

In the proposed model, the super-resolution of the image using a shallow neural network method is presented. By using shallow neural network the proposed get faster in computation and using 3 layers in the network which helps to reduce the data loss while Convolution operation is used. To create low-resolution image the data has transformed the input (training) images by applying random flipping, cropping and resizing the images into 256 x 256. This proposed system is better to previously proposed CNN models. With an upscaling factor of 3, the proposed system gives the better output image.

References

[1] Kim, Jiwon, and Lee, Jung Kwon, and Lee, Kyoung Mu. "Accurate Image Super-Resolution Using Very Deep Convolutional Networks."In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.

[2] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, Yun Fu. Residual Dense Network for Image Super-Resolution. 2018

[3] Saifuddin Hitawala, Yao Li, Xian Wang, and Dongyang Yang. Image Super-Resolution Using VDSR-ResNeXt and SRCGAN 2018

[4] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. TPAMI, 2015.

[5] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In ECCV. 2014.

[6] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In ICCV, 2009.

[7] Dong, Chao, and Loy, Chen Change, and He, Kaiming, and Tang, Xiaoou. Image Super-Resolution Using Deep Convolutional Networks, In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295307, Feb. 2016.

[8] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang. Deep laplacian pyramid networks for fast and accurate super-resolution. In CVPR, 2017.

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

[10] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee. Enhanced deep residual networks for single image super-resolution. In CVPRW, 2017.

[11] W. Shi, J. Caballero, F. Husz´ar, 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 CVPR, 2016.

[12] Y. Tai, J. Yang, and X. Liu. Image super-resolution via deep recursive residual network. In CVPR, 2017.

[13] C. G. Marco Bevilacqua, Aline Roumy and M.-L. A. Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In BMVC, 2012.

[14] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In ICLR, 2015

[15] J.Sun, Z.Xu, and H.-Y.Shum. Image super-resolution using gradient profile prior. In CVPR, 2008.

[16] H. Chang, D.-Y. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. In CVPR, 2004.

[17] C.-Y. Yang and M.-H. Yang. Fast direct super-resolution by simple functions. In ICCV, 2013.

[18] R.Zeyde, M.Elad, and.Potter. On single image scale-up using sparse-representations. In Curves and Surfaces, pages 711–730. Springer, 2012.

[19] Yang, Jianchao, and Wright, John, and Huang, Thomas, and Ma, Yi. Image Super-Resolution Via Sparse Representation, In Transactions on Image Processing, vol. 19, no. 11, pp. 28612873, Nov. 2010.

[20] J.-B. Huang, A. Singh, and N. Ahuja. Single image super-resolution using transformed self-exemplars.InCVPR,2015.