Pre-Project Report

On

Image Super Resolution using Deep Learning Approach

Submitted in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

IN

INFORMATION TECHNOLOGY

Submitted by

Shish Pal

IT-243/16

Saurabh Verma

IT-23/16

Under the supervision of

Dr. Shabir Ahmad Sofi

Department of Information Technology

National Institute of Technology Srinagar, J&K

TABLE OF CONTENTS

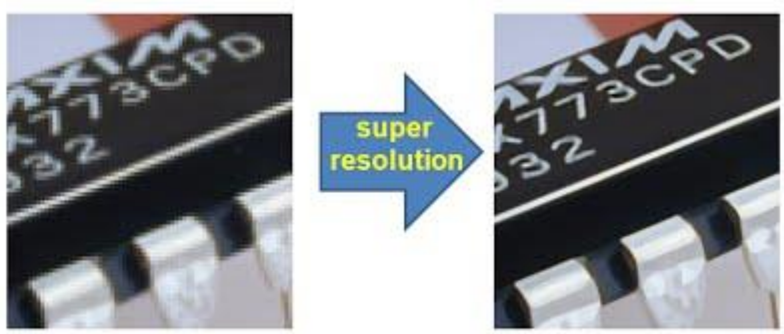
LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Motivation:

Our work is motivated by the advancement in deep learning algorithms for various computer vision problems e.g. face recognition[19], self-driving car, captcha recognition[20]. In this work, we are using deep learning for image super-resolution.

Image super-resolution has been a highly challenging task in computer vision for a long time. This task is even more difficult when we have a single low-resolution image to reconstruct a high-resolution image. In some Radar and Sonar imaging applications (e.g. magnetic resonance imaging (MRI), high-resolution computed tomography) Super-resolution can be very useful. The digital image is playing an exceptionally huge role in many aspects of our daily life like they are utilized in satellite TV, intelligent traffic monitoring, signature approval, recognition of handwriting on checks, and also in other fields of science and technology like geographical information system and astronomy. Because of the impact of the transmission channels and other factors on the images are unavoidably being corrupted by noise during the process of acquisition, compression, and transmission resulting in the deformation and loss of image data.



[Figure-1] A low resolution image kept besides its high resolution version

Image super-resolution has impacted almost all technical fields and plays a significant role in many fields such as medical imaging (Gamma-ray imaging, PET scan, X-Ray Imaging, Medical CT, UV imaging), astronomical imaging (removing noise from images captured by satellite), transmission and encoding, microscopic imaging, forensic science, image restoration, visual tracking, image registration, image segmentation, image classification and other areas where it is crucial to obtain the contents of the original image for strong performance. Traditional methods[5][13][15] used to recover the high-quality image are not efficient enough. Deep learning-based algorithms can help in converting low-quality images to high-quality images in a cost-efficient and more accurate manner.

1.2 **Problem Statement**

Image Super-Resolution is the process of reconstructing a high-resolution image using a low-resolution image. A low-resolution image is a single image input, high-resolution image is the ground truth and super-resolution is the predicted high-resolution image. In Computer Vision low-resolution images used for Machine learning and deep learning based solution is the down-sampled image with some blurring and noise added to them.

Deep learning uses algorithms that aim at learning the hierarchical representations of data. Neural network algorithms have shown prominent superiority over other traditional algorithms in many artificial intelligence domains, such as computer vision, speech recognition[18], and natural language processing. Generally, the strong capacity of Deep Learning to address substantial unstructured data is attributable to two main contributors: the development of efficient computing hardware and the advancement of sophisticated algorithms. Due to these reasons in this work, we are using deep learning based Convolution Neural Network for image super-resolution.

An image may have a “lower resolution” due to a smaller spatial resolution (i.e. size) or due to a result of degradation (such as blurring). We can relate the HR and LR images through the following equation: $LR = degradation(HR)$

Clearly, on applying a degradation function, LR image can be obtained from the HR image. But, to do the inverse, in the ideal case, yes! It can be done if exact degradation function is known. But, therein lies the problem. Usually degradation function is not known before hand. Directly estimating the inverse degradation function is an ill-posed problem. In spite of this, Deep Learning techniques have proven to be effective for Super Resolution.

CHAPTER 2
LITERATURE SURVEY

The super-resolution is used to get high-resolution images from low-resolution images. To obtain super-resolution images Hsieh S Hou[1] proposed a simple interpolation-based method. This interpolation-based method results in ringing and jagging artifacts. To recover these limitations Dai et.al[2] proposed an interpolation-based method with some prior information which resulted in improving the quality of the resultant image but was limited to a small up-scaling ratio.

Manifold based approaches(e.g. locally linear embedding) discover the mapping between low-resolution image patches and high-resolution image patches using local geometry[3]. In this approach, a large number of patches can be super-resolved by a small dataset but it is highly computationally complex and leads to loss of high frequency (textural and edges) information in resultant images. Yang et. Al[4] proposed a sparse representation based approach that provides improved results in terms of PSNR in comparison to other existing techniques. Inspired by the recent development in deep learning, authors Toru Nakashika, Chao Dong[6] used different deep learning algorithms to learn the end to end mapping[5] between low-resolution and high-resolution patches then make use of a learned model to super-resolved test images which provide the state-of-the-art results in image super-resolution.

Although much work has been done in both these areas separately, few advances have been made to achieve noise resilient image super-resolution. The traditional way to get noise resilient image super-resolution includes two steps pre-processing and then super-resolving the denoised image. This traditional method has limitations as during the denoising process, we lose some important textural/high-frequency details and the super-resolution process is not able to recover this loss. A different framework has been proposed by Ezequiel López-Rubio which uses a median filter[7] transform on a parallelogram to get noise-free image super-resolution. Sparse mixing estimator based approach is used by F. Qiu to get noise-free[8], super-resolution. Abhishek.et.al[9] has proposed an algorithm which takes a convex combination of frequency and an orientation-selective band of the de-noised and noisy HR images. The author shows that the high-frequency component of the noisy low-resolution image can be used to get textural information of the high-resolution image. These techniques do not provide many improved results in comparison to the conventional framework.

The majority of super-resolution algorithms[11],[12] focus on gray-scale or single-channel image super-resolution. For color images, the aforementioned methods first transform the problem to a different color space (YCbCr or YUV), and super-resolution is applied only on the luminance channel. There are also works attempting to super-resolve all channels simultaneously. For example, Kim and Kwon and Dai et al.[13] apply their model to each RGB channel and combined them to produce the final results. However, none of them has analyzed the super-resolution performance of different channels, and the necessity of recovering all three channels.

Convolutional neural networks (CNN) date back decades and deep CNNs have recently shown explosive popularity partially due to its success in image classification[14]. They have also been successfully applied to other computer vision fields, such as object detection face recognition and pedestrian detection. Several factors are of central importance in this progress: (i) the efficient training implementation on modern powerful GPUs[15] (ii) the proposal of the Rectified Linear Unit (ReLU) which makes convergence much faster while still presents good quality, and (iii) the easy access to an abundance of data (like ImageNet) for training larger models. Our method also benefits from this progress.

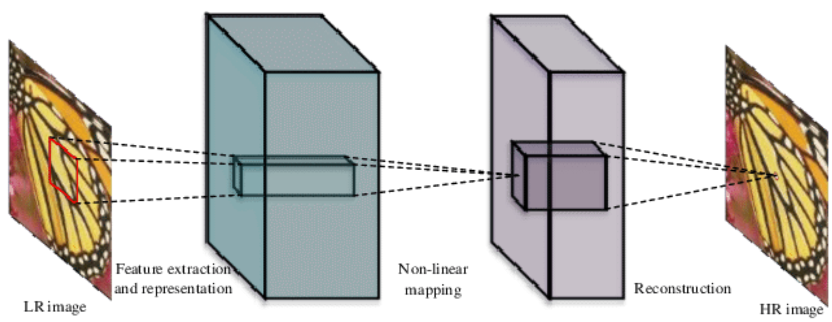
There have been a few studies of using deep learning techniques for image restoration. The multi-layer perceptron (MLP), whose all layers are fully-connected (in contrast to convolutional), is applied for natural image denoising and post-deblurring denoising. More closely related to our work, the convolutional neural network is applied for natural image denoising and removing noisy patterns (dirt/rain). These restoration problems are more or less denoising-driven. Cui et al[16]. propose to embed auto-encoder networks in their superresolution pipeline under the notion internal example based[17] approach. The deep model is not specifically designed to be an end-to-end solution since each layer of the cascade requires independent optimization of the self-similarity search process and the auto-encoder. On the contrary, our proposed Super-resolution convolutional neural network optimizes end-to-end mapping. Further, this is faster at speed. It is not only a quantitatively superior method but also a practically useful one.

CHAPTER 3
METHODOLOGY

super-resolution convolution neural network (SRCNN) is a deep convolutional neural network that learns the end-to-end mapping of low-resolution to high-resolution images. As a result, we can use it to improve the image quality of low-resolution images. To evaluate the performance of this network, we will be using three image quality metrics: peak signal to noise ratio (PSNR), mean squared error (MSE), and the structural similarity (SSIM) index. In brief, with better SR approach, we can get a better quality of a larger image even we only get a small image originally.

3.1 **SRCNN Operations:** The super-resolution convolution neural network (SRCNN) consists of the following operations:

- **Preprocessing**: Up-scales LR image to desired HR size.
- **Feature extraction**: Extracts a set of feature maps from the up-scaled LR image.
- **Non-linear mapping**: Maps the feature maps representing LR to HR patches.
- **Reconstruction**: Produces the HR image from HR patches.



[Figure- 2] SRCNN Network

3.2 **Performance metrics:** We will use metrics to evaluate the performance of our super resolution model:

3.2.1 **PSNR**

Peak Signal-to-Noise Ratio (PSNR) is commonly used objective metric to measure the reconstruction quality of a lossy transformation. PSNR is inversely proportional to the logarithm of the Mean Squared Error (MSE) between the ground truth image and the generated image.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(I(i) - \hat{I}(i))^2,$$

$$\mathrm{PSNR} = 10 \cdot \log_{10}(\frac{L^2}{\mathrm{MSE}}).$$

[Figure- 3]Calculation of PSNR

In the above formula, L is the maximum possible pixel value (for 8-bit RGB images, it is 255). Unsurprisingly, since PSNR only cares about the difference between the pixel values, it does not represent perceptual quality that well.

3.2.2 **SSIM**

Structural Similarity (SSIM) is a subjective metric used for measuring the structural similarity between images, based on three relatively independent comparisons, namely luminance, contrast, and structure. Abstractly, the SSIM formula can be shown as a weighted product of the comparison of luminance, contrast and structure computed independently.

$$\mathrm{SSIM}(I, \hat{I}) = [\mathcal{C}_l(I, \hat{I})]^{\alpha}[\mathcal{C}_c(I, \hat{I})]^{\beta}[\mathcal{C}_s(I, \hat{I})]^{\gamma},$$

In the above formula, alpha, beta and gamma are the weights of the luminance, contrast and structure comparison functions respectively.

CHAPTER 4
HARDWARES & SOFTWARES

Hardware:

Intel® Core™ i7-7500U CPU @ 2.70GHz × 4

8 GB Primary Memory.

Softwares & Libraries:

- Jupyter Notebook
- Python Libraries
  - Open CV
  - Keras
  - Matplotlib
  - Numpy
  - Scikit-image

CHAPTER 5
PROJECT CONTRIBUTIONS

Upon successful completion, this project will help in:

1. Taking good quality pictures from the low-cost cameras.
2. Normal SD TV content can also be converted to HD TV.
3. If we take an example of daily surveillance cameras it's impossible to put high-quality cameras everywhere so in this situation this will be very useful.
4. Improving the quality of medical images.
5. Help in improving the quality of earth remote sensing.
6. It can also help in astronomical observation.
7. Nowadays the biometric authentication system is used everywhere. So for this purpose, this model will also be useful.

REFERENCES

1. Hsieh S Hou and H Andrews, "Cubic splines for image interpolation and digital filtering", *Acoustics Speech and Signal Processing IEEE Transactions on*, vol. 26, no. 6, 1978.
2. Shengyang Dai, Mei Han, Wei Xu, Ying Wu and Yihong Gong, "Soft edge smoothness prior for alpha channel super resolution", *Computer Vision and Pattern Recognition 2007. CVPR'07. IEEE Conference on*, 2007.
3. Hong Chang, Dit-Yan Yeung, and Yimin Xiong, "Super-resolution through neighbor embedding", *Computer Vision and Pattern Recognition 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1, 2004.
4. Yang Jianchao, Wright John, Thomas S Huang, and Yi Ma, "Image super-resolution via sparse representation", *Image Processing IEEE Transactions on*, vol. 19, no. 11, 2010.
5. Toru Nakashika, Tetsuya Takiguchi, and Yasuo Ariki, "High-frequency restoration using deep belief nets for super-resolution", *Signal-Image Technology & Internet-Based Systems (SIT IS) 2013 International Conference on*, 2013.
6. Chao Dong, Chen Change Loy, Kaiming He and Xiaoou Tang, "Learning a deep convolutional network for image super-resolution", *Computer Vision-ECCV 2014*, 2014.
7. Ezequiel López-Rubio, "Superresolution from a single noisy image by the median filter transform", *SIAM Journal on Imaging Sciences*, vol. 9, no. 1, 2016.
8. F. Qiu, Y. Xu, C. Wang and Y. Yang, "Noisy image super-resolution with sparse mixing estimators", *Image and Signal Processing (CISP) 2011 4th International Congress on*, vol. 2, pp. 1081-1085, Oct 2011.
9. Abhishek Singh, Fatih Porikli and Narendra Ahuja, "Super-resolving noisy images", *Proceedings of the IEEE Conference on ComputerVision and Pattern Recognition*, 2014.
10. K. Zeng, J. Yu, R. Wang, C. Li and D. Tao, "Coupled deep autoencoder for single image super-resolution", *IEEE Transactions on Cybernetics*, no. 99, pp. 1-11, 2015.
11. Bevilacqua, M., Roumy, A., Guillemot, C., Morel, M.L.A.: Lowcomplexity single-image super-resolution based on nonnegative neighbor embedding. In: British Machine Vision Conference (2012)
12. Yang, J., Wang, Z., Lin, Z., Cohen, S., Huang, T.: Coupled dictionary training for image super-resolution. IEEE Transactions on Image Processing 21(8), 3467–3478 (2012)
13. Dai, S., Han, M., Xu, W., Wu, Y., Gong, Y., Katsaggelos, A.K.: Softcuts: a soft edge smoothness prior for color image superresolution. IEEE Transactions on Image Processing 18(5), 969–981 (2009)
14. He, K., Zhang, X., Ren, S., Sun, J.: Spatial pyramid pooling in deep convolutional networks for visual recognition. In: European Conference on Computer Vision, pp. 346–361 (2014)
15. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. pp. 1097–1105 (2012)

16. Cui, Z., Chang, H., Shan, S., Zhong, B., Chen, X.: Deep network cascade for image super-resolution. In: European Conference on Computer Vision, pp. 49–64 (2014)
17. Glasner, D., Bagon, S., Irani, M.: Super-resolution from a single image. In: IEEE International Conference on Computer Vision. pp. 349–356 (2009)
18. B. Jolad and R. Khanai, "An Art of Speech Recognition: A Review," 2019 2nd International Conference on Signal Processing and Communication (ICSPC), Coimbatore, India (2019)
19. D. Meena and R. Sharan, "An approach to face detection and recognition," 2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE), Jaipur, 2016
20. Y. Hu, L. Chen and J. Cheng, "A CAPTCHA recognition technology based on deep learning," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, 2018