Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 7, Issue 04, April -2020

Single Image Super Resolution using Deep Learning: A Survey

Yogeshvari Makwana[1], Prashant B.swadas [2], Pranay S. patel[3]

[1]*Department of Computer, BVM Engineering College*
[2]*Department of Computer, BVM Engineering College*
[3]*Department of Computer, BVM Engineering College*

ABSTRACT: *Single image super-resolution, which is used to restore high-resolution image from a single low-resolution image, is a difficult challenging problem in computer field. In recent times, dominant deep learning algorithms have been applied to Single image super resolution and have shown an highly efficient performance. In this paper, we surveyed deep learning-based super resolution method known as a super resolution convolutional neural network (SRCNN) that takes the low-resolution image as the input and outputs the high-resolution one. SRCNN has a non-complex structure yet provides high quality and fast speed. We get quick results for practical online usage. Hence, survey is carried out on different networks like Generative Adversarial Networks (GAN) and Convolutional Neural Network comparison between quality and speed.*

Keywords: *Image Super Resolution, SRCNN (Super Resolution Convolution neural network), SRGAN(Super Resolution Generative Adversarial Network)*

I INTRODUCTION

Nowadays, deep learning is exiting and interesting branch of machine learning. Deep learning is the most useful, supervised, time and cost saving resourceful machine learning approach[1]. It is an unlimited non - restricted learning approach. It is used in many areas like business, science and technology and government which has features like Object tracking, stock market analysis, biological image detection, recognition facial expression, hand writing recognition, computer vision, in depth adaptive testing and much more[3]. The rest of the paper is arranged as follows. In Section II we present overview of image super resolution., In Section III we present relevant pre-processing techniques of single image super resolution., we surveyed the methods like SRCNN (Super Resolution Convolutional Neural Network) and SRGAN (Super Resolution Generative Adversarial Network) for various Single Image Super Resolution tasks. In Section VI we have shown dataset details and results. We conclude this survey paper in Section VII.

II OVERVIEW OF IMAGE SUPER RESOLUTION

Super resolution is a class of method that increase the resolution within an image. Super resolution is a technique to restore high resolution images from one or more low resolution images. Super resolution of image is classified into two categories: one is single image super resolution and another one is multiple image super resolution[2].

Fig.1: Resolution of same image[5]

Compared with multi image super resolution, single image super resolution is widely accepted because of its high efficiency. Single image super resolution is a traditional problem in computer vision[4]. Single image super resolution is used for recovering a high resolution image from single low resolution image. In this survey, we attempt to give an overall review of recent Deep Learning based single image super resolution algorithms. We mainly focus on efficient neural network architectures designed for single image super resolution[5].

III REVIEW OF PRE-PROCESSING TECHNIQUES

In the low resolution image there are some irrelevant noise, so we have to improve the quality of the low resolution images by pre-processing methods. Bicubic interpolation is one of them[6]. Bicubic interpolation is also a convolutional operation, so it can be formulated as a convolutional layer. However, the output size of this layer is larger than the input size, so there is a fractional stride[1].

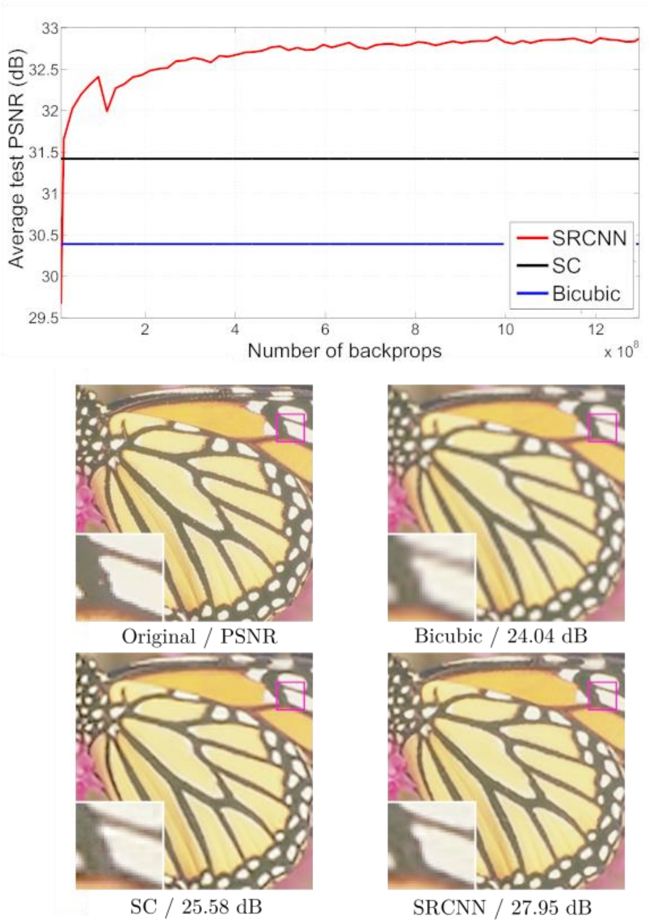


Fig.2: PSNR of different model for check quality of image [1]

IV SUPER RESOLUTION CONVOLUTIONAL NEURAL NETWORK (SRCNN)

The Convolution neural network has a long history. While going in depth study and research CNNs has shown exclusive popularity due to its success in image classification[7]. We can also see application to other computer vision fields, such as facial recognition, object/moving things detection[9]. CNN is widely used because it is design with simplicity and provide more speed for real time practically online usage. The block diagram of super resolution based on CNN is shown in Fig 2.

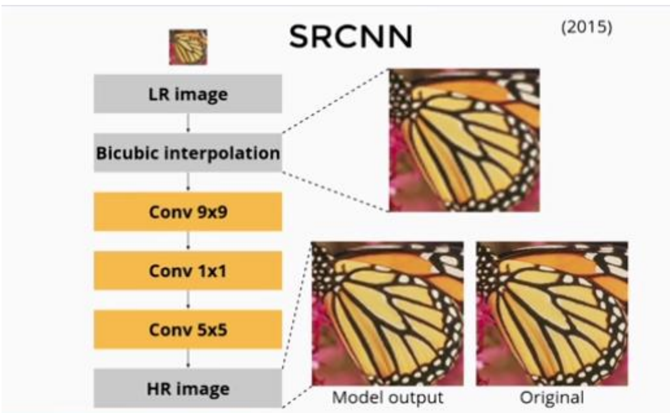


Fig.3: Convolutional neural network for Super Resolution [9]

We first pre-process the image to the mandatory (required) size using bicubic interpolation[8]. Bicubic interpolation is the only pre-processing technique that is perform in the SRCNN model. Overall the contributions of SRCNN are mostly in three steps:

(1)Patch extraction and representation: This process extracts patches from low resolution image and characterize each patch as a high dimensional vector. These vectors consists a set of feature maps where the numbers equals to the dimensionality of the vectors[10].

(2) Non-linear mapping: This process nonlinearly maps each high-dimensional vector onto a different high-dimensional vector. Each mapped vector is abstractly the representation of a high-resolution patch. These vectors involve another set of feature maps[11].

3) **Reconstruction**: This operation aggregate the above high-resolution patch-wise representations to produce the final high-resolution image. This image is predictable to be similar to the ground truth X[1].

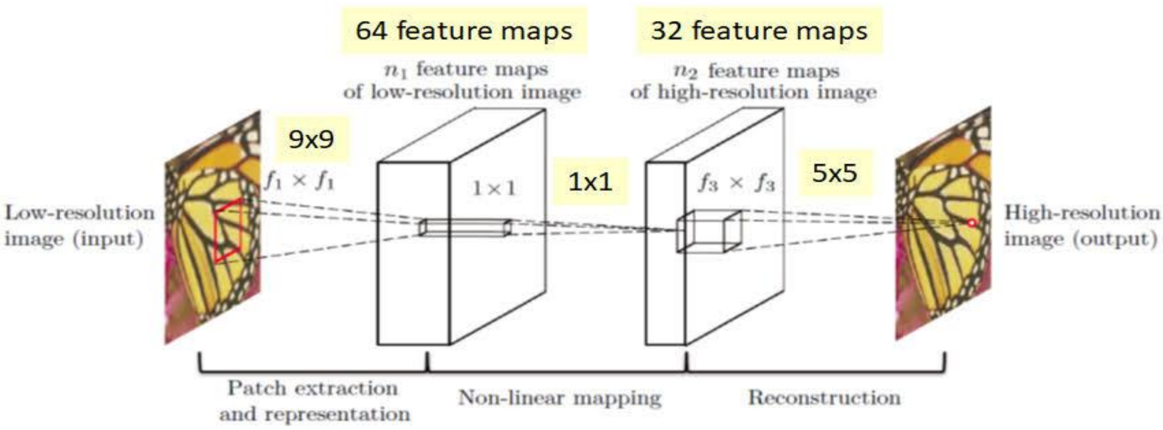


fig:4 Given a low-resolution image X, the first convolutional layer of the SRCNN extract a set of feature maps. The second layer maps these feature maps nonlinearly to high-resolution patch representations. The last layer combines the predictions within a spatial neighbourhood to produce the final high-resolution image F(X)[12],[1].

V SUPER RESOLUTION GENERATIVE ADVERSARIAL NETWORK (SRGAN)

For this first we have to understand what are GANs(Generative Adversarial Networks)

GANs: GANs having algorithmic architecture that uses two neural networks. They are class of AI algorithms mostly used in unsupervised machine learning[13],[15],[16]. They are used mostly in image generation, video generation and voice generation[14]. Let's understand how GAN works: One neural network, called the *generator*, generates new data instances, while the other, the discriminator, evaluates them for authenticity; i.e. the discriminator decide whether each instance of data that it reviews belongs to the actual training dataset or not. , the generator is creating new, synthetic images that it passes to the discriminator[17]. It does so in the hopes that they, too, will be deem authentic, even though they are fake. The goal of the generator is to generate passable hand-written digits: to lie without being caught. The goal of the discriminator is to identify images coming from the generator as fake.

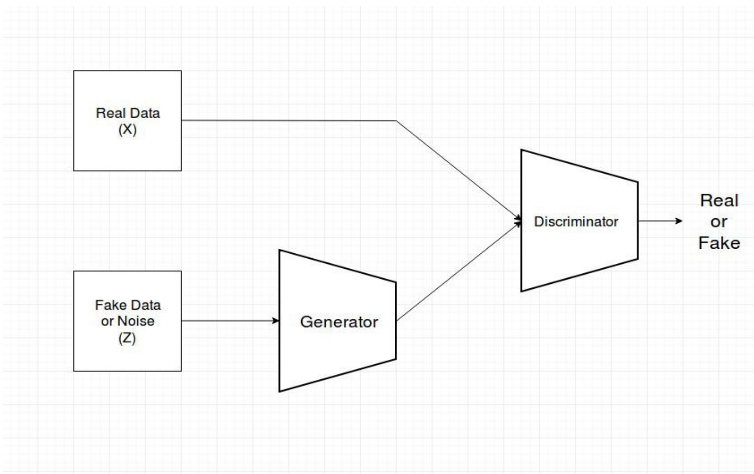


Fig 5: Basic architecture of GAN[5]

SRGAN : We have seen various methods for Single image super resolution. Those methods are fast and accurate as well. But still there is one problem which is not solved[18]. That is, how can we recover finer texture details from low resolution image so that image is not distorted. Recent work has largely focused on minimizing the mean squared reconstruction error[19]. The results have high peak signal-to-noise ratios(PSNR) means we have good image quality results, but they are often lacking high-frequency details and are perceptually unsatisfying as they are not able to match the fidelity expected in high resolution images[20]. Previous methods are trying to see similarity in pixel space which led to perceptually unsatisfying results or they produce blurry images[21]. So we need a stable model which can capture the perceptual differences between the model's output and the ground truth image.

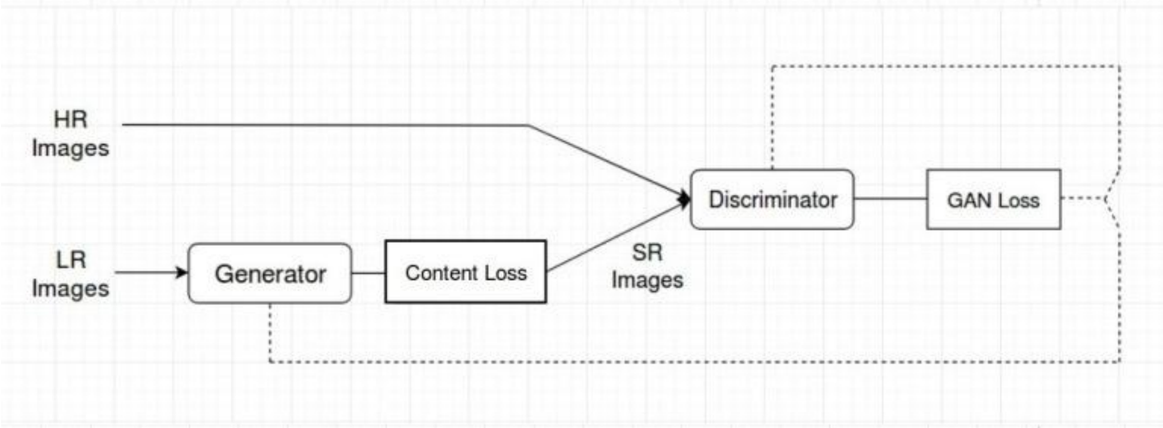


Fig 6:SRGAN Architecture[7]

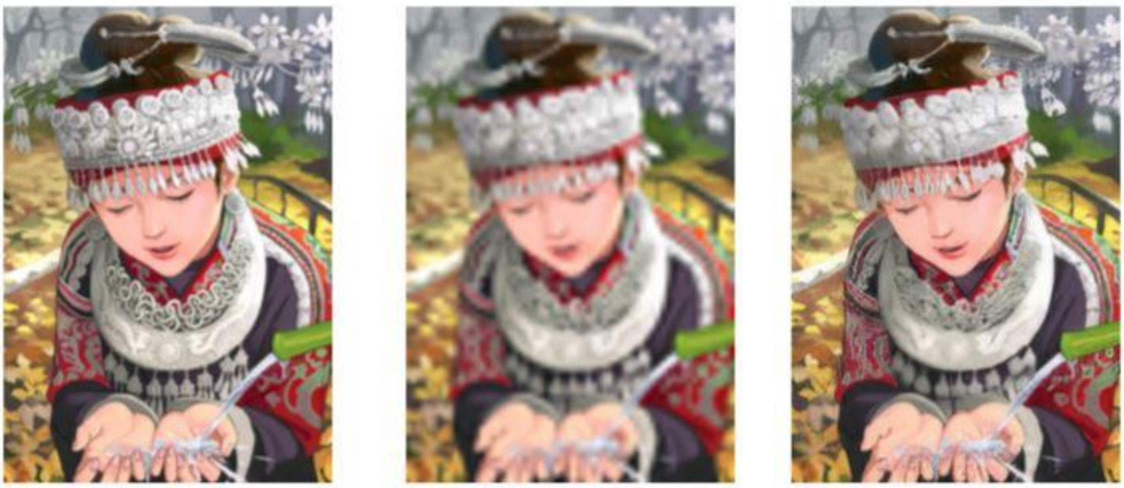
VI DATA SET DETAIL AND RESULT

Chih-Yuan Yang, Chao Ma1, and Ming-Hsuan Yang use two image sets as the HR ground truth data for evaluation. The first set contains 200 images from the Berkeley segmentation dataset [22], which is widely used for SISR evaluations [1,23]. All images are of 321×481 pixels covering diverse contents acquired in a professional photographic style. The second set contains 29 undistorted high-quality images from the LIVE1 dataset [25], which is widely used for image quality assessment [24]. The resolution of these images ranges from 480×720 to 512×768 pixels.

Chao Dong, Chen Change Loy and Kaiming He use the Set5 [2] as the validation set. They observe a similar trend even if we use the larger Set14 set [3].

Tero Karras, Timo Aila and Samuli Laine used to create the high-quality version of the CELEBA dataset, consisting of 30000 images in 1024 × 1024 resolution[4]. As a starting point, They took the collection of in-the-wild images included as a part of the original CELEBA dataset. These images are extremely varied in terms of resolution and visual quality, ranging all the way from 43×55 to 6732×8984.

Here, Results of bicubic and SRGAN compare with High Resolution image[5].



(a) HR | (b) bicubic(21.59dB/0.6423) | (d) SRGAN(21.15dB/0.6868)

VII CONCLUSION

In this paper, survey on various techniques for single image super resolution is done. Different techniques are used by various researchers to get high resolution image from single low resolution image. This paper talks about overview of how to convert single low resolution image into high resolution image.. According to this review we find some methods like SRCNN and SRGAN for getting more accurate and powerful results for single image super resolution. The future move towards, SRCNN, learns an end-to-end mapping among low- and high-resolution images, with small extra pre/post-processing beyond the optimization. With a lightweight structure, the SRCNN has achieved greater performance than the state-of-the-art methods. We have presented novel deep learning methods for single image super resolution.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, vol. 521, no. 7553, p. 436, 2015.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Image net classification with deep convolutional neural networks," in Proceedings of the Advances in Neural Information Processing Systems, 2012, pp. 1097– 1105.

[3] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath et al., "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," IEEE Signal Processing Magazine, vol. 29, no. 6, pp. 82–97, 2012.

[4] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in Proceedings of the International Conference on Machine Learning, 2008, pp. 160–167.

[5] Chao Dong, Chen Change Loy, Member, IEEE, Kaiming He, Member, IEEE, and Xiaoou Tang, Fellow, IEEE "Image Super-Resolution Using Deep Convolutional Networks " IEEE-31 Jul 2015

[6] Krupali Ramavat1, Prof. Mahasweta Joshi2, Prof. Prashant B. Swadas3 "A survey of Super resolution Techniques" 2016 International Research Journal of Engineering and Technology (IRJET) 2016

[7]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)

[8] Yang, C.Y., Ma, C., Yang, M.H.: Single-image super-resolution: A benchmark. In: European Conference on Computer Vision, pp. 372–386 (2014)

[9] Timofte, R., De Smet, V., Van Gool, L.: A+: Adjusted anchored neighborhood regression for fast super-resolution. In: IEEE Asian Conference on Computer Vision (2014)

[10] Dai, D., Timofte, R., Van Gool, L.: Jointly optimized regressors for image super-resolution. In: Eurographics. vol. 7, p. 8 (2015)

[11] M.S.Sajjadi, B.Sch¨olkopf, and M.Hirsch, "EnhanceNet:Singleimage super-resolution through automated texture synthesis," in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 4501–4510.

[12] K. Li, S. Peng, and J. Malik, "Super-resolution via conditional implicit maximum likelihood estimation," arXiv preprint arXiv:1810.01406, 2018.

[13] R. Mechrez, I. Talmi, F. Shama, and L. Zelnik-Manor, "Learning to maintain natural image statistics," arXiv preprint arXiv:1803.04626, 2018.

[14] K. Zhang, W. Zuo, and L. Zhang, "Learning a single convolutional super-resolution network for multiple degradations," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3262–3271.

[15] Z. Hui, X. Wang, and X. Gao, "Fast and accurate single image superresolution via information distillation network," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 723–731.

[16] 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 superresolution using an efficient sub-pixel convolutional neural network," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1874–1883.

[17] Timofte, R., De Smet, V., Van Gool, L.: A+: Adjusted anchored neighborhood regression for fast super-resolution. In: IEEE Asian Conference on Computer Vision (2014)

[18] Dong, C., Loy, C.C., He, K., Tang, X.: Learning a deep convolutional network for image super-resolution. In: European Conference on Computer Vision, pp. 184–199 (2014)

[19] D. Liu, Z. Wang, B. Wen, J. Yang, W. Han, and T. S. Huang, "Robust single image super-resolution via deep networks with sparse prior," IEEE Transactions on Image Processing, vol. 25, no. 7, pp. 3194– 3207, 2016. [20] D. Liu, Z. Wang, N. Nasrabadi, and T. Huang, "Learning a mixture of deep networks for single image super-resolution," in Proceedings of the Asian Conference on Computer Vision, 2016, pp. 145–156.

[21] W. Yang, J. Feng, J. Yang, F. Zhao, J. Liu, Z. Guo, and S. Yan, "Deep edge guided recurrent residual learning for image super-resolution," IEEE Transactions on Image Processing, vol. 26, no. 12, pp. 5895– 5907, 2017.

[22] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep Laplacian pyramid networks for fast and accurate super-resolution," in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 624–632.

[23] R. Dahl, M. Norouzi, and J. Shlens, "Pixel recursive super resolution," in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 5439–5448. [86] A. Singh and N. Ahuja, "Super-resolution using sub-band selfsimilarity," in Proceedings of the Asian Conference on Computer Vision, 2014, pp. 552–568.

[24] C. Deng, J. Xu, K. Zhang, D. Tao, X. Gao, and X. Li, "Similarity constraints-based structured output regression machine: An approach to image super-resolution," IEEE Transactions on Neural Networks and Learning Systems, vol. 27, no. 12, pp. 2472–2485, 2016.

[25] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, vol. 521, no. 7553, p. 436, 2015.

