iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IX **Month of publication:** September 2023

DOI: https://doi.org/10.22214/ijraset.2023.55631

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

Adaptive Filter in Single Image SRGAN

Atin Bera[1], Arya Bhattacharyya[2], Debmitra Ghosh[3]
[1, 2, 3]*JIS University, Kolkata, India*

Abstract: ***SRGAN is a generative adversarial network for super-resolution of a single picture. In this paper an adaptive filter in single image SRGAN which is a Super-Resolution Generative Adversarial Network (SRGAN), is a deep learning-based approach that is used to generate high-resolution images from low-resolution ones. The SRGAN model is based on the Generative Adversarial Network (GAN) architecture, which consists of two deep neural networks: a generator network and a discriminator network. The generator network takes a low-resolution image as input and tries to generate a high-resolution image that is similar to the original high-resolution image. The discriminator network, on the other hand, tries to distinguish between the high-resolution images generated by the generator network and the original high-resolution images. The SRGAN model has been shown to be very effective in generating high-quality, realistic images from low-resolution inputs, and it has been applied in various applications, including image and video upscaling, medical imaging, and satellite imaging.***
Keywords: SRGAN, ResNet, VGG19, CNN, ML

I. INTRODUCTION

One of the major problems in our today's life is accidents on foggy roads due to less amount visibility, which is occurred due to bad weather conditions. This project's main goal is to use an algorithm to improve visibility in foggy weather camera-based Advanced Driver Assistance systems (ADAS) and to deduce a rough brightness map with no fog. This algorithm makes more contrast and viewable seen by the single camera-based ADAS at the same time. The main issue is that contrast enhancement from a single blurry image is an incorrectly phrased challenge. In order to restore vision after a fog, it is necessary to estimate both the scene's depth map and brightness without fog. This means estimating two unknown parameters from a single image for each pixel. Regularization is required for this issue. This concept of an approximate depth map was improved in [1] by putting out a number of straightforward parametric geometric models specifically designed for road sceneries observed in front of a moving vehicle. By globally maximizing the scene depths for each type of model, the parameters are fitted to each view without resulting in any black pixels in the augmented image. The rigidity of the suggested geometric models is this approach's limitation. These models have some adverse situations like: The driver's perspective on the scene geometry along the entire road path must be approximated, which is difficult to do. A second solution to this issue was put forth around the same time in [2], which relied on the usage of color images with pixels that had a tint other than grey. Additionally, in our opinion, visibility enhancement algorithms must be able to process low-resolution images in order to transform intelligent vehicle applications into a function that is described in the visibility restoration problem in [3]. To improve the foggy roadside image, the SRGAN (Super Resolution Generative Adversarial Network) algorithm is used, which mainly combines pixel-wise MSE loss with a discriminator loss. The algorithms are applied to foggy images and results are compared with the images without fog. Here [38] also used GANs for unsupervised representation learning. In the paper[33] author uses the style for transfer and in [52] for iinpainting, describing the use of GANs to learn a mapping from one manifold to another. The squared error of the scattering networks and VGG19 [41] feature spaces is known by paper. [5]. The concept of this project is to make a low-resolution(LR) image into a high-resolution(HR) image. Using GANs(Generative Adversarial Networks) algorithm implemented by us. Super-resolution refers to the extremely difficult task of estimating a high-resolution (HR) image from its low-resolution (LR) equivalent (SR). From this project, we can easily make a lower-resolution image into a higher-resolution image. SR offered a wide range of applications and attracted a lot of interest from the computer vision research community . High upscaling factors, for which texture detail in the reconstructed SR images is often absent, make the underdetermined SR problem's ill-posedness especially apparent. The mean squared error (MSE) between the recovered HR image and the ground truth is typically the optimization goal of supervised SR algorithms. This is practical because the maximization of the peak signal-to-noise ratio (PSNR), a typical metric used to evaluate and compare SR methods, coincides with the minimization of MSE [50]. Nevertheless, because MSE (and PSNR) are established based on pixel-wise picture differences, their capacity to capture perceptually relevant differences, such as great texture detail, is severely constrained and the highest PSNR does not always reflect the perceptually better SR result. The recovered image is not photo-realistic in the sense mentioned because of the perceptual gap between the original and super-resolved images.

In this study, HR image, bicubic interpolation, a deep residual generative adversarial network optimized for a loss more sensitive to human perception, and a deep residual network optimized for MSE. In brackets are the PSNR and SSIM that correspond. So, a deep residual network has been proposed. (ResNet) with skip-connection and depart from MSE as the only optimization goal for a super-resolution generative adversarial network (SRGAN). In contrast to earlier research, our definition of a novel perceptual loss uses high-level feature mappings of the VGG network along with a discriminator that favors solutions that are perceptually difficult to differentiate from the HR reference images.

II. RELATED WORK

Two recent overview works on image SR [37, 50]. This article's focus is on single-image super-resolution, the initial techniques to address SISR were prediction-based. The SISR problem is oversimplified by certain filtering techniques, such as linear, and bicubic filtering, which often results in solutions with excessively smooth textures. Edge preservation has been the topic of several proposed methods [1, 34]. More effective methods typically rely on training data in order to create a sophisticated mapping between low and high-resolution image information. Several example-pair-based techniques rely on LR training patches that are known to have known HR counterparts.[16, 15] presented their early research. Compressed sensing-based solutions to the SR problem are related [51, 12, 56]. This paper [19] uses patch redundancies between scales in the image to drive the SR. In [26], where the author extends self-dictionaries by additionally allowing for tiny changes and shape variations, also uses this self-similarity paradigm. Convolutional sparse coding, as proposed by the researcher [23], increases consistency by analyzing the entire image as opposed to overlapping sections. The advantages of learning-based detail synthesis are merged with an edge-directed SR approach based on a gradient profile prior [42] and [44] to rebuild realistic texture detail without creating edge artifacts. A multi-scale dictionary is suggested by the author in [57] to identify the redundancy of comparable image patches at various scales[54] retrieval of associated HR photos from the web with similar content and their proposal of a structure-aware matching criterion for alignment to enable super-resolved landmark images.By identifying similar LR training patches in a low-dimensional manifold and pooling their associated HR patches for reconstruction, neighborhood embedding techniques upsample an LR image patch [45, 46]. By utilizing kernel ridge regression, In paper[30] highlights the propensity of neighborhood techniques to overfit and create a map of example pairings that is more inclusive. Trees, Random Forests, and the Gaussian Process Regression can all be used to address the regression problem[6], The most suitable regressors are chosen during testing after a large number of patch-specific regressors have been trained. Lately, SR methods based on convolutional neural networks (CNNs) have demonstrated exceptional performance. In [49].'s feed-forward network architecture, which is based on the learned iterative shrinkage and thresholding algorithm (LISTA) [21], the authors incorporate a sparse representation prior. To attain state-of-the-art SR performance, In [9, 10] used bicubic interpolation to upscale an input image and trained a three-layer deep fully convolutional network end-to-end. The ability of the network to directly train the upscaling filters was later demonstrated to be able to significantly improve performance in terms of accuracy and speed [11, 40, 47]. In this paper [29] introduced a very effective architecture with its deeply-recursive convolutional network (DRCN), which allows for long-range pixel dependencies while minimizing the number of model parameters. The research [28,5], which relies on a loss function closer to perceptual similarity to recover visually more convincing HR images, is particularly pertinent to this paper Following the success of the working paper [32], the state-of-the-art for many computer vision issues is currently established by specifically developed CNN architectures. Deep network architectures have been demonstrated to be challenging to train but have the potential to significantly improve the network's accuracy as they permit modeling mappings of extremely high complexity [41, 43]. Batch normalization is frequently employed to combat the internal covariate shift in order to effectively train these deeper network designs. Deeper network topologies have also been demonstrated to improve SISR performance, as shown, for example, in paper [29], which develops a recursive CNN and reports cutting-edge findings. The recently announced ideas of residual blocks [29] and skip-connections [30, 29] are another effective design decision that makes training deep CNNs easier. The identity mapping, which is straightforward in nature but potentially non-trivial to describe with convolutional kernels, is modeled by skip-connections instead of the network architecture. It has also been demonstrated that learning upscaling filters improves SISR accuracy and speed [11, 40, 47]. In contrast to 10], where used bicubic interpolation was to upscale the LR observation before feeding the image to the CNN, this is an improvement. The uncertainty involved in recovering lost high-frequency details like texture is difficult for pixel-wise loss functions like MSE to handle: minimizing MSE encourages finding pixel-wise averages of plausible solutions, which are typically excessively smooth and have poor perceptual quality [36, 28, 13, 5]. The difficulty of minimizing MSE is clearly understood. In [36] and [7], the authors used generative adversarial networks (GANs) [20] for the purpose of applying image generation to this problem.

To train a network that super-resolves face images with large upscaling factors (8), in the paper [13] the author combines adversarial training with loss functions based on Euclidean distances derived in the feature space of neural networks. It is demonstrated that the suggested loss enables the creation of visually superior images and can be applied to the inverse problem of decoding nonlinear feature representations, which is poorly posed. The use of features extracted from a pre-trained VGG network as opposed to low-level pixel-wise error estimates is suggested in the paper. [28] and [5] in a manner similar to this study. The loss function developed by the researchers is based on the Euclidean distance between feature maps taken from the VGG19 [41] network. Both super-resolution and artistic style transfer produced results that were more perceptually compelling [17, 18]. The author of the paper[33] recently examined the impact of contrasting and blending patches in VGG or pixel feature space. A strong framework for creating natural images with excellent perceptual quality is provided by GANs. The GAN method promotes the reconstructions to travel towards areas of the search space that have a high likelihood of having photorealistic images and are consequently nearer to the natural image manifold. The first very deep [29, 30] architecture is described in this study leveraging the idea of GANs to create a perceptual loss function for photo-realistic SISR. Contribute by us primarily: With the 16 blocks deep ResNet (SRResNet) that is optimized for MSE, it achieved a new state of the art for image SR with high upscaling factors (4), as evaluated by PSNR and structural similarity (SSIM). Here SRGAN is suggested which is a GAN-based network that is enhanced for a novel perceptual loss. Here, instead of using the MSE-based content loss, a loss-based function is calculated on the VGG network's feature maps [41], which are more resistant to changes in pixel space [33]. Using images from three publicly available benchmark datasets, confirm say that SRGAN is the new state of the art by a significant margin for the estimate of photo-realistic SR images with high upscaling factors (4).

III. PRELIMINARIES

Increasing the visibility or making the low-resolute image into a high-resolute visible image is the main aim of this project, where we clearly mentioned some kind of specific algorithms is used that are applied for making our project more efficient.

IV. SUPER RESOLUTION GAN

SRGANs are abbreviated as Super-Resolution Generative Adversarial Networks, which use a perceptual loss to convert a low-resolution image to a high-resolution image without losing much information. The main idea is to introduce a new loss function called Perceptual loss, which is a combination of both adversarial loss and content loss.

$$l^{SR} = l_X^{SR} + 10^{-3} l_{Gen}^{SR}$$

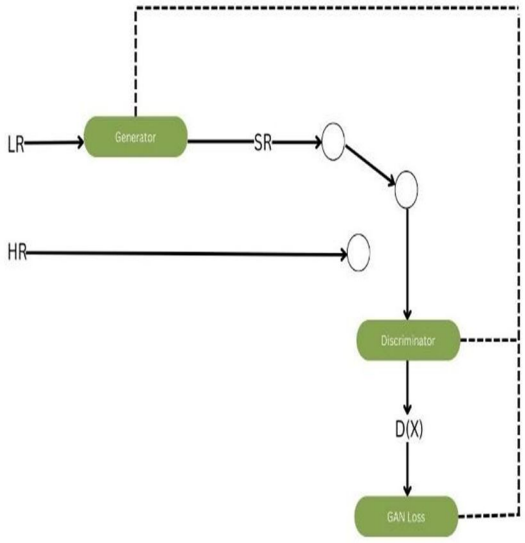


Figure 1. Architecture of SRGAN.

As journal papers are generally printed in black and white it is advisable that where ever possible the figures should be given in gray scale. After inserting the figures verify its readability by converting it to formats generally used for printing eg pdf.

V. ARCHITECTURE

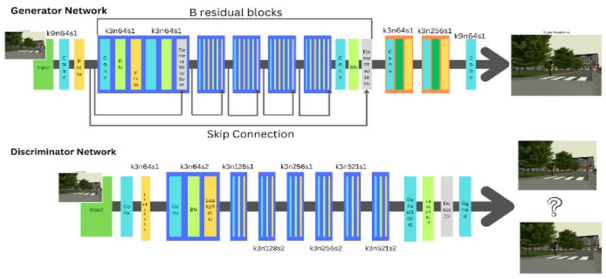


Figure 2. Architecture of Generator and Discriminator Network with corresponding kernel size (k), number of feature maps (n), and stride (s) indicated for each convolutional layer

Basically, GAN contains two parts, one is a generator and another one is a discriminator, where the generator produces some database of the probability distribution and the discriminator tries to guess whether data coming from a dataset of input dataset or generator.

VI. GAN ALGORITHM

GAN is a deep learning algorithm that is employed to produce artificial data that is comparable to a given dataset. A generator and a discriminator neural network make up the algorithm, and both are simultaneously trained.

When given a random noise vector as input, the generator network creates artificial data, while the discriminator network attempts to tell the difference between the artificial data produced by the generator and the real data from the dataset.

Feedback from the discriminator is given to the generator, which uses it to enhance output. The generator's objective is to produce fake data that, in the discriminator's opinion, can't be distinguished from genuine data.

The generator and discriminator networks compete with one another during iterative training. The discriminator finds it harder to discriminate between genuine and synthetic data as the generator gets better, and the quality of the synthetic data also gets better.

VII. METHOD

A. Data Description

The goal of SISR is to estimate a low-resolution input image I^{LR} into a high-resolution, super-resolved image I^{SR}. I^{LR} in this case, is a low-resolution variant of I^{HR}, its high-resolution counterpart. Only when training is taking place are the high-resolution photos accessible. I^{HR} is first subjected to a Gaussian filter, then a downsampling operation with a downsampling factor of r, to produce I^{LR} during training. I^{LR} is defined by a real-valued tensor of size WxHxC for an image with C color channels, and I^{HR}, $I^{SR}$by rWxrHxC, respectively. Our ultimate objective is to train a generating function G that calculates the matching HR counterpart for a given LR input image. A generator network is being trained as a feed-forward CNN $G_{\theta G}$ parametrized by θG to do this. Here, "G" specifies the weights and biases of an L-layer deep network, and was created by optimizing a loss function that is special to SRs, called. For training images I_n^{HR} , n=1, . . . ,N with corresponding I_n^{LR},n=1,...,N,we solve:

$$\hat{\theta}=\arg min_{\theta G}\frac{1}{N}\sum_{n=1}^{N} l^{SR}(G_{\theta G}(I_n^{LR}).I_n^{HR}) \qquad (3)$$

As a weighted combination of numerous loss components that model various desirable properties of the recovered SR image, A perceptual loss I^{SR} is explicitly built by us.

B. Adversarial Network Architecture

In keeping with paper [20], A discriminator network $D_{\theta D}$ is defined by us that alternately optimize with $G_{\theta D}$ in order to solve the adversarial min-max problem:

$$min_{\theta G}max_{\theta D}\mathrm{E}I^{HR}\sim p_{train}(I^{HR})[logD_{\theta D}(I^{HR})]+\mathrm{E}I^{LR}\sim p_{train}(I^{HR})[\log(1-G_{\theta D}(I^{HR}))] \qquad (2)$$

This formulation's fundamental notion is that it enables the training of a generative model G with the aim of deceiving a differentiable discriminator D that is trained to distinguish between genuine images and super-resolved images. By using this strategy, our generator can learn to produce results that are very close to actual photographs and challenging for D to categorize.
This encourages solutions that reside in the subspace, or manifold, of natural images and are perceptually superior. In contrast, like the MSE, pixel-wise error measurements are used to develop SR solutions.
A very deep generator network G, which is shown in Figure. 2, is composed primarily of B residual blocks with the same topology. We use the block arrangement suggested by [22], which was inspired by [28]. A batch-normalization layer is specifically employed by us [27] and ParametricReLU [25] as the activation function, followed by two convolutional layers with tiny 33 kernels and 64 feature maps. With the help of two
A discriminator network trained by us to distinguish actual HR photos from artificial SR data. In Fig 2, the architecture is displayed. Throughout the network, use LeakyReLU activation (= 0.2) and avoid max-pooling. The discriminator network is trained to resolve the maximizing problem in Equation 2. It comprises eight convolutional layers, identical to the VGG network, and 3x3 filter kernels that increase in number from 64 to 512 by a factor of 2 [41]. Stepped convolutions are used to reduce the image resolution every time the number of features doubles. To determine the chance of classifying the data, two thick layers, the 512 feature maps that result, and a final sigmoid activation function are used. MSE.

C. *Perceptual Loss Function*

This perceptual loss function defines the effectiveness of the generator network depending on SR. From [28] and [5] a loss function is constructed by us that gives a solution concerning perceptually relevant qualities, even though l^{SR} is frequently represented based on the MSE [10, 40]. The weight of the total content loss (l_X^{SR}), which is how we define the perceptual loss,

$$l^{SR}=l_X^{SR}+10^{-3}l_{Gen}^{SR} \quad (4)$$

D. *Content loss*

The formula for calculating the pixel-wise MSE loss is as follows:

$$l_{MSE}^{SR}=\frac{1}{r^2WH}\sum_{x=1}^{rW}\sum_{y=1}^{rH}(I_{x,y}^{HR}-G_{\theta_G}(I^{LR})_{x,y})^2 \quad (5)$$

The majority of cutting-edge methods for image SR rely on this as their most popular optimization goal [10, 40]. However, despite obtaining an especially high PSNR, MSE optimization problem solutions frequently lack high-frequency content, resulting in perceptually unpleasant results with excessively smooth textures . We build on the theories of [17], [5], and [28] and employ a loss function that is more closely related to perceptual similarity in place of relying on pixel-wise losses. Based on the ReLU activation layers of the pre-trained 19-layer VGG network mentioned in [41], we define the VGG loss. With i,j the feature map acquired by the J^{th} convolution (after activation) within the VGG19 network before the i^{th} max pooling layer, which was taken as supplied by us. The feature representations of a reconstructed image $G_{\theta_G}(I^{LR})$ and the reference image I^{HR} are used to determine the VGG loss:

$$l_{VGG/i,j}^{SR}=\frac{1}{W_{i,j}H_{i,j}}\sum_{x=1}^{W_{i,j}}\sum_{y=1}^{H_{i,j}}(\phi_{i,j}(I^{HR})_{x,y}-\phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2 \quad (6)$$

Here $W_{i,j}$ and $W_{i,j}$ outline the specific feature maps' sizes within the VGG network.

E. *Adversarial Loss*

The generative element of our GAN is added to the perceptual loss in addition to the content losses already mentioned. To trick the discriminator network, this encourages the network to favor solutions that are based on a variety of natural images. Based on the probabilities of the discriminator $D_{\theta_D}(G_{\theta_G}(I^{LR}))$ overall training samples, the generative loss l_{Gen}^{SR} is defined as:

$$l_{Gen}^{SR}=\sum_{n=1}^{N} -\log D_{\theta_D}(G_{\theta_G}(I^{LR})) \quad (7)$$

In this case, $D_{\theta_D}(G_{\theta_G}(I^{LR}))$ represents the likelihood that the reconstructed image $G_{\theta_G}(I^{LR})$ is an actual HR image. Instead of minimizing $\log[1 - D_{\theta_D}(G_{\theta_G}(I^{LR}))]$ [20], we minimize $\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$ for improved gradient behavior.

VIII. EXPERIMENTS

A. *Data and similarity measures*

We perform the experiment using the FRIDA dataset, all experiments are performed with a factor of four times between LR and HR images. This results in an image pixel reduction of a factor of 16.

All reported PSNR [dB] and SSIM [48] metrics were computed on the y-channel of center-cropped images with each border removed by a 4-pixel wide strip using the data package. The online material is supplementary to [26] and was used to create super-resolved pictures for the reference methods, including the nearest neighbor, bicubic.

B. Training Details And Parameters

Using an NVIDIA GeForce RTX 3050 GPU, we trained all networks using a random sample of 350000 photos from the ImageNet database [39]. Compared to the test images, these pictures are different. By employing a bicubic kernel and a downscaling factor of 4, we were able to derive the LR images from the HR images (BGR, C = 3). We randomly select 16 images, 96 x 96 HR submerges of various training images for each mini-batch. Because the generator model is of arbitrary size we can use it to create images of any size convolutional. We scaled the HR picture range to [-1, 1] and the LR input image range to [0, 1]. To determine the MSE loss, pictures with an intensity range of [-1, 1] were used. To produce VGG losses on par with the MSE loss, VGG feature maps were additionally rescaled by a factor of 11.25%. This is equal to adding a rescaling factor of 0.006 to Equation 5. We utilize [31] with β_1 = 0.9 for optimization. The SRResNet networks were trained using 10^6 update iterations and a learning rate of 10^4 Adam [31], with β_1 = 0.9, is used for optimization. With 10^6 update iterations and a learning rate of 10^{-4}, the SRResNet networks were trained. The trained MSE-based SRResNet network is initialization for the generator when training the actual GAN to prevent undesirable local maxima. The training process for all SRGAN variants included update iterations at a learning rate of 10^{-4} and further 10^5 iterations at a slower rate of 10^{-5}. We alternate the generator and discriminator network updates, which is similar to [20]'s (k = 1). The 16 residual blocks in our generator network are all the same (*B*= 16). To get a result that deterministically depends exclusively on the input during test time, the batch-normalization update is turned off [27].

C. Performance Of The Final Networks

We assess how well SRResNet and SR-GAN perform compared to NN, bicubic interpolation, and four cutting-edge techniques. In the supplemental material, examples of super-resolved images created using SRResNet and SRGAN are shown. They demonstrate that SRGAN significantly outperforms all reference methods and establishes a new benchmark for photo-realistic image SR.

IX. DISCUSSION AND FUTURE WORK

Figure 3: Foggy image (LR) to HR image

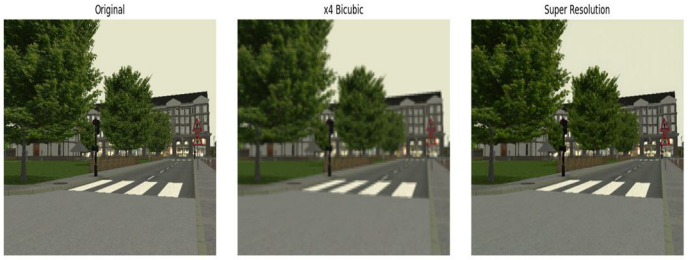


Figure 4: LR to HR image.

The perceived quality of super-resolved images rather than computing efficiency was the main focus of this effort. Contrary to [40], the provided model is not designed for real-time video SR. However, preliminary research on the network architecture indicates that thinner networks may be able to offer very effective substitutes at a negligible loss in performance quality. Unlike [10], we discovered deeper network designs to be advantageous. We hypothesize that the ResNet architecture significantly affects the effectiveness of deeper networks. We discovered that deeper networks (B > 16) can improve SRResNet performance even more but at the expense of longer training and testing times (cf. supplemental material).

Additionally, we discovered that SRGAN variations of deeper networks are getting harder to train because high-frequency artifacts start to show up. The choice of content loss is crucial if photo-realistic solutions to the SR problem are desired. The perceptually most compelling findings in this study were obtained using $l^{SR}_{VGG/5.4}$, which we attribute to the ability of deeper network layers to represent aspects of higher abstraction [55, 53, 35] outside of pixel space. The major distinction between super-resolved images without the adversarial loss and photo-realistic images, according to our hypothesis, is that feature maps of these deeper layers concentrate solely on the content, whereas the adversarial loss concentrates on texture details. We also point out that the best loss function varies depending on the situation. For instance, methods that hallucinate finer detail may not be as suitable for surveillance or medical purposes. It is difficult and will be the subject of future research to rebuild text or structured sceneries in a perceptually convincing manner [26]. Further enhancing photo-realistic image SR outcomes will be the creation of content loss functions that characterize picture spatial content while being more resistant to changes in pixel space.

X. CONCLUSIONS

When assessed using the commonly used PSNR measurement, the deep residual network (SRRes-Net), we have developed a new state of the art on publicly available benchmark datasets. We also generate different types of low-resolution images, a situation that was never considered previously in our domain. To improve visibility in such difficult situations, take help from this proposed algorithm. We have outlined some of the drawbacks of this PSNR-focused image super-resolution and presented SRGAN, which improves the content loss function by training a GAN to account for an adversarial loss. We have established through extensive MOS testing that SRGAN reconstructions for large upscaling factors are significantly more photo-realistic than reconstructions made using cutting-edge reference methods.

Author's Contribution

The authors confirm their contribution to the paper as follows: Study conception: Atin Bera, Arya Bhattacharyya; data collection: Debmitra Ghosh; Analysis and interpretation of results: Atin Bera, Arya Bhattacharyya, and Debmitra Ghosh; Draft manuscript preparation: Atin Bera, Arya Bhattacharyya Author. All authors reviewed the results and approved the final version of the manuscript.

XI. ACKNOWLEDGEMENT

We are deeply grateful to all those who contributed to the success of this research project First and foremost, we would like to thank our primary supervisor Debmitra Ghosh, for their guidance, support, and encouragement throughout the entire process. Their mentorship and expertise were invaluable in helping us shape the direction of our research and bring our ideas to fruition. We would also like to express our gratitude to the members of our research team, Atin Bera and Arya Bhattacharyya, who provided valuable input, insights, and assistance at every stage of the project. Their contributions were critical to the success of this research, and we are deeply grateful for their hard work and dedication. Without their generous contributions, this project would not have been possible

REFERENCES

[1] N. Hautiere, J.-P. Tarel, and D. Aubert, "Towards fog-free in- ` vehicle vision systems through contrast restoration," in IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'07), Minneapolis, Minnesota, USA, 2007
[2] R. Tan, N. Pettersson, and L. Petersson, "Visibility in bad weather from a single image," in Proceedings of the IEEE Intelligent Vehicles Symposium (IV'07), Istanbul, Turkey, 2007, pp. 19–24.
[3] S. G. Narashiman and S. K. Nayar, "Interactive deweathering of an image using physical model," in IEEE Workshop on Color and Photometric Methods in Computer Vision, Nice, France, 2003
[4] S. Borman and R. L. Stevenson. Super-Resolution from Image Sequences - A Review. Midwest Symposium on Circuits and Systems, pages 374–378, 1998. 2
[5] J. Bruna, P. Sprechmann, and Y. LeCun. Super-resolution with deep convolutional sufficient statistics. In International Conference on Learning Representations (ICLR), 2016. 2, 3, 5
[6] D. Dai, R. Timofte, and L. Van Gool. Jointly optimized regressors for image super-resolution. In Computer Graphics Forum, volume 34, pages 95–104, 2015. 2
[7] E. Denton, S. Chintala, A. Szlam, and R. Fergus. Deep generative image models using a laplacian pyramid of adversarial networks. In Advances in Neural Information Processing Systems (NIPS), pages 1486–1494, 2015. 3
[8] S. Dieleman, J. Schluter, C. Raffel, E. Olson, S. K. Snderby, ¨D. Nouri, D. Maturana, M. Thoma, E. Battenberg, J. Kelly, J. D.Fauw, M. Heilman, diogo149, B. McFee, H. Weideman, takacsg84, peter de rivaz, Jon, instagibbs, D. K. Rasul, CongLiu, Britefury, and J. Degrave. Lasagne: First release., 2015. 6
[9] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In European Conference on Computer Vision (ECCV), pages 184–199. Springer,2014. 3, 6, 8

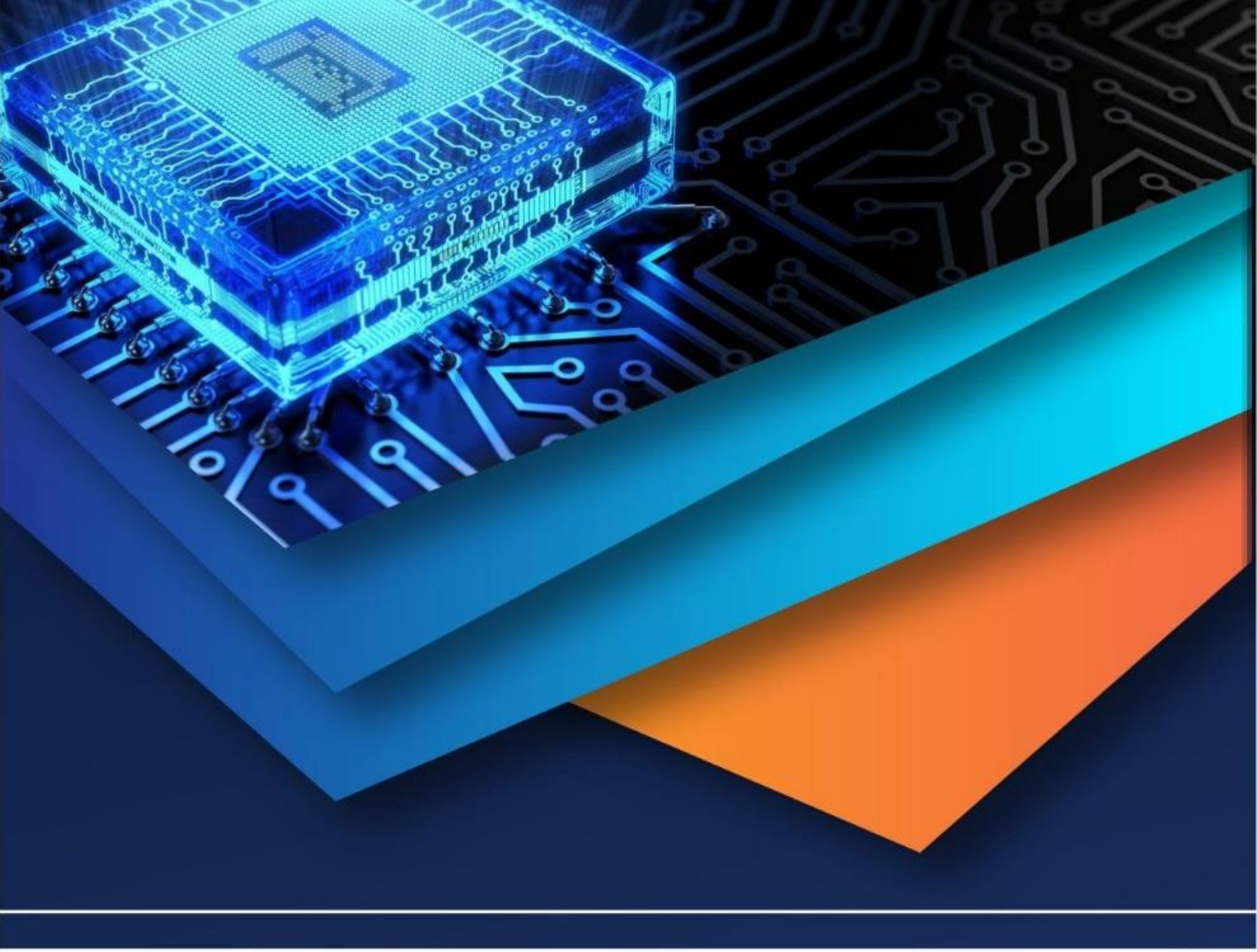
[10] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2):295–307, 2016. 3, 5, 7
[11] C. Dong, C. C. Loy, and X. Tang. Accelerating the super-resolution convolutional neural network. In European Conference on Computer Vision (ECCV), pages 391–407. Springer, 2016. 3
[12] W. Dong, L. Zhang, G. Shi, and X. Wu. Image deblurring and superresolution by adaptive sparse domain selection and adaptive regularisation. IEEE Transactions on Image Processing, 20(7):1838–1857,2011. 2
[13] A. Dosovitskiy and T. Brox. Generating images with perceptual similarity metrics based on deep networks. In Advances in Neural Information Processing Systems (NIPS), pages 658–666, 2016. 3
[14] J. A. Ferwerda. Three varieties of realism in computer graphics. In Electronic Imaging, pages 290–297. International Society for Optics and Photonics, 2003. 2
[15] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based superresolution. IEEE Computer Graphics and Applications, 22(2):56–65,2002. 2
[16] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low level vision. International Journal of Computer Vision, 40(1):25–47,2000. 2
[17] L. A. Gatys, A. S. Ecker, and M. Bethge. Texture synthesis using convolutional neural networks. In Advances in Neural Information Processing Systems (NIPS), pages 262–270, 2015. 3, 5
[18] L. A. Gatys, A. S. Ecker, and M. Bethge. Image Style Transfer Using Convolutional Neural Networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2414–2423, 2016. 3
[19] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In IEEE International Conference on Computer Vision (ICCV), pages 349–356, 2009. 2
[20] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In Advances in Neural Information Processing Systems (NIPS), pages 2672–2680, 2014. 3, 4, 6
[21] K. Gregor and Y. LeCun. Learning fast approximations of sparse coding. In Proceedings of the 27th International Conference on Machine Learning (ICML-10), pages 399–406, 2010. 3
[22] S. Gross and M. Wilber. Training and investigating residual nets, online at http://torch.ch/blog/2016/02/04/resnets. html. 2016. 4
[23] S. Gu, W. Zuo, Q. Xie, D. Meng, X. Feng, and L. Zhang. Convolutional sparse coding for image super-resolution. In IEEE International Conference on Computer Vision (ICCV), pages 1823–1831. 2015. 2
[24] P. Gupta, P. Srivastava, S. Bhardwaj, and V. Bhateja. A modified psnr metric based on hvs for quality assessment of colour images. In IEEE International Conference on Communication and Industrial Application (ICCIA), pages 1–4, 2011. 1
[25] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In IEEE International Conference on Computer Vision (ICCV), pages 1026–1034, 2
[26] J. B. Huang, A. Singh, and N. Ahuja. Single image super-resolution from transformed self-exemplars. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 5197–5206, 2015. 2,6, 8
[27] S. Ioffe and C. Szegedy. Batch normalisation: Accelerating deep network training by reducing internal covariate shift. In Proceedings of The 32nd International Conference on Machine Learning (ICML), pages 448–456, 2015. 3, 4, 6
[28] J. Johnson, A. Alahi, and F. Li. Perceptual losses for real-time style transfer and super- resolution. In European Conference on Computer Vision (ECCV), pages 694–711. Springer, 2016. 2, 3, 4, 5, 7
[29] J. Kim, J. K. Lee, and K. M. Lee. Deeply-recursive convolutional network for image super-resolution. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. 3, 6,
[30] K. I. Kim and Y. Kwon. Single-image super-resolution using sparse regression and natural image prior. IEEE Transactions on Pattern analysis and Machine Intelligence, 32(6):1127–1133, 2010.
[31] D. Kingma and J. Ba. Adam: A method for stochastic optimization. In International Conference on Learning Representations (ICLR),2015. 6
[32] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (NIPS), pages 1097–1105, 2012. 3
[33] C. Li and M. Wand. Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2479–2486, 2016. 3, 4
[34] X. Li and M. T. Orchard. New edge-directed interpolation. IEEE Transactions on Image Processing, 10(10):1521–1527, 2001. 2
[35] A. Mahendran and A. Vedaldi. Visualising deep convolutional neural networks using natural pre-images. International Journal of Computer Vision, pages 1–23, 2016. 7, 8
[36] M. Mathieu, C. Couprie, and Y. LeCun. Deep multi-scale video prediction beyond mean square error. In International Conference on Learning Representations (ICLR), 2016. 3
[37] K. Nasrollahi and T. B. Moeslund. Super-resolution: A comprehensive survey. In Machine Vision and Applications, volume 25, pages 423–1468. 2014. 1, 2
[38] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. In International Conference on Learning Representations (ICLR), 2016. 3, 4
[39] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. Imagenet large scale visual recognition challenge. International Journal of Computer Vision, pages 1–42, 2014. 6
[40] 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. 3, 4, 5, 6, 7, 8
[41] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations (ICLR), 2015. 2, 3, 4, 5
[42] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum. Image super-resolution using gradient profile prior. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1–8, 2008. 2
[43] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1–9, 2015. 3

[44] Y.-W. Tai, S. Liu, M. S. Brown, and S. Lin. Super Resolution using Edge Prior and Single Image Detail Synthesis. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2400– 2407, 2010. 2
[45] R. Timofte, V. De, and L. Van Gool. Anchored neighborhood regression for fast example-based super-resolution. In IEEE International Conference on Computer Vision (ICCV), pages 1920–1927, 2013. 2
[46] R. Timofte, V. De Smet, and L. Van Gool. A+: Adjusted anchored neighbourhood regression for fast super-resolution. In Asian Conference on Computer Vision (ACCV), pages 111–126. Springer, 2014. 2
[47] Y. Wang, L. Wang, H. Wang, and P. Li. End-to-End Image SuperResolution via Deep and Shallow Convolutional Networks. arXiv preprint arXiv:1607.07680, 2016. 3
[48] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4):600–612, 2004. 1,6
[49] Z. Wang, D. Liu, J. Yang, W. Han, and T. Huang. Deep networks for image super-resolution with sparse prior. In IEEE International Conference on Computer Vision (ICCV), pages 370–378, 2015. 3
[50] C.-Y. Yang, C. Ma, and M.-H. Yang. Single-image super-resolution: A benchmark. In European Conference on Computer Vision (ECCV), pages 372–386. Springer, 2014. 1, 2
[51] J. Yang, J. Wright, T. Huang, and Y. Ma. Image super-resolution as sparse representation of raw image patches. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1–8, 2008. 2
[52] R. Yeh, C. Chen, T. Y. Lim, M. Hasegawa-Johnson, and M. N. Do. Semantic Image Inpainting with Perceptual and Contextual Losses. arXiv preprint arXiv:1607.07539, 2016. 3
[53] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson. Understanding Neural Networks Through Deep Visualization. In International Conference on Machine Learning - Deep Learning Workshop 2015, page 12, 2015. 7, 8
[54] H. Yue, X. Sun, J. Yang, and F. Wu. Landmark image superresolution by retrieving web images. IEEE Transactions on Image Processing, 22(12):4865–4878, 2013. 2
[55] M. D. Zeiler and R. Fergus. Visualising and understanding convolutional networks. In European Conference on Computer Vision (ECCV), pages 818–833. Springer, 2014. 7, 8
[56] R. Zeyde, M. Elad, and M. Protter. On single image scale-up using sparse-representations. In Curves and Surfaces, pages 711–730. Springer, 2012. 2, 6
[57] K. Zhang, X. Gao, D. Tao, and X. Li. Multi-scale dictionary for single image super-resolution. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1114–1121, 2012. 2

Atin Bera has already completed his higher secondary education from Naktala High School (H.S) and is currently pursuing his B.Tech. at Jis University; currently, he is in his third year. He is interested in the fields of artificial intelligence and Machine learning. Previously, he wrote some papers on Artificial Intelligence and Machine Learning.

Arya Bhattacharyya completed his Higher Secondary from K.V. Barrackpore Air Force and is currently engaged in his B.Tech degree at JIS University. At present, he is doing research in the fields of Artificial intelligence and Machine Learning. In the future, it is planned to do a deep dive in technology and Image Processing.

Debmitra Ghosh is currently working as Assistant Professor in the department of Computer Science Engineering in JIS University.He is interested in the fields of artificial intelligence and Machine learning. Previously, he wrote some papers on Artificial Intelligence and Machine Learning.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY