Generative Adversarial Network based Image Super-Resolution to Enhance the shadow-less Image

19AD798 - Dissertation

Project Report - Review 2

Guided By: Dr. Rajathilagam B **Submitted on:** 18th June 2021

Submitted by:
Jishnu P (CB.EN.P2AID19017)

ABSTRACT

The presence of shadow is unavoidable while dealing with the outdoor images in a variety of computer vision applications. Shadows always trouble computer vision tasks like visual navigation, object detection and tracking. In order to unveil the information occluded by shadow, it is essential to remove the shadow. This is a two-step process which involves shadow detection and shadow removal. However, the field of shadow detection greatly improved compared to the shadow removal task, but the field of shadow removal has been difficult because it needs to be restored after removing the shadow.

While reconstructing region occluded by the shadow, the lack of quality in that area seems to be a challenging task. To address this issue, Image super-resolution is introduced. Single image super-resolution is a domain in image processing that aims at recovering a high-resolution image from a single low-resolution image. The need of high resolution image is essential for the computer vision applications for better performance in pattern recognition and analysis of images. In this paper, the low quality of the shadow-less image generated using the shadow removal GAN is improved using the Generative Adversarial network based Image Super-resolution model. By applying the image Super-resolution to the shadow-less image, we can improve the quality of the image as well as the overall performance of the shadow removal process. SRD (shadow removal dataset) , BSD 100(Berkeley Segmentation Dataset), DIV2K dataset are used as the datasets and the performance of the model evaluated using the metrics PSNR(Peak Signal to Noise ratio) and SSIM (Structural Similarity Index).

1. CHALLENGES AND MOTIVATION

- The shadow detection phase has greatly improved , but the field of shadow removal has been difficult because it needs to be restored after removing the shadow.

- The shadow removed region seems to be poor in quality compared to the non-shadow region in the shadow-less image.
- After the introduction of Generative Adversarial Networks (GAN) in 2014[1], the computer vision domain has taken leap at various tasks.
- GAN based Image Super-Resolution method is efficient for enhancing the low quality images.

2. LITERATURE SURVEY

There are various ways of enhancing image quality. One of the most commonly used technique is interpolation. This is easy to use but this leads to distorted image or reduces the visual quality of image. Most common interpolation methods produce blurry images, i.e. bi-cubic interpolation. More sophisticated methods exploit internal similarities of a given image or, use datasets of low- resolution images and their high- resolution counterparts to effectively learn a mapping between them.

Image super-resolution using generative adversarial networks introduced by Christian Ledig[2]. The approach uses the generative adversarial network with perceptual loss which is a combination of the adversarial loss and the content loss. BSD100, set5, set14 benchmark datasets were used to evaluate the performance and the corresponding metrics are PSNR = 25.16 & SSIM = 0.6688. The deep learning model introduced by C.Dong[3] named SRCNN was the first deep learning method to outperform traditional ones. It is a Convolutional Neural Network consisting of only 3 convolution layers: patch extraction and representation, non- linear mapping and reconstruction.This method has improved metrics PSNR = 25.70 & SSIM = 0.7184, which is better than the earlier method.

To improve the performance of the model, J.Kim[4] introduced very deep convolutional neural network model for more accurate image super-resolution. This method uses the similar structure as the SRCNN with the depth of 20, which makes the network more deeper and produces more accurate results and PSNR = 27.29 & SSIM = 0.725. Enhanced Super-Resolution Generative Adversarial Networks by Wang X[5] introduced the Residual-in-Residual Dense Block (RRDB) without batch normalization as the basic network building unit in the generative adversarial network(GAN). The model has the performance metrics PSNR = 24.83 & RMSE = 15.15 . Another method using the Generative Adversarial Network for Image Super-Resolution[6] with Combined Texture Loss improved the performance of the super-resolution method with metrics PSNR = 27.67 & SSIM = 0.764 .

2.1 INFERENCES

From the literature survey, we could find that the existing models user for the super-resolution are unable to detect and remove the shadows in the input images efficiently. The shadow also enhanced after the super-resolution. The metrics used for evaluating the performance are PSNR and SSIM. These metrics can be improved further by introducing more efficient models.

3. PROBLEM DEFINITION

Development of an Image Super-Resolution model with Generative Adversarial Network(GAN) which can improve the quality of the shadow-less image and improve the performance of higher level computer vision tasks in surveillance systems.

3.1. OBJECTIVES OF WORK

- Use Generative Adversarial Network for enhancing the quality of the given low resolution image.

- ➢ Use this image super-resolution module with the shadow removal module to enhance the quality of the shadow-less image. - "Enhanced Shadow Removal module"
- ➢ Evaluate the performance of the proposed model using the evaluation metrics and compare with existing models.

3.2 NOVELTY OF PROPOSED APPROACH

- ➢ Ensuring the quality of the shadow free image generated by the shadow removal module by introducing the super-resolution module in the pipeline.
- ➢ Introducing enhanced architecture for generator and discriminator and Combined loss for GAN training.

4. BASIC FLOW OF PROPOSED APPROACH

The image super-resolution technique is very important in the shadow-removal process since it improves the resolution of the shadow-less image and make it suitable for the higher-level computer vision tasks.

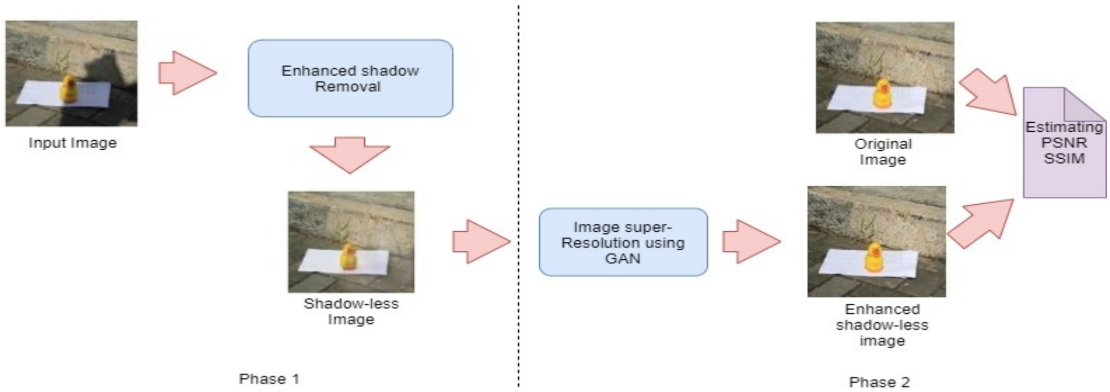

Fig 1: Basic flow of proposed approach

The basic flow diagram of the proposed approach illustrated in figure 1. The shadowed input image is fed to the enhanced shadow removal module, which is a modified condition based generative adversarial network and generates the shadow-less image for the given shadowed image. In order to improve the quality of the shadow-less image, it is given as the input to the generative adversarial network based image super-resolution module and corresponding super-resolution image is generated. The evaluation metrics such as the PSNR and SSIM will be evaluated on the resulting image.

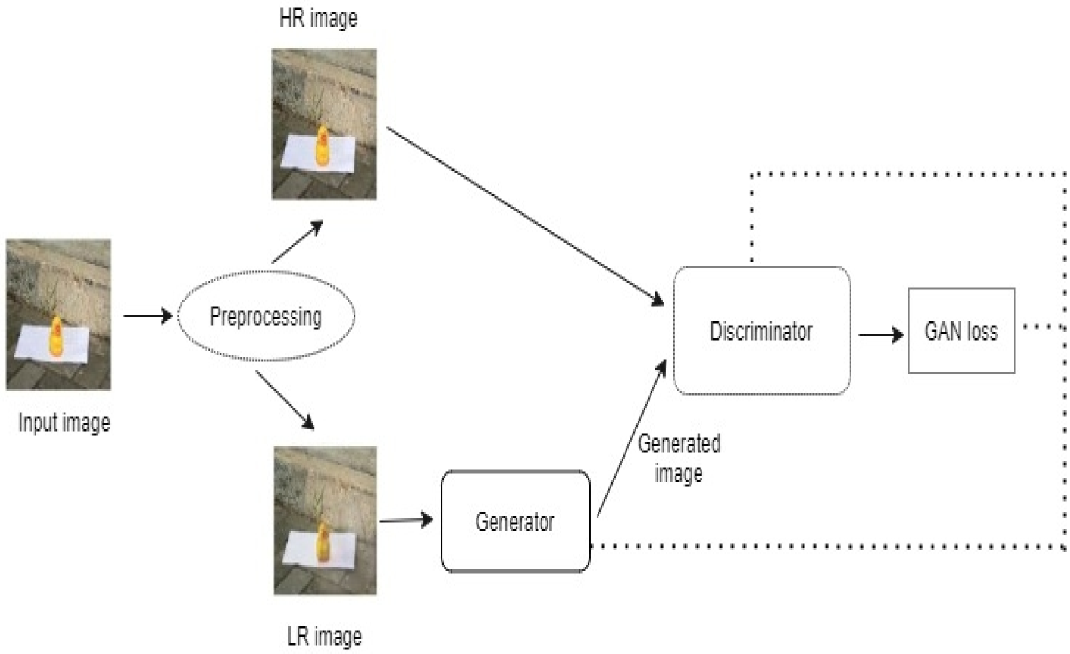

Fig 2: Basic architecture of Image super-resolution module

The basic architecture of the image super resolution module is illustrated in figure2. The input image will be converted to HR(High Resolution) image and the LR(Low Resolution) image during the preprocessing stage using interpolation techniques such as bicubic interpolation. LR image will be given as the input to the generator and produces the SR(Super Resolution) Image called as the generated image. The discriminator will distinguish between the HR image and the SR image and ensure that the GAN generates the Super resolution image of the given Input image. The training of generator and Discriminator will be carried out using the GAN loss.

4.1 ARCHITECTURE OF GENERATOR

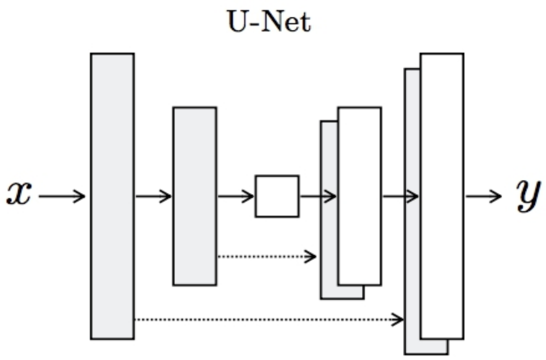

Fig 3: U-Net Architecture of Generator

Fig 3 Shows the overall abstract architecture of the Generator. The generator is an encoder-decoder model using a U-Net architecture[16]. The encoder and decoder of the generator contain convolutional, batch normalization, dropout, and activation layers. The encoder encodes the information in the given input image and the context information generated in the bottleneck is used for reconstructing the image using the decoder block. Skip connections are used between corresponding encoder and decoder layers for improving the quality of the image generated by the generator.

4.2 ARCHITECTURE OF DISCRIMINATOR

Architecture of the Discriminator is illustrated in Fig 4. The discriminator is a deep convolutional neural network that performs image classification. It takes both the source image (generated image) and the target image (shadow-less image) as input and predicts the likelihood of whether the shadow-less image is real or a fake translation of the generated image.

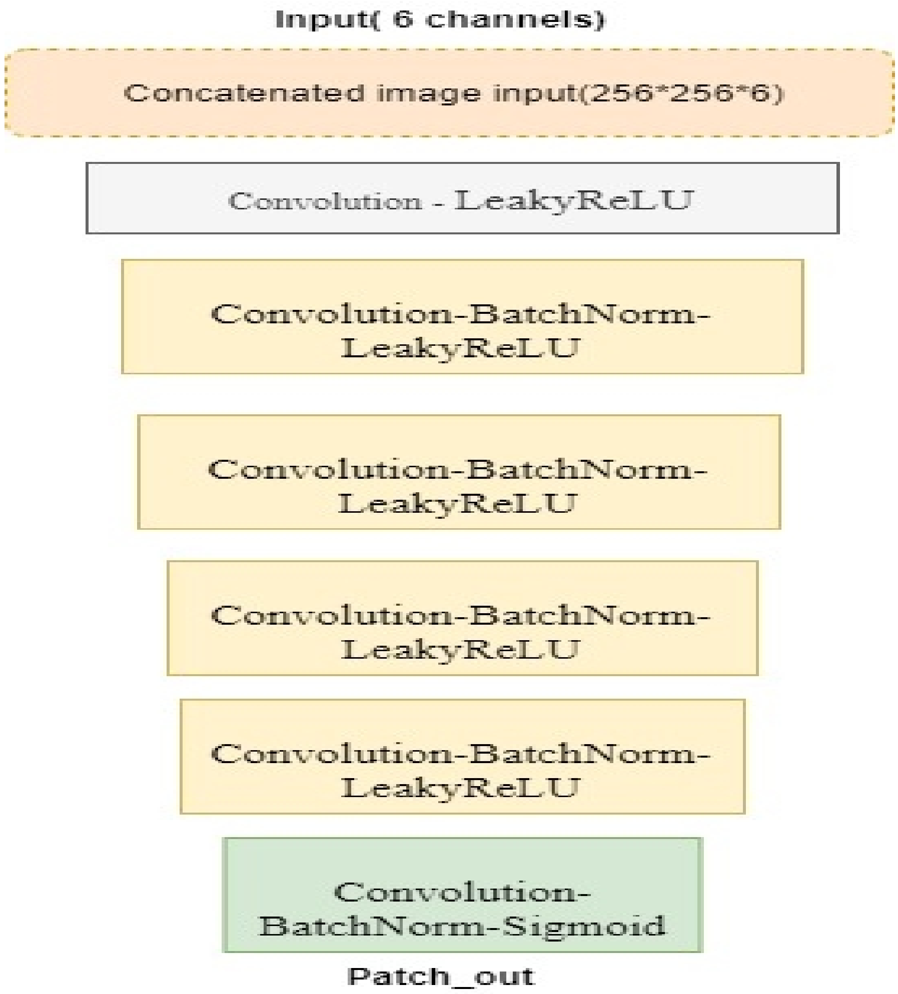

Fig 4: Architecture of Discriminator

The discriminator model is trained in the same way as a traditional GAN model. Adversarial training is used for training the discriminator model.

5. ALGORITHM DESIGN FOR PROPOSED SYSTEM

Adversarial training is used for training the discriminator model. The training of generator is too slow compared to the discriminator. This will lead to the issues in the GAN such as Vanishing gradient, Mode collapse, Non - convergence. The obvious solution is to balance generator & discriminator training to avoid over- fitting.

Algorithm 1: Enhancing Low Resolution image using GAN

Input : The LR and HR image
Output : Super-resolution(SR) image
d_epochs - discriminator training epochs
bs - batch size

for number of training epochs **do**
  for d_epochs **do**
    * sample minibatch of bs samples from LR images
    * sample minibatch of bs samples from HR images
    * Update the discriminator using adversarial loss

  end for
    * sample minibatch of bs samples from LR images
    * Update generator using combined loss
    Combined loss = λ1 * reconstruction loss + λ2 * adversarial loss
  end for

Fig 5: Algorithm design for enhancing the LR image.

Combined Loss(1) introduced to balance the training of Generator and Discriminator. Giving more importance to the reconstruction loss(2) rather than the adversarial loss(3) during the Generator training will reduce the effect of the fast training of discriminator on the Generator.

$$\textbf{\textit{Combined loss}} = \lambda 1 * \textbf{\textit{reconstruction loss}} + \lambda 2 * \textbf{\textit{adversarial loss}} \quad (1)$$

Where,

$\lambda 1, \lambda 2 = 100, 1$

Loss function between the generated fake image and the shadow-less image is called as the reconstruction loss.

$$L_{reconstruction} = \sum_{i=1}^{n} |y_{true} - y_{predicted}| \quad (2)$$

Where,

Ypred = The predicted value for the ith pixel.

Ytrue = The observed(actual) value for the ith pixel

n = Total number of pixels.

Adversarial loss (Binary Cross Entropy loss)

$$Ladversarial(G, D) = E_{xy}[logD(x, y] + E_{xy}[log(1 - D(x, G(x, y)))] \quad (3)$$

Where,

G - Generator , D - Discriminator

x - shadow-less image , y - shadowed image

6. DATASETS

BSD100 and DIV2K are the two datasets used in this project. Both datasets are the benchmark datasets used for the single image super resolution.

- **BSD 100** (Berkeley Segmentation Dataset) is a dataset used frequently for image denoising and super-resolution.
- Composed of 1100 large variety of images ranging from natural images to object-specific such as plants, people, food.
- Size = 575 MB
- Image format = .png

Fig 6: Sample images in BSD 100 dataset

- **DIV2K** is a dataset used frequently for image super-resolution.
- Size = 3.2GB
- Image format = .png
- Training = 800 images

Fig 7: Sample images in DIV2K dataset

6.1 PREPROCESSING

- Low Resolution(LR) images are generated by applying BICUBIC interpolation technique on High Resolution (HR) images.
- Saved the dataset as BSD(HR&LR).npz

7. TOOLS FOR IMPLEMENTATION

- Programming Language

  * Python 3.8

- Numerical Computation
  - * Tensorflow 1.13.2
  - * Numpy 1.19.x
- Visualization tools
  - * Matplotlib 3.3.2

8. EVALUATION METRICS

PSNR and SSIM are the two metrics used to evaluate the performance of the proposed model.

PSNR (Peak signal to noise ratio)

* This ratio is used as a quality measurement between the original and a compressed image.

* Higher the PSNR, the better the quality of the compressed or reconstructed image.

$$PSNR(dB) = 10 * \log(\frac{255^2}{MSE})$$

SSIM

* SSIM is used as a metric to measure the perceptual similarity between two given images.

* Value ranges from -1(very different) to +1(very similar)

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

9. APPLICATIONS OF PROPOSED MODEL

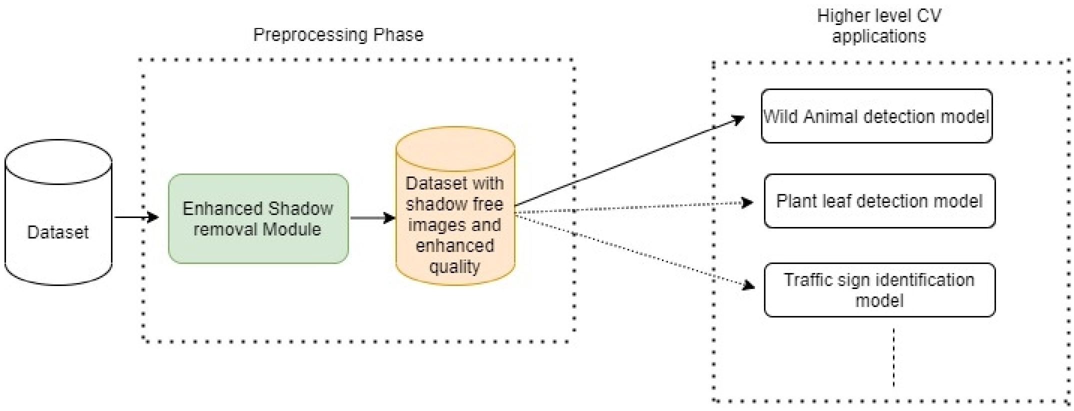

Fig 8: Illustration of the application of our model

Our model can be used in the preprocessing step of all the higher level computer vision applications as shown in Fig8.

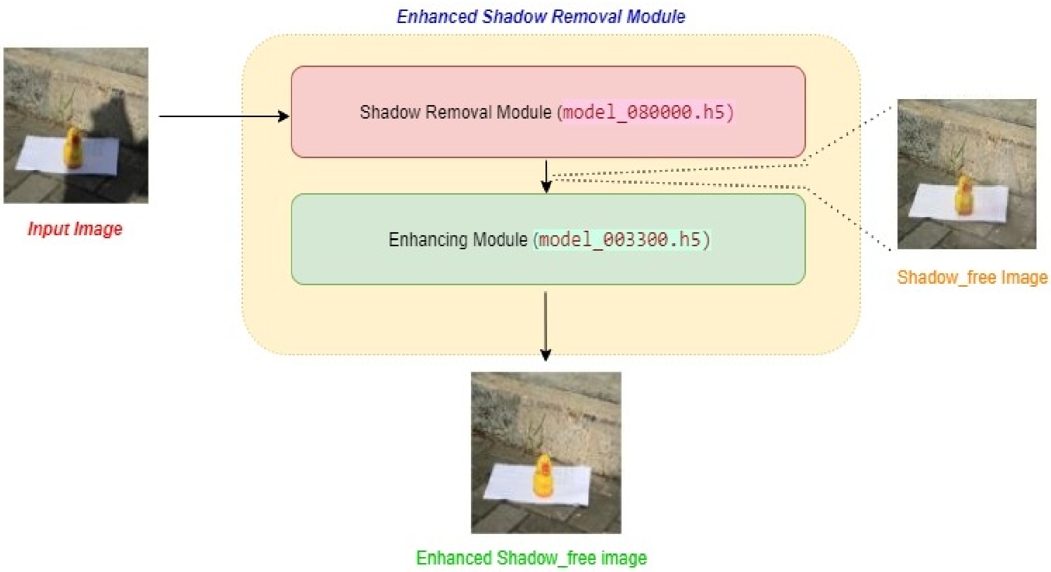

Fig 9: Illustration of the final module

The enhanced shadow removal module is illustrated in Fig9. The trained model for shadow-removal and image super-resolution are added in the pipeline in the order respectively. The resultant module will produce the enhanced shadow-free image for the given input image in 43seconds.

10. RESULTS AND DISCUSSION

The sample output of the proposed model after training for is illustrated in Fig 10. In this image its clear that the generator generated more realistic and super-resolution images for the given low resolution images.

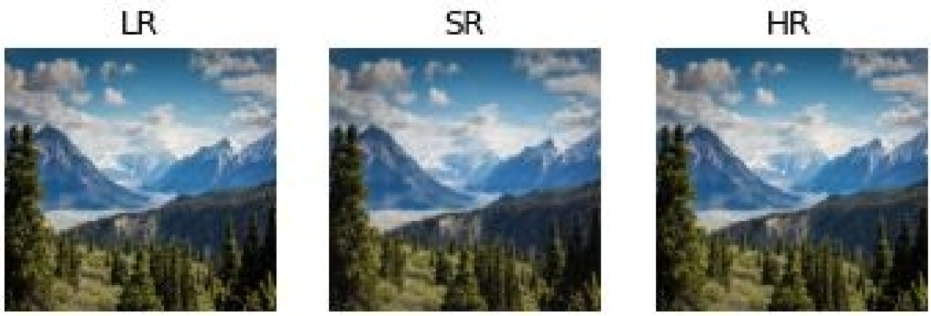

Fig 10: sample output of super-resolution module

The generator and discriminator loss is illustrated in the Fig 12. The discriminator loss is low compared to the generator and the generator loss reduces over time and it indicates that the generated started generating super resolution images which are almost similar to the HR images given to the Discriminator as the ground truth image.

Fig 11: sample output of enhanced shadow removal module

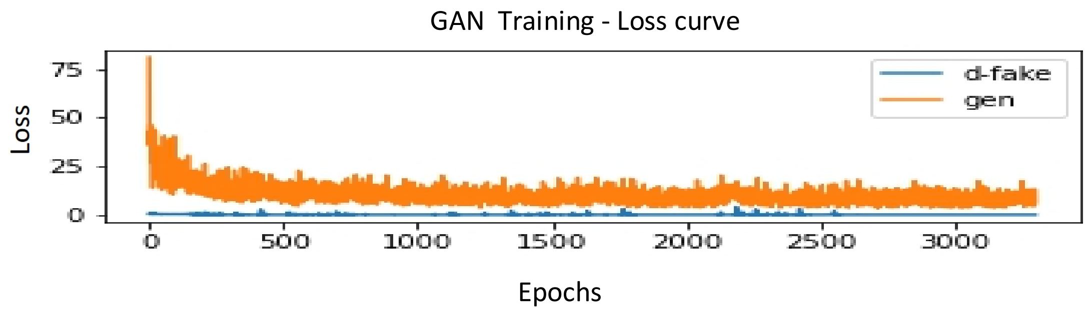

Fig 12: sample output of enhanced shadow removal module

Enhanced shadow-removal module output is illustrated in the Fig11. We can see that the low resolution image generated by the shadow removal module further enhanced by the image enhancement module. The Enhanced shadow-removal module guarantees that the model always produce the enhanced shadow-less images for the given shadowed images.

Table 1 : Performance analysis of Image Enhancement module

Model	PSNR(dB)	SSIM
SRGAN[2]	25.16	0.668
ESRGAN[5]	27.66	….
GAN combining Texture loss[6] (2020)	27.99	0.778
Our Model	72.67	0.798

The performance analysis of the image enhancement module is give in the Table 1. Our model performs well compared to all other standard algorithms and shows huge improvement in the PSNR value.

Table 2 : Performance Analysis of Enhanced Shadow removal module

Model	RMSE	PSNR	SSIM
ST-GAN[7]	7.47	30.66	…..
RIS-GAN[8]	6.67	31.64	…..
Our model	1.997	53.38	0.722

Performance analysis of the enhanced shadow removal module is given in the Table 2.

Our model performs well compared to the standard algorithms. SSIM introduced as the new evaluation metric and shows good results.

11. CONCLUSION

In this project, we have introduced Generative Adversarial Network based image enhancement module and Enhanced Shadow removal module. Initially, we use basic conditional GAN model on BSD100 dataset and analyzed the areas of improvements. Secondly, we modified the architecture and introduced a combined loss for training the model. The image enhancement module trained successfully and evaluated the performance. Then the module combined with shadow removal module to produce enhanced shadow-free images. The proposed model works well and produces good results compared to the standard algorithms.Tuned the parameters by conducting repeated experiments and identified the appropriate set of parameters for the model. From the experiments, it was seen that the proposed model performs better than the existing model.

As a future enhancement, hyper parameter tuning is a time-consuming task and also requires more domain knowledge. In addition to future enhancement, Neuroevolution techniques can be used to tune the hyper-parameters to identify the best set of parameters for the model and thereby improving the performance of our model.

12. REFERENCES

1. [1]Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, “Generative Adversarial Networks”, International Conference on Neural Information Processing Systems, 2014, https://arxiv.org/abs/1406.2661

2. [2]Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi, “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network” , Computer Vision and Pattern Recognition , 2017,https://arxiv.org/abs/1609.04802v5

3. [3]J. Kim, J. K. Lee and K. M. Lee, "Accurate Image Super-Resolution Using Very Deep Convolutional Networks," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1646-1654, doi: 10.1109/CVPR.2016.182.

4. [4]C. Dong, C. C. Loy, K. He and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295-307, 1 Feb. 2016, doi: 10.1109/TPAMI.2015.2439281.

5. [5]Wang X. et al. (2019), “ ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks.” In: Leal-Taixé L., Roth S. (eds) Computer Vision – ECCV 2018 Workshops. ECCV 2018,Springer, Cham. https://doi.org/10.1007/978-3-030-11021-5_5

6. [6]Jiang Y, Li J. Generative Adversarial Network for Image Super-Resolution Combining Texture Loss. Applied Sciences. 2020; 10(5):1729. https://doi.org/10.3390/app10051729

7. Jifeng Wang, Xiang Li2, Jian Yang, “Stacked Conditional Generative Adversarial Networks for Jointly Learning Shadow Detection and Shadow Removal”,CVPR 2017, https://doi.ieeecomputersociety.org/10.1109/CVPR.2018.00192

8. Ling Zhang,1 Chengjiang Long,2∗ Xiaolong Zhang,1 Chunxia Xiao3, “RIS-GAN: Explore Residual and Illumination with Generative Adversarial Networks for Shadow Removal”, AAAI 2020,https://doi.org/10.1609/aaai.v34i07.6979

9. L. Qu, J. Tian, S. He, Y. Tang and R. W. H. Lau, "DeshadowNet: A Multi-context Embedding Deep Network for Shadow Removal," IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017,https://ieeexplore.ieee.org/document/8099731

10. Sidorov, Oleksii, “Conditional GANs for Multi-Illuminant Color Constancy: Revolution or Yet Another Approach?”,The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2018,https://arxiv.org/abs/1811.06604

11. Wang, H. Xu, Z. Zhou, L. Deng and M. Yang, "Shadow Detection and Removal for Illumination Consistency on the Road," in IEEE Transactions on Intelligent Vehicles, 2020, https://ieeexplore.ieee.org/document/9068460

12. Wang, L. Deng, Z. Zhou, M. Yang and B. Wang, "Shadow detection and removal for illumination consistency on the road, "https://ieeexplore.ieee.org/document/8304275

13. Yun, Heejin & Kang Jik, Kim & Chun, Jun-Chul. “Shadow Detection and Removal From Photo-Realistic Synthetic Urban Image Using Deep Learning”, Computers, Materials & Continua2019, https://www.techscience.com/cmc/v62n1/38123

14. Xiaodong Cun, Chi-Man Pun, Cheng Shi, “ Towards Ghost-free Shadow Removal via Dual Hierarchical Aggregation Network and Shadow Matting GAN” AAAI 2020, https://arxiv.org/abs/1911.08718

15. M. Mirza, S. Osindero, "Conditional Generative Adversarial Nets", In Arxiv 2014. http://arxiv.org/abs/1411.1784

16. Olaf Ronneberger, Philipp Fischer, Thomas Brox,”U-Net: Convolutional Networks for Biomedical Image Segmentation”MICCAI 2015, https://arxiv.org/abs/1505.04597