Image Super Resolution using Generative Adversarial Networks

Dinesh Adhithya

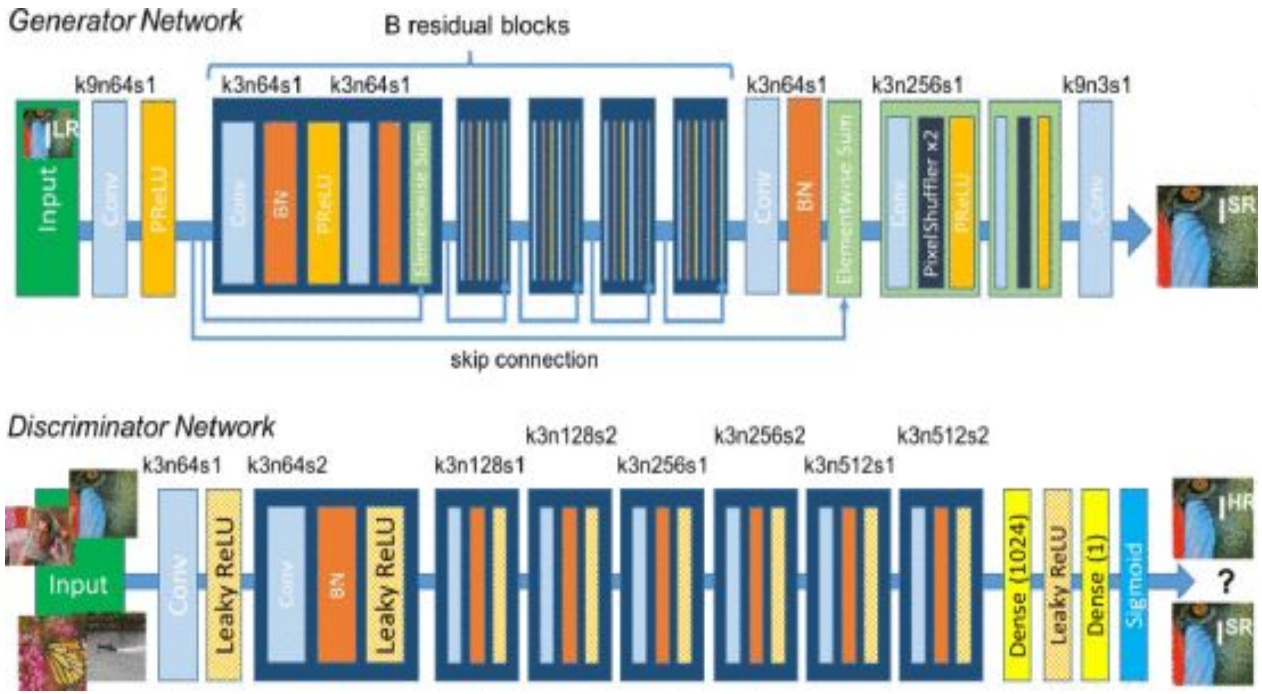
Introduction

- The aim of this work is to use deep learning techniques to improve the resolution of images used for semantic segmentation tasks.
- CNN (Convolutional neural network) based models outperform previously developed models and are a good choice for this work as well.
- Recently GAN based models have been proposed for this task and make use of conventional RMSE reconstruction error and adversarial loss.
- GAN's make use of a discriminator and Generator approach to construct high resolution images from low resolution images.This work makes use of residual networks , GAN's and transfer learning approaches to upsample images of high resolution from low resolution images.

Approach

- Make use of GAN’S , A generator architecture generates takes a low resolution image and up samples it into high resolution image.
- The discriminator predicts which of the two high resolution images is fake and real.
- The adversarial loss , of the discriminator is back propagated to the generator along mean square loss.
- The two models teach each other as a min-max game and the generator can be used to upsample images further on.

Model


Generator Network
B residual blocks
k9n64s1
k3n64s1
k3n64s1
k3n64s1
k3n256s1
k9n3s1
Input
LR
Conv
PReLU
Conv
BN
PReLU
Element-wise Sum
Conv
BN
Element-wise Sum
Conv
PixelShuffler x2
PReLU
Conv
SR
skip connection
Discriminator Network
k3n64s1
k3n64s2
k3n128s1
k3n128s2
k3n256s1
k3n256s2
k3n512s1
k3n512s2
Input
Conv
Leaky ReLU
Conv
BN
Leaky ReLU
Dense (1024)
Leaky ReLU
Dense (1)
Sigmoid
HR
?
SR

Experiments and Results

- **Metrics**

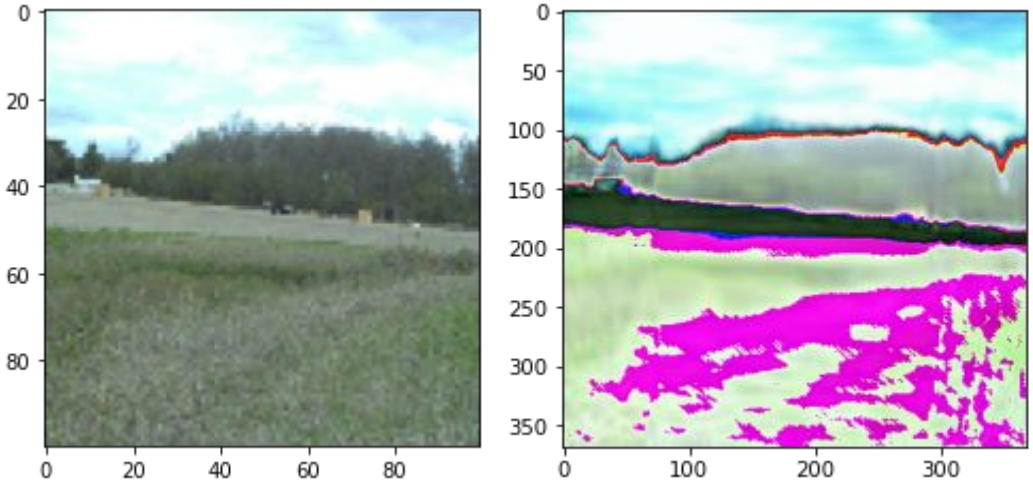
  PSNR = 20*log10(255) -10*log10(MSE)

- **Model Training**

  Trained on NVIDIA GEFORCE RTX 3060
  17,779 Rellis-3D images used for training
  Tested on RUGD dataset

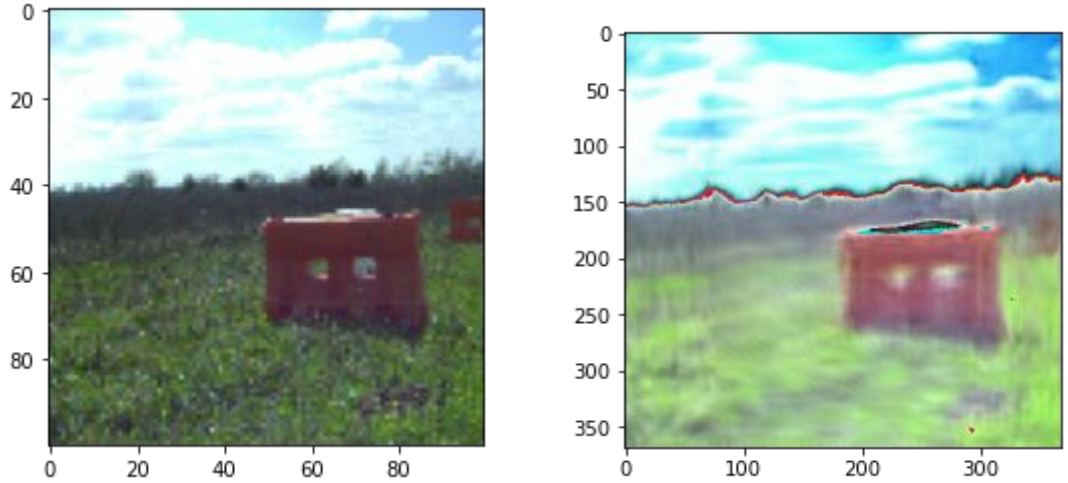
Model Performance

- SResNet model for model prediction with an PSNR score of **25.04** and final loss at **0.054** during training.



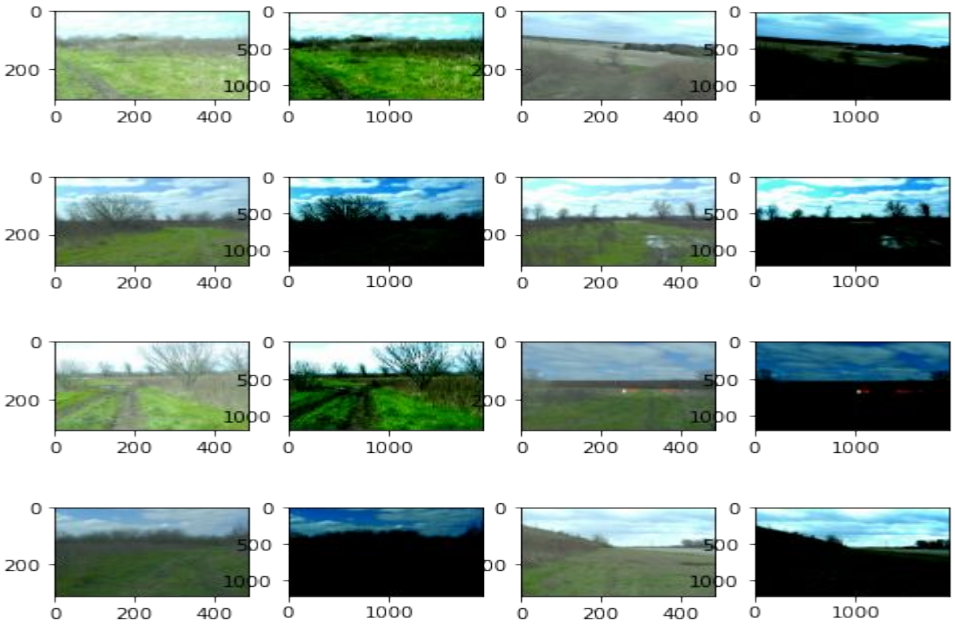
Model Performance

- SR GAN without the adversarial loss
- On the 100*100 low resolution images dataset the model performed with PSNR of **28.98** , with final loss at **0.00877**.The below image shows the results of the model.



Model Performance

- This approach makes use of **SRGAN model** with adversarial loss



Model Performance

- Transfer learning approach trained on the Rellis-3D dataset yielded images with loss 0.00012 with PSNR value of 35.5.

Conclusion

- Although we achieve considerable results using GAN'S , training and fine tuning of such models is a difficult task and divergence of such models makes them difficult to use.

- The predictions of the SRGAN can later be used to have a single unified dataset of images of both RUGD and RELLIS-3D image dataset for image segmentation tasks.

References

1. K. S. Viswanath, Kartikeya Singh, Peng Jiang, P. B. Sujit, and Srikanth Saripalli. Offseg: A semantic segmentation framework for off-road driving. 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE), pages 354–359, 2021.

2. Christian Ledig, Lucas Theis, Ferenc Husz´ar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network.
3. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 105–114, 2017.

4. C. Dong, C. C. Loy, K. He, and X. Tang. Image super resolution using deep convolutional networks, 2014.

5. J. Bruna, P. Sprechmann, and Y. LeCun. Super-resolution with deep convolutional sufficient statistics. CoRR, abs/1511.05666, 2015.