

Image Super Resolution Techniques Applied on Satellite Imagery

Chitipat Phentmune^{1,2} Hien Doan Thi^{2,3} Hoang Dieu Vu^{2,4}
Donghyun Ahn^{1,2} Hyunji Cha^{1,2} Sungwon Han^{1,2} Meeyoung Cha^{2,1}
¹KAIST, South Korea ²IBS, South Korea ³UET, Vietnam ⁴HUST, Vietnam

Abstract

This paper proposes an idea to apply deep learning models that generate super-resolution (SR) images for satellite pictures. We newly define the task by exploiting images of different zoom levels and demonstrate the results of implementing several recent models using over 12,080 samples. These models use satellite images at zoom level 16 and produce high-resolution images at zoom level 17 (i.e., four times the area of zoom level 16). We find that EDSR and RDN show significant improvement over Bicubic interpolation for images of urban built-up, but generate blurred output for other kinds of landcover such as vegetation. GAN-based methods create sharper images with details, although they achieve low PSNR scores. We confirm that different SR models are suitable for handling each landcover type on satellite imagery.

1. Introduction and Background

Recent studies have demonstrated how daytime satellite imagery can be used to extract meaningful economic and developmental insights over large geographical areas. As the zoom level increases by a unit, its resolution increases by two times, and the size of data that needs processing increases by four times (2×2). High-resolution images, however, are costly to obtain and often restricted in their usage, limiting their potential applications. Furthermore, as existing satellites wear out, the quality of images taken by those satellites tend to degrade. Hence the ability to reconstruct a high-resolution (HR) satellite image from a low-resolution (LR) one — *Super-Resolution (SR) technique* — could benefit the industry tremendously.

Some approaches use residual learning to connect the LR and HR images or between layers such as EDSR [3], yet these models suffer from vanishing and exploding gradients. Other methods like RDN [4] utilize DenseNet to mitigate gradient vanishing and handle the association of residual learning and dense block. Some approaches adopt generative adversarial networks (GAN). For instance, DeblurGAN [1] uses a perceptual loss to capture textual details. However, GAN-based models produce lower PSNR (peak signal to noise ratio), and they often introduce undesired artifacts under deeper networks. Models like EDSR [3] adopt the GAN loss [2] yet exploits the trade-offs between the PSNR value and realistic textual details.

Motivated by advances in SR techniques, our work newly defines the problem of constructing super-resolution land covers of broad geographic surfaces based on low-resolution images by exploiting the zoom-level structure in satellite imagery. We release a new dataset `SI_RAW` that includes 12,080 pairs of low and high-resolution satellite images collected from ArcGIS.¹ Our data comprise landcover images taken at both zoom level 16 (i.e., low-resolution) and zoom level 17 (i.e., high-resolution). We implemented and tested the following models: EDSR, DeblurGAN, RDN, and EDSR&GAN.

2. Methods and Findings

The experiment was conducted with a scaling factor of 2 at zoom level 16 and 17, which are reasonable to examine landcovers (e.g., water, vegetation, barren land, built-up). From the data, we discarded all cases where the LR and HR pictures were taken at different time periods and thereby increases the chance of observing differences in landcover, such as a new buildup structure. To ensure data quality, we utilized only those pairs that had gaps in PSNR scores more than 24. This process discarded 36% of LR-HR image pairs.

Overall, the EDSR and RDN models showed significant improvement over Bicubic interpolation in terms of the PSNR and SSIM scores (see Table 1). The PSNR scores were 30.82, 31.30, and 31.29 for Bicubic, EDSR, and RDN models, respectively. Visually, sample images of an urban built-up in Fig. 1 demonstrate that EDSR and RDN produce sharper and

¹We will release the implementation codes for four SR models via GitHub.

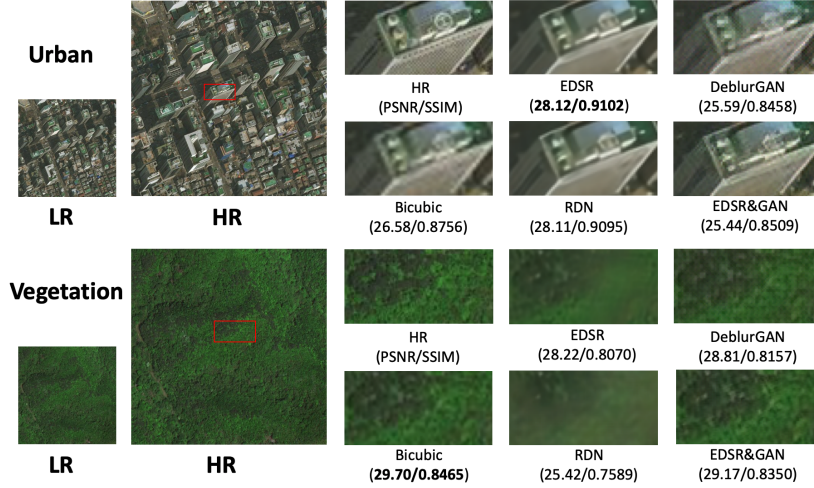


Figure 1. Visual results of super-resolution model with scale factor x2

Method	PSNR	SSIM
Bicubic	30.82	0.8860
EDSR	31.30	0.8934
RDN	31.29	0.8936
DeblurGAN	30.01	0.7459
EDSR&GAN	29.96	0.8692

Average PSNR and SSIM values for the scale factor x2. The best performance is shown in bold.

Table 1. Comparison result

more detailed texture than the Bicubic method. EDSR and RDN achieve nearly identical PSNR and SSIM scores. In terms of perceptual quality, EDSR and RDN produce a similar level of sharpness for urban areas with slightly better performance for RDN. RDN, however, shows blurry regions for images of vegetation landcover compared to EDSR.

In the GAN-based methods, DeblurGAN produces small square artifacts throughout the image, and EDSR&GAN generates abnormal and unpleasing noise in the urban picture. Both ways achieve lower PSNR and SSIM scores than the Bicubic interpolation.

3. Concluding Remark

The recent deep learning models developed for super-resolution (SR) techniques can be successfully applied to satellite images to enhance their quality. The PSNR-based methods, EDSR and RDN, can work well in urban areas but have overly smoothed artifacts in the vegetation landcover. The GAN-based models, DeblurGAN and EDSR&GAN, give a better perceptual result in vegetation without producing blurry areas. However, they showed unpleasing noise for the urban built-up landcover. We conclude that PSNR and SSIM alone are not a good measure for testing SR techniques for satellite images. This was demonstrated in the example where GAN-based methods showed perceptually better results although having lower PSNR scores than Bicubic interpolation. This calls for more accurate methods for measuring super-resolution performance should be researched further. To gain further improvement, one may utilize higher resolution images (e.g., a scaling factor of 4 or more) or a better architecture.

References

- [1] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8183–8192, 2018.
- [2] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *proc. of the CVPR*, 2017.
- [3] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee. Enhanced deep residual networks for single image super-resolution. In *proc. of the CVPR*, 2017.
- [4] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu. Residual dense network for image super-resolution. In *proc. of the CVPR*, 2018.