

# ESRGAN-BASED DEM SUPER-RESOLUTION FOR ENHANCED SLOPE DEFORMATION MONITORING IN LANTAU ISLAND OF HONG KONG

Zherong Wu<sup>1,2</sup>, Peifeng Ma<sup>1,2,\*</sup>

<sup>1</sup> Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong, PR China - iseis@cuhk.edu.hk

<sup>2</sup> Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen 518063, PR China - info@cuhkri.org.cn

**KEY WORDS:** ESRGAN, DEM, Super-Resolution, Slope Deformation

## ABSTRACT:

Monitoring, evaluating and understanding the slopes by Interferometric Synthetic Aperture Radar (InSAR) technology are critical for both human economy and natural environment. However, the resolution limitation of existing digital elevation model (DEM) in the slope areas causes the DEM phase residues and atmospheric effects promoted, which will influence the interpret accuracy of InSAR results. In this study, we propose a novel two-step ESRGAN-based DEM SR method to effectively recover high-resolution DEM from the original version. Firstly, we pretrain an ESRGAN with a large number of natural images. Based on it, we transfer the learnt knowledge into the DEM problem and fine-tune the DEM SR network. The recovered DEMs are utilized as the reference data to improve slope deformation monitoring and enhance the accuracy of InSAR estimation, especially in the mountainous areas with cloudy and rainy weather. Experiments indicate that the proposed method can achieve better results than the traditional methods and works in phase simulation, which is one of the key step of InSAR deformation monitoring.

## 1. INTRODUCTION

Landslide is an economically harmful and life-threatening hazard in many parts of the world, especially in the mountainous areas with cloudy and rainy weather (Dai, Lee, & Ngai, 2002; Iverson, 2000). As a type of microwave remote sensing (RS) technology, Interferometric Synthetic Aperture Radar (InSAR) has been widely used in landslide monitoring for its characteristic of high sensitivity of dynamic changes, high spatial resolution and wide coverage (Ma, Lin, Lan, & Chen, 2015). However, slopes often occur in areas with steep terrain, where the high-resolution and high-accuracy digital elevation models (DEMs) are usually confidential. Besides, the complexity of the spatial and temporal distribution of atmospheric phases increase the monitoring difficulties obviously. Therefore, a large number of DEM phase residues are left in the estimated deformation phase. Moreover, the phase delay caused by the vertical stratification of atmosphere is related to elevation, which will also affect the accuracy of repeat track interferometry. Therefore, it is especially useful to enhance the DEM quality and further improve the accuracy of InSAR slope monitoring.

Xu extended image SR in the DEM scene by proposing a nonlocal-based method (Xu et al., 2015). For the last five years or so, interest in the study of deep learning (DL) based single image super-resolution (SISR) methods has skyrocketed (Dong, Loy, He, & Tang, 2014; Kim, Kwon Lee, & Mu Lee, 2016; Tai, Yang, & Liu, 2017). Recent DL-based methods have achieved significant improvements both quantitatively and qualitatively. Among these methods, the Super-Resolution Generative Adversarial Network (SRGAN) pioneeringly augments the content loss function with an adversarial loss by training a generative adversarial network (Ledig et al., 2016). In order to further improve the visual quality, the SRGAN architecture has been improved to derive an Enhanced SRGAN (ESRGAN)

(Wang et al., 2018), which is capable of obtaining more realistic and natural textures.

This study introduces ESRGAN creatively into DEM super-resolution (SR) to address the problem of generally limited DEM resolution in the landslide areas. Section 2 introduces the general information of study area. Our DEM SR method and some details of network training are illustrated in Section 3. Section 4 analyses the effect test experiments and relevant results. Finally, the conclusion and future research are discussed.

## 2. STUDY AREA AND DATASETS

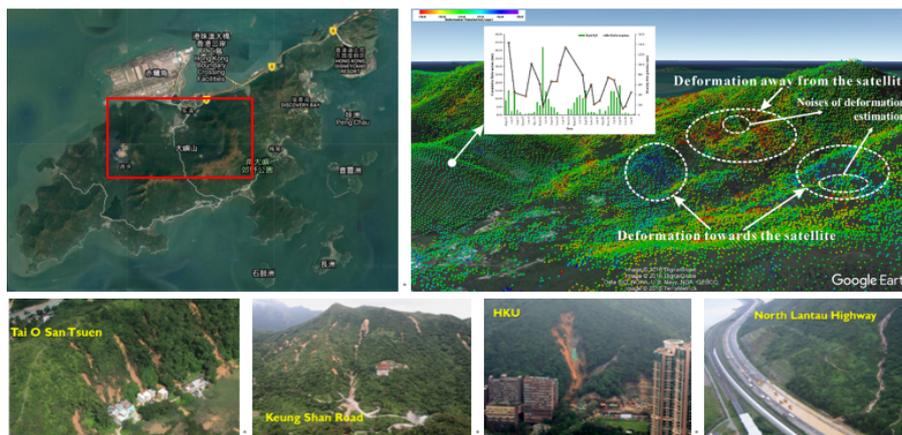
### 2.1 Study area

Lantau Island is located in the southwest of Hong Kong, with a total area of about 143 km<sup>2</sup>, as shown in *Figure 1*. It is the largest outlying island in Hong Kong. Due to the steep surface, there are only a small amount of plains on the slope near the sea, and human exploiting acts are rare in this region. As a result, the original natural state has been nearly completely preserved.

The main bedrock in Lantau area is volcanic rock and tuff seriously weathered to form a deeper eluvium, which is usually covered by younger colluvial and alluvial surface materials. The oldest rocks are sandstone and siltstone with smaller outcrops. Mixed forest of multiple tree species generates at the foot of the hillside. In the middle of the slope, there are dense shrubs and weeds. The outcrop of bedrock usually appears in the peak or the slope of more than 40 degrees. The microclimate in this region is mainly subtropical monsoon climate, which is characterized by torrid summers with high humidity and chill dry winters. The rainfall intensity is great with intensive storms and hurricanes.

---

\* Corresponding author



**Figure 1** Top left: Optical image of Lantau Island, Hong Kong Top right: Mean velocity from ALOS-1/PALSAR-1, Lantau Island, direction of satellite line of sight  
Bottom: Landslide impact of June 2008 rainstorm in Hong Kong

## 2.2 Datasets

### 2.2.1 Shuttle radar topography mission (SRTM)

The data of SRTM (Shuttle Radar Topography Mission) is mainly spearheaded jointly by the U.S. National Aeronautics and Space Administration (NASA) and the U.S. National Geospatial-Intelligence Agency (NGA). The radar images collected cover an area of 119 million km<sup>2</sup>. According to the resolution, SRTM data can be divided into SRTM1 and SRTM3, with the corresponding resolutions of 30m and 90m respectively (Jarvis, Reuter, Nelson, & Guevara, 2008). The elevation models derived from the SRTM data can be download freely and their file format is widely supported. SRTM DEM has been one of the most used data in geosciences since its release. This study mainly collects SRTM DEM data with the resolution of 30m.

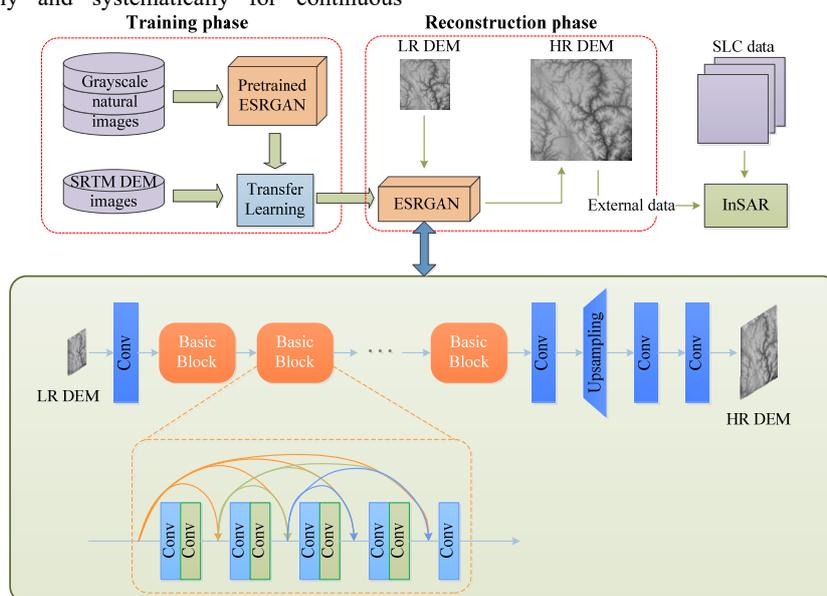
### 2.2.2 Sentinel-1 IW Single Look Complex (SLC) data

The Sentinel-1 mission (Torres et al., 2012) is funded by the funded by the Council of Europe (UC) and designed by the European Space Agency (ESA) for the goal of rapidly monitoring the environmental conditions and natural disasters around the world, comprised of a constellation of two polar-orbiting satellites. The Sentinel-1 SAR instrument and short revisit time will greatly advance capabilities of user data acquisition and provide data routinely and systematically for continuous

mapping of the Earth. Moreover, SENTINEL data is free of charge. The IW SLC product contains one image per sub-swath and per polarisation channel, with a total of three or six images. Each sub-swath image consists of a series of bursts, where each burst was processed as a separate SLC image.

## 3. METHODOLOGY

Inspired by the success of ESRGAN in image SR, this study introduces ESRGAN creatively into DEM super-resolution (SR) reconstruction. Nevertheless, directly utilizing image ESRGAN to the case of DEM is uncomplicated in theory but hard in practice. Most natural images are of 8-bit sizes, whose grey range is usually between 0 and 255, while the range of most DEM images exceeds 255 greatly (Xu et al., 2019). Thus, if training an ESRGAN for DEM SR directly, we need a considerable amount of DEM samples consist of vastly diverse height values. However, available high-resolution DEM training samples are insufficient at present. In order to address this issue, we proposed a novel method that includes two steps (**Figure 2**), pretraining the ESRGAN with abundant high-resolution natural images and fine-tuning the network with small quantities of SRTM DEM samples, detailed in 3.1 and 3.2.



**Figure 2** Pipeline of the proposed method and the ESRGAN for DEM SR

### 3.1 DEM SR network architecture

The designed DEM super-resolution network is based on the ESRGAN, as shown in **Figure 2**. The network is a GAN-based network consist of a feed-forward CNN  $G_{\theta_G}$  as the generator network and a discriminator network  $D_{\theta_D}$  to discriminate the original images and the reconstructed HR images.

As the optimization objective in generative adversarial nets (Goodfellow et al., 2014), the DEM SR network is trained to solve the min-max problem:

$$\min_{\theta_G} \max_{\theta_D} E_{I^{HR} \sim p_{train}(I^{HR})} [\log D_{\theta_D}(I^{HR})] + E_{I^{LR} \sim p_G(I^{LR})} \left[ \log(1 - D_{\theta_D}(G_{\theta_D}(I^{LR}))) \right] \quad (1)$$

The general idea behind is to train a generative model  $G$  with the optimization goal of deceiving pass a differentiable discriminator  $D$  that is improved iteratively to distinguish the super-resolved DEMs from real DEMs.

B Residual-in-Residual Dense Blocks (RRDB) with identical layout are at the core of the generator network  $G$ , which is shown in

**Figure 3** are. The layout is firstly proposed in (Gross & Wilber, 2016). Specifically, there are two convolutional layers with small  $3 \times 3$  kernels and 64 feature maps followed by ParametricReLU (He, Zhang, Ren, & Sun, 2015) as the activation function. We increase the resolution of the input image with two trained sub-pixel convolution layers as proposed by Shi et al. [48].

As to the architecture of the discriminator network (

**Figure 3**), it follows the guidelines summarized by (Radford, Metz, & Chintala, 2015). Besides, LeakyReLU activation ( $\alpha = 0.2$ ) is utilized to avoid max-pooling throughout the network. The network is consist of 8 convolutional layers with  $3 \times 3$  filter kernels, whose number increases from 64 to 512 gradually (Simonyan & Zisserman, 2014). The resulting 512 feature maps are followed by two dense layers and a sigmoid function finally to obtain a discriminant probability.

### 3.2 Training method and transfer learning

Based on the idea of perceptual loss proposed in SRGAN, we set the loss for the generator as:

$$L_G = \mu L_{percep} + \lambda L_G + \eta L_1 \quad (2)$$

Where  $L_1 = \mathbb{E}_{x_i} \|G(x_i) - y\|_1$  is the content loss that indicates the 1-norm distance between reconstructed result  $G(x_i)$  and the ground-truth  $y$ .  $L_G$  is the adversarial loss for generator. Specifically,  $L_{percep}$  is constrained on features before activation. And  $\mu, \lambda, \eta$  are the coefficients to balance different loss terms.

The high-resolution DEM samples cannot be acquired easily. If we train the DEM SR model from scratch, it may be hard to approach the global optimal solution. Inspired by the success of transfer learning, we apply the knowledge obtained from the natural images to DEM SR, that is, the network will be pretrained with the sets comprised by natural grayscale images, and fine-tuned with limited SRTM DEM data to serve the DEM SR.

### 3.3 Flowchart of the proposed method

#### Algorithm 1 ESRGAN-based DEM SR network

<b>Input:</b>	The low-resolution DEM $I_l$
<b>Output:</b>	The high-resolution DEM $I_h$
1:	Pretrain the enhanced super-resolution generative adversarial network using the grayscale natural images.
2:	Transfer the learnt weights as the initial value for DEM SR network.
3:	Fine-tune the model with DEM samples with various terrain characteristics.
4:	Produce the recovered high-resolution DEM by minimizing the error according to Eq. (1).

The flowchart of the proposed method is listed in Algorithm 1.

## 4. EXPERIMENT SETUP

### 4.1 Training details

The training process contains two stages. First during the pre-training phase, we train the network based on the  $L_1$  loss. The initial value of learning rate is set to 0.0004, and is decayed by the factor of 0.5 after 30 epochs. The batch size is set to 32. The threshold of flatness is 0.0 and 0.15. Typically, 300 epochs are sufficient for ESRGAN training using the mini-batch gradient descent method. There are 500 steps in per epoch.

In the transfer learning phase, we train the generator employing the model above as the initialization. The network is trained according to the loss function in Equation (2) with  $\mu = 0.5, \lambda = 0.005, \eta = 0.1$ .

adjust the initial learning rate to 0.0001. For optimization, we use Adam with  $\beta_1 = 0.9, \beta_2 = 0.999$ . We implement the model with the TensorFlow framework and train them using the NVIDIA GeForce RTX 2080 graphic processing units (GPUs) and free GPUs provided by Google Colaboratory.

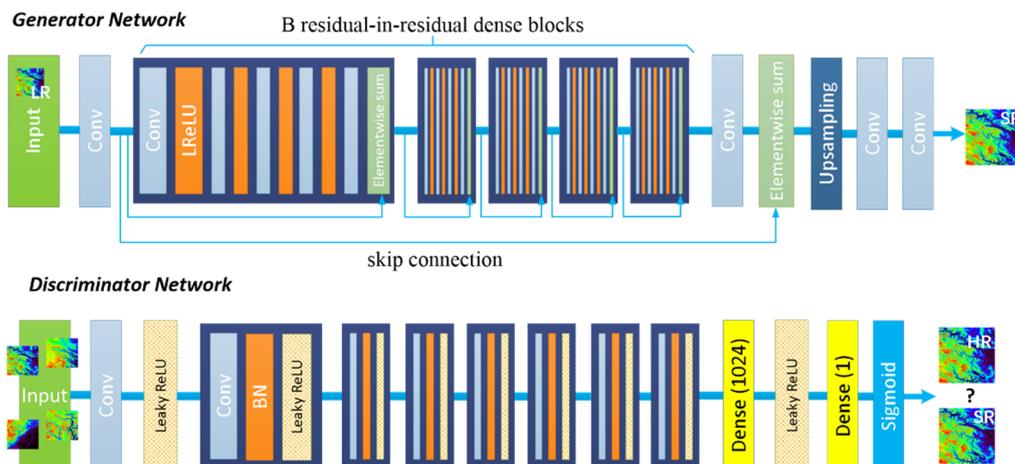


Figure 3 Architecture of Generator and Discriminator network

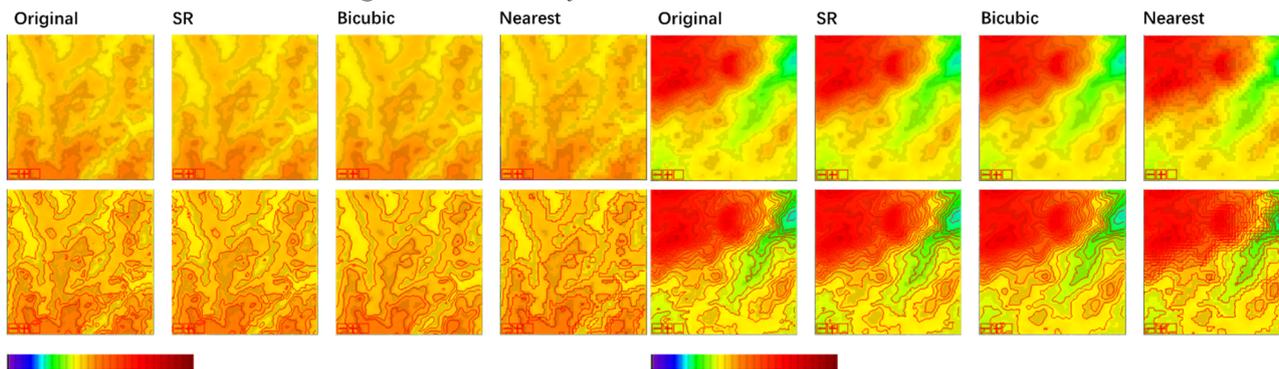


Figure 4 Qualitative results of ESRGAN-based DEM SR

## 4.2 Training and testing data

The pre-designed network is trained with the DIV2K datasets, which includes 900 natural image samples. Among them, 800 images are treated as training sets and left 100 images are used for validation. The original high-resolution images are down-sampled to get low-resolution versions by bicubic algorithm. Each sample consists of a high-resolution image and corresponding low-resolution one.

Then for transfer learning, SRTM1 DEM are used as training sets. We specially select DEM samples with different terrain characteristics and ensure that the number of multiple samples is as balanced as possible. The data is of 30-meter resolution and the size is 3601×3601. For network training, we divided them into overlapped 512×512 patches. It is worth mentioning that we amplified the samples with high elevation during patch sampling as a result of the lack of this type of samples.

## 5. EXPERIMENT AND DISCUSSION

### 5.1 Comparison with traditional methods

To verify the effectiveness of the SR algorithm, we compare our method with other traditional algorithms (e.g., bicubic, Nearest neighbor) in details. It can be observed from **Figure 4** that our proposed DEM SR network outperforms previous down-sampling approaches in both sharpness and details. Besides, our method also performs better in the restoration of the contours.

We also conduct experiments to quantitatively evaluate the ESRGAN-based method by the evaluation index, such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM).

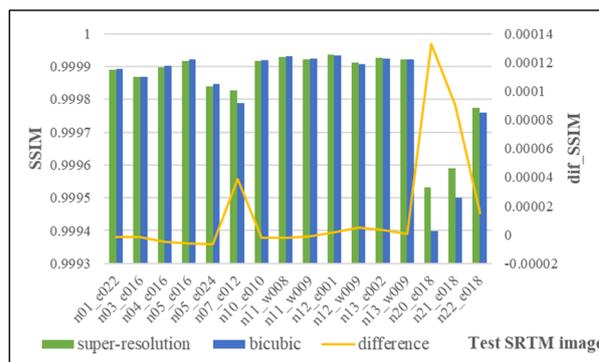
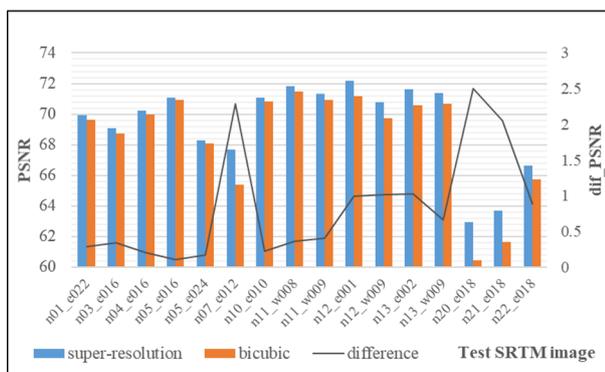


Figure 5 Quantitative evaluation on the SRTM test set

As **Figure 5** shows, the SR results outperforms the traditional bicubic method in the indices of PSNR and SSIM overall, but inferior slightly in some test images.

Besides, the absolute difference between the SR result and the ground truth of each SRTM DEM test patch is calculated. We also obtain the histograms of the difference images. **Figure 6** is one example of the results, which indicates that the difference image is zero-mean-value normal distribution approximately.

According to the above experiments, we can draw the conclusion that the results of the ESRGAN-based method are better than that of the traditional up-sampling methods, such as bicubic, nearest-neighbor, etc.

### 5.2 Effects on slope deformation monitoring

The proposed ESRGAN-based DEM SR method was utilized to provide detailed, high-resolution (HR) DEMs as the external reference data and further improve the robustness of deformation monitoring. Based on the HR DEM images down-sampled with multi bandwidths, we obtained a series of DEM samples under different resolutions and conducted comparative tests to gain the DEMs of the optimal resolution afterwards, which were introduced to enhance the slope deformation monitoring. The improvement of the slope deformation monitoring owing to the reconstructed high-resolution DEMs was explored in the mountainous Lantau Island area in Hong Kong. Both the theoretical evaluation (Bert, 2006) and comparison with in-situ data showed that the accuracy of monitoring results which were calculated based on the reconstructed DEMs were better than results with the original DEMs.

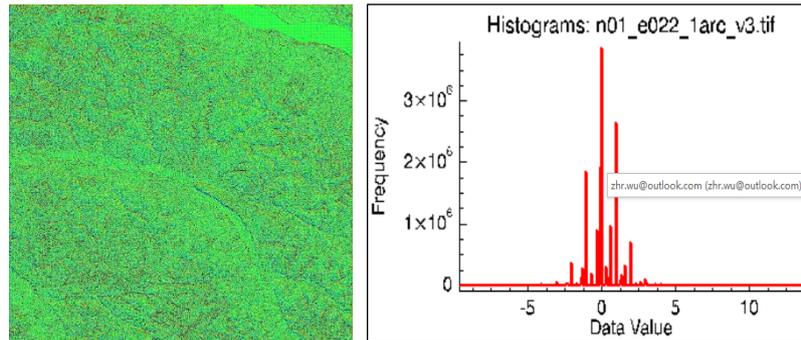


Figure 6 The difference image and histogram of the test DEM patch

## 6. CONCLUSION

In this study, we propose a ESRGAN-based DEM SR method. Based on the information learnt from external natural images and the original low-resolution DEM, this method can effectively reconstruct a DEM with more details and relatively high-resolution. Furthermore, we use it in the slope deformation monitoring and enhance the accuracy of phase simulation in InSAR processing. Meanwhile, our method also achieves a faster convergence speed especially when a large collection of high-resolution DEMs is confidential. The enhanced slope deformation detection method can assist the relevant government departments to enhance the warning capabilities of landslide disaster. To our knowledge, this is the first time to propose an ESRGAN-based SR method for the widely used SRTM DEM data, which may also benefit the fields of hydrology, meteorology, geology and engineering construction. However, various issues still need to be addressed. One issue to be resolved is applying the method to more potential landslide zones. Besides, the mutual supplementation of multi-sources information is another considerable way to achieve better DEM SR results. And the better match and error control of high-resolution DEM and Synthetic Aperture Radar (SAR) data will also be a key issue in our future work.

## ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (41601356, 41971278, 41671378) and the Research Grants Council (RGC) of Hong Kong (CUHK14504219).

## REFERENCES

Bert, M. K. (2006). Radar Interferometry: Persistent Scatterers Technique. *The Netherlands: Springer*.

Dai, F., Lee, C., & Ngai, Y. Y. (2002). Landslide risk assessment and management: an overview. *Engineering geology*, 64(1), 65-87.

Dong, C., Loy, C. C., He, K., & Tang, X. (2014). *Learning a deep convolutional network for image super-resolution*. Paper presented at the European conference on computer vision.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative adversarial nets*. Paper presented at the Advances in neural information processing systems.

Gross, S., & Wilber, M. (2016). Training and investigating residual nets. *Facebook AI Research*, 6.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Delving deep into rectifiers: Surpassing human-level performance on imagenet classification*. Paper presented at the Proceedings of the IEEE international conference on computer vision.

Iverson, R. M. (2000). Landslide triggering by rain infiltration. *Water resources research*, 36(7), 1897-1910.

Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). Hole-filled SRTM for the globe Version 4. *available from the CGIAR-CSI SRTM 90m Database*, 15, 25-54.

Kim, J., Kwon Lee, J., & Mu Lee, K. (2016). *Accurate image super-resolution using very deep convolutional networks*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2016, September 01, 2016). *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Ma, P., Lin, H., Lan, H., & Chen, F. (2015). Multi-dimensional SAR tomography for monitoring the deformation of newly built concrete buildings. *ISPRS Journal of Photogrammetry and Remote sensing letters*, 106, 118-128.

Radford, A., Metz, L., & Chintala, S. J. a. p. a. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks.

Simonyan, K., & Zisserman, A. J. a. p. a. (2014). Very deep convolutional networks for large-scale image recognition.

Tai, Y., Yang, J., & Liu, X. (2017). *Image super-resolution via deep recursive residual network*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., & Brown, M. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*, 120, 9-24.

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Change Loy, C. (2018). *ESRGAN: Enhanced super-resolution generative adversarial networks*. Paper presented at the

Proceedings of the European Conference on Computer Vision (ECCV).

Xu, Z., Chen, Z., Yi, W., Gui, Q., Hou, W., & Ding, M. (2019). Deep gradient prior network for DEM super-resolution: Transfer learning from image to DEM. *ISPRS Journal of Photogrammetry and Remote sensing letters*, 150, 80-90.

Xu, Z., Wang, X., Chen, Z., Xiong, D., Ding, M., & Hou, W. (2015). Nonlocal similarity based DEM super resolution. *ISPRS Journal of Photogrammetry and Remote sensing letters*, 110, 48-54.