

A GENERATIVE ADVERSARIAL NETWORK APPROACH FOR SUPER-RESOLUTION OF SENTINEL-2 SATELLITE IMAGES

F. Pineda^{1,2}, V. Ayma¹, C. Beltran¹

¹Pontificia Universidad Católica del Perú, Department of Engineering,
Artificial Intelligence Research Group (IA-PUCP), Lima, Perú
(fpineda@pucp.edu.pe, vaaymaq@pucp.pe, cbeltran@pucp.edu.pe)

²Universidad Nacional del Altiplano, Puno, Perú (ferpineda@unap.edu.pe)

KEY WORDS: PeruSat-1, Sentinel-2, Super-Resolution, GAN

ABSTRACT:

High-resolution satellite images have always been in high demand due to the greater detail and precision they offer, as well as the wide scope of the fields in which they could be applied; however, satellites in operation offering very high-resolution (VHR) images has experienced an important increase, but they remain as a smaller proportion against existing lower resolution (HR) satellites. Recent models of convolutional neural networks (CNN) are very suitable for applications with image processing, like resolution enhancement of images; but in order to obtain an acceptable result, it is important, not only to define the kind of CNN architecture but the reference set of images to train the model. Our work proposes an alternative to improve the spatial resolution of HR images obtained by Sentinel-2 satellite by using the VHR images from PeruSat1, a peruvian satellite, which serve as the reference for the super-resolution approach implementation based on a Generative Adversarial Network (GAN) model, as an alternative for obtaining VHR images. The VHR PeruSat-1 image dataset is used for the training process of the network. The results obtained were analyzed considering the Peak Signal to Noise Ratios (PSNR) and the Structural Similarity (SSIM). Finally, some visual outcomes, over a given testing dataset, are presented so the performance of the model could be analyzed as well.

1. INTRODUCTION

According to (Singh, 2018) there is an increasing demand for very high-resolution (VHR) images, which are the main resource to perform sophisticated applications in different research areas, such as computer vision, health, remote sensing, among others. Conventional models for super-resolution (SR) techniques, such as linear and cubic interpolations, splines, throws, and others, face some problems when processing fine details such as curves, edges, and textures, in images with high-frequency changes among their pixels (Singh, 2018). Deep Learning techniques raise as an alternative to these classic methods (Zhao et al., 2017), (Dong et al., 2016); more specifically, according to (Ledig et al., 2017), some of those complex issues can be solved by applying joint approaches between Generative Adversarial Networks (GAN) combined with Convolutional Neural Networks (Goodfellow et al., 2014).

Our approach for the SR model presents an alternative for improving the spatial resolution from the HR images obtained by Sentinel-2 satellite, preserving their spectral distribution; and the SR model was previously trained with a dataset of VHR images taken from the PeruSat-1 satellite. We take in consideration the model proposed by (Ledig et al., 2017), which besides integrates some pre-processing stages into the GAN network implementation so it can be able to process high-resolution (HR) satellite images.

The remainder of this work is organized as follows, Section 2 presents an overview on the satellites used and the fundamentals on Super-Resolution approaches; then, the dataset and the methodology are described in Section 3; the experimental design and results are presented in section 4, and finally, the conclusions and future work recommendations are discussed in section 5.

2. FUNDAMENTALS

In this section, we present the main concepts used in this work, describing the most important characteristics of the satellites used, and mainly describing the basis of the super-resolution approach.

2.1 PeruSat-1 Satellite

PeruSat-1 is an observation satellite (PerúSAT-1, satélite de Observación de la Tierra, 2018), launched on September 2016, and it is considered as the satellite with the highest spatial resolution in the region; their images can be used in different areas such as planning, agriculture, forestry, geology, disaster risk management, among others (Bartolom et al., 2017).

According to its governmental specifications (Comunicaciones, 2017), PeruSat-1 has a 2.8 meters resolution in its spectral bands; bands 1,2 and 3 correspond to blue, green, red spectra and band 4 correspond to the near-infrared (NIR); in the panchromatic spectral band, it reaches up to 70 centimeters spatial resolution, with a revisiting period of 21 days. The images covers an area corresponding to a swath width of 14.5 km and a radiometric resolution of 12 bits.

2.2 Sentinel-2 Satellite

Sentinel-2 is a European mission composed of two twin satellites (ESA, 2015), Sentinel-2A and Sentinel-2B, both with a frequency of five days revisiting period. Sentinel-2 is used in services such as monitoring, disaster management, security, climate change, and others. It can gather 13 spectral bands, 4 bands with a 10 meters resolution (bands 2, 3, 4 corresponding to the visible spectrum blue, green, red and band 8 corresponding to NIR); 6 bands with a 20 meters resolution (bands 5, 6, 7,

8a, 11 and 12), and 3 bands with a 60 meters resolution (bands 1, 9 and 10)

Figure 1 presents the spatial resolution differences from both PeruSat-1 and Sentinel-2 satellites, where it can be seen that Sentinel-2 images describe the scene with a lower spatial resolution compared with PeruSat-1 images, also presenting some issues when dealing with borders, textures, and some other details that are visualized when zooming into the images.

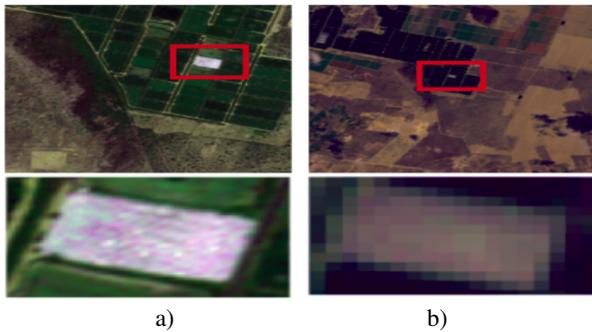


Figure 1. Example of the spatial resolution between PeruSat-1 and Sentinel-2. a) PeruSat-1 satellite image with 2.8 meters resolution per pixel, b) Sentinel-2 satellite image with 10 meters resolution per pixel

2.3 Super-Resolution (SR) Model

Super-resolution techniques are used to generate high resolution images from low resolution ones, and many solutions were proposed to improve the performance of the SR models since the past decades (Irani, Peleg, 1990). Huang and Yang (Huang, Yang, 2010) explore various SR alternatives such as those based on image observation models, processing at frequency domain and interpolation.

As an alternative to these classic SR models, we propose a deep learning approach based on a Generative Adversarial Network (GAN). GAN was introduced by Goodfellow et al. (Goodfellow et al., 2014), it proposes a new framework for the development of generative models via an adversarial process where two models are trained simultaneously; the first, a generative model (Generator) that captures the distribution of the data and creates a sample, and the second, a discriminative model (Discriminator) working as a probabilistic classifier, that estimates the probability that such a sample should be considered as a VHR or HR image. The generative model aims to maximize the probability that the discriminative model makes a mistake.

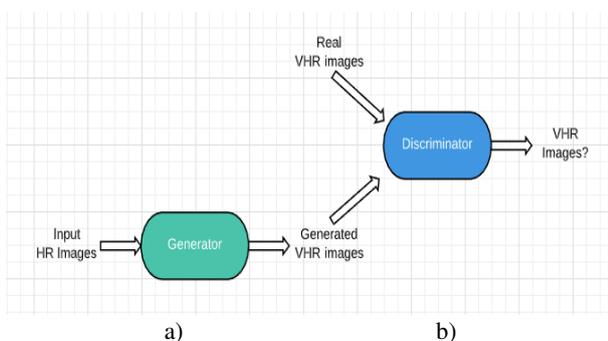


Figure 2. Overview of a Generative Adversarial Network

Figure 2 shows the GAN process, where an HR satellite image was used as input for the generator, which is trained to generate and increase the resolution to that image, obtaining a VHR

image. At the same time, the discriminator is trained as a classifier that must decide if the image received from the generator, or from the dataset with real VHR images, is a VHR or HR image indeed.

According to (Johnson et al., 2016) high-quality images can be generated using perceptual loss function; thus, for our GAN model approach we used the perceptual loss function proposed by (Ledig et al., 2017), using the features of a pre-trained VGG-19 network, where the perceptual loss function is the sum of two components, the content loss, l_x^{sr} , and an adversarial loss, such as:

$$l^{sr} = l_x^{sr} + 10^{-3}l_{Gen}^{sr} \quad (1)$$

3. METHODOLOGY

In this section, we describe the pre-processing steps that were performed on the satellite images; then we define the dataset based on VHR and HR images taken from the PeruSat-1 and Sentinel-2 satellites, respectively. Then, we present the methodology adopted in our work to perform the SR processes using a GAN network instantiation.

3.1 Pre-Processing stage

As a preliminary step to the experiments, some pre-processing steps were applied to the Sentinel-2 and PeruSat-1 satellite images to create the dataset used in our work. As shown in figure 3, the first step was to obtain the Sentinel-2 satellite images of the Copernicus platform and the PeruSat-1 images of the CONIDA COF catalog; as a second step, the PeruSat-1 (10512x10512) and Sentinel-2 (2255x2255) satellite images were divided into batches of 512x512 pixels. As a final step, the patches were manually selected discarding segments with the presence of clouds.

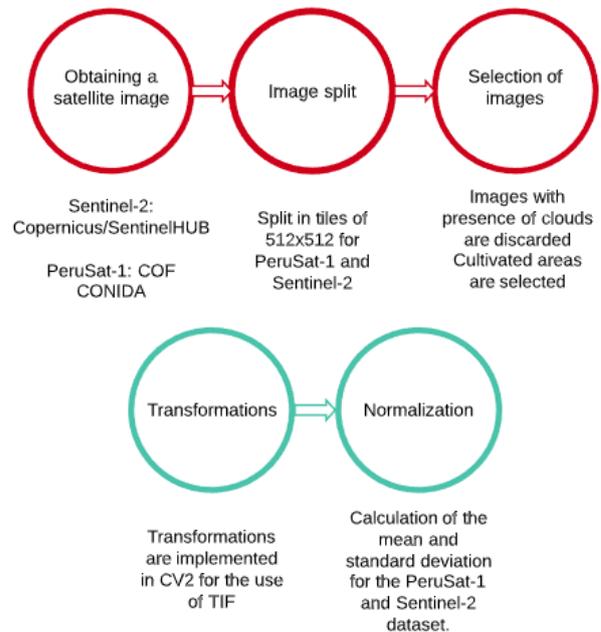


Figure 3. Pre-processing of satellite images of PeruSat-1 and Sentinel-2

Two additional processes were performed for handling the images, the implementation of some transformations such as crop,

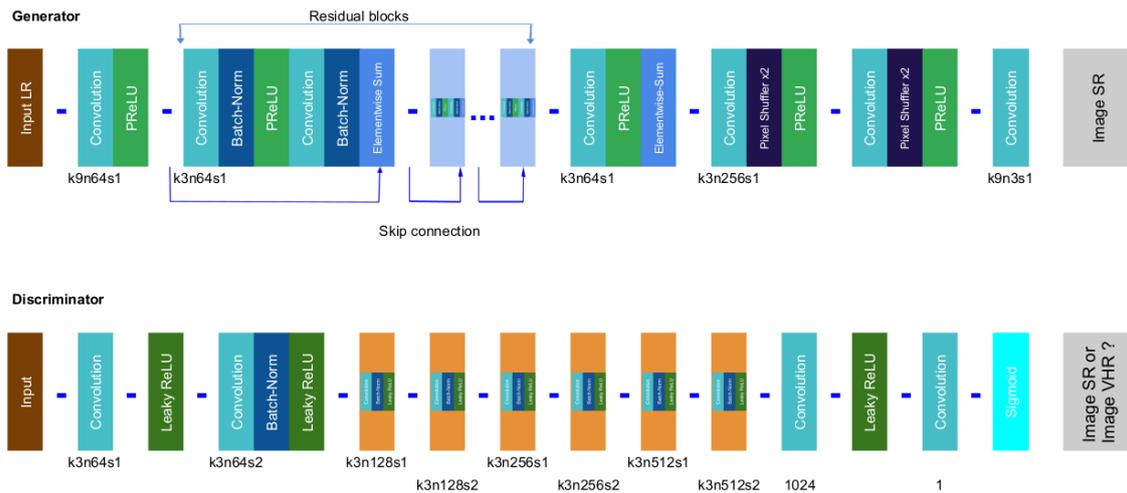


Figure 4. Generator and discriminator of SRPeruSatGAN architecture

size reduction, use of tensors, among others, and the calculation of the mean and standard deviation for normalization of datasets.

Considering that the images from the PeruSat-1 dataset were saved as TIFF files, an additional processing work was performed using OpenCV (Bradski, 2000) for the transformation, interpolation, scaling and visualization processes, thus preserving the 12-bit radiometric resolution provided by PeruSat-1 satellite, instead of the 8-bits resolution as used in previous works (Pineda et al., 2019), where, despite that using a 8-bits representation allowed the decreasing of the training times, it had the major disadvantage of losing information due to the lower radiometric resolution used in this process..

3.2 Dataset

The PeruSat-1 dataset was generated using 41 satellite images of the northern coast of Peru. The images were taken between March and July 2019, and were distributed through the satellite image portal of the Peruvian space agency (PORTAL DE IMÁGENES SATELITALES DE LA AGENCIA ESPACIAL DEL PERÚ, 2019), which allows the access to the ground observation catalog images acquired by PeruSat-1 satellite. Each image obtained includes a panchromatic image of 0.7 meters of resolution and a multispectral image with 2.8 meters of resolution. For the work, the 4 multispectral bands from the images were used.

Each satellite image was divided into smaller tiles of 512x512 pixel size, obtaining a total of 810 tiles from the 41 satellite images, into a TIF format, each including the red, green, blue and NIR bands. The images collected correspond to agricultural crop areas only, following the recommendations described in (Kawulok et al., 2019), and the tiles with presence of clouds and errors were discarded. This dataset was called PeruSat-1.

Similarly, Sentinel-2 satellite images were used to create a new dataset, 6 satellite images were used to obtain 80 tiles, each with a size of 512x512 pixels. The images correspond to the north coast of Peru, we considered bands 2, 3, 4 and 8 from this satellite, since they closely match to the spectral bands used in the PeruSat-1 dataset. As in the previous case from PeruSat-1, the selected areas correspond to agricultural cultivation fields. The data set created was called Sentinel-2.

3.3 SR Architecture

Our SR model is based on the work of (Ledig et al., 2017), and propose to use 5 identical residual blocks in the generator, each block consisting of two convolutional layers with kernels of size 3x3 and features maps of 64, stride of 1, followed by batch normalization and parametric ReLU as activation function(Xu et al., 2015). The improvement of the input image resolution is achieved through two layers of sub pixel convolution.

The task of classifying the images and deciding whether it is a real or a generated image is performed by the discriminator. The discriminator used in this approach consists of 8 blocks, each block is made up of a convolution layer, batch normalization, and leaky ReLU as an activation function, the final layers of the discriminator consist of two convolutional layers and a sigmoid function that allows classification, as shown in Figure 4 where the architecture is presented.

4. EXPERIMENTS AND RESULTS

This section describes the experiments performed and the results that were achieved by our approach, which proposes to apply SR based on a GAN model to increase the spatial resolution of Sentinel-2 images with the reference of images from PeruSat-1 satellite.

For the experiments, a Lambda Dual device was used, which is a GPU workstation equipped with two RTX 2080 Ti, an Intel i9-9820X CPU with up to 10 Cores, 64 GB memory, an HDD of 2 TB. The environment include pre-installed Ubuntu, TensorFlow, PyTorch, Keras, CUDA, cuDNN, python 3.6, raster OpenCV, Pycharm and PyTorch 1.0 packages, so the settings could be immediately booted up and easily configured.

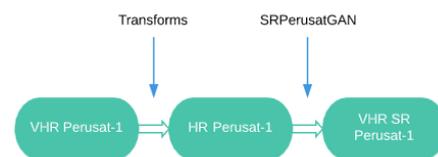


Figure 5. Stages of training PeruSat-1

For training our GAN model we used a resized version of the VHR PeruSat-1 images as input for the generator, emulating the

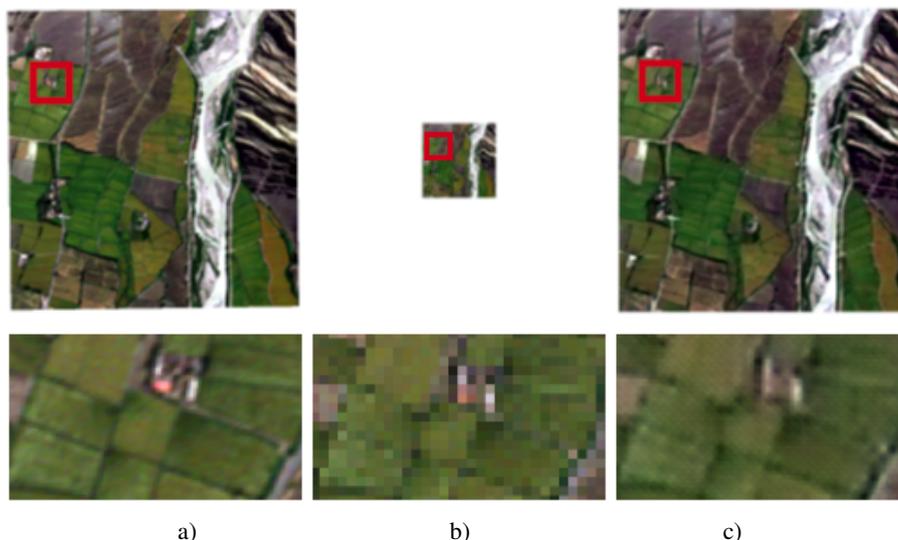


Figure 6. PeruSat-1 images with 12 bit of radiometric resolution, a) VHR Ground truth PeruSat-1, with size of 512x512, b)HR PeruSat-1, with size of 128x218, c) VHR Super-Resolution GAN PeruSat-1, with size of 512x512

resolution of the HR Sentinel-2 satellite images; in this way the generator phase was trained to deliver VHR images based on synthetic HR PeruSat-1 images, afterwards, in the generalization, is possible to increase the resolution of HR Sentinel-2 images to VHR ones, with resolutions close to those of PeruSat-1 satellite.

As previously stated, we used a 12-bit radiometric resolution image from PeruSat-1 satellite, thus maintaining their spectral characteristics in all their 4 spectral bands; and bands 1, 2, and 3 from 800 images of PeruSat-1 dataset were used for training the model. For this purpose, the transformations, crops, and interpolations were performed with OpenCV in Pytorch. Figure 5 show a VHR image used as the input image for the generator, this image was obtained by cropping segments of 128x128 size from each image of the PeruSat-1 dataset, these segments are reduced in a factor of 4 to segments of 32x32 using bicubic interpolation, making the PeruSat-1 images similar to those Sentinel-2 ones. The training of the model used 100 epochs, batch size of 48 and a scaling factor of 4. As output of the generator, emulated VHR PeruSat-1 images were obtained, with a size of 128x128.

For testing the outcomes on the trained network, 10 VHR images from the PeruSat-1 dataset were used. Figure 6 shows one of that images used for training the GAN model, and the outcome achieved by these approach. The original VHR batch of 512x512 pixels was reduced in a factor of 4 by an interpolation process, then, this new reduced image (HR) served as input for the trained GAN model, and the output returned a VHR image of 512x512 pixel size, emulating the original resolution of PeruSat-1 satellite.

We used two metrics to assess our approach, the first, the Peak Signal to Noise Ratio (PSNR) which is a metric used for the measurement of the accuracy on the spatial resolution; and the second, the Structural Similarity (SSIM), used for the measurement of the accuracy on the spectral resolution. According to (Horé, Ziou, 2010) SSIM measures the similarity between two images, and is considered to be correlated with the quality perception of the human visual system (HVS); regarding PSNR, higher values of this metric provides a higher image

Table 1. Metrics obtained with the test images of the PeruSat-1 dataset

Image	PSNR	SSIM
1	31.80127497	0.9582959373
2	32.1626894	0.9585022863
3	31.52290907	0.9616699915
4	30.86250225	0.9342486552
5	30.63083192	0.926805629
6	33.80766751	0.9775816312
7	28.92647074	0.9649785011
8	27.00884483	0.9658817027
9	31.18153502	0.9610201757
10	29.55104008	0.9273331245
Mean	30.74557658	0.9536317634
St.dev.	1.886191927	0.0176510896

quality, whilst smaller values implies higher numerical differences between the images being compared.

Table 1 presents the outcomes achieved by our SR approach, when evaluating its performance on the previously mentioned VHR images from the PeruSat-1 dataset. It shows the outcomes achieved on the PSNR and SSIM metrics when comparing each of the 10 original VHR PeruSat-1 images against its VHR emulated image version generated with our SR approach. As it can be seen in the table, all the values for both the PSNR and SSIM metrics, are stable and define a mean and standard deviation of 30.7456 and 1.8861 for the PSNR; and 0.9536 and 0.0177 for the SSIM, respectively. It is worth to notice that among the images on the dataset, the one that achieved the best results correspond to the sixth, which was the one presented in Figure 6 as well.

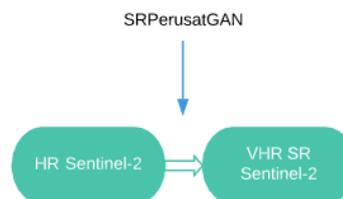


Figure 7. Stages of SR Sentinel-2 using the SRPerusatGAN

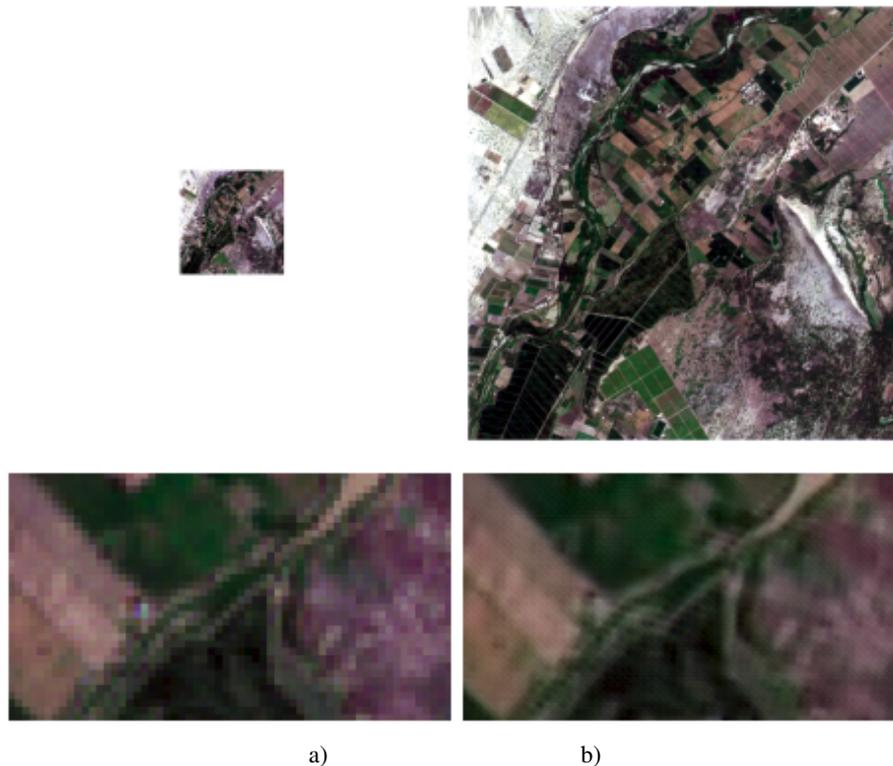


Figure 8. a) Sentinel-2 image with a size of 512x512, b) Sentinel-2 image with super-Resolution GAN, with a size of 2048x2048

From a qualitative point of view, once our SRPeruSatGAN model was trained, we applied it to improve the resolution of one of the images from the Sentinel-2 dataset, and the process we follow is that described in Figure 7, where instead of using a resized version of a VHR image from PeruSat-1 dataset, this time a HR Sentinel-2 image (512x512) was used directly as the input for our model, obtaining a VHR (2048x2048) emulated version of that HR image, as shown in Figure 8; and whereby visually assessing it can be seen that applying our SR approach actually improves the spatial resolution of the Sentinel-2 HR image, and also preserves their spectral characteristics, which becomes a major issue for upcoming processes as those of segmentation or classification tasks.

5. CONCLUSIONS

This work presents the application of Super-Resolution technique based on a generative adversarial network approach to improve the spatial resolution of Sentinel-2 (10m x pixel) satellite images in its blue, green, and red bands. To achieve our objective, we developed a GAN model based on the architecture proposed by (Ledig et al., 2017), which was trained with images from PeruSat-1 (2.8m x pixel) satellite.

To evaluate the effectiveness of our model, we performed a downsampling from a set of ten PeruSat-1 images, to obtain resolutions similar to those of Sentinel-2 satellite, and then we use that images as input to the new model. Our results based on the PSNR and SSIM metrics, presented in Table 1, show us the confidence of the model in order to apply it in the task of improving Sentinel-2 satellite images.

The use of 12-bits radiometric resolution from the PeruSat-1 images for training the proposed network allows an actual improvement in the spatial resolution when applied over

Sentinel-2 HR images, approximating their resolution to those of PeruSat-1 images; likewise, our approach allows to maintain the spectral characteristics of the original Sentinel-2 images.

To acquire VHR satellite images is very difficult, thanks to the Peruvian Spatial Agency (CONIDA) who are in charge of the Peruvian Satellite (PeruSat-1), we got access to work with 41 VHR Satellite images. From these images, we generated 512x512 small tiles in order to create our dataset. The size of 512 x 512 pixels tiles, was defined after different previous experiments, finding that image tiles of 1024 x 1024 needs more GPU memory, thus increasing the training times; and in the other hand, small image tiles of 128 x 128 pixel, lead to some losses of the spatial characteristics, which are important to preserve for further applications, like image segmentation.

The time to train our model with the PeruSat-1 dataset was approximately 5 hours using one Nvidia RTX 2080 card. After that, the process of applying SR to a sentinel-2 image with the trained PeruSatGAN model takes 1 or 2 seconds with the same graphics card. In order to enhance the training time we can use more graphics cards in parallel.

Given the auspicious results obtained in this work, we are very encouraged to develop some other implementations focusing on the improvement of the spatial resolution toward the enhancement of each individual spectral band from Sentinel-2 satellite images, thus proposing a new alternative to classic pansharpener processes, which will also increase the spectral information of the VHR images from PeruSat-1 satellite by integrating new information provided by those new synthetic layers, which will extend the actual capabilities of PeruSat-1 satellite.

ACKNOWLEDGEMENTS

The authors would like to thank the support of Artificial Intelligence Laboratory at Pontificia Universidad Católica del Perú, CONIDA (National Aerospace Research and Development Commission), PRONABEC (National Scholarship and Educational Credit Program), and FONDECYT (National Fund for Scientific, Technological Development and Technological Innovation) under the financing agreement No. 131-2018-FONDECYT-SENCCO.

REFERENCES

- Bartolom, S. A. N., Huarochir, D. E., Kp, M., Huarochir, D. E., 2017. Perusat-1 Y Clasificación De La Carretera Central. 000(511).
- Bradski, G., 2000. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.
- Comunicaciones, M. d. T. y., 2017. PERUSAT-1 Y CLASIFICACIÓN DE LA CARRETERA CENTRAL. 000(511).
- Dong, C., Loy, C. C., He, K., Tang, X., 2016. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2), 295–307.
- ESA, 2015. SENTINEL-2 User Handbook. 1–64. https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative Adversarial Networks. 1-9. <http://arxiv.org/abs/1406.2661>.
- Horé, A., Ziou, D., 2010. Image quality metrics: PSNR vs. SSIM. *Proceedings - International Conference on Pattern Recognition*, 2366–2369.
- Huang, T., Yang, J., 2010. Image super-resolution: Historical overview and future challenges. *Super-resolution imaging*, 19—52.
- Irani, M., Peleg, S., 1990. Super Resolution From Image Sequences. *Ieee*, 115–120.
- Johnson, J., Alahi, A., Fei-Fei, L., 2016. Perceptual losses for real-time style transfer and super-resolution. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9906 LNCS, 694–711.
- Kawulok, M., Piechaczek, S., Hrynczenko, K., Benecki, P., Kostrzewa, D., Nalepa, J., 2019. On training deep networks for satellite image super-resolution. 2–5. <http://arxiv.org/abs/1906.06697>.
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., Shi, W., 2017. Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 105–114.
- PerúSAT-1, satélite de Observación de la Tierra, 2018.
- Pineda, F., Ayma, V., Aduviri, R., Beltran, C., 2019. Super Resolution approach using Generative Adversarial Network models for improving Satellite Image Resolution. *Proceedings - 6th International Conference on Information Management and Big Data*.
- PORTAL DE IMÁGENES SATELITALES DE LA AGENCIA ESPACIAL DEL PERÚ, 2019.
- Singh, S., 2018. Enhanced Deep Image Super-Resolution. *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 1207–1211.
- Xu, B., Wang, N., Chen, T., Li, M., 2015. Empirical Evaluation of Rectified Activations in Convolutional Network. <http://arxiv.org/abs/1505.00853>.
- Zhao, L., Sun, Q., Zhang, Z., 2017. Single Image Super-Resolution Based on Deep Learning Features and Dictionary Model. *IEEE Access*, 5(Cccv), 17126–17135.