

## Enhancing SRTM digital elevation models with deep-learning-based super-resolution image generation

Leonardo Assumpção Moreira<sup>1</sup> - ORCID: 0000-0002-4172-0165

Livia Moreira Poelking<sup>2</sup> - ORCID: 0000-0002-1294-7678

Hideo Araki<sup>1</sup> - ORCID: 0000-0001-7116-6102

<sup>1</sup>Universidade Federal do Paraná, Programa de Pós-graduação em Ciências Geodésicas, Curitiba - PR, Brasil.

E-mails: leonardoamoreira@gmail.com, haraki@ufpr.br

<sup>2</sup>None, Engenharia Química, Cambridge - Cambridgeshire, United Kingdom.

E-mail: livia.poelking@gmail.com

Received in 16<sup>th</sup> September 2021.

Accepted in 04<sup>th</sup> September 2022.

### Abstract:

Digital elevation models are responsible for providing altimetric information on a surface to be mapped. While global models of low and medium spatial resolution are available open source by several space agencies, the high-resolution ones, which are utilized in scales 1:25,000 and larger, are scarce and expensive. Here we address this limitation by the utilization of deep learning algorithms coupled with SISR techniques in digital elevation models to obtain better spatial quality versions from lower resolution inputs. The development of a GAN-based methodology enables the improvement of the initial spatial resolution of low-resolution images. A dataset with different pairs of digital elevation models was created with the objective of allowing the study to be carried out, promoting the emergence of new research groups in the area as well as enabling the comparison between the results obtained. It has been found that by increasing the number of iterations the performance of the generated model was improved and the quality of the generated image increased. Furthermore, the visual analysis of the generated image against the high- and low-resolution ones showed a great similarity between the first two.

**Keywords:** Digital Elevation Model; Generative Adversarial Network; Image Super Resolution; Machine Learning; Deep Learning; Neural Networks; Cartographic Production.

---

**How to cite this article:** MOREIRA LA, POELKING LM, ARAKI H. *Enhancing SRTM digital elevation models with deep-learning-based super-resolution image generation. Bulletin of Geodetic Sciences.* 28(4): e2022021, 2022.



This content is licensed under a Creative Commons Attribution 4.0 International License.

# 1. Introduction

Brazil, despite being the fifth-largest country in territorial extension on the planet, faces great challenges to carry out and keep all its basic systematic mapping up to date. And as the years go by, the demand for access to geospatial data to carry out different activities, whether professional or recreational, is increasing. The emergence and development of new technologies and methodologies make it possible for professionals in the fields of geosciences to meet this increasing demand (LUNARDI et al., 2012).

Digital elevation models (DEM) are important in cartographic production because they are responsible for providing altimetric terrain information such as dimensioned points and contour lines. Presently there is an availability shortage of DEM for the realization of cartographic and related products for scales greater than 1:50,000, including cadastral scales. Despite being extremely useful for municipalities and companies in carrying out their inherent activities, these are the scales that present the biggest cartographic emptiness in Brazil.

The means to solve this problem at present are photogrammetric flight contracting, slant-view satellite image acquisition (such as WorldView and Pleiades), and radar image acquisition (such as TanDEM-X, RADARSAT and Cosmo-SkyMed). All of them have better spatial resolution than Shuttle Radar Topography Mission (SRTM) but are high-investment engineering products.

According to Valeriano & Rosseti (2012), the SRTM was obtained in the year of 2000 through interferometry by the SIR-C/X-SAR sensor generating a DEM with 30 meters of spatial resolution. Nevertheless, this resolution was not available for South America. Only in 2013 did NASA make such data available. Other products of this nature have been made available over the years and have brought with them a significant improvement in spatial resolutions. Regarding the studies to increase the spatial resolution of DEM, the most known interpolation methods are Triangular Irregular Network (TIN), Inverse Distance Weighting (IDW), Topo to Raster, Kriging, Natural Neighbor, Bilinear, Bicubic and others.

Currently deep learning has succeeded in a wide area of applications, emerging as the field of machine learning with the largest and fastest spread among uses in traditional or even more modern areas (ALOM et al., 2019). One area which has greatly benefited from these advances is the Single Image Super-Resolution (SISR). SISR is an important area of research in the field of computer vision and has found many practical applications in problems in other areas. SISR aims to generate a high-resolution (HR) image departing from a single low-resolution (LR) image (CHEN et al., 2021). Modernly, the state-of-the-art methods in super-resolution of images use several concepts of artificial intelligence (AI) and deep learning to approach the problem and have yielded good results.

Cheon et al. (2018) used a generative adversarial network (GAN) to increase the resolution of digital images by seeking to develop a model with a balance between perception and distortion. They based their work on the Residual Network with Enhanced Upscaling Module for Super-resolution by Kim and Lee (2018) and obtained a good performance for both perception and distortion in the training of their proposed model, therefore being effective in applications of super-perceptive resolution.

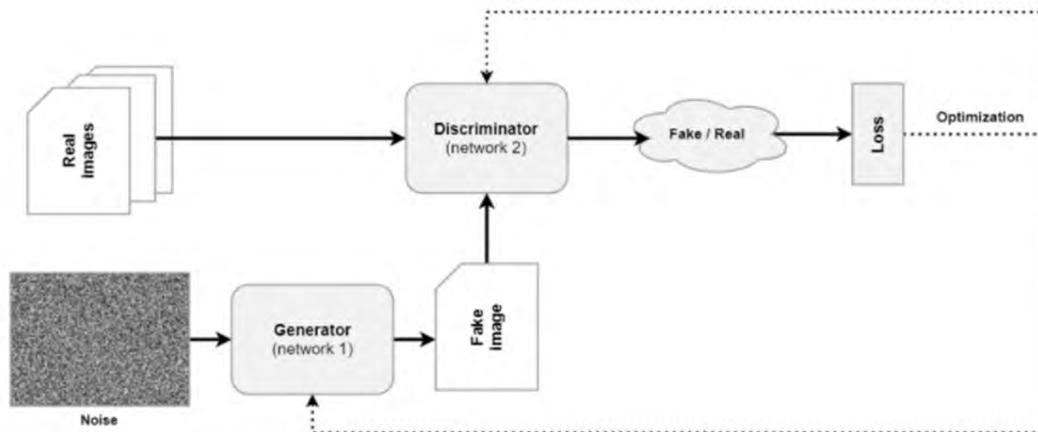
In this work we aim to develop an alternative for the generation of a digital elevation model with better spatial resolution using single image super-resolution technique through Generative Adversarial Network. The model proposed here to produce a digital elevation model with super-resolution has been used to obtain super-resolution color photographs (GOODFELLOW et al., 2014).

In brief, this research aims to contribute to the academic environment through new approaches to the generation of information from orbital altimetry data, a subject closely related not only to geodesy, but also to several other related areas such as cartography, remote sensing, photogrammetry, among others. Among the benefits to the area of geodetic sciences that the proposed methodology may bring, are the reduction of costs in the production of high spatial resolution DEM as well as the generation of high-resolution DEM for regions where only low spatial resolution models are available.

## 2. GAN – Theoretical Review and Model Proposal

A Generative Adversarial Network (GAN) is a class of machine learning systems invented by Ian Goodfellow in 2014 (GOODFELLOW et al., 2014) in which two neural networks compete against each other in a game (in the game theory sense, in the form of a zero-sum game). This technique generates new data with the same statistics as the training set (DHAKAL, 2012). For example, a GAN trained on color (RGB) photographs can generate new photographs with many realistic features that look at least superficially authentic to human observers. Although originally proposed as a form of generative model for unsupervised learning, GAN has been proven useful for semi-supervised learning, fully supervised learning, and reinforcement learning.

Following up the original GAN, the Super-Resolution Generative Adversarial Network (SRGAN) algorithm has been developed. The SRGAN combines deep neural networks with a GAN to learn how to generate upscaled images (Figure 1). During training, a high-resolution image is first downsampled into a lower resolution image and input into a generator. The generator then tries to upsample that image into a super-resolution (SR) image. The discriminator is used to compare the generated super-resolution image to the original high-resolution image. The GAN loss from the discriminator is then back propagated into both the discriminator and generator. The discriminator is mainly composed of convolution, batch normalization and parameterized ReLU (PReLU) layers (LEDIG et al., 2017).



Source: <https://www.machinecurve.com/index.php/2021/03/23/generative-adversarial-networks-a-gentle-introduction/> (2021).

**Figure 1:** Schematic of a generative adversarial network.

Although the SRGAN has stimulated new creations and brought new ideas capable of generating realistic textures during the super-resolution of a single image, undesirable artifacts were often observed. To further improve the visual quality, three main components were added to the SRGAN, namely (i) network architecture, (ii) adversarial loss, and (iii) perceptual loss, to derive an Enhanced SRGAN (ESRGAN). In addition, two modifications were made to the generator structure: the removal of all batch normalization (BN) layers and the replacement of the original basic block with the proposed Residual-in-Residual Dense Block (RRDB), which combines multilevel residual network and dense connections (WANG et al., 2018). The idea of relativistic GAN was also introduced to let the discriminator predict the relative realness instead of the absolute value. Unlike the standard discriminator in SRGAN, which estimates the probability that an input image is real and natural, the relativistic discriminator attempts to predict the probability that a real image is relatively more realistic than a fake one (WANG et al., 2018).

Furthermore, ESRGAN presents a more effective loss of perception ( $L_{\text{percep}}$ ) by restricting resources before activation rather than after activation as practiced in SRGAN. Based on the idea of being closer to the perceptual similarity, the perceptual loss extended to the SRGAN has been proposed. The perceptual loss has been previously defined in the activation layers of a pre-trained deep network, where the distance between two activated features is minimized. Contrary to convention, resources were used before the activation layers, which overcome two disadvantages of the original design (WANG et al., 2018). Perceptual loss was also improved using the features before activation which provided stronger supervision for gloss consistency and texture recovery. Benefiting from these improvements, the ESRGAN achieved consistently better visual quality with more realistic and natural textures than SRGAN (WANG et al., 2018).

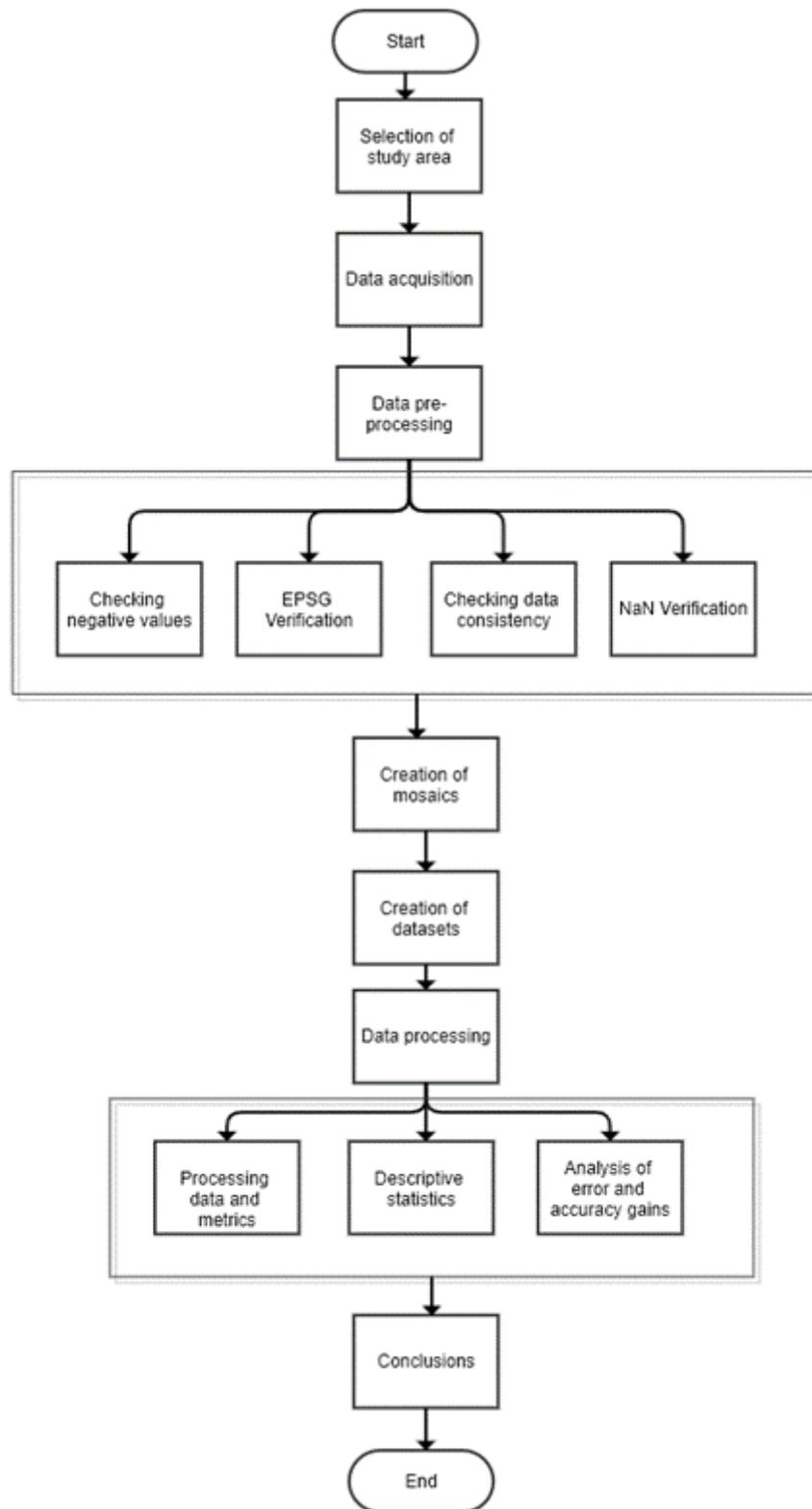
Single Image Super-Resolution has attracted increasing attention in the AI research community. Since the pioneering work of single image Super-Resolution deep Convolutional Neural Network (SRCNN), deep Convolutional Neural Network (CNN) approaches have brought a prosperous development. Various network architecture projects and training strategies have continuously improved super-resolution performance, especially the peak signal-to-noise ratio (PSNR) value. However, these PSNR-driven approaches tend to produce over-smoothed results without sufficient high-frequency detail, as the PSNR metric fundamentally disagrees with the subjective assessment of human observers (WANG et al., 2018).

### 3. Methodology

This section describes the methodology used in this work to achieve the results presented below. The work proposed here is based in the ESRGAN with modifications described in subsection 3.4 to adapt the algorithm to DEMs. The proposed algorithm is referred to as MDTESRGAN. The programming language used in the developments carried out in this work was the Python 3.7 interpreter.

#### 3.1 Synthesis of the Pipeline

In order to synthesize the adopted procedures, a flowchart with the used pipeline has been produced (Figure 2). In the pipeline, the generation of the digital elevation model with the MDTESRGAN algorithm is called “data processing”. The sub-sections 3.2 to 3.5 describe the elements presented in this flowchart.



Source: The Author (2021).

**Figure 2:** Synthesis flowchart of the pipeline followed in this work.

## 3.2 Study Area

The selection of the study area for this work follows two-fold criteria: it should not only allow the use of several digital elevation models, but also provide a qualitative evaluation of different types of terrain models. In this context, the study area selected for the present work is the municipality of Monte Castelo in the State of Santa Catarina due to the variations of its relief features. The State of Santa Catarina covers an area of 95.346 km<sup>2</sup> in the south of Brazil and has recently had its territory mapped at a scale of 1:50,000, producing digital terrain models of 1-meter spatial resolution, available to download on the internet, which will be used later in the MDTERGAN algorithm evaluation.

## 3.3 Data

To carry out this research, the following data has been selected: DEM SRTM of 90 meters and DEM SRTM of 30 meters (<https://earthdata.nasa.gov>). All these selected digital models are available for download and use on the internet and have global coverage (Table 1).

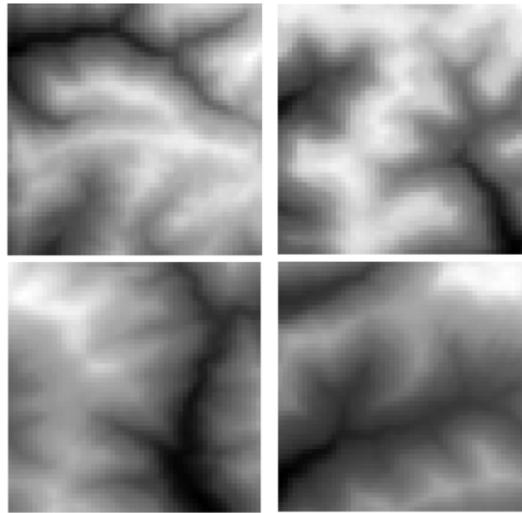
**Table 1:** Table of product specifications of SRTM.

Products specifications	Characteristics
Projection	Geographic
Horizontal Datum	WGS84
Vertical Datum	EGM96 (Earth Gravitational Model 1996)
Vertical Units	Meters
Spatial Resolution	1 arc-second for global coverage (~30 meters)
	3 arc-seconds for global coverage (~90 meters)
Raster Size	1-degree tiles

Source: <https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-non> (2021).

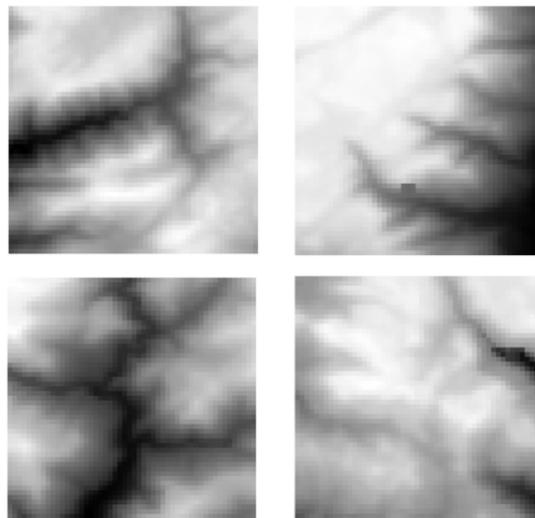
### 3.3.1 Dataset

For the accomplishment of the present work, the following dataset has been created: DEM SRTM of 90 meters (as LR images) and DEM SRTM of 30 meters (as HR images or ground truth), with 10 pairs of training images and 4 pairs of validation images (Figure 3 and Figure 4). The criterion used to partition the 14 image pairs was the proportion of 70% of the samples for training the algorithm and 30% of the samples for its validation. The images used were cropped in the following dimensions: 156 pixels by 156 pixels.



Source: The Author (2021).

**Figure 3:** 4 low-resolution samples generated for the dataset used to perform the MDTERSGAN training.

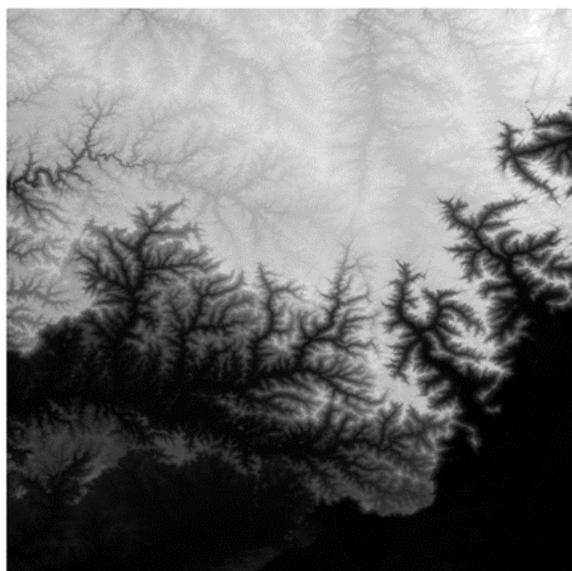


Source: The Author (2021).

**Figure 4:** 4 low-resolution samples generated for the dataset used to perform the MDTERSGAN validation.

The elevation data used is SRTM 1 Arc-Second Global, which provides worldwide coverage of void-filled data at a resolution of 1 arc second (30 meters) and grants free distribution of this high-resolution global dataset. It was used SRTM Version 3.0.

The high-resolution dataset was built from samples of the S30W051.hgt file (Figure 5), with 30-meter resolution pixels. As for the coordinate systems, the image has the EPSG 4326 and the WGS84 geographic coordinate system. It has dimensions of 3601 pixels wide by 3601 pixels high. The data type is Int16. It has only 1 data band. The metadata type of the GDAL driver is SRTMHGT File Format.



Source: The Author (2021).

**Figure 5:** The DEM SRTM file S30W051.hgt image with 30 meters resolution viewed in QGIS.

The low-resolution dataset was built from samples of the SH-22-X-C.tif file (Brazil mapping grid) with pixels of 90 meters of resolution. As for coordinate systems, the image has the EPSG 4326 and the WGS84 geographic coordinate system. It has dimensions of 1800 pixels wide by 1200 pixels high. The data type is UInt16. It has only 1 data band. The GDAL driver metadata type is GeoTIFF.

In relation to the dataset built with samples from DEM SRTM 90 meters and DEM SRTM 30 meters, considering the region with the respective images selected for their construction, it has been found that using the resampled LR image to calculate the average difference between the LR and HR images, a discrepancy value of 0.0801 meters was obtained and the difference of standard deviation presented is 0.2311 meters.

### 3.3.2 Pre-processing of the dataset data

The images that make up the datasets are preprocessed to verify:

- a) the coordinate systems. Due to the existence of different coordinate systems in the selected image pairs, the generated samples may have different portions of terrain surfaces, making it impossible to obtain the super-resolution of the desired low-resolution image.
- b) the existence of negative pixel values and the existence of not-a-number (NaN) instances. Those may cause inconsistent results in the runs and must be resolved in the most appropriate way before the construction of the datasets.

Also, a statistical analysis of the high- and low-resolution images from which the dataset samples will be extracted is recommended because it allows the verification of inconsistencies between them. Here the statistical analysis of the images used to collect the dataset samples indicates that the arithmetic mean of the high-resolution image is 489.1774 meters, while the arithmetic mean of the low-resolution image is 490.1813 meters. As for the standard deviation, the high-resolution image presented a value of 384.1116 meters, and the low-resolution image presented a value of 383.8805 meters. This result is expected, since the two analyzed DEMs correspond to the same terrain relief.

### 3.4 Processing with the MDTESRGAN Algorithm

The pairs of images generated for the dataset will be used to perform the processing in the MDTESRGAN algorithm. The MDTESRGAN algorithm has been developed adapting the ESRGAN algorithm by changing the inputs and outputs. ESRGAN uses files with extension .PNG colored images with 3 bands (RGB) as input and output. To run files with altimetric information, it was necessary to change the input and output to .TIF images with only 1 band. Another necessary modification was the image scaling (spatial and radiometric resolution), so that the MDTESRGAN algorithm could receive different combinations of digital models.

The ESRGAN algorithm offers the possibility to upscale the image by 2x, 4x or 8x. In order to adapt the ESRGAN into MDTESRGAN, the scaling factor had to be defined. Thus, a scaling factor of 4x between low- and high-resolution images has been set for the MDTESRGAN.

### 3.5 Evaluation Metrics

The measures used to perform quality control of the results from the experiments here reported are PSNR, Structural Similarity Index Measure (SSIM), Mean Squared Error (MSE), Naturalness Image Quality Evaluator (NIQE), and Root Mean Squared Error (RMSE).

The PSNR is a measure defined as the ratio between the maximum energy of a signal and the noise that affects its faithful representation (SCIKIT-IMAGE, 2021). A higher PSNR means smaller noise and because of that generally indicates that the reconstruction is of higher quality.

The SSIM is a metric used to measure the similarity between two images. The resulting SSIM index is a decimal value between -1 and 1, with a value of 1 occurring only when the two data sets are identical and therefore indicating perfect structural similarity (IMATEST, 2021; SCIKIT-IMAGE, 2021).

The NIQE is an image quality score. It compares a given image to a standard model calculated from images of natural scenes, where a lower score obtained indicates better perceptual quality (GITHUB, 2021).

The MSE is a measure of the quality of a given estimator. The MSE values of two statistical models can be used to measure how well they explain a given set of observations. The value of MSE is always positive or greater than zero. A value close to zero represents a better quality of the regression model. In the absence of noise, the MSE is zero. The RMSE is the square root of the MSE and allows the results of the metric to be analyzed in the same dimension as the data (SCIKIT-IMAGE, 2021).

### 3.6 Analysis of variation of, and gains in discrepancy and accuracy

While experimental errors represent the difference between the measured value and the true value of the physical quantity, i.e., a deviation between an observed value and its actual value. Discrepancy, on the other hand, can be defined as disagreement or inequality, when a given measure is compared with another. It is the difference between a measured value and a reference value. Therefore, in this work we use the discrepancy to quantify the success of the methodology and accuracy as the degree of consistency of the measured quantity with its mean. Hence the research employed the analysis of the discrepancy and accuracy variation of the training images of the runs performed. The analyses included (i) the discrepancy variation of the image set, (ii) the calculated arithmetic mean of the discrepancies, and (iii) the variation of image accuracy as a function of the calculated arithmetic mean

of the accuracies. This procedure aims to verify the behavior of the pixels of the generated images compared to the high-resolution image.

The gains in discrepancy and accuracy are calculated by subtracting the discrepancy and accuracy of the high-resolution image by the value of the image produced by the MDTESRGAN algorithm. This procedure intends to evaluate the effectiveness of the selected methods and criteria. When positive, these values imply that the methodologies applied to the data caused an improvement in the results.

### 3.7 Regression analysis of the pixels and discrepancies

A regression analysis has been performed aiming to estimate the number of iterations necessary to reduce the discrepancies between the generated and low-resolution images and reach a correspondence between the pixels of the generated image compared to the high-resolution one, i.e., the generated and high-resolution images would have 100% corresponding pixels.

### 3.8 Hardware resources

To perform the MDTESRGAN processing, the following hardware equipment has been used (Table 2):

**Table 2:** Technical specifications of the processing hardware

Components	Specifications
Brand	DELL
Physical Memory (RAM) Installed	16 GB
Graphics card	Geforce RTX 2060
Operational system	Microsoft Windows 10 Home Single Language
Model	G5 5590
System type	PC based in X64

Source: Dell (2021).

## 4. Analysis of Results

### 4.1 The processing runs 1 to 3

In this work, three processing runs were performed with a variation in the number of iterations among them. The first run was performed with 10,000 iterations and is referred to as run 1. The second run has 50,000 iterations and is referred to as run 2. And the third run has 100,000 iterations and is referred to as run 3. Such conformation of runs has been selected so that it is possible to observe the behavior of the algorithm in relation to the parameter iteration number, as well as to understand the model in the variation of the results of each run.

The 10 training sets are used to define the model parameters as well as to perform in-sample analysis of the models generated in each run. The 4 validation sets are used to perform the Peak Signal to Noise Ratio (PSNR) metric calculations of the runs and to perform the losses calculations. The method used to partition the dataset in this article was the Holdout Method, considering  $p=70\%$ .

For the analysis of the processing metrics, it has been found that the implementation of five different metrics for evaluating the generated images added greater analysis power to the research. It is noteworthy that among the metrics used, the first two (PSNR and SSIM) are directly proportional values, while the last three (NIQE, MSE and RMSE) are inversely proportional values.

The following table presents the metric values of 10 images generated as a result of the training runs 1 to 3. For each processing run, the metric values are presented in the form of the mean value with the respective variance in parentheses (Table 3).

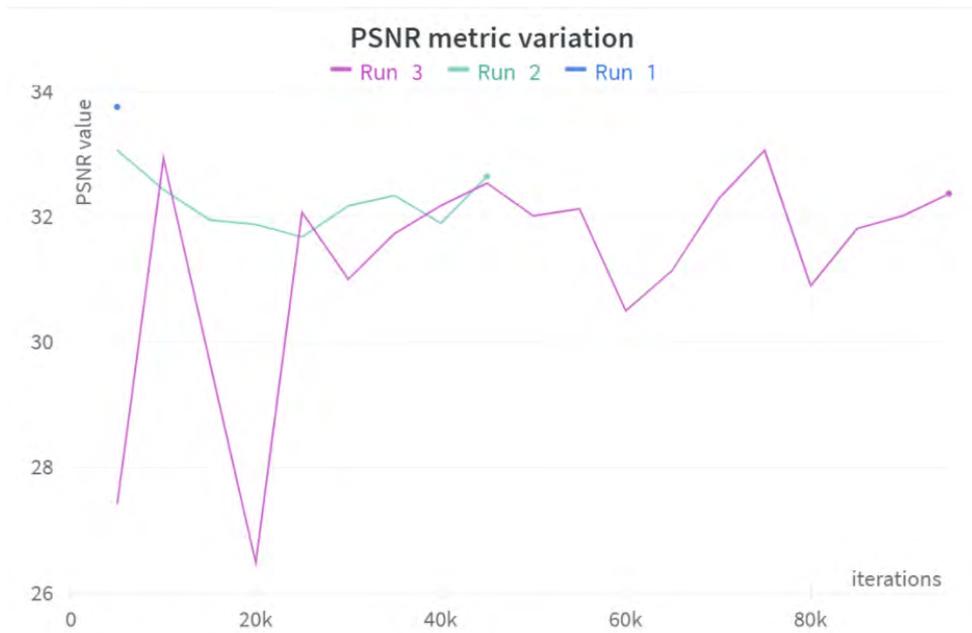
**Table 3:** Metric values of 10 images generated after the training runs 1 to 3.

Run	Metrics				
	PSNR	SSIM	NIQE	MSE	RMSE
1	18.7161 (+7)	0.7308 (+0.06)	13.7076 (+3)	30.2063 (+7)	5.4845 (+0.6)
2	22.7660 (+3)	0.8538 (+0.09)	12.9371 (+3)	18.96552 (+7.4)	4.3300 (+0.8)
3	34.3536 (+1.2)	0.9802 (+0.09)	12.8189 (+3.1)	4.8965 (+0.7)	2.2115 (+0.2)

Source: The Author (2021).

These results in Table 3 show that the run 3 with 100,000 iterations performed better than the shorter runs in all five metrics.

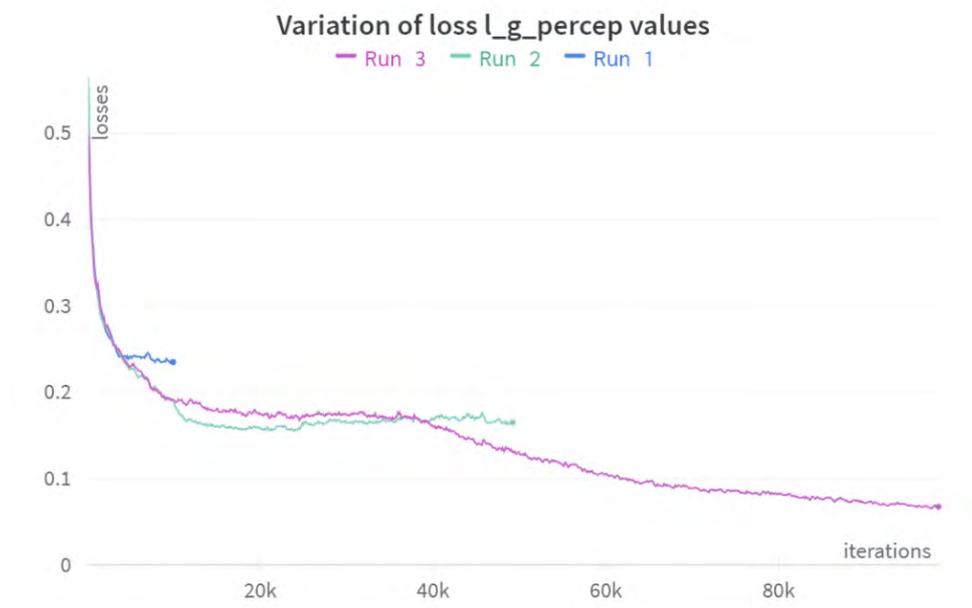
At the end of the runs with the validation dataset, the values of the PSNR metrics were calculated leading to a PSNR of 33.754 for the run 1, a PSNR of 32.646 for the run 2, and a PSNR of 32.374 for the run 3. So that the PSNR decreases with an increasing number of iterations, and the run 1 with 10,000 iterations performed best, with the highest PSNR value. Nevertheless, if one follows the PSNR as the iterations progress (Figure 6), one observes an oscillatory variation of the PSNR values, thus not occurring a continuous growth or decline over time.



Source: The Author (2021).

**Figure 6:** PSNR metric variation during the validation runs.

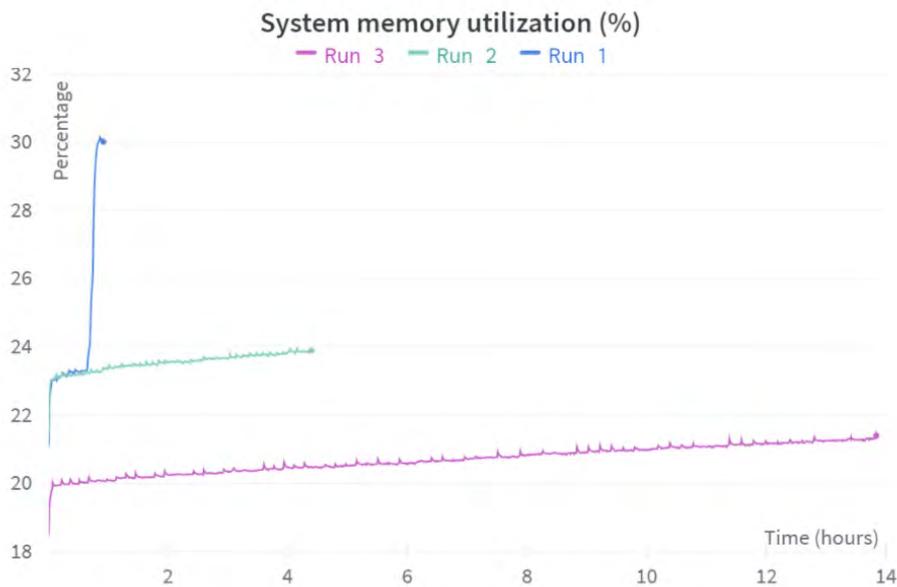
When the loss of perception ( $l_g$ \_percep on loss) is compared, they all start from an initial maximum and from there on show a tendency to approach zero. It is noticed that the run with more iterations comes closer to zero than the ones with fewer steps (Figure 7).



Source: The author (2021).

**Figure 7:** Variation of loss  $l_g$ \_percep values.

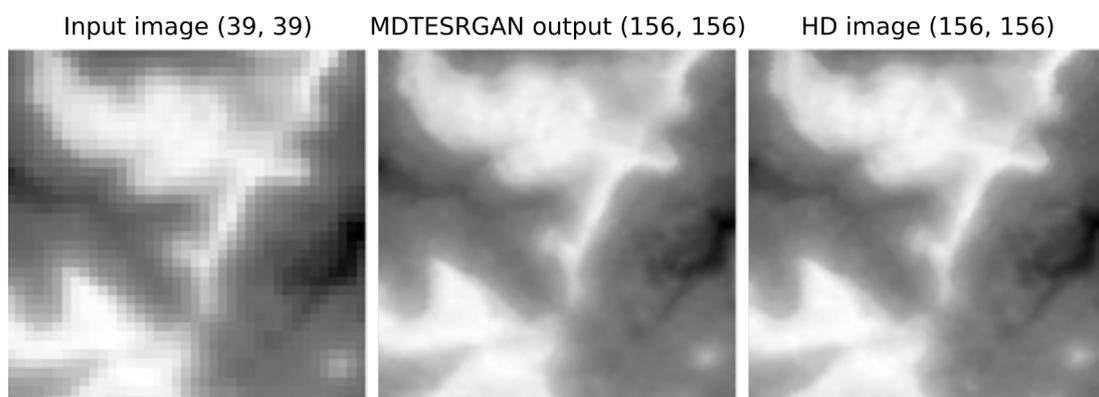
Regarding the system memory usage during the processing runs, it was found that run 1 reached its peak of about 30% of the total memory. The run 2 reached values of about 24%, while the run 3 reached values of about 21% (Figure 8).



Source: The Author (2021).

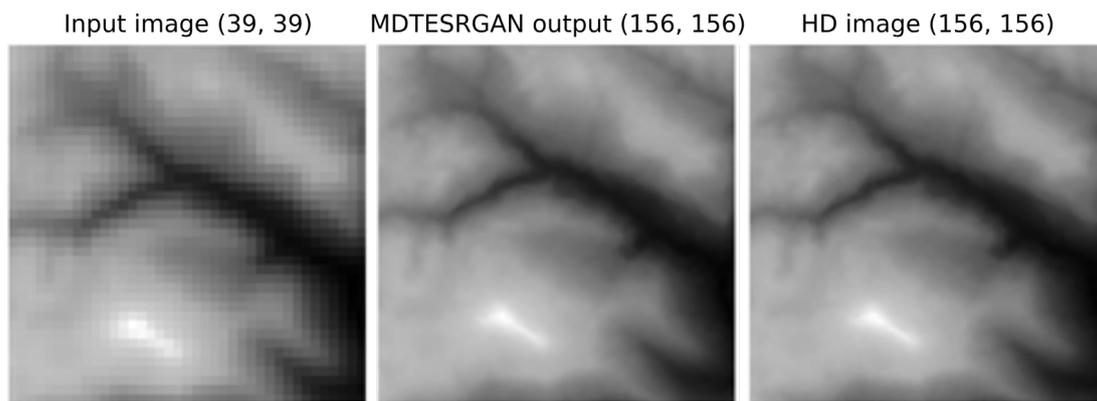
**Figure 8:** System memory utilization during processing in percentage.

Also, the visual analysis will enable the perceptive comparison of the generated image with the original low-resolution image and the high-resolution target. In this analysis, the 3 images were placed side by side for comparison: the low-resolution image on the left, the one generated in the center, and the high-resolution image on the right (Figure 9, Figure 10, and Figure 11). It can be seen that all the generated images analyzed showed great similarity to the respective high-resolution ones.



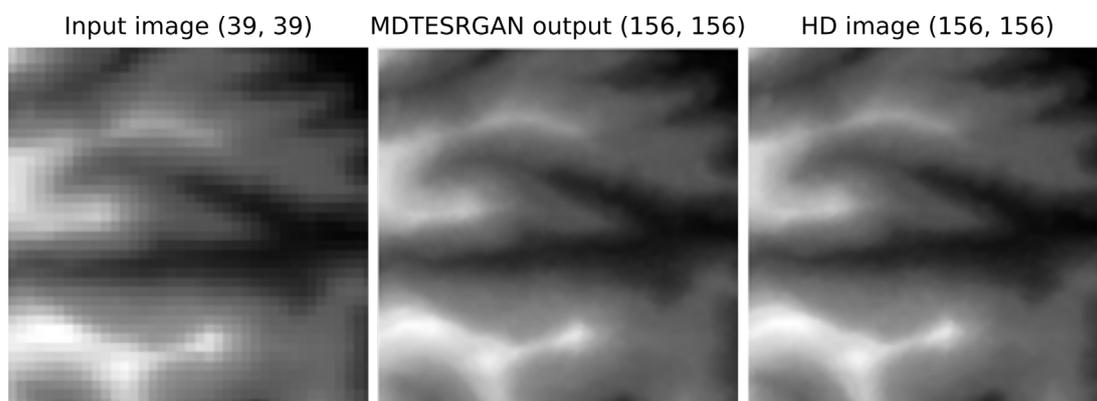
Source: The Author (2021).

**Figure 9:** Example of the visual analysis of the generated image (middle), the high-resolution image (right) and the low-resolution image (left) of image 1.



Source: The Author (2021).

**Figure 10:** Example of the visual analysis of the generated image (middle), the high-resolution image (right) and the low-resolution image (left) of image 2.



Source: The Author (2021).

**Figure 11:** Example of the visual analysis of the generated image (middle), the high-resolution image (right) and the low-resolution image (left) of image 3.

In order to give an idea how typical interpolation methods would score in this problem, we provide results of different metrics (PSNR, SSIM, MSE and RMSE) from various interpolation methods (Table 4). MDTEsRGAN with 100,000 iterations has shown better results than all other typical interpolation methods used in this benchmark.

**Table 4:** Comparison of the metrics of the results of different interpolators.

Interpolators	PSNR	SSIM	MSE	RMSE
low-resolution	20.6918	0.5561	554.4939	23.5477
bilinear	22.4056	0.6824	373.6995	19.3313
bicubic	22.2540	0.6687	386.9754	19.6717
lanczos	22.2501	0.6680	387.3231	19.6805
MDTEsRGAN 10,000	20.5421	0.5710	573.9508	23.9573
MDTEsRGAN 50,000	22.9123	0.6249	332.5414	18.2357
MDTEsRGAN 100,000	36.5822	0.9711	14.2841	3.7794

Source: The Author (2021).

## 4.2 Comparison of the validation images

We use 3 images from among the validation images (within each processing run) to perform the statistical analyses.

As for the analysis of the images from runs 1 to 3 with respect to the total discrepancy gain and total accuracy gain, it has been verified that the performance of the images from the run 3 with 100,000 steps were superior to the performance of the other two runs with fewer iterations (Table 5).

**Table 5:** Table of image analysis regarding the gains in discrepancy and accuracy.

Statistics	Image	10,000 iterations (m)	50,000 iterations (m)	100,000 iterations (m)
Total discrepancy gain	1	4.0490	-4.7587	-0.9670
	2	-1.6603	-7.2561	5.0453
	3	-1.0014	-0.5625	0.9980
Total accuracy gain	1	0.2216	4.7500	18.7790
	2	-0.3821	15.2333	28.4568
	3	0.6021	18.0758	29.1543

Source: The Author (2021).

As for the percentage discrepancy gain and percentage accuracy gain, the performance of images 2 and 3 of the 100,000 iterations had the greatest gain compared to the other two images (Table 6).

**Table 6:** Table of image analysis regarding percentage discrepancy gains and percentage accuracy gains.

Statistics	Image	10,000 iterations (%)	50,000 iterations (%)	100,000 iterations (%)
Percent discrepancy gain	1	4.9320	-579.6436	-117.7937
	2	-29.8086	-130.2742	90.5819
	3	-87.4543	-49.1280	87.1564
Percent accuracy gain	1	0.0094	20.1221	79.5529
	2	-1.1214	44.7087	83.5189
	3	1.7858	53.6084	86.4645

Source: The Author (2021).

Considering the gain-in-discrepancy and gain-in-accuracy analysis of the validation images, image 2 of the run 3 with 100,000 iterations showed a discrepancy gain of 5.0453 meters relative to its corresponding low-resolution image, and the image 3 of the 100,000 iterations showed an accuracy gain of 29.1543 meters, being the best in their respective aspects among the images analyzed. In proportional terms, these results were equivalent to 90.5819% and 86.4645%, respectively, of the reference values. In the same analysis for the test images, the image from run 3 showed a discrepancy gain of -0.3822 meters relative to its corresponding low-resolution image.

In the analysis of the discrepancy and accuracy of the generated images compared to the high-resolution ones, the performance of the images from run 3 with 100,000 iterations were superior to the performance of the other two runs with fewer iterations (Table 7).

**Table 7:** Table of image analysis regarding discrepancy and accuracy.

Statistics	Image	10,000 iterations (m)	50,000 iterations (m)	100,000 iterations (m)
Discrepancy	1	-3.2280	-5.5796	1.7880
	2	-7.2302	-12.8259	0.5246
	3	-2.1464	-1.7076	0.1471
Accuracy	1	23.3841	18.8558	4.8267
	2	34.4545	18.8390	5.6155
	3	33.1161	15.6424	4.5639

Source: The Author (2021).

As for the analysis of the gain of the generated image relative to the low-resolution image in number of pixels, the performance of the images of the run 3 with 100,000 iterations were superior to the performance of the other two runs with fewer iterations (Table 8). The Image 3 in the run 3 scored 2375 coincident pixels, which is a higher score than the other two analyzed images. If a standard deviation of 2 meters were considered for the pixel values, this number would increase to 10608 pixels.

**Table 8:** Gain of the generated image in relation to the LR image in number of pixels.

Statistics	Image	10,000 iterations (pix)	50,000 iterations (pix)	100,000 iterations (pix)
Gain of Gen Img in rel to LR Img	1	-10	28	1815
	2	-18	160	1636
	3	2	532	2375

Source: The Author (2021).

In the analysis of pixels with smaller discrepancies in the images generated in relation to the low-resolution image, it is verified that the performance of the images from run 3 with 100,000 iterations were superior to the performance of the other two runs with fewer iterations (Table 9).

**Table 9:** Analysis of pixels with smaller discrepancies.

Statistics	Image	10,000 iterations (pix)	50,000 iterations (pix)	100,000 iterations (pix)
Pixels with smaller discrepancies	1	1,2662	13,164	21,128
	2	1,2069	15,308	21,931
	3	1,2968	19,077	22,580

Source: The Author (2021).

Considering the reduction of pixel altitude discrepancy values of the generated images compared to the low-resolution images, the image 3 of the run 3 showed the highest number of pixels with discrepancy reduction in the generated image among the evaluated images, with 22580 of minor discrepancies out of a total of 24336 possible. If a standard deviation of 2 meters were considered for the pixel values, this number would increase to 23407 pixels.

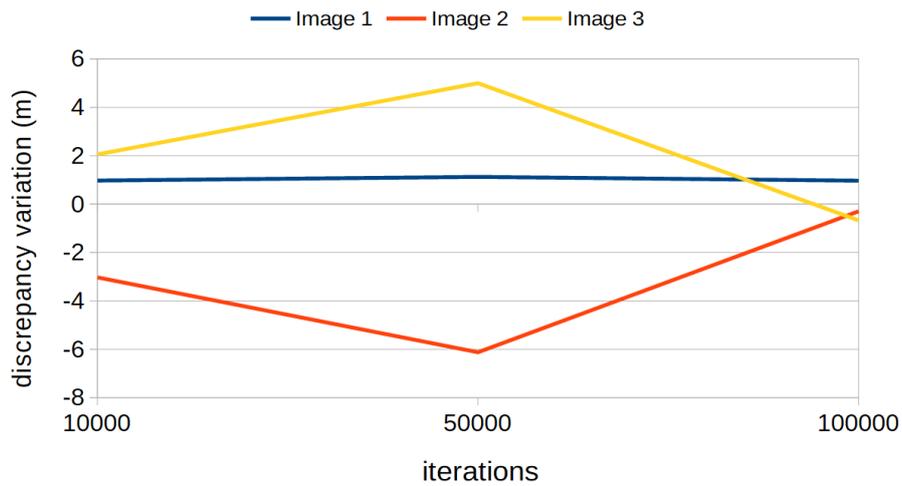
### 4.3 Analysis of discrepancy and accuracy variation

The analysis of the variation in discrepancy and accuracy of the validation images aims to show the constancy of the behavior of the pixels in each of the generated images analyzed in relation to the high-resolution reference images (Table 10, Table 11, Figure 12, and Figure 13).

**Table 10:** Discrepancy of the validation images.

Image	Discrepancy		
	10,000 iterations (m)	50,000 iterations (m)	100,000 iterations (m)
1	-3.2280	-5.5796	1.7880
2	-7.2302	-12.8259	0.5246
3	-2.1464	-1.7076	0.1471
Average	-4.2015	-6.7044	0.8199

Source: The Author (2021).



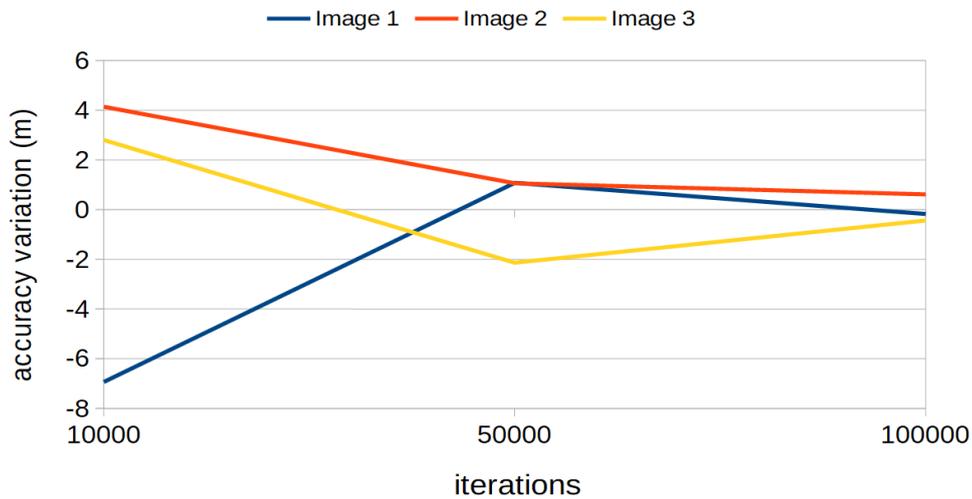
Source: The Author (2021).

**Figure 12:** Discrepancy variation of the validation images.

**Table 11:** Accuracy of the validation images.

Image	Accuracy		
	10,000 iterations (m)	50,000 iterations (m)	100,000 iterations (m)
1	23.3841	18.8558	4.8267
2	34.4545	18.8390	5.6155
3	33.1161	15.6424	4.5639
Average	30.3182	17.7791	5.0020

Source: The Author (2021).



Source: The Author (2021).

**Figure 13:** Accuracy variation of the validation images.

#### 4.4 Regression analysis of the pixels and discrepancies

The linear regression analysis of the pixel-by-pixel coincidence aims to estimate the number of iterations required for the processing of the DEMs with the MDTERGAN algorithm to achieve a pixel-by-pixel equivalence between the model-generated image and the high-resolution target (Table 12).

**Table 12:** Estimation of the number of iterations needed to reach pixel equivalence.

Image	Percentage of equal pixels	Estimate of the number of iterations
1	100	134,866.67
2	100	116,100.00
3	100	110,966.67
Average	100	120,644.45

Source: The Author (2021).

The linear regression analysis of discrepancy reduction, on the other hand, aims to estimate the number of iterations required for DEM processing with the MDTERGAN algorithm to achieve a pixel equivalence with discrepancy reduction between the model-generated image and the low-resolution image (Table 13).

**Table 13:** Estimation of the number of iterations needed to reach the 100% of minor discrepancies.

Image	Percentage of minor discrepancies	Estimate of the number of iterations
1	100	142,500.00
2	100	140,625.00
3	100	121,800.00
Average	100	134,975.00

Source: The Author (2021).

Considering the regression analysis, it was found that the 100% coincident pixel rate for image 3 would be achieved with a run with 110967 iterations. With similar methodology, to achieve 100% of the minor pixel discrepancies for the same image, the estimate would be 121,800 iterations.

## 5. Conclusions

The present work sought to create a dataset for the effective insertion of digital elevation models in super-resolution image studies, to develop a learning machine using Generative Adversarial Network to obtain an increase in spatial resolution of such digital elevation models and to analyze the new models generated by this algorithm.

The work presented an analysis of the increase in spatial resolution obtained in the experiments and verified the potential of using digital elevation models generated by Generative Adversarial Networks in cartographic production for the extraction of altimetric data.

Considering the analysis and statistical comparisons presented in this work, it has been found that increasing the number of iterations is favorable to the performance of the generated model and to the quality of the generated image. This indicates that the proposed methodology is constructive and fulfills the SISR task by generating a high-resolution DEM image departing from a single low-resolution DEM image. Furthermore, a regression analysis indicated that a full pixel equivalence may be reached in a finite number of iterations.

The only indicator which does not corroborate this tendency was the PSNR for the validation dataset, where the run with fewer iterations had a higher PSNR value than the runs with more iterations. Even though, an oscillatory variation for the PSNR over time has been observed, this conflicting result may also be an indication that, despite the PSNR being the scoring metrics in the GAN algorithm, the PSNR alone may not be a sufficient quality indicator, justifying the search for alternatives to the traditional PSNR-driven approaches.

Also, the MDTESRGAN showed a substantial improvement in the quality metrics compared to the traditional interpolation methods in the benchmark runs. Hence the development of super-resolution digital elevation models is timely, given the technological advances in the areas of artificial intelligence, orbital sensors, and computational resources. In this context, future research is planned to further the investigations presented here, including studies with DEMs of different spatial resolutions.

## AUTHOR'S CONTRIBUTION

All authors contribute equally.

## REFERENCES

- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Asari, V.K. (2018). The history began from AlexNet: A comprehensive survey on deep learning approaches. *arXiv preprint*. arXiv:1803.01164 [cs.CV]. <https://doi.org/10.48550/arXiv.1803.01164>
- Chen, H., He, X., Qing, L., Wu, Y., Ren, C., Zhu, C. E. (2021). Real-World Single Image Super-Resolution: A Brief Review. *Information Fusion*. v. 79, p. 124-145. arXiv:2103.02368. 18 p. <https://doi.org/10.48550/arXiv.2103.02368>

- Cheon, M., Kim, J., Choi, J., Lee, J. (2018). Generative Adversarial Network-based Image Super-Resolution using Perceptual Content Losses. *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. <https://doi.org/10.48550/arXiv.1809.04783>
- Dalmolin, Q., Santos, D. R. D. (2014). Ajustamento de observações para geotecnologias. Edition Revised. Curitiba.
- Dhokal, B. (2021). Diving into different GAN architectures. *Towards Data Science*. <https://towardsdatascience.com/diving-into-different-gan-architectures-a96d05c03c5c>
- Gemael, C. (1994). Introdução ao ajustamento de observações: aplicações geodésicas. Ed. of UFPR. Curitiba, PR. 319 p. ISBN 8585132922 (broch.).
- Github. (2021). Izvorski, A. NIQE: Natural Image Quality Evaluator. <https://github.com/aizvorski/video-quality/blob/master/niqe.py>
- Goodfellow, I. J., Pouget-abadie, J., Mirza, M., Xu, B., Warde-farley, Ozairy, D. S., Courville, A., Bengio, Y. (2013). Generative Adversarial Nets. *Veterinary Immunology and Immunopathology*. vol. 155, no. 4. arXiv:1406.2661. <https://doi.org/10.48550/arXiv.1406.2661>
- Imatest. (2021). SSIM: Structural Similarity Index. *Imatest*. <https://www.imatest.com/docs/ssim>
- Kim, J., Lee, J. (2018). Deep Residual Network with Enhanced Upscaling Module for Super-Resolution. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Salt Lake City, UT. pp. 913-9138. doi: 10.1109/CVPRW.2018.00124.
- Ledig, C., Theis, L., Husz, F., Caballero, J., Cunningham, A.; Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., Wenzhe S. (2017). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *Proceedings of the IEEE conference on computer vision and pattern recognition*. arXiv:1609.04802. <https://doi.org/10.48550/arXiv.1609.04802>
- Lunardi, O. A., Penha, A. L. T., Cerqueira, W. (2012). O Exército Brasileiro e os Padrões de Dados Geoespaciais para a INDE. *IV Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação*. Recife – PE.
- Scikit-Image. (2021). Module: metrics. *Scikit-image image processing in python*. <https://scikit-image.org/docs/dev/api/skimage.metrics.html>
- Valeriano, M. D. M., Rossetti, D. D. F. (2012). Topodata: Brazilian full coverage refinement of SRTM data. *Applied Geography*. Elsevier. <https://doi.org/10.1016/j.apgeog.2011.05.004>
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., Loy, C. C. (2018). ESRRGAN: Enhanced Super-Resolution Generative Adversarial Networks. *Proceedings of the European conference on computer vision (ECCV) workshops*. arXiv:1809.00219. <https://doi.org/10.48550/arXiv.1809.00219>