

## Article

# Single-Image Super-Resolution of Sentinel-2 Low Resolution Bands with Residual Dense Convolutional Neural Networks

Luis Salgueiro <sup>1</sup> , Javier Marcello <sup>2</sup>  and Verónica Vilaplana <sup>1,\*</sup> 

<sup>1</sup> Department of Signal Theory and Communications, Universitat Politècnica de Catalunya (UPC), 08034 Barcelona, Spain; luis.fernando.salgueiro@upc.edu

<sup>2</sup> Instituto de Oceanografía y Cambio Global, IOCAG, Unidad Asociada ULPGC-CSIC, 35017 Las Palmas de Gran Canaria, Spain; javier.marcello@ulpgc.es

\* Correspondence: veronica.vilaplana@upc.edu

**Abstract:** Sentinel-2 satellites have become one of the main resources for Earth observation images because they are free of charge, have a great spatial coverage and high temporal revisit. Sentinel-2 senses the same location providing different spatial resolutions as well as generating a multi-spectral image with 13 bands of 10, 20, and 60 m/pixel. In this work, we propose a single-image super-resolution model based on convolutional neural networks that enhances the low-resolution bands (20 m and 60 m) to reach the maximal resolution sensed (10 m) at the same time, whereas other approaches provide two independent models for each group of LR bands. Our proposed model, named Sen2-RDSR, is made up of Residual in Residual blocks that produce two final outputs at maximal resolution, one for 20 m/pixel bands and the other for 60 m/pixel bands. The training is done in two stages, first focusing on 20 m bands and then on the 60 m bands. Experimental results using six quality metrics (RMSE, SRE, SAM, PSNR, SSIM, ERGAS) show that our model has superior performance compared to other state-of-the-art approaches, and it is very effective and suitable as a preliminary step for land and coastal applications, as studies involving pixel-based classification for Land-Use-Land-Cover or the generation of vegetation indices.

**Keywords:** Sentinel-2; super-resolution; convolutional neural network; deep learning



**Citation:** Salgueiro, L.; Marcello, J.; Vilaplana, V. Single-Image Super-Resolution of Sentinel-2 Low Resolution Bands with Residual Dense Convolutional Neural Networks. *Remote Sens.* **2021**, *13*, 5007. <https://doi.org/10.3390/rs13245007>

Academic Editors: Igor Yanovsky and Jing Qin

Received: 11 October 2021

Accepted: 7 December 2021

Published: 9 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Managed by the European Space Agency (ESA), the Sentinel-2 satellites play an important role in today's remote sensing as they provide multispectral optical imagery that can be used for several applications such as land cover-land use monitoring, change detection, vegetation and soil analysis, and mapping of physical variables, among others. Some of its characteristics are its considerable surface coverage, the high revisit time [1] and the possibility of getting the data for free (available at [2]), democratizing images for research, as well as launching free and commercial products, making it increasingly useful for Earth observation data.

Each satellite provides 13 bands: 4 high-resolution (HR) bands with 10 m/pixel, 6 low-resolution (LR) bands with 20 m/pixel, and 3 very low-resolution (VLR) bands with 60 m/pixel. Due to the spectral variety and the wide swath, covering a surface of 290 km, the satellites produce nearly 23 TB/day of data, resulting in a vast amount of data that needs to be stored [3]. Besides storage issues, other reasons for designing sensors at different scales were transmission bandwidth constraints or band selection that do not require HR images, among others [4].

Spatial resolution is a fundamental parameter for the analysis of remote sensing imagery. It is defined as the minimum distance in which two separated objects are distinguishable, and depends on several factors, for instance altitude, distance, and the quality of the instruments [5]. Another relevant feature, in line with spatial resolution, is the Ground Sampling Distance (GSD), which is the surface of the Earth represented by a pixel [6].

Many applications require the detail of images to be at the highest possible resolution to obtain its best performance [7,8]. Unfortunately, the excellent resolution achieved in the visible and near-infrared (VIS-NIR) spectral bands by Sentinel-2 may not be enough in some applications, especially to make use of the information given by the LR and VLR bands located in the infrared (IR) and Short-wave IR (SWIR) spectrum. These bands are suitable for a wide range of applications, such as environmental studies [9,10] and the production of land cover maps [11,12] which motivates the study of algorithms to enhance the LR bands' spatial resolution.

To overcome these constraints, several approaches were proposed to improve Sentinel-2's low-resolution bands and, thus, increase the quality of the images beyond its sensor limitation [7]. Among those methods, we found pansharpening techniques, which need a panchromatic band, and Super-Resolution (SR) algorithms, which have gained popularity over the last few years due to their great capacity for producing a HR image given a LR image [13]. To generate the HR imagery, SR can be achieved based on one (Single-Image Super-Resolution) or a sequence of LR images (Multi-Image Super-resolution) [14].

This paper proposes the use of Single-Image Super-Resolution techniques to enhance the spatial quality of the LR and VLR bands, and reach the resolution of HR bands. With the advent of Deep Neural Networks (DNN), these image processing techniques have often been applied in remote sensing [15]. In particular, in this work, we propose a Convolutional Neural Network (CNN) based model that increases the resolution of both sets of images.

Many authors tackle the challenge of enhancing the 20 m and 60 m bands to reach the maximal resolution of 10 m. Some works present analytical methods and others [3,4,16] two independent DNN models, tailored for super-resolving each band group.

As opposed to previous approaches, in this work we develop a single CNN model that takes the 10 m, 20 m, and 60 m bands as input images and jointly produces the corresponding SR bands at maximum resolution for 20 m and 60 m bands. The proposed model reuses the super-resolved 20 m bands to produce the 60 m super-resolution, delivering excellent qualitative and quantitative results, improving state-of-the-art techniques that tackle the same problem. We have divided the training into two stages: first we train the model for super-resolving the 20 m bands, and then for the 60 m bands. We also point out different applications that demonstrate the potential improvement by using the SR images obtained by our model.

Note that this work specifically addresses the problem of enhancing the spatial resolution of the low resolution bands of Sentinel 2. A different issue concerns improving the Sentinel-2 spatial resolution beyond 10 m. In this case, multisensor data are needed and, in general, expensive and very high resolution imagery is required to train and validate the models. A pixel size of 10 m may be enough for many local and regional applications but, due to the economic implications, we preferred to address both problems independently: on the one hand, improving the resolution of the 20/60 m bands down to 10 m and, on the other hand, when higher resolution is required, super-resolving these channels to higher spatial resolutions, as included in our previous work [17].

The rest of the paper is organized as follows. In Section 2, we present a review of related works that address the same problem. A technical description of Sentinel-2 and the dataset used is presented in Section 3.1. In Section 3.2 we describe the proposed model and the methodology used for training it, along with the evaluation metrics used. Section 4 shows the results where in Section 4.1 we made a comparison with other models and in Section 4.2 we put forward several applications that can benefit from our model. In Section 5, we present a discussion about our results and, finally, Section 6 provides concluding remarks.

## 2. Related Works

In this section, for simplicity, we will refer to LR when addressing both low resolution Sentinel-2 bands (LR and VLR).

One of the first options for spatial enhancement is interpolation, such as linear and bicubic. Interpolation is simple and fast, but the resulting image is often blurry and with low-quality [4].

Some platforms often carry two instruments on-board, a high-resolution sensor called panchromatic (PAN), as well as a low-resolution multi-spectral (MS) sensor. The panchromatic band has higher spatial resolution than the MS bands but a wider spectral bandwidth usually overlapping the spectral range of the MS. A common practice is to use this HR band to enhance the LR MS channels using pansharpening techniques [18]. This is not the case with Sentinel-2 satellites, as they do not carry a panchromatic sensor on-board. Instead, the multispectral instrument on-board Sentinel-2 provides data at three different spatial resolutions in the VIS-NIR and SWIR, whose spectral ranges do not overlap. In any case, some pansharpening techniques were proposed for super-resolving the LR bands using individual 10 m bands or a synthetic panchromatic band [19–24].

Some analytical methods were also proposed to improve the Sentinel-2 LR bands. For instance, Wang et al. [25] presented a fusion algorithm that extends two common approaches: component substitution and multi-regression analysis. They fuse HR and LR bands to produce SR bands using a method called area-to-point regression kriging (ATPRK), which is computationally efficient and better preserves the spectral information. This method was previously applied to the MODIS satellite and adapted for Sentinel-2. In [26], Brodu proposed to super-resolve the LR bands combining band-inherent information, to maintain spectral coherency, and independent geometric information common to all bands.

However, lately, machine learning approaches have shown to outperform analytical methods. Lanaras et al. [4] proposed a CNN with skip connections between feature maps (resblocks), mainly focusing on detail differences between the LR and HR bands rather than learning a direct mapping. In this manner, the model learns to sharpen the spatial resolution of LR bands by combining the information of HR and LR bands. The LR bands are first upsampled with bicubic interpolation and concatenated with HR bands before entering the network.

Zhang et al. [27] proposed to use a Generative Adversarial Network (GAN) [28] for super-resolving the LR bands. The generator uses a similar approach as in [4] but is enhanced with more residual blocks, and trained with adversarial training. Zhu et al. [16] used a similar methodology but, instead of using resblocks, a channel-attention mechanism [29] was proposed to better exploit the interdependence among channels of the feature maps and to let the network focus on more informative details. On the other hand, Zhang et al. [3] proposed a model combining resblocks with self-attention mechanisms. In addition, they proposed a distributed training procedure suitable for high-performance environments, such as supercomputers, that can achieve state-of-the-art results, speeding-up the training process while maintaining the loss of performance at minimum values for both models. All these models improved the baseline performance established by [4] using a dataset proposed for that purpose.

Inspired by Single Image Super-Resolution (SISR) models, Liebel et al. [7] adapted a Super-Resolution CNN (SRCNN) [30] to work with the HR and LR bands of Sentinel-2. Wagner et al. [8] followed a similar approach, adopting a Very Deep Super-Resolution model (VDSR) [31] and Palsson et al. [32] proposed a modified version of the generator of Super-Resolution Generative Adversarial Network [33]. They were more interested in assessing the effects of hyper-parameters rather than obtaining optimal results. Gargiulo et al. [34] proposed a CNN modified from a model, originally designed for pansharpening, which was fast in inference mode. Wu et al. [35] proposed a Parallel Residual Network that processes each band independently, in parallel with resblocks, before fusing the feature maps and adding the output to the bicubic upsampling of the LR band that is super-resolved. These models used different datasets, hindering the comparison with previous models.

To the best of our knowledge, all works that tackle the sharpening of both sets of LR Sentinel-2 bands use two independent CNN models, one to enhance the 20 m bands and the other for the 60 m bands. In this work, we propose a model that can produce the SR of all the LR bands (20 m and 60 m) using only one single network architecture. Our proposal splits the training into two stages, first for the 20 m bands and then for the 60 m bands, using Residual in Residual Dense Blocks (RRDB) [36] as core blocks, which enhance resblocks proposed in [4] and adopted in other works, by reusing feature maps from all previous layers. We start training using a group of RRDBs that mainly focus on super-resolving the 20 m bands. Then, we expand our model by adding more RRDBs blocks that focus on the 60 m bands. More details about the proposal are discussed in Section 3.2.

### 3. Materials and Methods

#### 3.1. Dataset

The Copernicus Sentinel-2 mission comprises a constellation of two polar-orbiting satellites, Sentinel-2A launched in 2015 and Sentinel-2B launched in 2017. Both satellites fly in the same orbit with a phase of 180 degrees, producing high revisit frequency (around 5 days) and are equipped with a Multi-Spectral Instrument that records images at the nadir in 13 multi-spectral bands with 3 different spatial resolutions (10 m, 20 m, and 60 m) covering spectral frequencies ranging from the visible to the shortwave infrared. Technical details of each band can be seen in Table 1.

**Table 1.** Spatial and spectral characteristics of Sentinel-2 bands. Source: [37]

Spectral Band	S2A Central Wavelength (nm)	S2A Bandwidth * (nm)	S2B Central Wavelength (nm)	S2B Bandwidth * (nm)	Spatial Resolution GSD (m)
B1: Coastal Aerosol	442.7	21	442.3	21	60
B2: Blue	492.4	66	492.1	66	10
B3: Green	559.8	36	559.0	36	10
B4: Red	664.6	31	665.0	31	10
B5: Red-edge 1	704.1	15	703.8	16	20
B6: Red-edge 2	740.5	15	703.8	15	20
B7: Red-edge 3	782.8	20	779.7	20	20
B8: Near-IR	832.8	106	833.0	106	10
B8A: Near-IR narrow	864.7	21	864.0	22	20
B9: Water Vapor	945.1	20	943.2	21	60
B10: SWIR-Cirrus	1373.5	31	1376.9	30	60
B11: SWIR-1	1613.7	91	1610.4	94	20
B12: SWIR-2	2202.4	175	2185.7	185	20

\* The Bandwidth (nm) is measured at Full Width Half Maximum (FWHM).

The main applications of the LR and VLR bands are devoted to environmental studies, vegetation and land cover mapping, discrimination of snow, ice, and clouds, as well as the retrieval of water-vapor, cirrus, and aerosol information, thus enabling a wide range of earth observation applications [38].

As indicated, Sentinel-2 images are freely available at [2]. A dataset of Sentinel-2 Level-1C images for sharpening the LR and VLR bands to 10 m was proposed in [4], covering diverse regions of the world and spanning different climate zones, land cover, and biome types. In this work, for comparison purposes, we use a dataset composed of 60 images (45 for training and 15 for testing), performing experiments using the same train-test split. Due to the unavailability of ground truth images (i.e., images at a higher resolution than the original), we assume that spatial details are self-similar and scale-invariant between all bands, as considered in previous works [4,16,35]. Thus, to generate pairs of input-target images for training and testing, we applied Wald's protocol [39], where images are down-

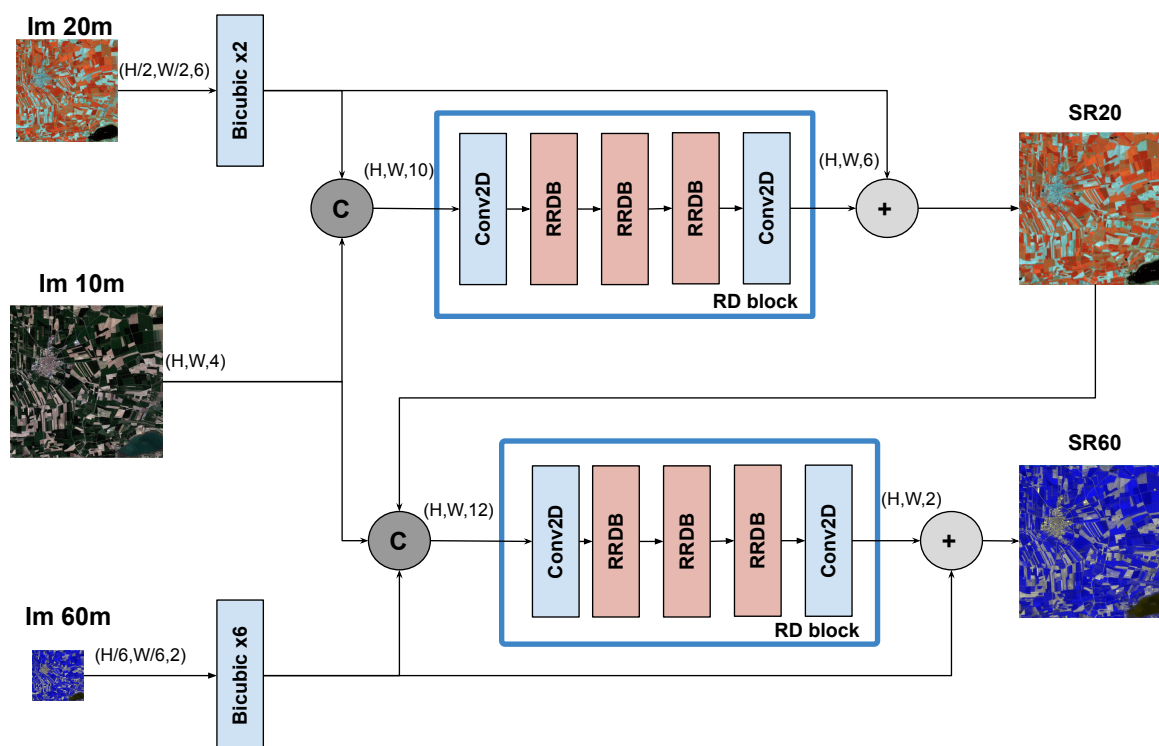


sampled, applying a Gaussian filter first, and, then, scaled in accordance to the desired scaling factor.

### 3.2. Proposed Model

#### 3.2.1. Network Architecture

The proposed model, named Sen2-RDSR for Sentinel-2 Residual Dense Super-Resolution, is shown in Figure 1. The model takes the 10 m, 20 m, and 60 m bands as input and produces super-resolved images for the 20 m and 60 m bands at 10 m of GSD.

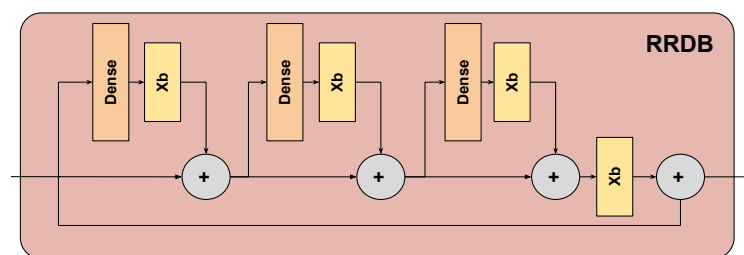


**Figure 1.** Sen2-RDSR model. The model produces two outputs, SR of 20 m bands and SR of 60 m bands. The RD blocks are composed of a convolutional layer, three RRDB blocks, and a convolutional layer that reconstructs the image with the corresponding output channels.

The architecture is formed by two branches. One that produces the SR of the 20 m bands (SR20), using as input the original 20 m and 10 m images, and a second branch that generates the SR of the 60 m bands (SR60) from the original 60 m, the 10 m bands, and the super-resolved 20 m bands obtained by the first branch.

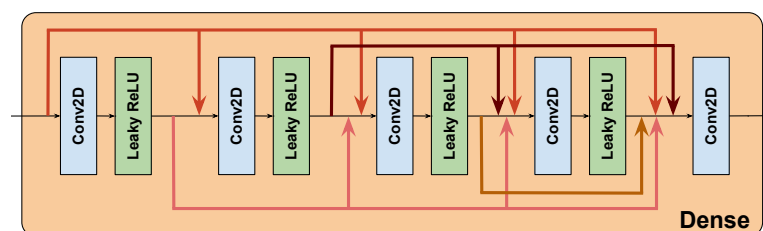
The two branches are composed of a Residual Dense block (RD) that includes a series of Residual in Residual Dense Blocks (RRDB) with shared weights. Residual in Residual Dense Blocks were first proposed in [36] for SISR tasks.

The RRDB performs the extraction of feature information of high complexity, aiming to recover details in LR images. The RRDB is a combination of three Dense Blocks (Figure 2), where the features are reused, allowing a synergistic effect and boosting the recovery of residual details. In addition, a long skip-connection in the same block maintains coherence with the input image. The scalar  $x_b$ , usually a value between 0 and 1, acts as a residual scaling: feature maps are multiplied by this value and scaled-down essentially to maintain the stability during training [40]. The value of  $x_b$  is fixed for all dense blocks. Several remote sensing applications have adopted this block structure due to its great performance in enhancing low-resolution images [17,18,41,42].



**Figure 2.** Residual in Residual Dense Block (RRDB). Each 2D convolutional layer has 32 feature maps, a  $3 \times 3$  kernel with stride of 1.

The Dense Block [43] in a RRDB (Figure 3) is a combination of five convolutional 2D layers that are connected in a feed-forward manner, where each layer's input comes from its preceding layers. In this way, features are learned more effectively and the performance is improved by using the hierarchical information from all previous layers. Each convolutional layer has 32 filters with  $3 \times 3$  kernels, stride 1, and Leaky-ReLU as activation function.



**Figure 3.** Dense Block. A RRDB is composed of three dense blocks scaled by  $Xb$  and combined, as depicted in Figure 2.

In the first branch, the LR (20 m bands) image is first upsampled using bicubic interpolation to match the size of the HR bands (10 m bands). Then, all bands are concatenated (C) and processed by a 2D convolutional layer that acts as a shallow feature extractor. This layer has 128 filters with  $3 \times 3$  kernels and stride 1. Then, the feature maps pass throughout three RRDB blocks with shared weights, and finally, are reconstructed with a 2D convolution that produces the output image with the corresponding number of channels, matching in this case, the number of 20 m bands. Next, the output features are added to the bicubic interpolated LR bands, thus reducing the spectral distortion in the SR bands.

The architecture of the second branch is similar to the first one but introduces the super-resolved 20 m bands as well, which are concatenated (C) with the HR (10 m) bands and the bicubic interpolated VLR (60 m) bands. The feature maps are processed by the sequence of three RRDB blocks and a convolutional layer and the result is added to the bicubic interpolated VLR image to obtain the final SR image with minimal spectral distortion.

### 3.2.2. Training Details

The training is done in two stages. First, we train the model to generate the super-resolution of the 20 m bands and then, on a second stage, we train the model for super-resolving the 60 m bands. In the second stage, the weights of all layers learned in the first stage are frozen to avoid changes during the training. Note that the training is done in two stages but inference is performed for the 20 m and 60 m bands at once.

In the first training stage, after applying Wald's protocol, bands with 20 m and 40 m (originally at 10 m and 20 m, respectively) are used as inputs, while in the second stage, bands with 60 m, 180 m, and 360 m (initially at 10 m, 20 m, and 60 m, respectively) are used. In both stages, the original bands are used as target images. After the generation of the input-target image pairs, from each image we select 8000 random patches of  $32 \times 32$  pixels to train the SR20 branch and 500 random patches to train the SR60 branch. We use a 90–10% split for creating the training-validation subsets. For testing, each image of the

test set is cropped into non-overlapping patches of  $32 \times 32$  pixels for the SR20 model and  $192 \times 192$  for the SR60 model. Each image is input to the model and, finally, reconstructed to compute the quantitative metrics.

Models are trained for 500 epochs with early-stopping using Adam optimizer and L1-norm as loss function. In each training stage, only the corresponding SR output is considered for calculating the loss. The learning rate is  $2 \times 10^{-4}$  in SR20 and  $5 \times 10^{-5}$  in SR60, with cosine annealing as the learning rate scheduler. Gradient clipping [44] is used and the scaling factor in the RRDB block is set to  $x_b = 0.2$  for maintaining stability in training. A batch size of 64 is set for training the models due to memory restrictions, using two Nvidia RTX-2080 GPUS with 11 GB of GRAM, configured with Pytorch-Lightning.

### 3.3. Quantitative Metrics

Several metrics are considered for the quantitative evaluation of the SR bands: RMSE, SRE, SAM, PSNR, SSIM, and ERGAS. In the following, we denote  $Y$  as the target image,  $X$  as the SR image, both with  $B$  channels and  $(H, W)$  as spatial dimensions,  $\mu$  and  $\sigma$  are the mean and standard deviation, respectively, and  $E$  is the expected value.

- Root Mean Square Error (RMSE): measures the mean error in the pixel-value space.

$$\text{RMSE}(X, Y) = \sqrt{E[(X - Y)^2]} \quad (1)$$

- Signal to Reconstruction Ratio Error (SRE) [4]: measures the relative error in reference to the power of the signal, in dB, where the higher, the better ( $n$  is the number of pixels).

$$\text{SRE}(X, Y) = 10 \log_{10} \frac{E[X]^2}{\|Y - X\|_1^2 / n} \quad (2)$$

- Spectral Angle Mapper (SAM) [45]: measures the spectral fidelity between two images. It is expressed in radians, where smaller angles represent higher similarities.

$$\text{SAM}(X, Y) = \arccos \left( \frac{X \cdot Y}{\|X\|_2 \|Y\|_2} \right) \quad (3)$$

- Peak Signal to Noise Ratio (PSNR): it is one of the standard metrics used to evaluate the quality of a reconstructed image. Here,  $\text{MaxVal}$  takes the maximum value of  $Y$ . Higher PSNR, generally, indicates higher quality.

$$\text{PSNR}(X, Y) = 20 \log_{10} \left( \frac{\text{MaxVal}(Y)}{\text{RMSE}(X, Y)} \right) \quad (4)$$

- Structural Similarity (SSIM) [46]: measures the similarity of two images by considering three aspects: luminance, contrast, and structure. SSIM takes in consideration the mean ( $\mu$ ) and variance ( $\sigma$ ) of the images, where a value of 1 corresponds to identical images. Constants  $C_1 = k_1 L$  and  $C_2 = k_2 L$  are values that depend on the dynamic range ( $L$ ) of pixel values ( $k_1 = 0.01$  and  $k_2 = 0.03$  are used by default).

$$\text{SSIM}(X, Y) = \frac{(2\mu_X \mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \quad (5)$$

- Erreur relative globale adimensionnelle de systhese (ERGAS) [47]: measures the quality of the reconstructed image considering the scaling factor ( $S$ ) and the normalized error per each channel ( $B$ ). Lower values imply higher quality.

$$\text{ERGAS}(X, Y) = \frac{100}{S} \sqrt{\frac{1}{B} \sum_{j=1}^B \left[ \frac{\text{RMSE}(X_j, Y_j)}{E(Y_j)} \right]^2} \quad (6)$$

In addition to the quantitative metrics, for visual quality comparison we use, as well, false color composites of bands {B7,B6,B5} and {B8A,B11,B12} for the 20 m bands, {B9,B9,B1} for 60 m bands and the true RGB {B4,B3,B2} for the 10 m bands.

## 4. Results

### 4.1. Super-Resolution Results

Tables 2 and 3 show the average results for both SR tasks on the test set. We obtained both tables in each corresponding training stage, where Wald's protocol [39] was also applied to the test set considering the proper scaling factor. We compare our results with the bicubic upsampling and two state-of-the-art models that use the same dataset, DSen2 [4] and Zhang et al. [3]. As can be noticed in Table 2, we outperformed both models in the four metrics considered and tied in the other two, with an improvement of 0.61 in RMSE, 0.19 in SRE and 0.2 in PSNR with the second best model. In the case of 60 m bands, the results in Table 3 show that we outperformed in three metrics, tied in two, and lost in one (the PSNR metric). We had a 1.11 decrease in RMSE and an increase of 0.16 in SRE, but the difference in PSNR was 0.84.

**Table 2.** Results of the sharpening of 20 m bands.

	RMSE	SRE	SAM	PSNR	SSIM	ERGAS
Bicubic	125.68	26.44	1.21	45.82	0.82	3.33
DSen2 [4]	35.85	35.94	0.78	55.54	<b>0.93</b>	1.07
Zhang et al. [3]	34.99	36.19	<b>0.75</b>	55.77	<b>0.93</b>	1.03
Sen2-RDSR	<b>34.38</b>	<b>36.38</b>	<b>0.75</b>	<b>55.94</b>	<b>0.93</b>	<b>1.02</b>

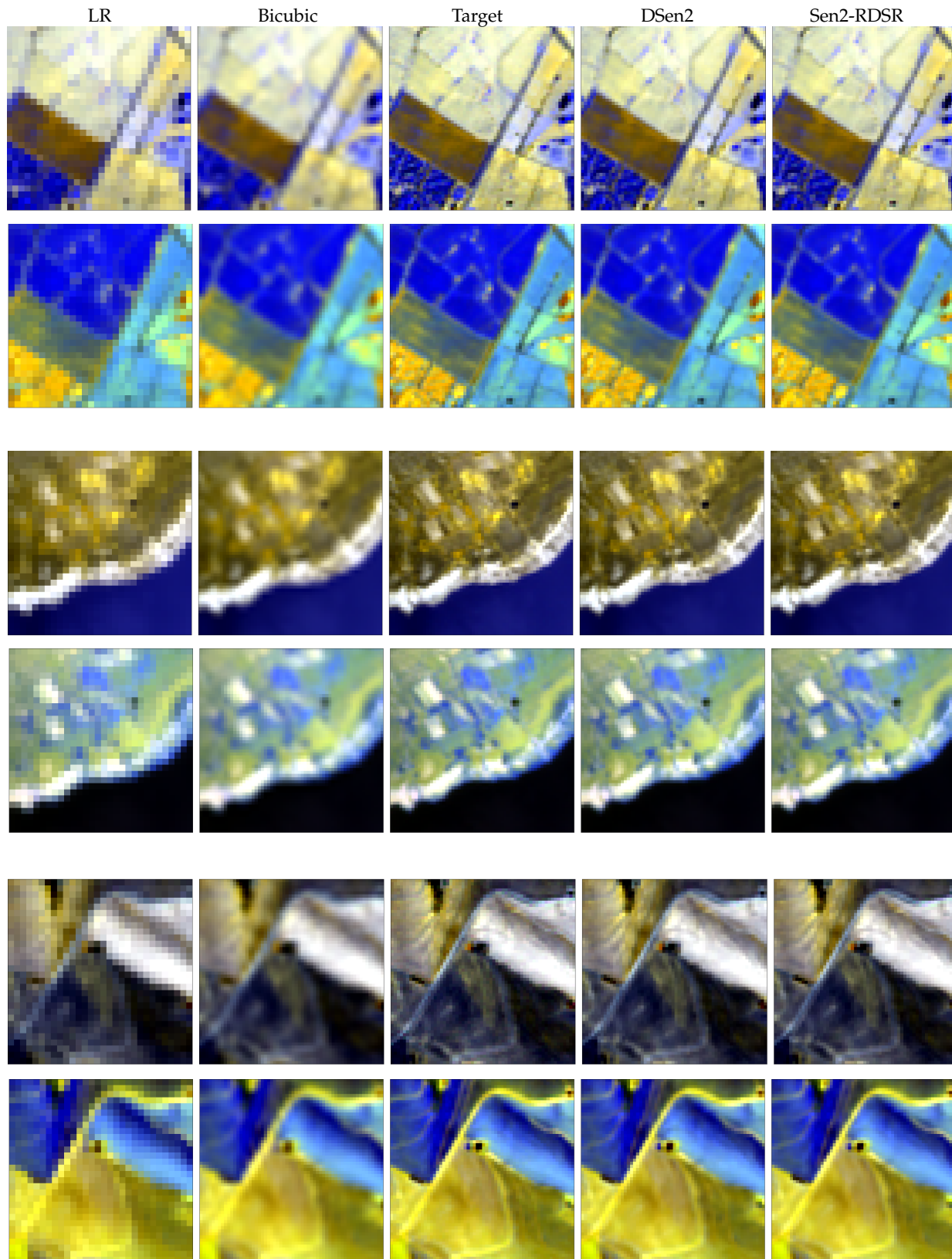
**Table 3.** Results of the sharpening of 60 m bands.

	RMSE	SRE	SAM	PSNR	SSIM	ERGAS
Bicubic	162.16	19.77	1.78	37.66	0.35	2.43
DSen2 [4]	28.11	34.47	0.36	52.49	0.89	1.38
Zhang et al. [3]	26.80	34.98	<b>0.34</b>	<b>52.94</b>	<b>0.90</b>	1.29
Sen2-RDSR	<b>25.69</b>	<b>35.14</b>	<b>0.34</b>	52.10	<b>0.90</b>	<b>0.41</b>

Visual comparisons are also presented in Figures 4 and 5, where some patches with a different sort of spatial and spectral information were used from the test set, the DSen2 images were obtained using its public repository (<https://github.com/lanha/DSen2/tree/master/models>). It is worth noting that, visually, the SR images are very similar to the target, demonstrating the low error achieved in Tables 2 and 3. To help understand the differences between the models, we plotted the absolute errors between the results and the targets in Figures 6 and 7.

If we compare the absolute error between the bicubic upsampling and the DL models, there is a significant difference, which supports the idea of using DL algorithms for super-resolution. In general, most of the error maps present dark-blue areas that correspond to a small range of values between 0 and 50 (where the original reflectance values range between 0 and more than 10.000). To provide a more detailed analysis, in Tables 4 and 5 we show the quantitative data at band level for RMSE and SRE, since they were the only metrics available in the other models.

From Table 4 our model outperforms in three of the six bands considering the RMSE, with a decrease for the B5, B8A, and B11. The decay is noticeable for bands 8A and B11, where the drop was more than 25% for B8A and more than 30% for B11. However, the big difference with the second-best model was in B12, where the difference was around 24%. Regarding the SRE metric, our model consistently outperforms in all bands. Analyzing the results on Table 5 we obtained improvements in both bands for RMSE, but with SRE we outperformed in B1 but lost in B9.



**Figure 4.** Sharpening results for Sentinel-2 20 m bands. Images correspond to different crops from the test set. Odd rows show the false-color composite {B7,B6,B5} as RGB and even rows show the false-color composite {B12,B11,B8A}. The first column corresponds to the original LR 20 m bands, the second column is the bicubic interpolation, the third column is the target image, the fourth column is the result from DSen2 [4], and the fifth column is our result. The size of the images is  $30 \times 30$  pixels for LR and  $60 \times 60$  for the rest.

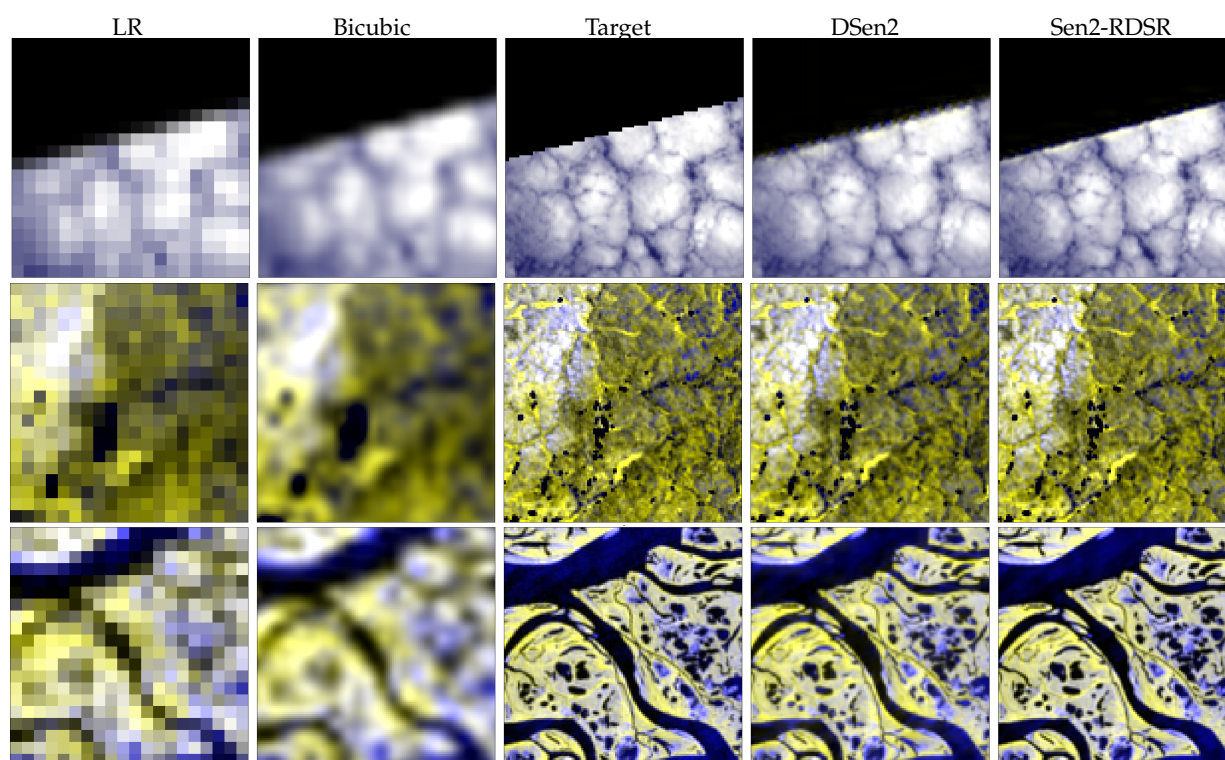


**Table 4.** Results of the sharpening of 20 m bands (per band).

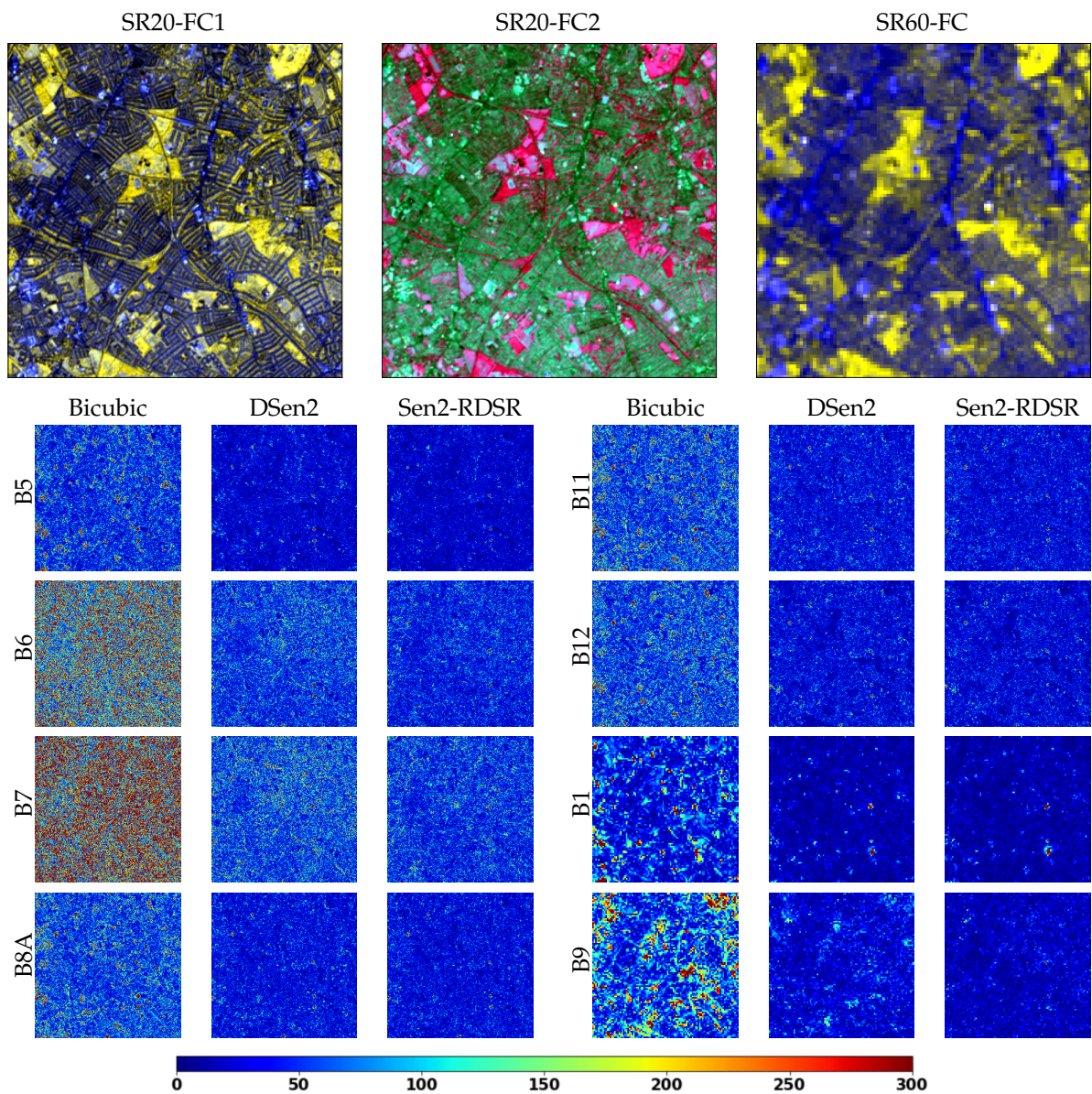
	B5	B6	B7	B8A	B11	B12
RMSE						
Bicubic	101.23	133.35	153.96	87.37	74.14	162.34
DSen2 [4]	27.74	32.68	36.07	38.02	36.22	34.55
Zhang et al. [3]	27.48	<b>32.27</b>	<b>35.58</b>	37.46	35.56	<b>33.68</b>
Sen2-RDSR	<b>26.98</b>	35.95	41.28	<b>27.62</b>	<b>24.78</b>	42.01
SRE						
Bicubic	25.42	25.89	25.66	26.80	24.44	25.81
DSen2 [4]	36.15	36.33	36.37	36.49	36.45	35.97
Zhang et al. [3]	36.26	36.44	36.49	36.62	36.66	36.22
Sen2-RDSR	<b>36.46</b>	<b>36.96</b>	<b>36.87</b>	<b>36.76</b>	<b>37.26</b>	<b>36.76</b>

**Table 5.** Results of the sharpening of 60 m bands (per band).

	B1	B9	B1	B9
RMSE		SRE		
Bicubic	169.89	146.97	22.36	16.98
DSen2 [4]	29.28	27.51	37.25	34.44
Zhang et al. [3]	27.60	26.18	37.77	<b>34.95</b>
Sen2-RDSR	<b>26.72</b>	<b>23.59</b>	<b>37.88</b>	32.41

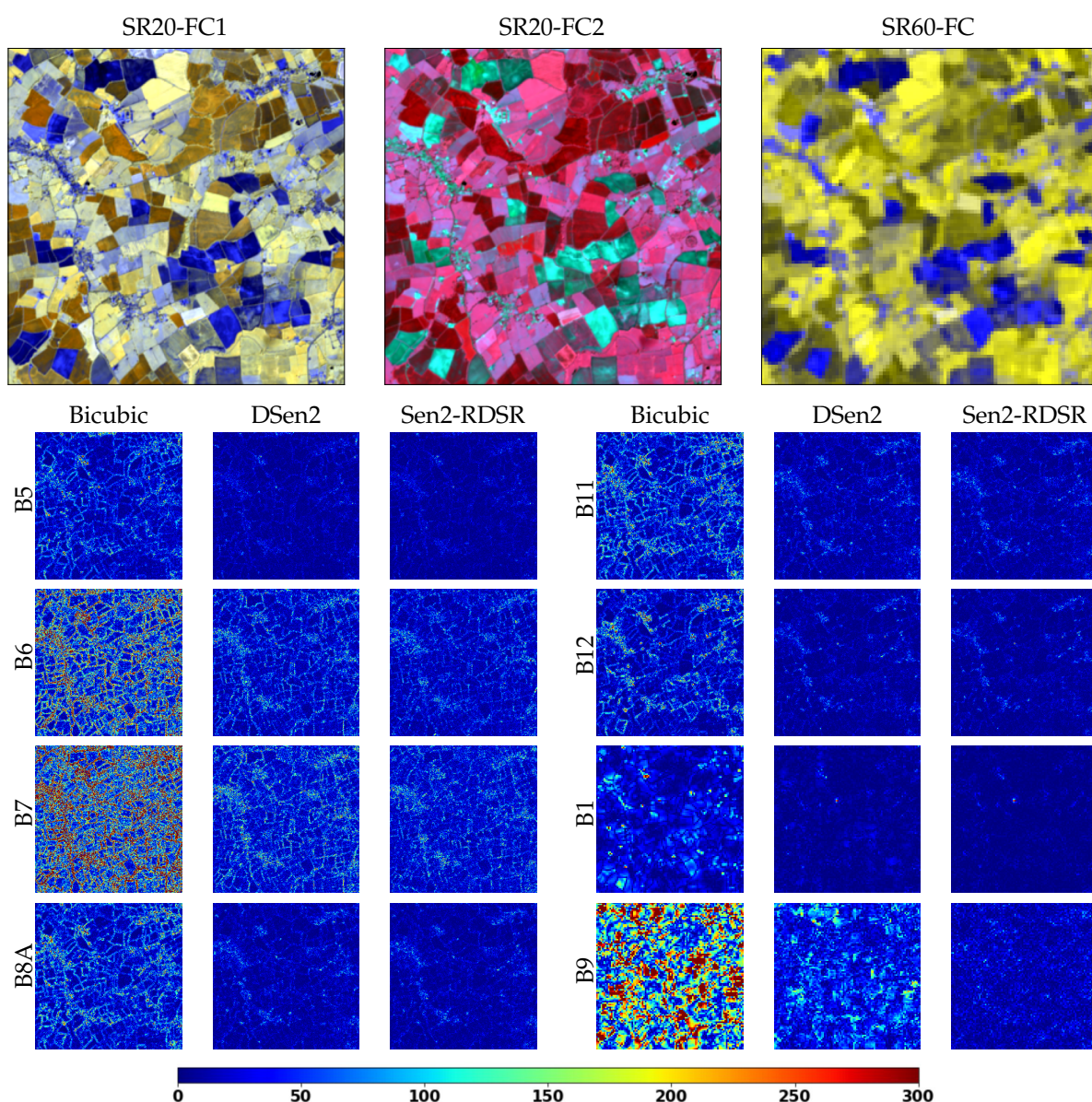


**Figure 5.** Sharpening results for the Sentinel-2 60 m bands. Images correspond to crops from the test set with the false-color composite {B9,B9,B1} as RGB. The first column corresponds to the original VLR 60 m bands, the second column is the bicubic interpolation, the third column is the target image, the fourth column is the result from DSen2 [4], and the fifth column is our result. The size of the images is  $20 \times 20$  pixels for LR and  $120 \times 120$  for the rest.



**Figure 6.** Comparison of the absolute difference between the target and the SR image per band. A small region of  $100 \times 100$  pixels was selected to show the difference between the models. The first row shows the false color composite for SR20 FC1: {B5, B6, B7}, FC2: {B12, B11, B8A} and SR60 FC: {B9, B9, B1} images. The remaining rows show the error maps for the bicubic interpolation, DSen2 (images were generated from its public GitHub repository), and our results.





**Figure 7.** Comparison of the absolute difference between the target and the SR image per band. A small region of  $100 \times 100$  pixels was selected to show the difference between the models. The first row shows the color composite for SR20 FC1:{B5,B6,B7}, FC2:{B12,B11,B8A} and SR60 {B9,B9,B1} images. The remaining rows show the error maps for the bicubic interpolation, DSen2 (images were generated from its public GitHub repository), and our results.

#### 4.2. Applications

Having shown the excellent performance of our approach for super-resolving Sentinel-2 bands, we present some applications that can benefit from the proposed SR model. It is important to point out that the analysis presented in this section is simply intended to visually demonstrate the benefits of working with the enhanced bands and not to include a detailed quantitative study.

In general, using more spectral channels allows a better determination of the objects spectral signature [48]; therefore, in the context of semantic segmentation, this feature could lead to segmentation results that are more accurate [49–51]. Usually, fully annotated ground-truth images are necessary for training deep learning models. However, dense annotations are time-consuming and expensive to generate, and, for some remote sensing applications, labeling requires expert knowledge or on-the-ground surveys. An alternative solution, at least for proof of concept studies, as the ones presented in this work, is to rely

on Support Vector Machines (SVM). This machine learning technique performs well, even when trained with few and sparse labeled data [52–55]. In this section, SVM was applied to show the benefits of using Sentinel-2 sharpened 20 m and 60 m bands in different scenarios.

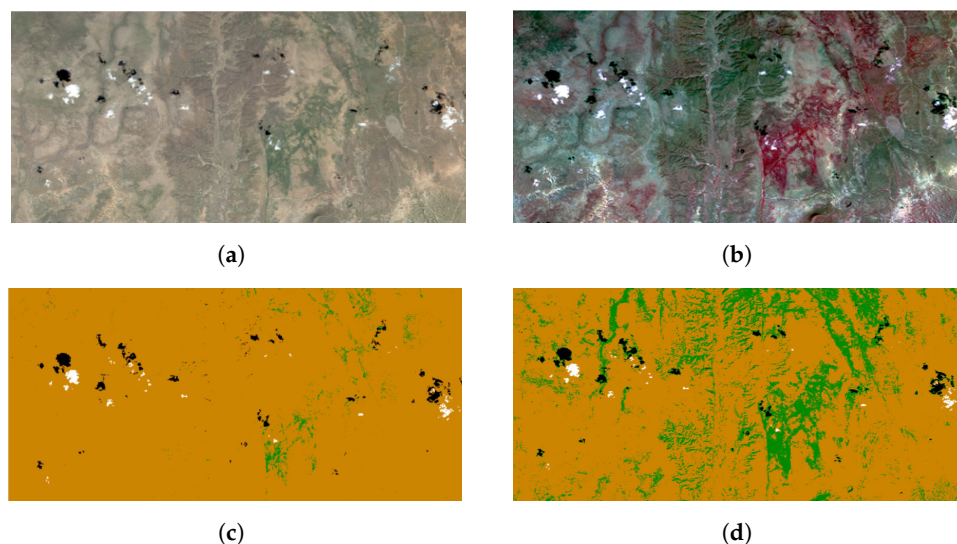
To illustrate the convenience of using the sharpened 20 m bands, we have trained a SVM classifier to perform Land-Cover Land-Use (LULC) classification on the image shown in Figure 8a. The Sentinel-2 image belongs to the dataset mentioned in Section 3.1 and corresponds to 30 December 2016. To generate the SVM maps, a SVM with radial basis kernel was trained using 200 pixel per class.

Four classes can be appreciated in the selected scene: vegetation (green), sand (orange), clouds (white), and shadows (black). Segmentation maps obtained only using the 10 m bands and a combination of 10 m plus the SR 20 m bands are presented in Figure 8c,d. Reddish pixels in Figure 8b correspond to vegetation areas. As we can see, adding the information provided by the SR of 20 m bands can be very useful for more discriminative results, especially in vegetation zones, as the 20 m bands were specially designed to monitor vegetation covers.

Another application considered that highlights the benefits of improving the GSD of the Sentinel bands are map indices, which are useful for environmental studies. In Wang and Qu [56], the Normalized Multi-band Drought Index (NMDI), which takes into account the soil moisture background, was proposed to monitor potential drought conditions by using three specific 20 m bands of Sentinel-2 (B8A, B11, B12) (Equation (7)). This index uses the difference between two liquid-water absorption bands in the shortwave-infrared region as a measure of water sensitivity in vegetation and soil.

$$\text{NMDI} = \frac{B8A - (B11 - B12)}{B8A + (B11 - B12)} \quad (7)$$

NMDI is commonly used in agriculture [57], fire monitoring [58], forest analysis [59], etc. We have chosen a small agriculture zone to better show the performance of super-resolution (Figure 9c). We can observe from the NMDI maps that vegetation zones contrast better from the soil if the maps are obtained with the super-resolved bands.



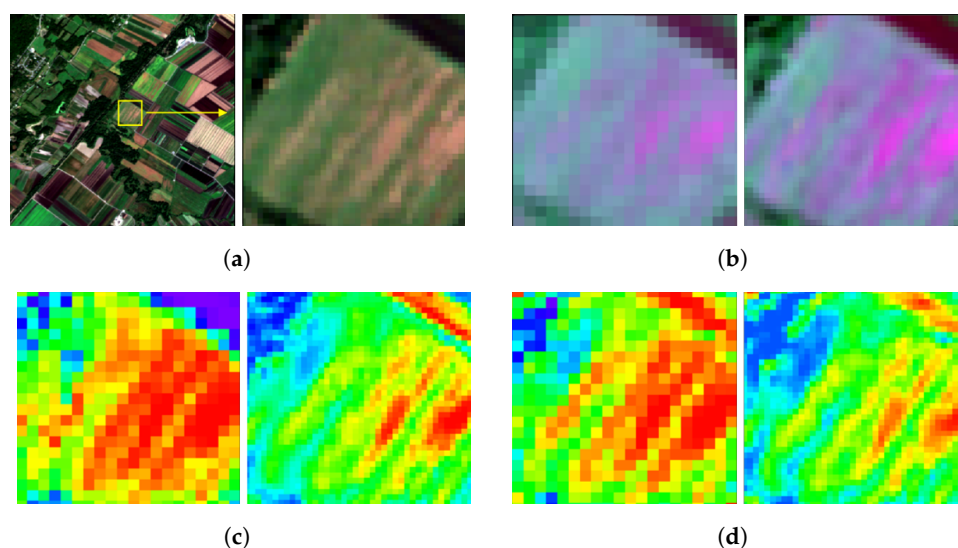
**Figure 8.** Comparison of segmentation maps obtained with only the 10 m GSD bands and with a combination of 10 m and SR 20 m bands. (a) True Color 10 m GSD (RGB); (b) False color 20 m GSD (B7, B6, B5); (c) Segmentation using only 10 m bands; (d) Segmentation using 10 m and 20 m.

Other interesting indices for vegetation studies are those that make use of the Red Edge bands of Sentinel-2 [60–62]. The NDVI-RE, see Equation (8), is a modification of the traditional NDVI (Normalized Difference Vegetation Index) that uses the Red-Edge

bands, so it is more affected by the chlorophyll content and leads to a more accurate map, especially in drier zones [62].

An example is shown in Figure 9d where the 20 m Red Edge bands B5 and B7 are used. Index maps are colored using a rainbow palette where red areas represent less vigorous vegetation. Comparing the maps with the true color image in Figure 9a, we can see that red areas correspond to sandy (dry) regions and are better outlined than using the original 20 m bands.

$$\text{NDVI-RE} = \frac{B7 - B5}{B7 + B5} \quad (8)$$



**Figure 9.** Comparison of vegetation indices using the 20 m bands with and without super-resolution (Sentinel-2 image of 28 September 2017, New York, U.S.A.). (a) Agriculture zone (RGB) (left: full scene, right: zoom-in); (b) Agriculture zone [B12,B8A,B11] (Left: 20 m GSD, right: SR (10 m GSD)); (c) NMDI Index (left: original bands, right: SR bands); (d) NDVI-RE index (left: original bands, right: SR bands).

On the other hand, remote sensing plays an important role for the effective management and monitoring of coastal areas [63]. To show the benefits of using the SR 60 m bands, we have selected a coastal area and the seafloor map has been obtained using the SVM algorithm, as well. Specifically, the 60 m coastal blue band {B1} can be very helpful to derive bathymetry and benthic mapping data in shallow waters due to its excellent penetration capability in clear waters.

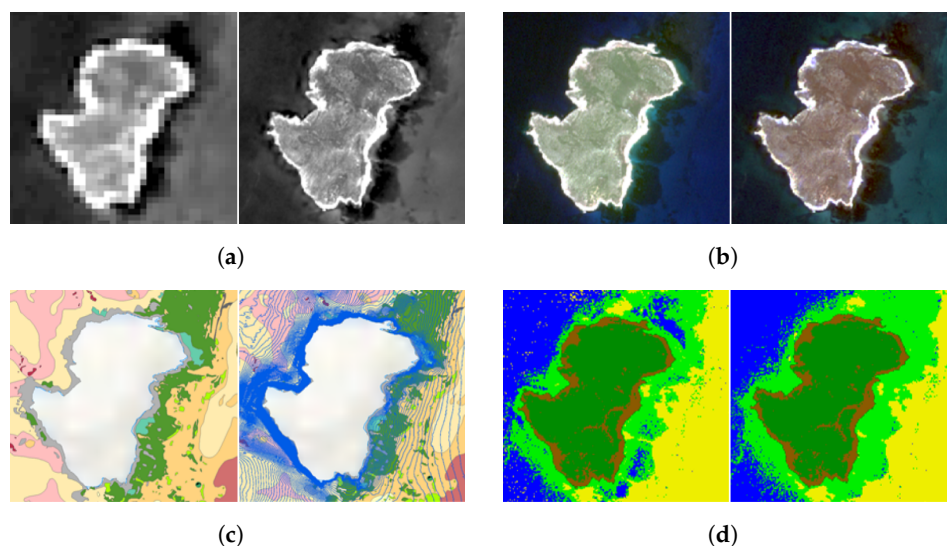
Figure 10 illustrates the results achieved in this scenario using a Sentinel-2 image from 29 June 2019 of the area of Cabrera Island, Spain. The remarkable spatial improvement in the original coastal channel can be appreciated in Figure 10a. A color composite using the SR coastal band is also included in the right column of Figure 10b to better appreciate the sea bottom (in all the images, the same enhancements have been used, applying a Gaussian histogram equalization with a 10% on brightness increase).

A reference benthic map for the visual analysis, provided by the Spanish Institute of Oceanography [64], is included in Figure 10c. The green areas correspond to different densities of *Posidonia oceanica* seagrass, gray areas relate to photophilic algae on rocky substrates, while the remaining colors identify different types of soils. Bathymetric information is also provided in the right column to demonstrate the complexity of the scene, with water depths up to 30 m on the right side of the image and much higher depths on the other side.

In particular, we chose the radial basis kernel but SVM parameters were not fine tuned and only a mean of 500 pixel per class (0.05% of the image area) were considered during the training. The derived SVM seafloor maps are displayed in Figure 10d. The Sentinel-2 original 10 m bands {B2, B3, B4, B8} have been considered to generate the map on the left



side while the result adding the SR coastal band is provided on the right side. The same training regions have been used in both cases. As expected, the coastal band does not affect the land classification (vegetation in dark green and soil in brown); however, some improvements are visible in water areas (vegetation in green lime color, sand in yellow, and deepwater areas in blue) thanks to the inclusion of the coastal channel at 10 m resolution.



**Figure 10.** Comparison of classification maps with SVM using the SR blue coastal band with and without super-resolution; (a) Coastal band (left: original, right: SR band); (b) Composites (left: RGB {B4,B3,B2}, right: {B3,B2,B1}); (c) Reference benthic map (left: original, right: with 1 m isobaths); (d) SVM maps (left: 10 m original bands, right: adding the SR coastal band).

## 5. Discussion

A comparative analysis with pansharpening techniques was not included, mainly due to the lack of a panchromatic band in Sentinel-2. In any case, a preliminary study was done with a few specific images of the dataset using classical pansharpening algorithms, in which a 10 m band acted as the panchromatic channel. The NIR band had the best performance among the available HR bands but, as expected, the spectral distortion was quite obvious in the LR bands, due to the lack of spectral overlapping between the HR and the LR bands. That said, depending on the application, we could apply pansharpening algorithms, but this would require a further comprehensive analysis to study the potential of newer and advanced pansharpening approaches.

Compared with the pansharpening paradigm, we can identify two main differences in the proposed approach. First, the idea of using all the HR bands available to super-resolve the LR bands as opposed to one band and, second, that these HR bands do not necessarily need to overlap the LR bands. We make use of this first key-difference for super-resolving the 60 m bands, where we use the 10 m bands combined with the super-resolved 20 m bands as HR bands for super-resolving at such high scaling factor. The SR 20 m bands were obtained with the RD-Blocks adjusted in the first stage of the training, in contrast with other approaches that used the bicubic upsampling of these bands.

It is important to note that our comparative analysis includes two approaches that have used the same dataset for training and testing the models. We acknowledge the existence of other works, as stated in Section 2, that have tackled the same problem but only provided the results for 20 m bands or, in other cases, have used other images to provide its benchmarks. However, the visual comparison, presented in Figures 4 and 5, shows that our model can super-resolve the LR and VLR bands with similar performance as a state-of-the-art model such as DSen2 [4], although, it can be quite difficult to spot the difference even for a well-trained eye. Analyzing the performance of the 60 m bands, in Figure 5 we can see that the reconstruction using a bicubic interpolation presents obvious lack of details, in contrast with the similarity between the SR images and target. Even

for this scaling factor, the models can properly transfer the high-frequency information, keeping spectral distortion at minimum.

We also provide a visual comparison of the errors per band in Figures 6 and 7 where the details are easier to identify. Our model obtained more dark blue areas than those presented by the DSen2 model [4].

If we analyze the errors of each band presented by the bicubic images, we can detect a decrease in performance for bands 6, 7, and 9. These errors are attenuated by the DL algorithms, and the results have proven the idea that high-frequency details obtained from the HR bands can be remarkably combined.

Inspecting Figures 6 and 7 in further detail, regarding the 20 m bands, we observe more red spot areas (absolute errors of more than 250) in DSen2 bands than in our model. We also observe, in general, better performance in the Red-Edge bands, e.g., bands 5 and 8A, although it performs well with the SWIR bands also, even being spectrally away from the influence of the HR bands. The good performance with band B8A can help generate better NMDI maps, see Equation (7), which are often used for vegetation analysis.

Furthering the analysis of the 60 m bands, the absolute errors presented in Figures 6 and 7 showed fewer red spots areas than DSen2 (errors bigger than 250), and especially in the B9 band, where there are less light blue areas as well (errors between 100 and 150).

Tables 4 and 5 present the quantitative metrics per band where our model prevails in general. Looking at the RMSE results for the 20 m bands, it is noticeable how our model performs well with bands B5, B8A, and B11, where the influence of the HR band, especially in the confluence of the red (B4) and NIR band (B8) contributes to generate a good margin for bands B8A and B11. On the other hand, as expected, B12 has a significant drop in performance compared to the other two benchmark models. Regarding the RMSE results of 60 m bands, our model improves both bands consistently, again thanks to the spectral proximity of the HR bands of Sentinel-2.

Considering the applications, Land-Use-Land-Cover (LULC) is one of the main areas of interest in remote sensing, providing fundamental information to support regional and local policies for sustainable development, risk prevention and assessment, environmental protection, etc. The interest in applying automatic semantic segmentation (pixel-based classification) to remote sensing imagery has recently increased [65], especially with the use of Sentinel-2 satellites, because they are free of charge, have a high revisit time and spectral variety, thanks to the development of deep convolutional neural networks for accurate semantic segmentation. Furthermore, other relevant land applications are environmental studies (vegetation, desertification, fires, etc.) or agriculture [57]. In these scenarios, the use of spectral indices provides a powerful tool to monitor and map such areas. In this context, Sentinel-2 offers the opportunity to address all the previous applications thanks to the multispectral channels available, specially the 20 m resolution band located in the Red Edge and SWIR.

On the other hand, Sentinel-2 provides three 60 m channels that are mainly useful for atmospheric correction purposes. In fact, the cirrus band (Band 10) is not supplied in the Sentinel-2 product as it does not contain surface information. However, the coastal aerosol (Band 1), located in the lower blue part of the spectrum, can be useful as well in coastal applications, thanks to its water penetration capability. As shown, the application of the SR model proposed can help create better land and coastal maps, compared to using the original bands, opening the possibility for local studies that require a high spatial resolution.

It is important to highlight that the goal of the analysis included in Section 4.2 was just to visually demonstrate the practicality of our model in different scenarios and not to provide a detailed quantitative analysis or to achieve optimal mapping. For land areas, the enhanced 20 m bands proved their benefits. In coastal scenarios, enhancing the 60 m coastal channel can also be a good alternative to reach water depths over 30 m, particularly

in order to put together more accurate bathymetric and seabed maps at higher water depth and spatial resolution.

## 6. Conclusions

Sentinel-2 satellites provide multi-spectral images with different ground sampling distances that enable a variety of studies and analyses of the Earth's surface. Due to the physical limitation of sensors and to minimize storage of data, only a small set of bands are provided with maximal spatial resolution. In this work, we have proposed a fully convolutional neural network that sharpens Sentinel-2 20 m and 60 m bands to 10 m GSD to get all bands at the maximal sensor resolution. Our SR model uses residual learning and dense connections that have proven their ability for enhancing low-resolution images in SISR tasks. The CNN model is formed by two branches, one for the SR of 20 m bands and the other for the 60 m bands, which are trained separately in two stages. At inference time, however, all bands are super-resolved in the same forward pass.

Quantitative and qualitative results have shown that our method performs better than state-of-the-art models that tackle the same problem. We have also shown that land applications such as LULC mapping and vegetation analysis could benefit from using the sharpened 20 m bands, and coastal studies could improve the quality of bathymetric and seafloor mapping using the SR 60 m bands.

**Author Contributions:** Conceptualization, L.S., J.M. and V.V.; data curation, L.S.; methodology, L.S., J.M. and V.V.; software, L.S. and J.M.; supervision J.M. and V.V.; validation, L.S., J.M. and V.V.; formal analysis, L.S., J.M. and V.V.; resources, L.S., J.M. and V.V.; writing—original draft preparation, L.S., J.M. and V.V.; writing—review and editing, L.S., J.M. and V.V.; funding acquisition, L.S., J.M. and V.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the Spanish Agencia Estatal de Investigación (AEI) under projects ARTEMISAT-2 (CTM2016-77733-R) and PID2020-117142GB-I00 of the call MCIN/AEI/10.13039/501100011033).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data can be downloaded from [2].

**Acknowledgments:** L.S. would like to acknowledge the BECAL (Becas Carlos Antonio López) scholarship for the financial support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

ATPRK	Area to Point Regression Kriging
CNN	Convolutional Neural Network
DNN	Deep Neural Network
ESA	European Space Agency
GAN	Generative Adversarial Network
GSD	Ground Sampling Distance
HR	High-Resolution
IR	Infrared
LULC	Land Use Land Cover
LR	Low-Resolution
MS	Multi-Spectral
NDVI	Normalized Difference Vegetation Index
NDVI-RE	Normalized Difference Vegetation Index Red-edge
NMDI	Normalized Multi-band Drought Index

PAN	Panchromatic band
PSNR	Peak Signal to Noise Ratio
RMSE	Root Mean Square Error
RRDB	Residual in Residual Dense Block
SAM	Spectral Angle Mapper
Sen2-RDSR	Sentinel-2 Residual Dense Super-Resolution
SISR	Single-Image Super-Resolution
SR	Super-Resolution
SRCNN	Super-Resolution Convolutional Neural Network
SRE	Signal to Reconstruction Ratio Error
SSIM	Structural Similarity
SVM	Support Vector Machine
SWIR	Short-Wave Infrared
SR20	Super-Resolution of 20 m bands
SR60	Super-Resolution of 60 m bands
TB	TeraByte
VDSR	Very Deep Super-Resolution
VIS-NIR	Visible and Near Infrared
VLR	Very-Low Resolution

## References

1. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36.
2. Copernicus Open Access Hub. European Space Agency. Available online: <https://scihub.copernicus.eu/dhus/#/home> (accessed on 21 March 2021).
3. Zhang, R.; Cavallaro, G.; Jitsev, J. Super-Resolution of Large Volumes of Sentinel-2 Images with High Performance Distributed Deep Learning. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 617–620. doi:10.1109/IGARSS39084.2020.9323734.
4. Lanaras, C.; Bioucas-Dias, J.; Galliani, S.; Baltasavias, E.; Schindler, K. Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 305–319.
5. Alparone, L.; Aiazzi, B.; Baronti, S.; Garzelli, A. *Remote Sensing Image Fusion*; CRC Press: Boca Raton, FL, USA, 2015.
6. Lillesand, T.; Kiefer, R.W.; Chipman, J. *Remote Sensing and Image Interpretation*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
7. Liebel, L.; Körner, M. Single-image super resolution for multispectral remote sensing data using convolutional neural networks. *ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 883–890.
8. Wagner, L.; Liebel, L.; Körner, M. Deep residual learning for single-image super-resolution of multi-spectral satellite imagery. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *4*. <https://doi.org/10.5194/isprs-annals-IV-2-W7-189-2019>.
9. Toming, K.; Kutser, T.; Laas, A.; Sepp, M.; Paavel, B.; Nöges, T. First experiences in mapping lake water quality parameters with Sentinel-2 MSI imagery. *Remote Sens.* **2016**, *8*, 640.
10. Kolokoussis, P.; Karathanassi, V. Oil spill detection and mapping using sentinel 2 imagery. *J. Mar. Sci. Eng.* **2018**, *6*, 4.
11. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.R.; Murayama, Y.; Ranagalage, M. Sentinel-2 data for land cover/use mapping: a review. *Remote Sens.* **2020**, *12*, 2291.
12. Pedrayes, O.D.; Lema, D.G.; García, D.F.; Usamentiaga, R.; Alonso, Á. Evaluation of Semantic Segmentation Methods for Land Use with Spectral Imaging Using Sentinel-2 and PNOA Imagery. *Remote Sens.* **2021**, *13*, 2292.
13. Anwar, S.; Khan, S.; Barnes, N. A deep journey into super-resolution: A survey. *ACM Comput. Surv. (CSUR)* **2020**, *53*, 1–34.
14. Arefin, M.R.; Michalski, V.; St-Charles, P.L.; Kalaitzis, A.; Kim, S.; Kahou, S.E.; Bengio, Y. Multi-image super-resolution for remote sensing using deep recurrent networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 206–207.
15. Tsagkatakis, G.; Aidini, A.; Fotiadou, K.; Giannopoulos, M.; Pentari, A.; Tsakalides, P. Survey of deep-learning approaches for remote sensing observation enhancement. *Sensors* **2019**, *19*, 3929.
16. Zhu, X.; Xu, Y.; Wei, Z. Super-Resolution of Sentinel-2 Images Based on Deep Channel-Attention Residual Network. In Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 628–631.
17. Salgueiro Romero, L.; Marcello, J.; Vilaplana, V. Super-Resolution of Sentinel-2 Imagery Using Generative Adversarial Networks. *Remote Sens.* **2020**, *12*, 2424. doi:10.3390/rs12152424.
18. Zhou, C.; Zhang, J.; Liu, J.; Zhang, C.; Fei, R.; Xu, S. PercepPan: Towards unsupervised pan-sharpening based on perceptual loss. *Remote Sens.* **2020**, *12*, 2318.
19. Kaplan, G. Sentinel-2 Pan Sharpening—Comparative Analysis. *Proceedings* **2018**, *2*, 345. doi:10.3390/ecrs-2-05158.
20. Vaiopoulos, A.; Karantzas, K. Pansharpening on the narrow VNIR and SWIR spectral bands of Sentinel-2. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 723.



21. Du, Y.; Zhang, Y.; Ling, F.; Wang, Q.; Li, W.; Li, X. Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sens.* **2016**, *8*, 354.
22. Gašparović, M.; Jogun, T. The effect of fusing Sentinel-2 bands on land-cover classification. *Int. J. Remote Sens.* **2018**, *39*, 822–841.
23. Armannsson, S.E.; Ulfarsson, M.O.; Sigurdsson, J.; Nguyen, H.V.; Sveinsson, J.R. A Comparison of Optimized Sentinel-2 Super-Resolution Methods Using Wald's Protocol and Bayesian Optimization. *Remote Sens.* **2021**, *13*, 2192. doi:10.3390/rs13112192.
24. Brook, A.; De Micco, V.; Battipaglia, G.; Erbaggio, A.; Ludeno, G.; Catapano, I.; Bonfante, A. A smart multiple spatial and temporal resolution system to support precision agriculture from satellite images: Proof of concept on Aglianico vineyard. *Remote Sens. Environ.* **2020**, *240*, 111679. doi:10.1016/j.rse.2020.111679.
25. Wang, Q.; Shi, W.; Li, Z.; Atkinson, P.M. Fusion of Sentinel-2 images. *Remote Sens. Environ.* **2016**, *187*, 241–252.
26. Brodu, N. Super-resolving multiresolution images with band-independent geometry of multispectral pixels. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4610–4617.
27. Zhang, K.; Sumbul, G.; Demir, B. An Approach To Super-Resolution Of Sentinel-2 Images Based On Generative Adversarial Networks. In Proceedings of the 2020 Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), Tunis, Tunisia, 9–11 March 2020; pp. 69–72.
28. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *arXiv* **2014**, arXiv:1406.2661.
29. Zhang, Y.; Li, K.; Li, K.; Wang, L.; Zhong, B.; Fu, Y. Image super-resolution using very deep residual channel attention networks. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 286–301.
30. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 295–307.
31. Kim, J.; Lee, J.K.; Lee, K.M. Accurate image super-resolution using very deep convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1646–1654.
32. Palsson, F.; Sveinsson, J.R.; Ulfarsson, M.O. Sentinel-2 image fusion using a deep residual network. *Remote Sens.* **2018**, *10*, 1290.
33. Ledig, C.; Theis, L.; Huszar, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; Shi, W. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 22–25 July 2017.
34. Gargiulo, M.; Mazza, A.; Gaetano, R.; Ruello, G.; Scarpa, G. Fast super-resolution of 20 m Sentinel-2 bands using convolutional neural networks. *Remote Sens.* **2019**, *11*, 2635.
35. Wu, J.; He, Z.; Hu, J. Sentinel-2 Sharpening via parallel residual network. *Remote Sens.* **2020**, *12*, 279.
36. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; Change Loy, C. Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018.
37. MultiSpectral Instrument (MSI) Overview. Available online: <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument> (accessed on 26 November 2021).
38. Sentinel-2 User Handbook. Available online: [https://sentinel.esa.int/documents/247904/685211/Sentinel-2\\_User\\_Handbook](https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook) (accessed on 21 March 2021).
39. Wald, L.; Ranchin, T.; Mangolini, M. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 691–699.
40. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, USA, 4–9 February 2017; Volume 31.
41. Chen, H.; Zhang, X.; Liu, Y.; Zeng, Q. Generative adversarial networks capabilities for super-resolution reconstruction of weather radar echo images. *Atmosphere* **2019**, *10*, 555.
42. Romero, L.S.; Marcello, J.; Vilaplana, V. Comparative study of upsampling methods for super-resolution in remote sensing. In Proceedings of the Twelfth International Conference on Machine Vision (ICMV 2019), Amsterdam, The Netherlands, 16–18 November 2019; Volume 11433, p. 114331J.
43. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 22–25 July 2017; pp. 4700–4708.
44. Pascanu, R.; Mikolov, T.; Bengio, Y. On the difficulty of training recurrent neural networks. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 1310–1318.
45. Yuhas, R.H.; Goetz, A.F.; Boardman, J.W. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In Proceedings of the Summaries 3rd Annu. JPL Airborne Geosci Workshop, Pasadena, CA, USA, 1–5 June 1992; Volume 1, pp. 147–149.
46. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612.
47. Wald, L. *Data Fusion: Definitions and Architectures: Fusion of Images of Different Spatial Resolutions*; Presses des MINES: Paris, France, 2002.
48. Camps-Valls, G.; Tuia, D.; Gómez-Chova, L.; Jiménez, S.; Malo, J. Remote sensing image processing. *Synth. Lect. Image Video Multimed. Process.* **2011**, *5*, 1–192.



49. Tarabalka, Y.; Chanussot, J.; Benediktsson, J.A. Segmentation and classification of hyperspectral images using watershed transformation. *Pattern Recognit.* **2010**, *43*, 2367–2379.
50. Signoroni, A.; Savardi, M.; Baronio, A.; Benini, S. Deep learning meets hyperspectral image analysis: A multidisciplinary review. *J. Imaging* **2019**, *5*, 52.
51. Li, S.; Song, W.; Fang, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Deep learning for hyperspectral image classification: An overview. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6690–6709.
52. Moliner, E.; Romero, L.S.; Vilaplana, V. Weakly Supervised Semantic Segmentation For Remote Sensing Hyperspectral Imaging. In Proceedings of the ICASSP 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 4–8 May 2020; pp. 2273–2277.
53. Medina Machín, A.; Marcello, J.; Hernández-Cordero, A.I.; Martín Abasolo, J.; Eugenio, F. Vegetation species mapping in a coastal-dune ecosystem using high resolution satellite imagery. *GIScience Remote Sens.* **2019**, *56*, 210–232.
54. Maulik, U.; Chakraborty, D. Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 33–52.
55. Sheykhmousa, M.; Mahdianpari, M.; Ghanbari, H.; Mohammadimanesh, F.; Ghamisi, P.; Homayouni, S. Support Vector Machine vs. Random Forest for Remote Sensing Image Classification: A Meta-analysis and systematic review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 6308–6325. doi:10.1109/JSTARS.2020.3026724.
56. Wang, L.; Qu, J.J. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophys. Res. Lett.* **2007**, *34*, L20405.
57. Pan, H.; Chen, Z.; Ren, J.; Li, H.; Wu, S. Modeling winter wheat leaf area index and canopy water content with three different approaches using Sentinel-2 multispectral instrument data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *12*, 482–492.
58. Pereira-Pires, J.E.; Aubard, V.; Ribeiro, R.A.; Fonseca, J.M.; Silva, J.; Mora, A. Semi-automatic methodology for fire break maintenance operations detection with Sentinel-2 imagery and artificial neural network. *Remote Sens.* **2020**, *12*, 909.
59. Cucca, B.; Recanatesi, F.; Ripa, M.N. Evaluating the Potential of Vegetation Indices in Detecting Drought Impact Using Remote Sensing Data in a Mediterranean Pinewood. In *International Conference on Computational Science and Its Applications*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 50–62.
60. Sun, Y.; Qin, Q.; Ren, H.; Zhang, T.; Chen, S. Red-edge band vegetation indices for leaf area index estimation from Sentinel-2/msi imagery. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 826–840.
61. Lin, S.; Li, J.; Liu, Q.; Li, L.; Zhao, J.; Yu, W. Evaluating the effectiveness of using vegetation indices based on red-edge reflectance from Sentinel-2 to estimate gross primary productivity. *Remote Sens.* **2019**, *11*, 1303.
62. Evangelides, C.; Nobajas, A. Red-Edge Normalised Difference Vegetation Index NDVI705 from Sentinel-2 imagery to assess post-fire regeneration. *Remote Sens. Appl. Soc. Environ.* **2020**, *17*, 100283.
63. Marcello, J.; Eugenio, F.; Gonzalo-Martín, C.; Rodríguez-Esparragon, D.; Marqués, F. Advanced Processing of Multiplatform Remote Sensing Imagery for the Monitoring of Coastal and Mountain Ecosystems. *IEEE Access* **2020**, *9*, 6536–6549.
64. IEO (Instituto Español de Oceanografía). Parque Nacional Marítimo-Terrestre del Archipiélago de Cabrera (Data Source). Available online: <http://www.ideo-cabrera.ideo.es/> (accessed on 13 April 2021).
65. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177.