HAL
open science

On The Role of Alias and Band-Shift for Sentinel-2 Super-Resolution

Ngoc Long Nguyen, Jérémy Anger, Lara Raad, Bruno Galerne, Gabriele Facciolo

To cite this version:

Ngoc Long Nguyen, Jérémy Anger, Lara Raad, Bruno Galerne, Gabriele Facciolo. On The Role of Alias and Band-Shift for Sentinel-2 Super-Resolution. 2023 International Geoscience and Remote Sensing Symposium (IGARSS 2023), Jul 2023, Pasadena, United States. pp.4294-4297, 10.1109/IGARSS52108.2023.10282805 . hal-04136470

HAL Id: hal-04136470

https://hal.science/hal-04136470v1

Submitted on 21 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON THE ROLE OF ALIAS AND BAND-SHIFT FOR SENTINEL-2 SUPER-RESOLUTION

Ngoc Long Nguyen [a] *Jérémy Anger* [a,b] *Lara Raad* [c] *Bruno Galerne* [d,e] *Gabriele Facciolo* [a]

[a] Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, France [b] Kayrros SAS
[c] Univ Gustave Eiffel, CNRS, LIGM, F-77454 Marne-la-Vallée, France
[d] Institut Denis Poisson, Université d'Orléans, Université de Tours, CNRS, France [e] Institut universitaire de France (IUF)

ABSTRACT

In this work, we study the problem of single-image super-resolution (SISR) of Sentinel-2 imagery. We show that thanks to its unique sensor specification, namely the inter-band shift and alias, that deep-learning methods are able to recover fine details. By training a model using a simple L_1 loss, results are free of hallucinated details. For this study, we build a dataset of pairs of images Sentinel-2/PlanetScope to train and evaluate our super-resolution (SR) model.

***Index Terms*—** Super-resolution, Sentinel-2, Alias

1. INTRODUCTION

The use of satellite imagery has become increasingly prevalent in a variety of fields, from environmental monitoring to urban planning. One such satellite is the Sentinel-2 constellation, which provides recurrent 10m/pixel resolution optical imagery. The high frequency revisit of Sentinel-2 makes it useful for monitoring temporal changes, such as the growth of crops or the spread of urban development. However, the relatively low spatial resolution can be a limitation for certain applications, such as identifying small objects or analyzing fine-scale features.

In this paper, we propose a deep learning approach for SISR of Sentinel-2 imagery. Unlike previous methods that aim for a x4 increase in spatial resolution, our work focuses on a x2 increase, which we argue is a more reasonable and practical choice. Additionally, we avoid using generative adversarial networks (GANs) in our method, as they have been known to introduce hallucinations and artifacts that can be undesirable for sensitive applications. Instead, we use an L_1 cost function, which has been shown to effectively preserve image details while minimizing distortion [1] (see Figure 1).

This study focuses on understanding what makes SISR of Sentinel-2 imagery possible. To this aim, we explore two

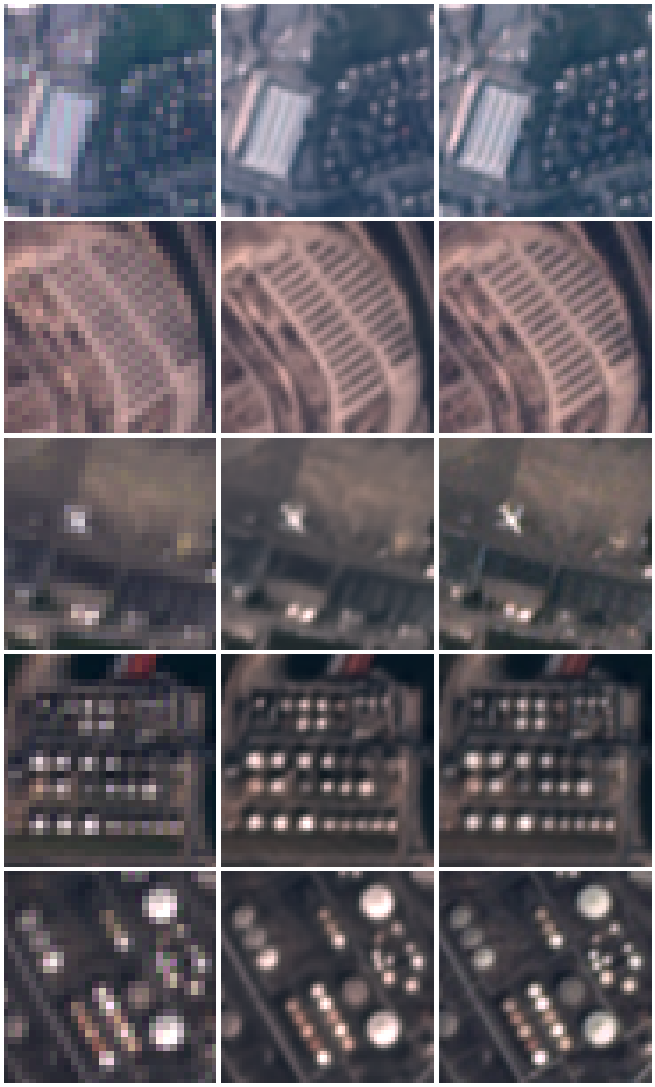
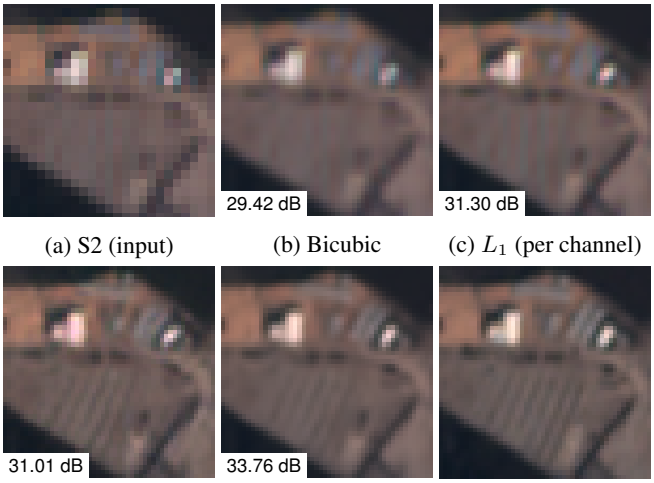


Fig. 1: SISR results obtained with the L_1 loss. We argue that the characteristic alias and band-shift are key for x2 SR of Sentinel-2 imagery.

This work was supported by a grant from Région Île-de-France, and DGA Astrid Maturation project SURECAVI ANR-21-ASM3-0002. This work was performed using HPC resources from GENCI–IDRIS (grants 2022-AD011012453R1, 2022-AD011012458R1 and 2022-AD011012472R1).



(a) S2 (input) (b) Bicubic (c) L_1 (per channel)

(d) GAN loss (e) L_1 loss (f) PS (ground-truth)

Fig. 2: SR results from different models. The per-channel result corresponds to one model per channel trained independently. The GAN and the L_1 loss are able to restore similar details.

unique characteristics of Sentinel-2: the alias and the inter-band shift and find that they enable the reconstruction of fine structures. It is worth noting that super-resolving the 10m bands of Sentinel-2 is a relatively new problem, and while we do not aim to achieve the best possible results, we provide an analysis of the specific features of Sentinel-2 imagery that are relevant for SR.

2. RELATED WORK

Early research on SISR of Sentinel-2 images focuses on pan-sharpening the lower-resolution (20m and 60m) bands to complete a uniform 10m GSD data cube [2, 3]. Recent trends are super-resolving the 10m bands of Sentinel-2 using other relevant very high-resolution satellites. For example, [4] generates low-resolution/high-resolution (LR-HR) image pairs from the PeruSat-1 satellite (2.8m GSD) to train a x4 SR model and use it to reconstruct fine textures in the Sentinel-2 10m bands. However, these techniques use a pre-determined degradation model, like bicubic downsampling, to create LR from HR. So when the input deviates from the pre-defined degradation model, the performance may drop substantially. To fill the gap between simulated and real-world remote sensing images, real HR satellites such as PlanetScope [5, 6], VENµS [7], and WorldView [8] are used directly to supervise the SR of Sentinel-2. Perceptual losses, such as GAN or high-level feature matching, are used in these works to produce sharp outputs.

Besides focusing on perceptual restoration, most past studies do not justify why the reconstruction of actual high-frequency details is feasible from a single multi-band image. Results reported in [9] suggest that alias and displacement between frames are crucial for exploiting complementary information in different frames (or different spectral bands in our case) and obtaining up-to-par SR performance.

3. METHOD

Our method is specifically designed for x2 SR of Sentinel-2 images. It is based on the ESRGAN architecture [10], adapted for a smaller network and a single-term loss. The model is trained on pairs of Sentinel-2 and PlanetScope images, suitable for x2 SR, as described in Section 4.

Architecture. Given that ESRGAN was developed for a x4 SR factor, we propose some adjustments for a factor of 2. We found that using only 8 RRDB blocks instead of 23 RRDB blocks [10] was enough to obtain satisfactory results while significantly reducing the training and inference time.

Cost function. The ESRGAN model [10], was initially trained on a set of HR natural images from the DIV2K dataset [11]. It uses a base model trained with a loss L_1, followed by a second training phase using a cost function that includes the relativistic discriminator loss [12], the perceptual loss [13], and the L_1 loss. However, when adapting the model for Sentinel-2 SR, we found that training the model on Sentinel-2/PlanetScope image pairs using only the L_1 loss instead of the complete loss function with perceptual terms resulted in a similar detail reconstruction (second row in Figure 2). This suggests that, thanks to the alias and inter-band shift present in Sentinel-2 imagery, the problem is better posed. Hence the L_1 loss is sufficient for successful x2 SR. This is further explored in Section 5.

4. S2/PS DATASET

For this study, we built a dataset of Sentinel-2 L1C and PlanetScope image pairs, referred to as the S2/PS dataset. The Sentinel-2 L1C images have a spatial resolution of 10m/px, while the PlanetScope images have a resolution of 3m/px. The latter were resampled to 5m/px and registered with the S2 coordinate system with bicubic interpolation. Since PlanetScope images tend to be well-sampled [14], the resampling step does not introduce a significant loss of information. The pairs are from images taken on the same day, a few hours apart, to minimize changes due to vegetation or human activity. The remaining changes in the images include the presence of clouds, differences in satellite perspective, and shadows. In this work, we use the PlanetScope images acquired with the PS2 instrument (Dove Classic), so we restrict our study to the three visible bands (blue, green, and red).

To prepare this dataset, we performed equalization of the mean and standard deviation for each band from each PlanetScope image to the corresponding Sentinel-2 image. The residual spatial shift between the downsampled PlanetScope

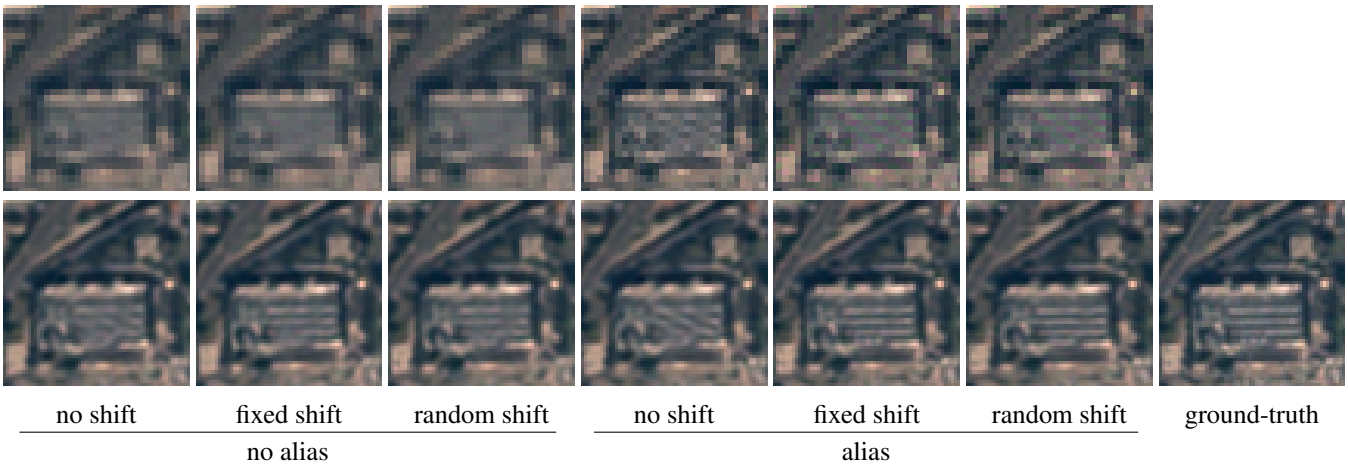


Fig. 3: Synthetic dataset. The top row shows simulated LR images corresponding to the six acquisition configurations, and the bottom row shows the SR results obtained with networks trained on these specific synthetic datasets. The bottom-right image is the ground-truth HR.

		PSNR (dB)		
		Test set	Train set	Val set
no alias	no shift	46.69	47.13	47.35
	fixed shift	47.20	47.38	47.57
	random shift	46.87	47.15	47.37
alias	no shift	46.67	47.29	47.51
	fixed shift	**49.30**	**49.25**	**49.54**
	random shift	<u>48.12</u>	<u>48.44</u>	<u>48.73</u>

Table 1: Shift and alias influence. Best PSNR in **bold** and second best <u>underlined</u>.

image and the Sentinel-2 one is estimated using the phase correlation algorithm. Then, the PlanetScope image is resampled using a third-order spline interpolation and introducing zeros where information was missing. Pairs with a phase correlation score below 0.55 were removed from the dataset.

We used 380 full-scene images and extracted up to 20 LR crops of size 200×200 from each image. The test set consists of 65 of these scenes (or 693 crops) and is geospatially disjoint from the train set (3680 crops). The validation set (406 crops) is selected from different dates. Figure 1 shows zoom-in crops from the dataset, where alias and inter-band shift are clearly visible in the S2 images.

5. EXPERIMENTS

This section presents experiments that empirically show that Sentinel-2 imagery is well-suited for the problem of SISR.

A surprisingly performant L_1 loss. We compare two models: one trained with the L_1 loss and one trained with the original ESRGAN loss, with the relativistic discriminator and feature similarity terms. Quantitatively, the average PSNR (over 12 bits) computed over the test set yields 42.21dB for the L_1 loss and 37.29dB for the ESRGAN loss. Visually, we observe that the results obtained with the L_1 loss are slightly smoother than those obtained with the ESRGAN loss, but do not contain any color artifacts. This can be seen in the second row of Figure 2. In addition, the details in the images generated by the L_1 model are much better than those obtained through bicubic interpolation. Overall, these experiments suggest that, in the case of Sentinel-2, the L_1 loss is an effective solution to increase the resolution by a factor 2 without risking the introduction hallucinated details [1].

Furthermore, as discussed in [6], GAN-like losses are particularly important when LR-HR image pairs are not well registered, which is not the case for our S2/PS dataset.

Additional results using the L_1 model on the test set are shown in Figure 1. We find that the network is able to resolve aliased patterns into high-frequency details, with strong fidelity to the ground-truth PlanetScope images. Given that the L_1 loss minimizes distortion [1], one can be confident that there are few hallucinated details. In ambiguous cases, the network will likely favor a blurry result instead of sharp, but potentially wrong details.

Cross-spectral information. We claim that our network exploits cross-spectral information to increase spatial resolution. To validate this hypothesis, we perform the following experiment: from the S2/PS dataset, we train three networks, each dedicated to super-resolving one specific spectral band, and only this band is given as input. On the test set, we observe a drop of 0.88dB in the PSNR, and we observe visually that the network is no longer able to resolve fine structures such as very high-frequency patterns. Even though the LR signal is aliased in each spectral band, the network no longer has the ability to perform a consistent, joint reconstruction of

the signal. This can be observed in the top-right image of Figure 2. A related observation was reported in [5] in which a network trained with both RGB and NIR bands performed better than just with RGB bands.

Aliasing and band-shift influence. We argue that the model described in Section 3 is able to exploit specific characteristics of the Sentinel-2 sensor, namely the presence of alias in each band and the inter-band shifts. The alias is due to a low spatial sampling with respect to the modulation transfer function (MTF) of the instrument [15], and the inter-band shifts originate from time delays between the acquisition of the lines of the different spectral bands [15]. Combined, these two aspects yield a configuration that is better-posed than standard SISR, and real information can be recovered under these acquisition specificities.

Next, we provide experimental evidence that the acquisition configuration of Sentinel-2 is indeed favorable to SISR. To this aim, we construct synthetic datasets using six different acquisition configurations: with and without alias, and with and without fix/random inter-band shifts. In each configuration, we use the PlanetScope images as ground-truth and we synthesize LR images according to each configuration. The presence of alias is controlled by the amount of blur introduced before downsampling. The shifts are +/-1 offsets applied to the bands before downsampling and then compensated by 0.5 offsets on the LR images. In each configuration, 0.1% Gaussian noise was added to match Sentinel-2 noise level. The first row of Figure 3 shows the effects of these configurations over the generated LR images. The configuration *with alias* and *random inter-band shift* is the most faithful simulation of Sentinel-2 imagery.

We train one network per scenario according to the same training details as in Section 5. Table 1 shows the PSNR (over 12 bits) over the train, validation and test sets for the different settings, and the bottom row of Figure 3 shows SR results. These results highlight that the combined presence of the inter-band shift and the alias allows the network to retrieve significant information from the signal. In contrast to the usual SISR scenario where little alias and no inter-band shift are present, our experiments assert that Sentinel-2 imagery is well-suited for SISR.

6. CONCLUSION

Our study aims to investigate the factors that enable SISR of Sentinel-2 imagery. While surprising at first, our extensive experiments on carefully designed synthetic datasets show that the resolution gain can be explained by the ability of the network to exploit the characteristics of Sentinel-2 imagery, namely alias and inter-band shift. We also validate our study by training a model for Sentinel-2 SR using a simple L_1 loss. Our model successfully increases the spatial resolution of Sentinel-2 images from 10m to 5m GSD, while minimizing distortion and avoiding the creation of false details.

7. REFERENCES

[1] Y. Blau and T. Michaeli, "The perception-distortion tradeoff," *CVPR*, 2018.

[2] C. Lanaras, J. Bioucas-Dias, S. Galliani, E. Baltsavias, and K. Schindler, "Super-resolution of sentinel-2 images: Learning a globally applicable deep neural network," *ISPRS*, 2018.

[3] M. Gargiulo, A. Mazza, R. Gaetano, G. Ruello, and G. Scarpa, "Fast super-resolution of 20 m sentinel-2 bands using convolutional neural networks," *Remote Sensing*, 2019.

[4] F. Pineda, V. Ayma, and C. Beltran, "A generative adversarial network approach for super-resolution of sentinel-2 satellite images," *ISPRS*, 2020.

[5] M. Galar, R. Sesma, C. Ayala, L. Albizua, and C. Aranda, "Super-resolution of sentinel-2 images using convolutional neural networks and real ground truth data," *Remote Sensing*, 2020.

[6] M. Zabalza and A. Bernardini, "Super-resolution of sentinel-2 images using a spectral attention mechanism," *Remote Sensing*, 2022.

[7] Julien Michel, Juan Vinasco-Salinas, Jordi Inglada, and Olivier Hagolle, "Sen2venμs, a dataset for the training of sentinel-2 super-resolution algorithms," *Data*, vol. 7, no. 7, pp. 96, 2022.

[8] L. Salgueiro Romero, J. Marcello, and V. Vilaplana, "Super-resolution of sentinel-2 imagery using generative adversarial networks," *Remote Sensing*, 2020.

[9] NL. Nguyen, J. Anger, A. Davy, P. Arias, and G. Facciolo, "Self-supervised multi-image super-resolution for push-frame satellite images," *CVPRW*, 2021.

[10] Wang X. et al., "Esrgan: Enhanced super-resolution generative adversarial networks," *ECCVW*, 2019.

[11] E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," *CVPRW*, 2017.

[12] A. Jolicoeur-Martineau, "The relativistic discriminator: a key element missing from standard gan," *arXiv*, 2018.

[13] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," *ECCV*, 2016.

[14] J. Anger, C. de Franchis, and G. Facciolo, "Assessing the sharpness of satellite images: Study of the planetscope constellation," *IGARSS*, 2019.

[15] F. Gascon et al., "Copernicus sentinel-2a calibration and products validation status," *Remote Sensing*, 2017.