ENHANCING THE SPATIAL RESOLUTION OF SENTINEL 3 SYNERGY THROUGH THE SUPER-RESOLUTION VIA REPEATED REFINEMENT METHOD

Maria Dekavalla[1], Dimitris Bliziotis[1], Valantis Tsiakos[1], Georgios Tsimiklis[1], Angelos Amditis[1]

[1]Institute of Communication and Computer Systems

ABSTRACT

Despite the abundance of open Earth Observation (EO) data from the Copernicus program and the GEOSS platform, their uptake in the context of Climate Change (CC) related applications is often limited due to inherent spatial and temporal resolution constraints. Super-resolution techniques aim to recover high spatial resolution images from degraded low spatial resolution images, but current approaches are being tested mainly for high spatial resolution imagery and small scaling factors. To fill this gap, this study investigates the ability of the diffusion-based image super-resolution via repeated refinement (SR3) method in enhancing the spatial resolution of the near-daily Sentinel-3 SYNERGY images from 300 to 75 meters by taking advantage of the higher spatial resolution of the Sentinel-2 MSI images. In quantitative and qualitative evaluations of results, SR3 provides convincing results and reveals its suitability in spatially enhancing Sentinel-3 SYNERGY images.

Index Terms— Sentinel-3 SYNERGY, super-resolution, large scaling factor, diffusion model

1. INTRODUCTION

The open EO data from the Copernicus program and GEOSS provide the required information to support Climate Change (CC) adaptation and mitigation policies. However, inherent spatial and temporal resolution constraints often limit their uptake in the context of CC-related applications. To address this issue, the H2020 EIFFEL[1] project will provide tools to enhance the spatial resolution of EO data from the Copernicus program and address the data requirements of five different CC adaptation and mitigation applications.

Two of the Copernicus missions that can support the envisioned CC-related applications are Sentinel-2 and Sentinel-3. They both carry multispectral sensors and are focused on global monitoring of the earth's surface. On the one hand, Sentinel-2 twin satellites include two identical satellites that carry the multispectral instrument (MSI), which provides a total of 13 spectral bands covering the wavelength region from 443 to 2190 nm of the electromagnetic spectrum, with a spatial resolution ranging from 10 to 60 m and a temporal resolution of 5 days. On the other hand, the Sentinel-3 mission satellites have a temporal resolution of 1.4 days and carry two multispectral radiometers on board: (i) the Ocean and Land Colour Imager (OLCI) and (ii) the Sea and Land Surface Temperature Radiometer (SLSTR). The OLCI captures the earth's surface using 21 bands in the spectral range between 390 and 1040 nm at 300 m spatial resolution. SLSTR provides 12 bands in the spectral range between 555 nm and 1200 nm at a spatial resolution of 500 m. ESA has also developed a synthetic product, named Sentinel-3 SYNERGY (SYN), which includes atmospherically corrected OLCI and SLSTR bands at 300 m spatial resolution [1]. However, such a spatial resolution is too coarse to provide sufficient detail for small extent areas of interest. While the 5-day revisit cycle of the fine spatial resolution Sentinel-2 and the cloud contamination can further degrade the temporal resolution and limit its application in detecting rapid surface changes crucial to some applications such as detecting intraseasonal ecosystem disturbance [2]. Therefore, there is a great need for data that have both the spatial resolution of Sentinel-2 and the temporal resolution of Sentinel-3 to support a wide range of monitoring applications.

There are several solutions that try to combine the fine spatial resolution of Sentinel-2 with the high temporal frequency of Sentinel-3 by taking advantage that both missions carry multispectral sensors and have similar wavelengths for four bands (i.e., blue, green, red, and NIR bands). It is possible to distinguish between two general approaches: (i) spatiotemporal data fusion and (ii) deep learning-based super-resolution. Spatiotemporal data fusion is a cost-effective way to combine the spatial information from high spatial resolution images with the temporal information from frequent but low spatial resolution images to generate images with high spatiotemporal resolution. Existing spatiotemporal data fusion algorithms can be divided into five categories: unmixing, weight function, Bayesian, deep-learning, and hybrid approaches. Although unmixing-based methods were proven to be more efficient in capturing gradual reflectance change and land cover type change in heterogeneous environments, they cannot capture rapid changes and small changes in land cover type which are invisible in the input low spatial resolution images [3]. The performance of the spatial-temporal fusion methods is based on the availability of fine spatial resolution images

[1] https://www.eiffel4climate.eu/

that are temporally close to the prediction date. Due to cloud contamination, sometimes very few effective Sentinel-2 images are available for fusion. Strong temporal changes from the time of the only available Sentinel-2 image to the prediction may lead to a small correlation of the only available Sentinel-2 with the ideal prediction of the Sentinel-2 image at the prediction time [4]. Moreover, spatiotemporal data fusion methods need to rebuild the model for each prediction, which is time-consuming [5]. When dealing with a large quantity of data, the computation time of a fusion algorithm can be extensive, greatly limiting its application in monitoring long-term and large-scale land surface dynamics. Whereas deep learning-based super-resolution methods can directly transfer trained models and consequently reduce the computational time.

Super-resolution techniques aim to enhance high spatial resolution images from degraded low spatial resolution images and have achieved great progress due to the recent advance of deep neural networks, including convolutional neural networks [6,7] and generative adversarial networks [8]. Yet the majority of the existing super-resolution techniques are simple regression-based methods with feedforward convolutional networks and are designed for augmenting the spatial resolution of high spatial resolution images (e.g. Sentinel-2 and Landsat 8/9) and for small scaling factors (x2, x3), where the spatial uncertainty of the super-resolution mapping can be compensated to some degree with the spatial details included at the low-resolution image [10]. Reconstruction of high spatial resolution images at large scaling factors (> x3) is more challenging due to the large spatial resolution gap and the severe information loss of the original medium/low spatial resolution images. In recent years, diffusion denoising probabilistic models (DDPM) have shown promising results in performing super-resolution tasks with large scaling factors on natural and face images [11]. However, they have not been systematically tested yet in EO images which include small features and complex scenes. To fill this gap, this study investigates the performance of the diffusion-based Super-Resolution via Repeated Refinement (SR3) model to enhance the 300 m spatial resolution of Sentinel-3 SYNERGY data to the target resolution of 75 meters by taking advantage of the higher spatial resolution of the Sentinel-2 MSI data.

2. SUPER-RESOLUTION VIA REPEATED REFINEMENT (SR3) MODEL

DDPMs are deep generative models that are trained to learn a data distribution using variational inference to produce samples matching the data after a finite time [12]. Their key idea is to first disrupt images by defining a forward Markovian diffusion process that gradually adds Gaussian noise to a high-resolution image over T iterations. And then training a U-Net architecture to remove various levels of noise from the noisy image. During inference, the U-Net is repurposed into a generative model by starting with Gaussian noise at the maximum level, then iteratively refining the noisy image by gradually attenuating noise and amplifying the image.

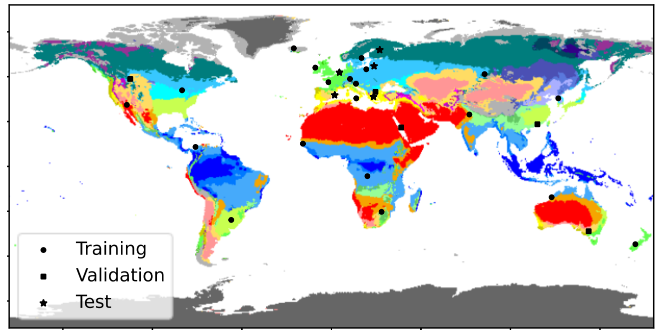

Fig. 1. Locations of Sentinel-2 and Sentinel-3 image pairs used for training, validation and testing. The locations are plotted over the Köppen-Geiger climate classification map [13].

SR3 [11] is an early application of DDPMs to super resolution conditional image generation by proposing a simple and effective modification to the U-Net architecture. Specifically, the SR3 model includes a U-Net architecture like the one found in DDPM but the residual blocks were replaced by residual blocks from BigGAN [14] and the skip connections were rescaled. The number of residual blocks and the channel multipliers at different resolutions were also increased. The low-resolution image is resampled to the target resolution and concatenated with a normal distribution noisy image to condition the model to create images that have the same distribution as the initial low-resolution image.

3. EXPERIMENTS

We evaluated the performance of SR3 model in enhancing the spatial resolution of Sentinel-3 images. For this purpose, we employed Sentinel-2 images as the input high-resolution images required for the forward diffusion process, and Sentinel-3 images as the low-resolution images employed as input in the reverse (denoising) process.

Effectively exploiting such inter-sensor synergies raises important challenges in terms of sensor alignment, and substantial spatial and spectral resolution differences [3, 10]. Each pair of Sentinel-2 and Sentinel-3 images must be atmospherically corrected, projected to a common coordinate system and co-registered to reduce the geometric errors due to the large spatial resolution difference between them. Considering that atmospherically corrected Sentinel-3 OLCI products are still not available, Sentinel-3 SYNERGY images were used instead, which include atmospherically corrected bands acquired from both OLCI and SLSTR sensors.

The training dataset includes pairs of Sentinel-2 and Sentinel-3 SYNERGY Level-2 images that were acquired on the same day and cover the same areas. The considered areas were randomly picked, aiming to include different types of land covers and climate zones to obtain data diversity. Fig. 1 shows the distribution of the images composing the dataset. Specifically, 14 image pairs were used for training and 5 image pairs for validation and testing, respectively.

The selected image pairs were obtained from the Copernicus Open Access Hub. Specifically, only the bands Oa4, Oa6, Oa8, and Oa17 bands of Sentinel-3 SYNERGY and the bands 2, 3, 4, and 8a of Sentinel-2 which have similar wavelengths were included in the training datasets. The Sentinel-3 images were reprojected onto the coordinate system of the corresponding Sentinel-2 counterparts to obtain the intersection between both products as output. Then both Sentinel-2 and Sentinel-3 images were resampled to 75 m with bilinear interpolation. And they were also co-registered using the Sentinel-2 images as spatial reference with the AROSICS inter-sensor registration method [15].

Each processed Sentinel-2 and Sentinel-3 image was split into 1000 random image patches of 128x128 pixels. To simplify training, only tiles that do not contain background, clouds and water surface pixels were selected for the training dataset. Considering that remote sensing images have more complex distributions and contain multiple categories of ground objects than natural images, it was decided to first train the model with the global training dataset and then retrain only with a smaller training dataset containing Sentinel-3 and Sentinel-2 images acquired from the same locations as the ones contained in the test dataset.

The model was trained with PyTorch for 1 million training steps and a batch size of 32. The Adam optimizer with a linear warmup schedule over 10k training steps, followed by a fixed learning rate of 1e-4 was used to be consistent with the initial implementation and demonstration of the diffusion models. Training was run on a NVIDIA GeForce RTX 3090 GPU, with 16 GB of RAM for approximately 5 days. For numerical stability, we divided the raw reflectance values of Sentinel-2 by 25000 before training.

4. RESULTS

Fig. 2 demonstrates 5 test Sentinel-3 natural color composite images that were super-resolved with the model trained during the first training phase with the global training dataset and the second training phase with a similar to the test images training dataset. The Sentinel-2 and Sentinel-3 images resampled to 75 meters spatial resolution are also included for comparison. These test image pairs were also used to perform the quality assessment of the DL-based super-resolution for Sentinel-3 SYNERGY images (Table 1).

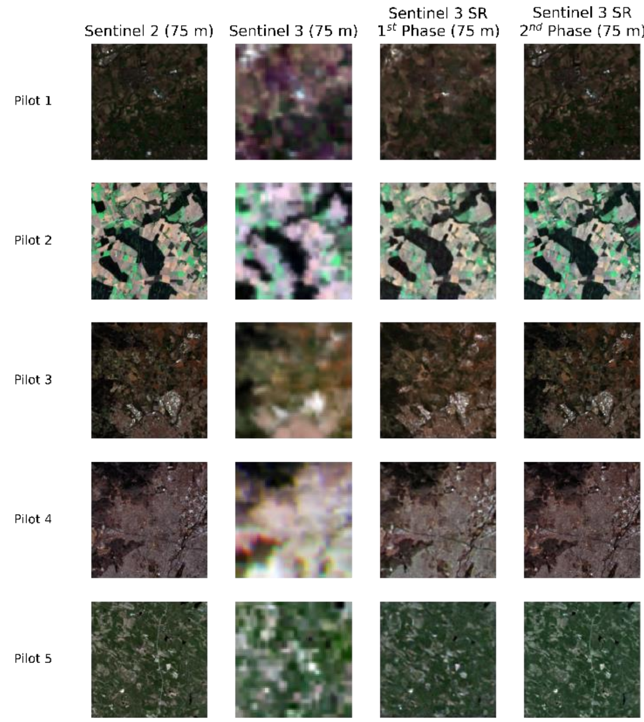

Fig. 2. Qualitative results for 5 test areas

The qualitative results demonstrated in Fig.2 show that both models trained with the global and the test-specific dataset can produce high-quality super-resolved Sentinel-3 images. The second training of the model with the test-specific dataset contributes to the generation of super-resolved images that resemble the reference Sentinel-2 images and achieved better quantitative results than the first training dataset (Table 1) that can be relatively accepted in most cases (i.e., ERGAS < 3 and SAM ≈ 0).

5. CONCLUSION & FUTURE WORK

This study investigated the performance of the novel diffusion-based super-resolution SR3 model in enhancing the spatial resolution of Sentinel-3 SYNERGY images. The model was first trained with a training dataset including randomly picked Sentinel-3 and Sentinel-2 images and then further trained with images similar to the test images. The two trained models produced qualitative consistent results. But the test-specific model achieved also acceptable scores in reference-based metrics (i.e., SAM ≈ 0, ERGAS < 3) and produced images with a higher level of detail. This can be explained by the fact that diffusion models create images that have the same distribution as the images they are trained on.

Although results are encouraging, further research is required to achieve additional improvements. Specifically, our future work will be directed towards investigating the performance of SR3 with bigger training datasets and for bigger scaling factors than 4, and modifying the model to achieve both qualitative and quantitative consistent results when super-resolving images that are out of the distribution of the training dataset.

6. ACKNOWLEDGEMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101003518.

Table 1. Quality assessment results

	Pilot 1		Pilot 2		Pilot 3		Pilot 4		Pilot 5	
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
SAM	0.13	0.03	0.05	0.03	0.14	0.03	0.06	0.02	0.18	0.07
SSIM	0.87	0.99	0.97	0.99	0.86	0.99	0.95	0.99	0.98	0.99
ERGAS	6.62	3.40	4.41	2.88	6.02	2.99	4.73	2.88	12.44	7.79
PSNR	36.19	49.44	44.01	48.93	36.85	49.86	40.92	49.04	47.11	52.82
RMSE	387.52	84.24	157.40	89.39	359.21	80.25	224.77	88.22	110.18	57.13

REFERENCES

[1] C. Henocq et al., "Olci/SLSTR SYN L2 algorithm and Products Overview," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018. doi:10.1109/igarss.2018.8517420

[2] F. Gao, J. Masek, M. Schwaller, and F. Hall, "On the blending of the landsat and MODIS surface reflectance: Predicting daily landsat surface reflectance," IEEE Transactions on Geoscience and Remote Sensing, vol. 44, no. 8, pp. 2207–2218, Aug. 2006, doi: 10.1109/TGRS.2006.872081.

[3] X. Zhu, E. H. Helmer, F. Gao, D. Liu, J. Chen, and M. A. Lefsky, "A flexible spatiotemporal method for fusing satellite images with different resolutions," Remote Sens Environ, vol. 172, pp. 165–177, Jan. 2016, doi: 10.1016/j.rse.2015.11.016.

[4] Q. Wang and P. M. Atkinson, "Spatio-temporal fusion for daily Sentinel-2 images," Remote Sens Environ, vol. 204, pp. 31–42, Jan. 2018, doi: 10.1016/j.rse.2017.10.046.

[5] H. Gao et al., "CuFSDAF: An Enhanced Flexible Spatiotemporal Data Fusion Algorithm Parallelized Using Graphics Processing Units," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, 2022, doi: 10.1109/TGRS.2021.3080384.

[6] C. Tuna, G. Unal, and E. Sertel, "Single-frame super resolution of remote-sensing images by convolutional neural networks," Int J Remote Sens, vol. 39, no. 8, pp. 2463–2479, Apr. 2018, doi: 10.1080/01431161.2018.1425561.

[7] L. Liebel and M. Körner, "Single-image super resolution for multispectral remote sensing data using convolutional neural networks," in International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, International Society for Photogrammetry and Remote Sensing, 2016, pp. 883–890. doi: 10.5194/isprsarchives-XLI-B3-883-2016.

[8] K. Jiang, Z. Wang, P. Yi, G. Wang, T. Lu and J. Jiang, "Edge-Enhanced GAN for Remote Sensing Image Superresolution," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 8, pp. 5799-5812, Aug. 2019,doi:10.1109/TGRS.2019.2902431

[9] M. R. UI Hoque, R. Burks, C. Kwan and J. Li, "Deep Learning for Remote Sensing Image Super-Resolution," 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 2019, pp. 0286-0292, doi: 10.1109/UEMCON47517.2019.8993047

[10] R. Fernandez, R. Fernandez-Beltran, J. Kang and F. Pla, "Sentinel-3 Super-Resolution Based on Dense Multireceptive Channel Attention," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 7359-7372, 2021, doi: 10.1109/JSTARS.2021.3097410.

[11] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, "Image Super-Resolution Via Iterative Refinement," IEEE Trans Pattern Anal Mach Intell, 2022, doi: 10.1109/TPAMI.2022.3204461.

[12] J. Ho, A. Jain, P. Abbeel, "Denoising diffusion probabilistic models," Advances in Neural Information Processing Systems, 33:6840–6851, 2020.

[13] H. Beck, N. Zimmermann, T. McVicar, et al., "Present and future Köppen-Geiger climate classification maps at 1-km resolution," Scientific Data, vol. 5, 180214, 2018, doi:10.1038/sdata.2018.214

[14] A. Brock, J. Donahue, K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," arXiv preprint arXiv:1809.11096, 2018.

[15] D. Scheffler, A. Hollstein, H. Diedrich, K. Segl, and P. Hostert, "AROSICS: An Automated and Robust Open-Source Image Co-Registration Software for Multi-Sensor Satellite Data," Remote Sensing, vol. 9, no. 7, p. 676, Jul. 2017, doi: 10.3390/rs9070676.