# SUPER-RESOLVING WASTE DUMPS FROM SPACE WITH DEEP LEARNING: ROMANIA REGION USING SENTINEL-2 AND SPOT6/7

*Teodora Selea, George Boldeanu, Fredrik Samujel Nistor, Mihaela Violeta Gheorghe*

GMV Innovating Solutions Romania

## ABSTRACT

An increase in the total quantity of waste produced on a global scale has severe repercussions for the ecosystem, including health risks for individuals and environmental damage. This increase is directly correlated with urban development in recent decades, making all member states of the European Union (EU) responsible for complying with waste management regulations. The detection of waste dumps represents the essential component of waste management, and satellite data provide the ability to monitor it. However, expensive, very high-resolution images are needed for proper identification. This paper proposes a super-resolution (SR) workflow to increase the readability of low-resolution but accessible satellite data (Sentinel-2). We assess the workflow for the specific use case of waste dump detection in Romania (including seven major cities). We also analyze several dataset pre-processing techniques and 18 popular Deep Learning (DL) models, providing valuable insight into super-resolution applications for waste management.

***Index Terms—*** super-resolution, waste management, deep learning, Sentinel-2, Spot 6/7.

## 1. INTRODUCTION

The increased quantity of waste has become a prominent concern at a global level. One reason waste accumulates is rising income in industrialized countries, which leads to increasing product consumption. Waste management was addressed at the EU level to reduce municipal waste to 10% at most. Therefore, action should be taken in conjunction with each municipality. Romania faces similar challenges as other EU countries, as waste detection on site is laborious or even inaccessible in some regions.

Earth Observation (EO) data offers a viable solution for waste detection due to its temporal and spatial resolution. Several research studies applied Artificial Intelligence (AI) techniques for waste detection, with good results obtained from high-resolution data. Illegal waste is the most difficult type of garbage dump to identify due to its small area and irregular shape. Nevertheless, it is a priority to detect to prevent further damage created by it.

However, the high cost and low temporal resolution of commercial high-spatial-resolution satellite imagery have made it challenging to develop and implement an operational waste identification/monitoring service based on satellite imagery. Therefore, generating synthetic high-resolution images using super-resolution techniques solves this issue. By improving the visual experience while decreasing pixelation, these models provide experts with actionable data, allowing them to distinguish more minor features on the ground and contributing to more effective decision-making.

The goal of this paper is to provide support for the actual task of waste detection, by super-resolving the images. We created the dataset around the use-case of interest, and it incorporates satellite images of urban periphery and agricultural areas, that include areas with waste dumps. Next, we conducted an analysis and established a benchmark of super-resolution methods applied to satellite data, focusing on the performance of the super-resolved waste areas. However, we argue that the results obtained in this paper may also be applicable in other contexts, with similar particularities of the data.

Even though numerous approaches have been proposed in the super-resolution field (especially in Computer Vision), the remote sensing (RS) community still lacks benchmark datasets and results, which are both greatly needed to assess the efficacy of super-resolution models on EO data. The current use-case alongside its data particularities, is not represented at the moment, to the best of our knowledge. The multispectral and temporal characteristics of EO data raise additional challenges compared to classic Red, Green, Blue (RGB) data. Therefore, specific guidelines and a baseline must be established to progress in the EO super-resolution domain. The main contributions of this paper are as follows:

1. We detail the steps of creating a regional dataset for SR, including two sensors, making it easy to reproduce or extend in any other area.
2. We assess the transformations needed to be applied on the input pair of LR-HR to optimise the performance of our models, resulting in valuable knowledge useful in similar super-resolution applications.
3. We provide a benchmark of 18 popular DL models for SR and create the first and most comprehensible benchmark on super-resolution using multisensor data.
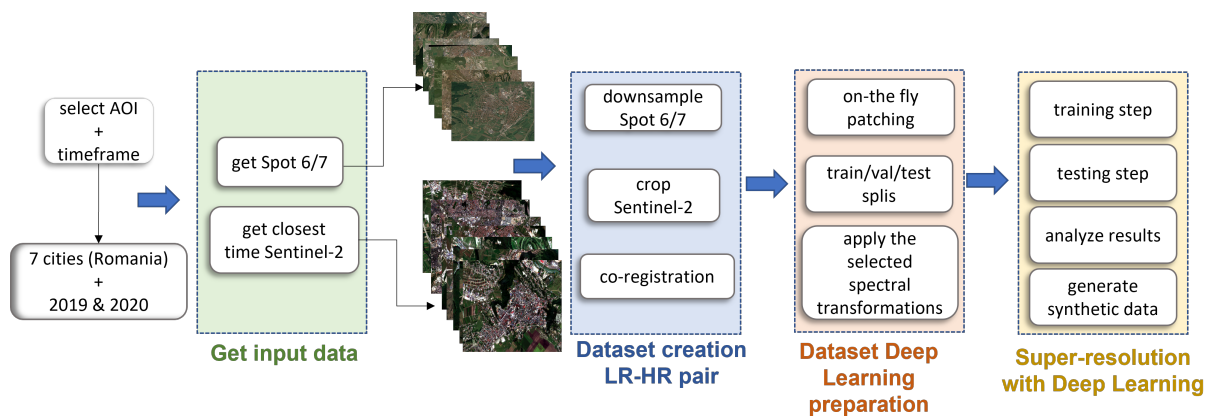
**Fig. 1**: Proposed workflow for applying super-resolution on Romania cities

4. We analyse the capabilities of the models to learn colour mappings across sensors and to super-resolve the input image for Romania and illegal waste dumps.

## 2. SUPER-RESOLUTION DATA

As with other DL methods, super-resolution algorithms require many sample data to train. Worldstrat[3] and MuS2[4] are the largest two-sensor datasets for SR that permit an up-sampling factor of 4. Both datasets use Sentinel-2 as their LR image but differ in the choice of HR data. Worldstrat uses Spot 6/7, and MuS2 is based on WordView-2. However, the gap for a super-resolution dataset remains, as Worldstrat and MuS2 are intended for noncommercial use only. Also, to our knowledge, there is no open-source dataset for a specific super-resolution use case ( e.g. waste detection) or a particular region (e.g. Romania). Therefore, super-resolution on satellite data remains a challenging task. The following sections detail how we created our two sensors dataset for Romania, focusing on illegal waste dumps.

### 2.1. Study Area

In this study, we aim to super-resolve the illegal waste dumps around the cities in Romania. The proposed study region includes seven cities in Romania (Alba Iulia, Brasov, Bucharest, Deva, Sibiu, Sighisoara and Suceava). As Spot-6/7 has no fixed rate of revisitation, we acquired a Spot-6/7 image for each city in 2020 to meet our requirements. Next, we downloaded a corresponding Sentinel-2 L2A tile for each HR image, with the closest acquisition time, to create our super-resolution dataset pairs of LR and HR.

To our knowledge, there are no open datasets for illegal waste dumps. Our experts have manually identified illegal waste dumps in two cities (Alba Iulia and Bucharest). In our experiments, we include all seven cities, as we argue that it provides additional training information, which improves the final performance of super-resolution.

### 2.2. Dataset Creation

Creating pairs of LR/HR images is challenging due to the different acquisition sensors. Most SR applications follow a downsampling technique applied to HR to obtain the LR image. This technique is applied in the Computer Vision (CV) and RS domains. Given the constraints of increased availability of LR input data instead of HR images, we proceed with the use-case of SR across sensors. The challenges in preprocessing the two data are 1)time matching; 2) spatial extend; 3) co-registration; 4) spectral differences. In the workflow proposed in Figure 1, we present the main steps to create pairs of LR / HR images, that also incorporate the areas with waste dumps. First, we applied pansharpening on the Spot 6/7 tile to obtain a 4-band (Red, Green, Blue, NIR) image at 1.5 m resolution. Second, we downsample the Spot data to 2.5 m spatial resolution. Doing so achieves the desired upsampling factor of 4 between the Sentinel-2 tile at 10 m and Spot 6/7. The Spot 6/7 image covers a much smaller area than the Sentinel-2 tile. Therefore, we cropped the Sentinel-2 images to the same spatial extent as their corresponding HR. Finally, since different sensors acquire the LR and HR images, we co-register each pair to align pixels properly. We used Sentinel-2 as the reference image for each pair because the Spot image is absent during the inference stage.

### 2.3. Data Pre-processing for Deep Learning

**Patching & Splits Creation**: Patching of the input data is commonly used in CV, especially when working with DL models, due to hardware memory limitations. For the experiments carried out in this paper, we patch the original data into $32 \times 32$-pixel size (LR image) and $128 \times 128$-pixel size (HR image) without overlap. We randomly split the patches into training, validation, and testing, keeping a proportion of 80%-10%-10% from the total patches for each city. By doing so, we ensure to capture the characteristics of every city in all the splits. Next, we redistribute the patches containing the

illegal waste dumps (Alba Iulia and Bucharest) to ensure they are present in all the splits (Table 1).

| City | No. train patches | No. val patches | No. test patches |
|---|---|---|---|
| Alba Iulia | 244 (1) | 31 | 31 (2) |
| Brasov | 307 | 38 | 39 |
| Bucharest | 1120 (6) | 140 (1) | 140 (1) |
| Deva | 156 | 20 | 20 |
| Sibiu | 374 | 47 | 47 |
| Sighisoara | 72 | 9 | 10 |
| Suceava | 153 | 19 | 20 |
| Total | 2426 (7) | 304 (1) | 307 (2) |

**Table 1**: Patches distribution across splits (train, validation and test) for each city. The number in parentheses represents the number of illegal dumps for each split.

**Spectral Transformations**: Our LR and HR images have different acquisition sensors, which results in a different band number with a distinct value range, even for the common bands. The SR model learns the mapping between the LR-HR images, including colour mapping and super-resolve from LR-HR. We have applied a normalisation step in all proposed experiments to help the models converge faster and achieve better results. Although standardisation or min/max normalisation are popular choices in CV, we apply linear normalisation using the 1st and 99th percentiles, with a range [0,1]. By doing so, the normalisation becomes less sensitive to the outliers present in satellite data. Next, we apply a histogram matching transformation to calibrate HR (Spot 6/7) to LR (Sentinel-2). Figure 2 illustrates the range of values for the Red channel and corresponding changes.

## 3. METHODS

### 3.1. Deep Learning for Super-Resolution

Super-resolution datasets are made up of pairs of a LR image and an HR image. Applying super-resolution on RS data usually involves a technique widely used in Computer Vision: synthetically creating an LR image by downsampling the HR image. However, this method is challenging to apply in real-world scenarios where LR data is primarily available and accessible (such as Sentinel-2), with missing HR data (such as Spot 6/7). Therefore, the deep learning network must learn how to super-resolve the LR image and the colour mapping between LR and HR.
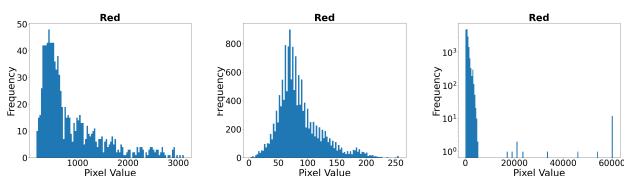


**Fig. 2**: Histogram from one training sample patch: a) LR b) HR c) HR after histogram matching (log scale)

Recent studies focused on assembling the advances in SR for hyper-spectral data [2], which shows the need to establish a benchmark for super-resolution in RS. However, multispectral data, such as Sentinel-2, have different properties, given fewer bands. Moreover, the previous reviews did not account for multi-sensor LR-HR. In the proposed experiments, we pair two different sensors (Sentinel-2 and Spot 6/7); therefore, the task becomes more challenging.
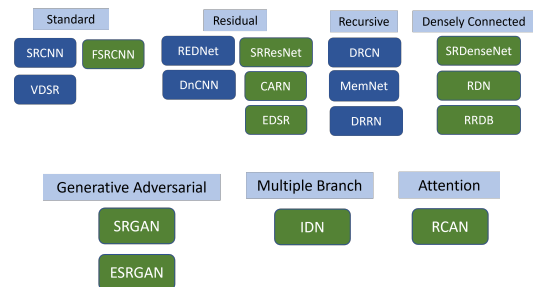


**Fig. 3**: Analysed models grouped by categories as in [1]

This paper proposes the first benchmark experiments on SR using 18 distinct models for single image super-resolution (SISR) applied to multi-sensor remote sensing data. Models differ by their main characteristics and are grouped into seven different categories (Figure 3), as in [1]. We include models from both categories of input handling: LR is first upsampled and then fed into the network or directly using the LR as input. The novelty of this work also comes from the experimental setting, which includes three types of channel combinations using the four common bands: Red (R), Green (G), Blue (B), Nir (NIR). First, we experiment with each channel individually to identify the possible problematic bands. This experiment is also based on the idea that most models were designed to work with one band (the Y channel from classic RGB images). Second, we experiment with the corresponding RGB channels and third, with the RGBNIR combination to assess the model's performance when the input consists of multiple bands. To our knowledge, this paper presents the first experimental evaluation of the necessity of including sensor calibration (e.g. histogram matching) as preprocessing.

## 4. RESULTS

Table 2 presents the results of the proposed experiments, including the 18 models, the three-channel-type combinations, and without or with histogram matching as pre-processing step. We present the results from both a visual perspective and with popular SR metrics (PSNR and SSIM). We argue that the results obtained in this study may be further applied in illegal waste dumps or generic super-resolution as we propose a visual and metric baseline.

Single-band experiments reveal that the NIR band is problematic for all models, achieving the lowest score among all bands. This indicates a larger difference between the repre-

| Model | Train param | Red lin norm | | Red hst+lin norm | | Green lin norm | | Green hst+lin norm | | Blue lin norm | | Blue hst+lin norm | | NIR lin norm | | NIR hst+lin norm | | True Color lin norm | | True Color hst+lin norm | | True Color NIR lin norm | | True Color NIR hst+lin norm | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |
| SRCNN | 57.3 K | 18.95 | 0.4337 | 19.9 | 0.4051 | 18.01 | 0.3572 | 19.06 | 0.3402 | 16.97 | 0.2499 | 19.31 | 0.2752 | 16.72 | 0.3873 | 14.57 | 0.2841 | 18.52 | 0.3721 | 20 | 0.3623 | 17.97 | 0.3736 | 16.84 | 0.3347 |
| FSRCNN | 12.8 K | 18.46 | 0.3962 | 19.49 | 0.3825 | 17.7 | 0.3305 | 18.92 | 0.3237 | 16.69 | 0.2237 | 18.86 | 0.2619 | 15.73 | 0.3214 | 13.92 | 0.2462 | 17.89 | 0.331 | 19.68 | 0.3404 | 17.17 | 0.3156 | 15.64 | 0.3051 |
| REDNet | 1.0 M | 18.95 | 0.4314 | 19.88 | 0.4072 | 17.82 | 0.3524 | 19.19 | 0.3464 | 16.99 | 0.2468 | 19.22 | 0.2735 | 16.68 | 0.3842 | 15.27 | 0.2878 | 18.33 | 0.3554 | 20.05 | 0.362 | 17.42 | 0.3627 | 15.8 | 0.3205 |
| VDSR | 664 K | 17.42 | 0.4028 | 19.44 | 0.3848 | 16.04 | 0.3378 | 18.29 | 0.3204 | 15.88 | 0.2306 | 18.99 | 0.2596 | 15.64 | 0.384 | 14.3 | 0.2837 | 18.29 | 0.3633 | 19.992 | 0.3598 | 17.57 | 0.3679 | 16.27 | 0.3319 |
| DRCN | 423 M | 16.56 | 0.2763 | 18.42 | 0.3129 | 16.8 | 0.2474 | 18.08 | 0.2733 | 16.13 | 0.1702 | 18.96 | 0.2311 | 14.88 | 0.2421 | 14.98 | 0.2169 | 16.91 | 0.2462 | 19.0013 | 0.3041 | 16.21 | 0.2426 | 17.29 | 0.3011 |
| DRRN | 297 K | 17.28 | 0.39 | 19.33 | 0.3881 | 15.96 | 15.96 | 20.53 | 0.1695 | 15.73 | 0.2251 | 19.03 | 0.2681 | 15.78 | 0.3794 | 13.19 | 0.2755 | 18.01 | 0.3529 | 21.8827 | 0.1716 | 6.92 | 0.0474 | 16.59 | 0.1705 |
| SRResNet | 1.5 M | 18.93 | 0.4294 | 19.79 | 0.383 | 18.1 | 0.3582 | 18.87 | 0.3261 | 17.19 | 0.2497 | 19.33 | 0.27 | 17.13 | 0.3793 | 15.26 | 0.2703 | 18.68 | 0.3577 | 20.039 | 0.3523 | 18.08 | 0.355 | 17.66 | 0.3408 |
| SRGAN | 6.7 M | 19.33 | 0.4386 | 19.92 | 0.3989 | 18.46 | 0.3635 | 18.72 | 0.328 | 17.38 | 0.2592 | 18.72 | 0.272 | 17.3 | 0.3812 | 16.34 | 0.288 | 18.78 | 0.3706 | 20.288 | 0.36218 | 18.17 | 0.3733 | 17.72 | 0.3368 |
| DnCNN | 557 K | 19.19 | 0.4356 | 20.22 | 0.3954 | 18.27 | 0.3638 | 18.19 | 0.3355 | 17.25 | 0.2568 | 19.28 | 0.2795 | 13.93 | 0.3827 | 15.52 | 0.2918 | 18.81 | 0.3656 | 20.177 | 0.36 | 18 | 0.3628 | 16.4 | 0.3508 |
| EDSR | 43 M | 19.22 | 0.442 | 20.11 | 0.4091 | 18.46 | 0.3723 | 19.25 | 0.3441 | 17.41 | 0.2627 | 19.44 | 0.274 | 16.54 | 0.392 | 15.89 | 0.304 | 18.94 | 0.3783 | 20.5562 | 0.3786 | 18.53 | 0.3888 | 18.31 | 0.38 |
| MemNet | 2.9 M | 19.09 | 0.4348 | 19.69 | 0.391 | 18.14 | 0.3657 | 19 | 0.3404 | 17.12 | 0.2584 | 19.23 | 0.2756 | 17.08 | 0.3816 | 15.45 | 0.2916 | 18.59 | 0.366 | 20.128 | 0.3625 | 18.09 | 0.3727 | 17.92 | 0.3729 |
| IDN | 552 K | 19.11 | 0.423 | 19.77 | 0.386 | 18.21 | 0.3456 | 18.86 | 0.309 | 17.13 | 0.2398 | 21.73 | 0.171 | 16.76 | 0.3146 | 15.38 | 0.2516 | 18.64 | 0.345 | 19.878 | 10.34 | 17.88 | 0.3384 | 17.09 | 0.31 |
| CARN | 1.1 M | 19.08 | 0.4269 | 19.55 | 0.395 | 18.26 | 0.3563 | 19 | 0.3316 | 17.25 | 0.2536 | 19.32 | 0.269 | 17.08 | 0.3733 | 15.46 | 0.2702 | 18.64 | 0.3521 | 20.0342 | 0.3475 | 17.9 | 0.3383 | 15.14 | 0.2377 |
| SRDenseNet | 2.7 M | 19.26 | 0.4413 | 20.23 | 0.4148 | 18.31 | 0.3707 | 20.23 | 0.3428 | 16.19 | 0.2521 | 19.39 | 0.2718 | 16.59 | 0.3857 | 15.2 | 0.281 | 18.82 | 0.3717 | 20.2104 | 0.3611 | 18.41 | 0.3758 | 16.67 | 0.3352 |
| RDN | 22 M | 19.3 | 0.4401 | 20 | 0.4118 | 18.33 | 0.3739 | 19.24 | 0.3416 | 17.45 | 0.263 | 19.48 | 0.2751 | 17.07 | 0.3897 | 16.23 | 0.304 | 18.93 | 0.3759 | 20.218 | 0.36478 | 18.4 | 0.3855 | 18.21 | 0.373 |
| RRDB | 16 M | 18.96 | 0.4383 | 19.94 | 0.4052 | 18.37 | 0.3672 | 19.18 | 0.3426 | 17.32 | 0.257 | 19.18 | 0.2757 | 15.43 | 0.3706 | 15.86 | 0.2954 | 18.77 | 0.3714 | 20.0883 | 20.0883 | 18.18 | 0.3791 | 17.88 | 0.36977 |
| ESRGAN | 31 M | 19.28 | 0.4427 | 19.9 | 0.409 | 18.33 | 0.3768 | 18.85 | 0.342 | 17.35 | 0.2637 | 19.417 | 0.28 | 17.13 | 0.3897 | 16.73 | 0.31 | 18.87 | 0.3824 | 20.191 | 0.3672 | 18.5 | 0.3888 | 18.12 | 0.37264 |
| RCAN | 15.9 M | 19.17 | 0.4365 | 19.43 | 0.38 | 18.42 | 0.3608 | 19.09 | 0.334 | 17.33 | 0.2603 | 19.45 | 0.2727 | 17.1 | 0.3832 | 17.5 | 0.3147 | 18.86 | 0.372 | 20.76 | 0.3634 | 18.47 | 0.3771 | 20.44 | 0.3728 |

**Table 2**: Experimental results after 20 epochs of training

sentation of data on these two bands between the LR Sentinel-2 and the HR Spot 6/7. A considerable increase in score was obtained for both NIR and blue bands after histogram matching. However, the NIR band still has a low score, and therefore should be further investigated. The RGB (True Color) channel experiments achieved the best results, with a considerable increase in score when histogram matching was applied. Even though on RGBNIR the models achieve a lower score than on RGB, probably due to the NIR channel, it is still an increase that the NIR-only experiments.
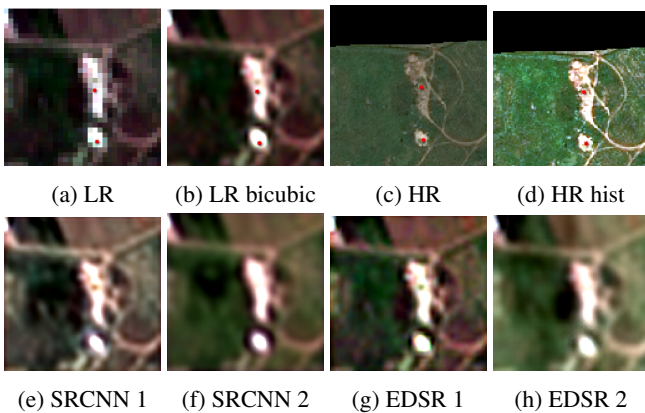
| (a) LR | (b) LR bicubic | (c) HR | (d) HR hist |
| --- | --- | --- | --- |
| (e) SRCNN 1 | (f) SRCNN 2 | (g) EDSR 1 | (h) EDSR 2 |

**Fig. 4**: Visual comparison on a testing patch

Figure 4 visually compares our experiments on a test patch. In the first row, we represent the input LR patch (Figures 4a, 4b) together with the original HR (Figure 4c) and its corresponding histogram-matched version (Figure 4d). The illegal dumps are represented by red stars. We observe the difference between the LR and HR colours and the changes in HR after matching it with LR. We illustrate the prediction of two popular models with different input requirements. SRCNN - LR bicubic, EDSR - LR. We denote with 1 the experiments with linear normalisation and 2 with additional histogram matching. As visible in the patches, histogram normalisation enables the models to learn a smoother texture, the results being visible closer to the HR. However, we argue that even the simplest models (SRCNN) can learn a good colour mapping between the LR and HR images.

## 5. CONCLUSIONS

In this paper, we propose an extensive benchmark of 18 SR models applied on dual-sensor data. Sentinel-2 and Spot 6/7, with various pre-processing techniques. We argue that the synthetic data generated in this paper may be further used for the waste detection use-case, as input for further semantic segmentation. In addition to this, the results serve as a benchmark for further development and revealed problematic bands, such as NIR, which require further techniques development.

## 6. ACKNOWLEDGEMENTS

## REFERENCES

[1] Saeed Anwar, Salman Khan, and Nick Barnes. A deep journey into super-resolution: A survey. *ACM Computing Surveys (CSUR)*, 53(3):1–34, 2020.

[2] C. Chen, Y. Wang, N. Zhang, Y. Zhang, and Z. Zhao. A review of hyperspectral image super-resolution based on deep learning. *Remote Sensing*, 15(11), 2023.

[3] Julien Cornebise, Ivan Oršolić, and Freddie Kalaitzis. Open high-resolution satellite imagery: The world-strat dataset–with application to super-resolution. *arXiv preprint arXiv:2207.06418*, 2022.

[4] Pawel Kowaleczko, Tomasz Tarasiewicz, Maciej Ziaja, Daniel Kostrzewa, Jakub Nalepa, Przemyslaw Rokita, and Michal Kawulok. Mus2: A real-world benchmark for sentinel-2 multi-image super-resolution, 2022.