Super-Resolution for Agriculture EO Services Project

Sara Verbič, Devis Peressutti, Sinergise Solutions d.o.o., a Planet Labs company.
Petra Recko Novak, Jure Gnilšek, Kristina Jajtič, Geodetski Zavod Celje d.o.o.
Zoltan Bartalis, European Space Agency

Aim

The overall objective of the SR4AGRI project is to enhance the 20 m and 60 m bands of Sentinel-2 to 10 m pixel size to improve the Area Monitoring System (AMS), where Sentinel-2 data is providing a basis for continuous monitoring of agriculture, forestry, other vegetation and urban areas for the purpose of ensuring sustainability of the ecosystems and assessing the carbon footprint of human's activities.

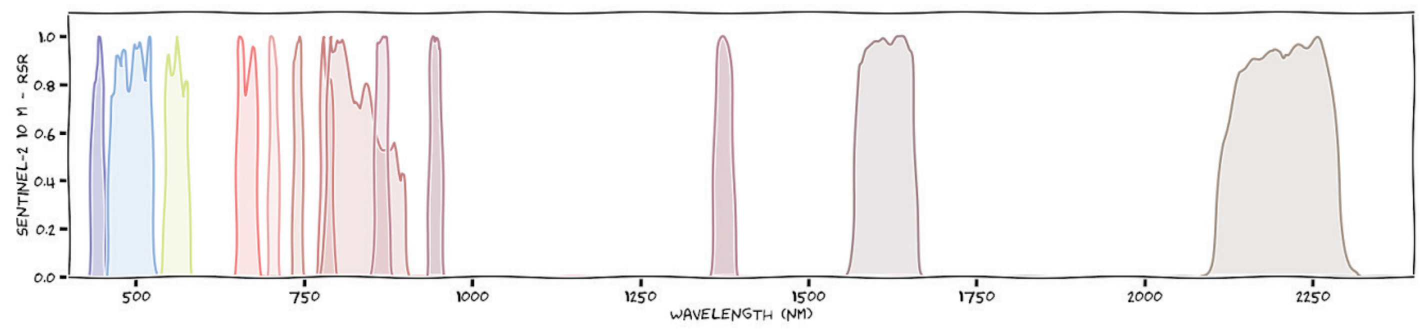

Figure 1. The outcome of SR4AGRI is to provide all Sentinel-2 bands at 10 m/pix.

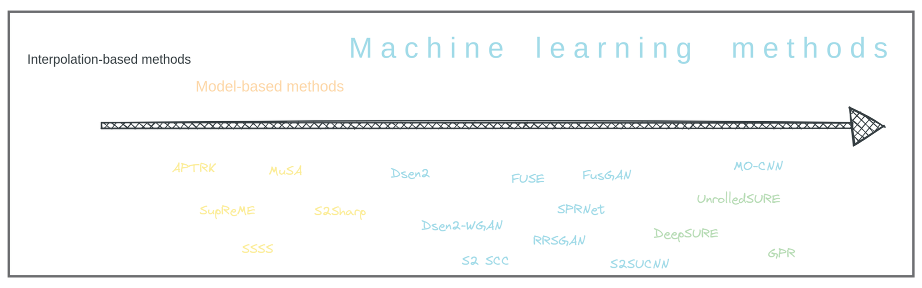

Figure 2. Reviewed methods for SSF of Sentinel-2 imagery. Methods shown in green are hybrid methods between model and ML-based methods.

Spatio-Spectral Fusion: SOTA

Problem specification

Super-resolution methods aim to artificially enhance the (spatial) resolution of an acquired image. Mathematically speaking, we denote with $I_{LR} \in \Re^{HxWxC}$ va low-resolution image of height H, width W, and with C channels, and with $I_{HR} \in \Re^{kHxkWxC}$ the corresponding high-resolution image with an improved resolution given by the scaling factor k. The aim of (single-image) super-resolution is to estimate a nonlinear function $f: \Re^{HxWxC} \rightarrow \Re^{kHxkWxC}$ that provides an estimate of I_{HR} starting from I_{LR}, as shown in eq. 1.

$$\hat{I}_{HR} = f(I_{LR}). \tag{1}$$

The most commonly used imaging model that describes the degradation process that generates a LR version of an HR image is shown below.

$$I_{LR} = (I_{HR} \otimes b)_{\downarrow k} + n, \tag{2}$$

Where $\otimes b$ denotes a convolution with a blurring kernel b, $\downarrow k$ is the downsampling operation by a factor k and n represents the noise term.

SOTA in CV/AI

A schematic representation of the main trends and advances in AI is shown in Figure 2. The top arrow indicates both time and complexity. The diagrams report the main advances in CV in the last few years, namely convolutional neural networks, generative adversarial networks, attention layers, and diffusion models. These advances have been adopted and adapted for the super-resolution task, resulting in the architecture names shown. The arrow at the bottom represents a recent but growing interest in adopting explainable AI (xAI) methods to accompany the output of SR algorithms, to increase their trustworthiness and uptake in downstream applications. For more details about advances in SR refer to the review papers from Moser et al. (2023), and Moser et al. (2024).

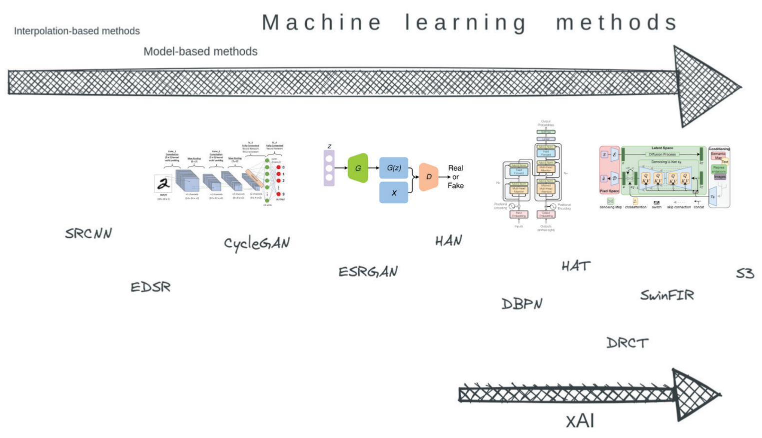

Figure 3: Main trends in CV/AI and proposed ML methods for super-resolution.

SOTA in Remote Sensing

For consistency with the literature, we adopt the notations and vocabulary suggested in Sdraka et al. (2022). SR4AGRI focuses on spatio-spectral fusion (SSF), which involves fusing images of different spatial and spectral resolutions to produce an image of the highest possible spatial resolution. An overview of super-resolution types for remote sensing is shown in Figure 4.

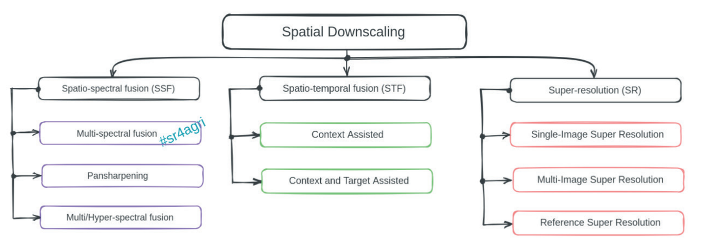

Figure 4. Three main groups of spatial downscaling techniques for remotely sensed imagery.

For SR4AGRI, both model-based, ML-based and hybrid models will be considered for the comparison and the selection of a best performing method. Selection of the methods will depend on: complexity of the method, availability of code or pseudo-code, demonstrated and solid improvements over other methods. Based on the results reported and the diverse nature of the algorithms, good candidates for selection are **ATPRK**, **MuSA**, **DSen2**, **FUSE**, **SPRNet**, **FuSGAN** and **DeepSURE**. More methods could be considered if time allowed, but it could also be possible that a lower number of methods will be evaluated due to complexity or reproducibility issues.

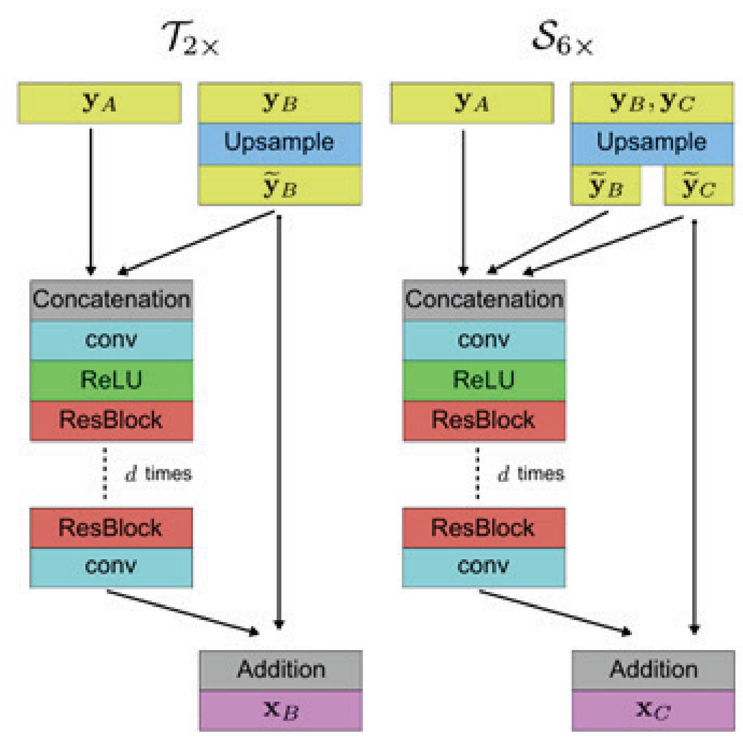

Figure 5. Architecture for DSen2 and VDSen2 proposed by Lanaras et al. 2018. Figure taken from the paper.

Evaluation metrics

A large number of metrics has been proposed and used in the literature, as reported in Figure 6. Metrics can be broadly categorized into two sets: 1) metrics that quantify their distortion with respect to the reference HR image, and 2) metrics that quantify the perceptual quality of the super-resolved images.

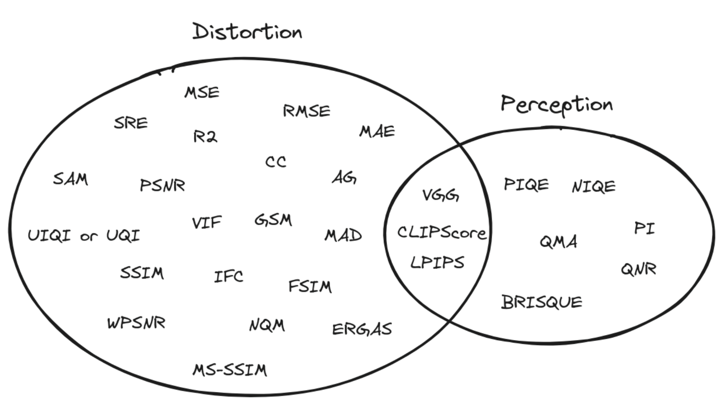

Figure 6. Metrics proposed to evaluate super-resolution algorithms. Distortion metrics require a reference for calculation, while perception metrics do not.

In their foundational work, Blau et al. (2018):

- prove that there exists an unattainable region in the perception-distortion, see Figure 7;
- show that is possible to improve either perceptual quality or distortion only at the expense of the other;
- suggest that super-resolution algorithms should be compared on the perception-distortion plane;
- show that GANs operate in proximity of the perception-distortion bound, since their training loss is a weighted average of distortion and perception losses

Figure 7. Distortion-perception plane. Algorithms on this plane can be directly compared.

Training/Validation dataset

Wald's protocol assumes that the performance of data fusion models is independent of the scale, provided that certain conditions hold. In their seminal work, Wald et al. (1997) suggest first degrading the input image according to a factor k, thus creating LR-HR image pairs, and proceed to design a model tasked to downscale it to the original resolution. The developed method can be transferred to downscale the original image into one of much higher resolution, as shown in Figure 8. Effectively, this method doesn't require images from other missions/sensors, nor other source of reference data, therefore enabling to scale the training data generation to large areas.

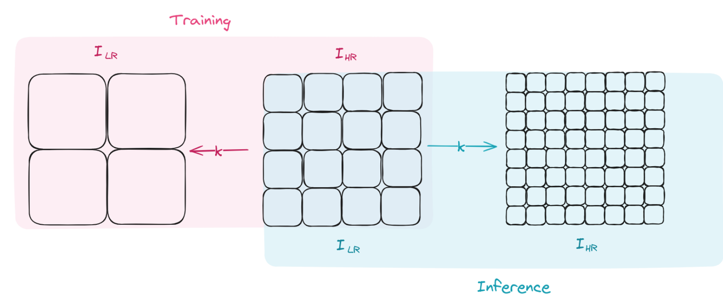

Figure 8. The Wald's protocol pipeline.

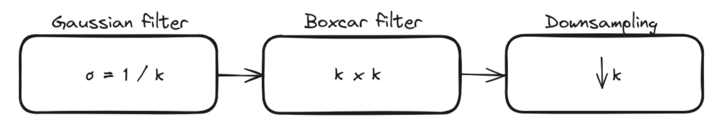

Figure 9. Degradation process for Sentinel-2 bands, proposed in Lanaras et al. (2018).

An alternative proxy at the target spatial resolution would provide a better and independent validation for the assessment of the SSF algorithms. The ideal proxy would be imagery with spectral responses for some bands similar to the native Sentinel-2 20 m and 60 m, with a GSD comparable to, or better, than 10 m, and with acquisition parameters, e.g. time of day, viewing angle, similar to Sentinel-2. PlanetScope from Planet Labs fulfill many of these criteria outlined above for being good proxies for I_{HR}, as shown in Figure 10.

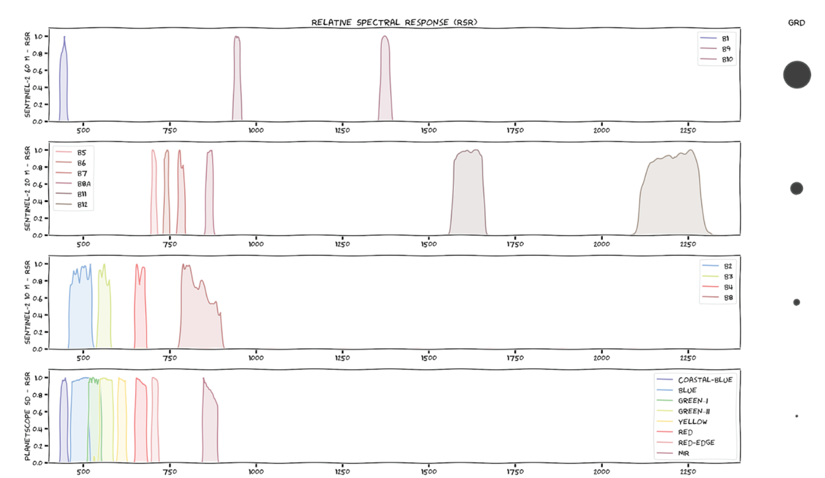

Figure 10. Relative spectral response of PlanetScope SuperDoves (bottom row) and Sentinel-2 bands at 10 m, 20 m, and 60 m GSD.

sr4agri: Requirements

What is AMS

The Area Monitoring Service (AMS) allows monitoring of agricultural land on a large scale, providing valuable insights into changes and agricultural activities, and enabling informed decision-making and resource management. AMS uses S2 data to identify bare and (partly) vegetated soil, vegetation characteristics, photosynthetic activity (an indication of vegetation / crop growth) and non-agricultural land cover (e.g. water, forest, built-up areas), see Figure 11.

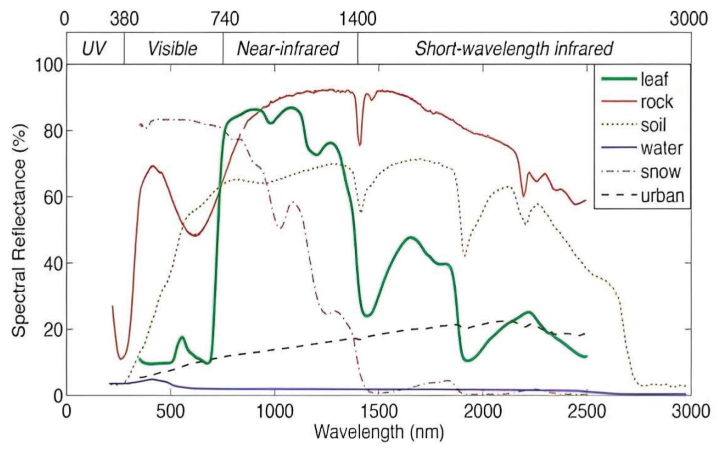

Figure 11. Reflectance spectra of the main land covers of Earth's surface.

User requirements

Using the database for remote sensing indices, we explored which of the plethora of remote sensing indices can be used with the S2 bands, and what is their field of application. Grouping the fields into two groups: agriculture and "other", we obtain the results shown in Figure 12.

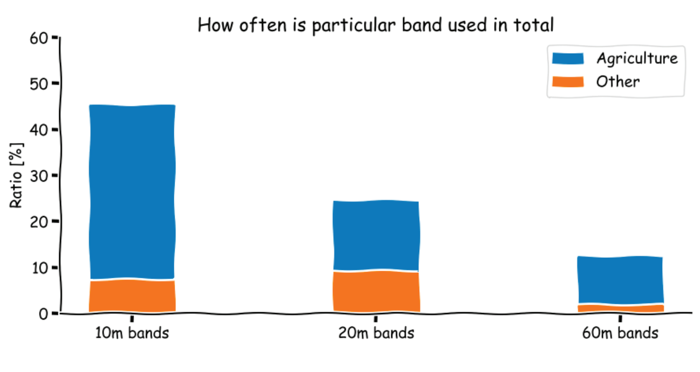

Figure 12. Percentage of usage of S2 bands for agricultural and other applications grouped by native resolution.

Distribution of the size of the arable polygons in Slovenia measured by the number of full S2 pixels at 10 m pixel size contained within the geometry. Polygons that contain 0 full S2 pixels, in black, cannot be monitored by the Area Monitoring service. Polygons containing one full pixel, in red, or up to 36 full 10 m pixels, in orange, will likely contain spatially mixed information for the S2 spectral bands at the native 20 m and 60 m GSD. These polygons would benefit most from SSF techniques.

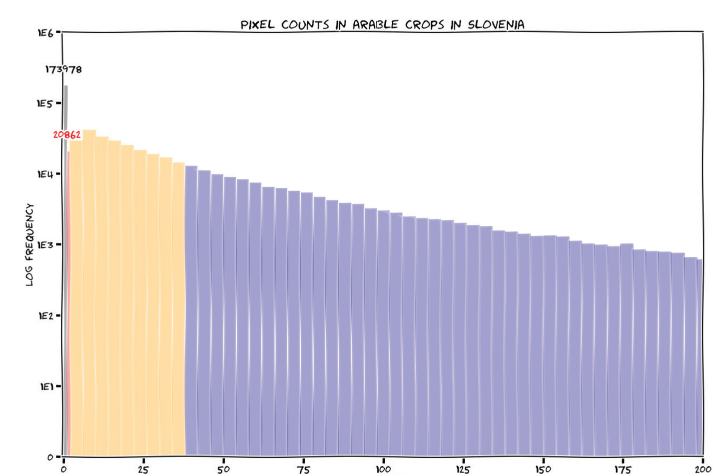

Figure 13. Percentage of polygons in Slovenia that would benefit from SSF techniques (in yellow).

The table below shows the percentage of agricultural parcels in Slovenia that can be reliably monitored within AMS.

# of pixels in poly	% of total arable polygons	% of total arable area	Description
0	29.6	3.6	These polygons are discarded and cannot be monitored with S2.
1	3.5	0.7	These polygons contain 1 full S2 pixel.
2-36	39.3	22.6	These polygons likely contained mixed spectral information for the 20m and 60m bands. Monitoring of these polygons would benefit from super-resolved spectral information.
36-	27.5	73.1	These polygons are large enough to not be affected by spatial mixing of the coarse S2 bands.

Technical requirements

We will divide the workload into three phases, following FAIR criteria:

- Comparison of state-of-the-art methods on a small scale dataset for selection of most promising method;
- Development of super-resolution method on a large scale dataset and validation of super-resolution performance based on super-resolution criteria;
- Validation and verification of the super-resolved bands in downstream agricultural applications using metrics pertinent to the downstream task.

Priority	Feature
Must-have	• Super-resolved bands must have high fidelity to native spectral information • Super-resolved bands must have higher spatial resolution than native bands
Should-have	• The must-have properties should generalize globally • The super-resolution algorithm generalizes over different land covers
Could-have	• The super-resolution algorithm is light-weight and requires little computational resources
Won't have (now)	• The super-resolution bands are accompanied with an uncertainty score/map

References

Blau, Y. and Michaeli, T., 2018. The perception-distortion tradeoff, Proceedings of the IEEE/CVF Conf. Comput. Vis. Pattern Recognit. arXiv: 1711.06077.

Lanaras, C., Bioucas-Dias, J., Galliani, S., Baltsavias, E., & Schindler, K., 2018. Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network, ISPRS J. Photogram. Remote Sens., vol. 146, pp. 305–319, doi: 10.1016/j. Isprsjprs.2018.09.018

Moser B. B., F. Raue, S. Frolov, S. Palacio, J. Hees and A. Dengel, 2023. Hitchhiker's Guide to Super-Resolution: Introduction and Recent Advances, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 8, pp. 9862-9882, doi: 10.1109/TPAMI.2023.3243794

Moser, B. B., Shanbhag, A. S., Raue, F., Frolov, S., Palacio, S., and Dengel, A., 2024. Diffusion Models, Image Super-Resolution And Everything: A Survey. doi: 10.48550/arXiv.2401.00736

Sdraka, M., et al., 2022. Deep Learning for Downscaling Remote Sensing Images: Fusion and super-resolution, IEEE Geoscience and Remote Sensing Magazine, vol. 10, no. 3, pp. 202-255. doi: 10.1109/MGRS.2022.3171836

Wald, L. et al., 1979. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images, Photogrammetric Engineering and Remote Sensing, vol. 63, pp. 691-699.

Acknowledgements

SR4AGRI is funded by the European Space Agency under the project AO/1-11503/22/NL/SC.

SINERGISE esa sentinelhub by PLANET LABS Geodetski zavod Celje d.o.o.