




Band Reconstruction Using a Modified UNet for Sentinel-2 Images

Iulia Coca Neagoe , Daniela Faur , *Member, IEEE*, Corina Vaduva , and Mihai Datcu , *Fellow, IEEE*

Abstract—Multispectral (MS) remote sensing images are of great interest for various applications, yet, quite often, an MS product exhibits one or more noisy bands, strip lines, or even missing bands, which leads to decreased confidence in the information it contains. Meeting this challenge, this article proposes a UNet-based neural network architecture to reconstruct a spectral band. The worst case scenario is considered, that of a missing band, the reconstruction being performed based on the available bands. Besides the comparison with state-of-the-art methods, both the qualitative and quantitative analyses are fulfilled considering several metrics: root-mean-square error, structural similarity index, signal-to-reconstruction error, peak-signal-to-noise ratio, and spectral angle mapper. The experiments focus on Sentinel-2 open data within the Copernicus program. Various patterns of urban areas, agricultural regions, and regions from North Pole or Kyiv, Ukraine are included in our dataset to prove the efficiency of band reconstruction regardless of land-cover diversity.

Index Terms—Band reconstruction, multispectral (MS) images, remote sensing, Sentinel-2 (S2), UNet.

I. INTRODUCTION

REMOTE sensing used in the scope of earth observation (EO) is among the most essential technologies in learning and understanding the earth's surface. The systems used to acquire information are complex and sensitive, but to effectively exploit them, we expect a flawless operation. However, for various reasons, such as extreme atmospheric conditions or physical degradation of some components of these sensors, important information may be missing in the acquired images. As the time passes since satellites were launched and fulfill their mission, there is a great risk of sensor's degradation. These degradations can mean noise in the form of lines, the lack of partial information, or even the total lack of a spectral band. Fig. 1 shows such an example, which emphasizes the

Manuscript received 15 October 2022; revised 13 February 2023 and 28 April 2023; accepted 4 May 2023. Date of publication 17 May 2023; date of current version 28 July 2023. This work was supported by the Romanian Ministry of Education and Research, National Research Council of Romania - The Executive Unit for the Financing of Higher Education, Research, Development and Innovation (UEFISCDI), within PNCDI III program under Grant PN-III-P4-ID-PCE-2020-2120. (Corresponding authors: Iulia Coca Neagoe; Daniela Faur; Corina Vaduva; Mihai Datcu.)

Iulia Coca Neagoe, Daniela Faur, and Corina Vaduva are with the Research Center for Spatial Information (CEOSpaceTech), Politehnica University of Bucharest (UPB), 060042 Bucharest, Romania (e-mail: neagoe.iulia@yahoo.ro; danielafaur@gmail.com; corina.vaduva@gmail.com).

Mihai Datcu is with the Research Center for Spatial Information (CEOSpaceTech), Politehnica University of Bucharest (UPB), 060042 Bucharest, Romania, and also with the Earth Observation Center (EOC), German Aerospace Center (DLR), 82234 Wessling, Germany (e-mail: mihai.datcu@dlr.de).

Digital Object Identifier 10.1109/JSTARS.2023.3276912

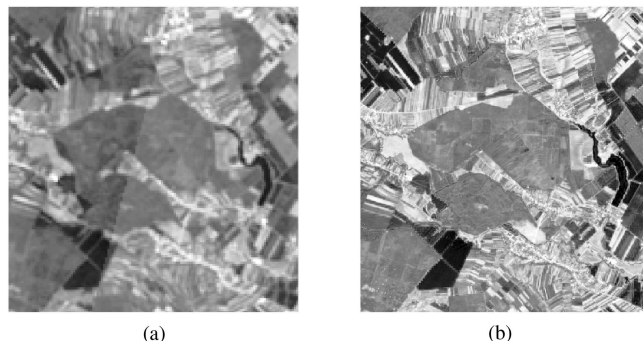


Fig. 1. Example of (a) corrupted band and (b) reconstructed band using the proposed method.

need of band recovery due to sensor-generated artifact in the acquired band. The recovered one does not contain any corrupt information because the complementary spectral information is clean and untainted.

Addressing the challenge of a missing or degraded spectral band, this article presents a method of valorizing the spectral information available in the other bands to predict it. We considered the worst case scenario, that is, the lack of any information relative to the missing band, while the reconstruction process uses the available complementary spectral bands.

The suggested method implements a convolutional neural network (CNN) architecture named U-Net, modified to fulfill a specific need: learn spectral and spatial information in order to better reconstruct the band. Designing a framework able to reconstruct any band of a Sentinel-2 (S2) product leads to a solution individually applied for each band. The following sections detail the proposed concept, the implementation aspects, the consideration of the sensor's spectral characteristics, and the experimental results. In the end, we provide the assessment of the results, draw conclusions, and highlight further perspectives.

II. RELATED WORK

Remote sensing instruments for recording information about the earth's surface are the main sources of data regarding the observation and understanding of the planet on which humans live. However, it infrequently happens that these sensors suffer certain physical degradations, which cause incomplete data acquisition. The missing data can be manifested both in the form of noises present on certain bands or stripe lines and the total absence of a spectral band.

The incompleteness of MS products affects numerous applications that rely on them. Hence, this challenge was widely investigated in an attempt to solve the missing information reconstruction.

Over time, different methods have been submitted to approach this subject, which can be differentiated according to the used basic principle: the information needed to perform the reconstruction. Thus, depending on the source of this information, four types of methods can be distinguished: spatial-based, spectral-based, temporal-based (multitemporal), and mixed (hybrid) methods [1].

A. Spatial-Based Methods

Spatial-based methods, also known as inpainting methods, are among the most classical and primitive because they use the remaining information to generate the missing information. The basic idea behind them is represented by the fact that both the missing data and the remaining data contain the same geometric and statistical structures [2]. Spatial correlation is the main hypothesis on which these methods are based.

Many of these solutions have been studied in the field of computer vision, but they can also be applied to multispectral remote sensing images. Over time, they have evolved from classical methods, based on algorithms and sequential transformations, to the use of neural networks. Zhang et al. [3] propose an alternative interpolation method for the local linear histogram-matching technique, namely the kriging geostatistical technique able to fill data gaps. The same interpolation has been used in [4] to retrieve aerosol optical depth over the pixels that could not be considered by the custom dedicated algorithm. The approach presented in [5] is based on the joint interpolation of image gray levels and gradient/isophote directions, smoothly extending in an automatic fashion the isophote lines into the holes of missing data. Maalouf et al. [6] introduce a Bandelet-transform-based inpainting technique to remove cloud-contaminated portions of a remotely sensed image and fill the missing data. Chan and Shen [7] introduce a nonlinear partial differential equation inpainting model based upon curvature-driven diffusions for non-textured images. Previous works use a maximum-*a-posteriori*-based algorithm for both destriping and inpainting problems [8] or multichannel nonlocal total variation for textured images and reconstruction of large-scale areas [9].

However, spatial-based methods can encounter problems such as unconnected edges, blur effect, or texture inconsistency.

B. Spectral-Based Methods

The methods of recovering a band based on the use of spectral information are dependent on the data available on the other spectral bands. Due to the characteristics of the sensors, in the multispectral and hyperspectral images, there is a lot of redundant information that can be used for useful purposes. However, there is a prerequisite: the bands needed for reconstruction must exist and be complete.

So far, few spectral-based methods have been proposed to recover the missing information or even an entire band. The first ones addressed the problem of Aqua MODIS images that had

band 6 with stripe noise and assumed the use of polynomial regression to retrieve this band from band 7 measurements [10], combining histogram matching to correct the detector-to-detector striping of the functional detectors with local least-squares fitting that restores the missing data based on a cubic polynomial derived from the relationship between bands 6 and 7 [11] or within-class local fitting that incorporates scene types and spectral band characteristics [12]. These methods included only two bands as sources of information or reconstruction for Aqua MODIS, while Gladkova et al. [13] described a quantitative image restoration algorithm that handles a small number of functioning detectors to train a restoration function that is based on a multivariate regression using the information in a spatial-spectral window around each restored pixel. For the same task, Li et al. [14] implemented a robust multilinear regression based on the spectral relations between working detectors in band 6 and all the other spectra, showing better results.

Later, methods based on neural networks revealed improved results. Rout [15] studied the ability of supervised and adversarial learning to address the task of missing band reconstruction with the sole supervision of existing spectral and spatial prior distribution. Their band recovery method modified the super-resolution solution developed by Rout et al. [16], based on deep CNNs that encompass two major learning mechanisms: global and local residual learning.

C. Temporal-Based Methods

Temporal-based methods imply the use of additional information obtained from images acquired at a short time interval over the same geographical area.

The most well-known approaches include temporal replacement [17], [18], [19], [20], [21], the use of temporal filters [22], [23], [24], or the temporal learning model [25], [26]. Zeng et al. [17] used multitemporal regression analysis and a regularization method to recover missing pixels for Landsat ETM imagery. Furthermore, based on the concept of utilizing temporal correlation of multitemporal images, in [18], a patch-based information reconstruction algorithm spatiotemporally segments a sequence of images into clusters containing several spatially connected components called patches and then clones information from cloud-free and high-similarity patches to their corresponding cloud-contaminated patches. In [19], missing measurements are reconstructed through an unsupervised contextual prediction process that reproduces the local spectrotemporal relationships between the considered image and an opportunely selected subset of the remaining temporal images, while Zhang et al. [20] handle missing data by creating appropriate covariates and then fitting a functional concurrent linear model on the resulting data. A straightforward method based on the Savitzky–Golay filter to smooth out noise caused by cloud contamination and atmospheric variability in NDVI time series is presented in [22], while a changing-weight filter approach for reconstructing a high-quality NDVI time series is presented in [24]. The reconstruction of area obstructed by clouds based on the compressive sensing theory introduced

in [25] enables finding sparse signal representations in under-determined linear equation systems, while, for the same task, Li et al. [26] advance two multitemporal dictionary learning algorithms, expanding the K-SVD and Bayesian algorithms, to make better use of the temporal correlations.

The biggest drawback of these methods lies in the ground changes appeared in the time elapsed to first accurate available data used for reconstruction. These changes may be due to new construction or natural hazards, such as flood or fire, or even problems that occurred during the acquisition of the image (observation conditions and atmospheric conditions).

D. Hybrid Methods

Since the three types of methods presented above have both advantages and disadvantages when used separately, there is also the possibility of combining them in order to obtain better results.

Hybrid methods for information recovery assume the use of additional information from different domains. Therefore, there can be several possibilities of combinations, i.e., spatiotemporal methods [17], [27] or spatio-spectral methods [28], based on the idea of making the prediction process learn from information available in the cloud-free neighborhood of contaminated areas for the contextual reconstruction of cloud-contaminated areas in multitemporal multispectral images. There are also methods, such as in [29], advancing a unified spatial–temporal–spectral deep CNN for reconstruction, that use information from all three domains. Solutions developed for pan sharpening (i.e., predicting pixels signatures at higher resolution) are also combining the spatial and spectral information. In [30], a four-layer CNN is proposed using a loss function without requiring a reference.

Later, methods based on convolutional networks [31], proposing a new approach to denoising, inpainting, and super-resolution of hyperspectral image data based on intrinsic properties of a CNN without any training, or generative networks [32], implementing a modified unsupervised CNN context generate model, were explored.

Although using the combined advantages of two or more methods would result in obtaining results with greater accuracy, it must be taken into account that, in any combination that includes the temporal domain, the existence of very recent preceding images is determined by too many uncontrollable factors.

In the process of reconstructing a band, the super-resolution effect can also be encountered if the recovery development involves all the spectral bands complementary to the band to be reconstructed, regardless of their spatial resolution. For example, Rout et al. [16] use multisensor bands as input information to reconstruct a band in the SWIR domain at a resolution that the target sensor does not have.

Brodu [33] presents *Superres*, a super-resolution method for S2 products based on exploiting both the local consistency between neighborhood pixels and the geometric consistency of subpixel constituents across multispectral bands in order to bring all the bands from 20 and 60 m/pixel down to 10 m/pixel. Starting from the highest resolution bands, band-dependent information

is separated from information that is common to all bands' geometry of scene elements. This model is then applied to unmix low-resolution bands, preserving their reflectance, while propagating band-independent information to preserve the subpixel details.

Lanaras et al. [34] introduce *DSen2* and *VDSen2*, state-of-the-art CNNs to perform end-to-end upsampling; these are two configurations of ResNet architecture, i.e., Deep Sentinel-2 and a Very Deep Sentinel-2, being trained with low resolution data. Thus, one has access to a virtually infinite amount of training data by downsampling real S2 products. They use globally sampled data over a wide range of geographical locations to obtain a network that generalizes across different climate zones and land-cover types and can super-resolve arbitrary S2 product without the need of retraining. Before [34], there are no significant results in applying deep learning for super-resolution. A comparatively shallow three-layer CNN architecture was originally designed for single-image (blind) super-resolution as introduced in [54]. They train pan-sharpening networks for Ikonos, GeoEye-1, and WorldView-2. Similarly, PanNet, a network introduced in [55], based on the high-performance ResNet architecture, was also applied on WorldView-2, WorldView-3, and Ikonos. Due to their performance and relevance as state-of-the-art approaches for our endeavor, the abovementioned methods *Superres* and *DSen2*, *VDSen2* will serve as baseline methods for comparison and validation.

Methods of synthetic generation of a band are also represented by those who seek to obtain multispectral images based on RGB ones. Rangnekar et al. [35] train a conditional adversarial network to learn an inverse mapping from a trichromatic space to 31 spectral bands within 400–700 nm, an aerial hyperspectral dataset. Similarly, Rodríguez-Suárez [36] focus on conditional generative adversarial networks (CGANs) to achieve the reconstruction of multispectral images from RGB images. Different regression network models (convolutional neuronal networks, U-Net, and ResNet) have been adapted and integrated as generators in the CGAN and compared in terms of performance for multispectral reconstruction. A very comprehensive work that reviews all the methods of multispectral image enhancement from the point of view of super-resolution, noise reduction, inpainting, or restoration is proposed by Tsagkatakis et al. [37].

III. PROPOSED CONCEPT

With the primary objective to fully exploit the spectral information in the interest of reconstructing a missing band of a multispectral image, this article proposes a method to extract that information from the concurrent spectral bands of the same product. This concept is based on the following premises.

- 1) Multispectral images represent a product describing a certain area from the earth's surface containing bands acquired at different wavelengths. Consequently, there is a spatial and spectral correlation between bands.
- 2) Deep neural networks have proven their ability to learn from the unique representations of various target attributes.

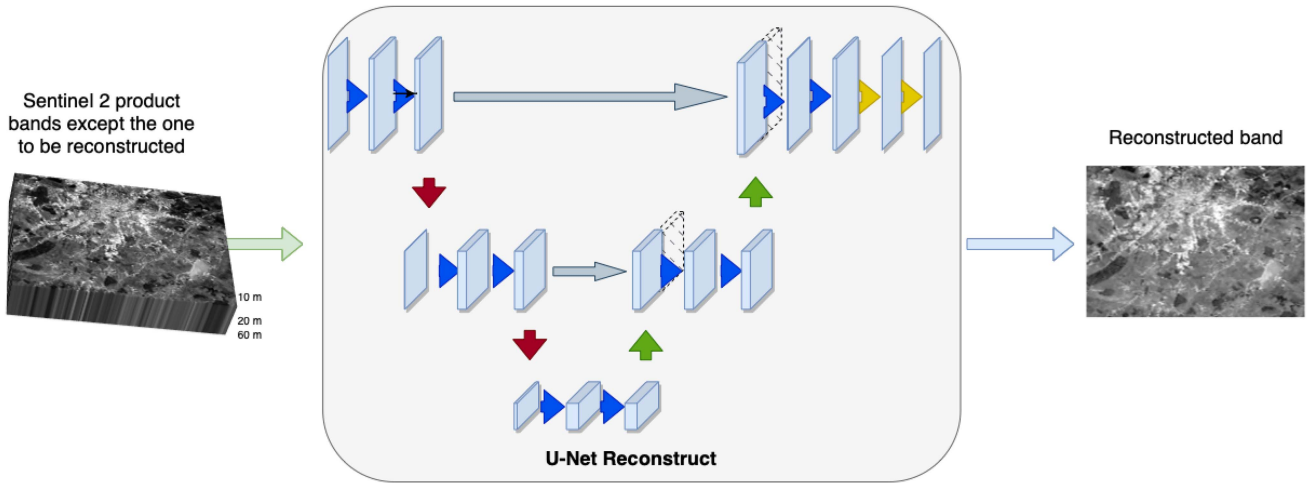


Fig. 2. Proposed method (UNetBRec) concept for band recovery. The CNN with U shape receives as input all the bands except the one to be reconstructed and obtains the predicted band at the output.

- 3) The worst case scenario involves the total absence of a spectral band, but the existence of complementary spectral information in the same multispectral image is sufficient, and no other supplementary data are required.

Starting from these premises, this article proposes a generalized method to reconstruct any missing or corrupted band from a multispectral image. The only condition is the integrity of the complementary spectral information.

Fig. 2 illustrates the general overview of the method. The CNN, UNetBRec, receives as input all the spectral bands except the one to be reconstructed and returns as output a single band with the same width and height. During the training process, the network performs a comparison between ground truth and the generated band to adjust its parameters and obtain a better result. The method is generalized for each band, the single difference between the trained models being the bands received as input and the one used to compare the output. In the case of an L2A S2 product, there is a set of 12 trained networks; each of them is used to recover a specific band.

The proposed concept is designed to prevent the need for additional information, other than the one available in the multispectral product for which one of the bands is to be recovered. Consequently, our solution is able to produce the missing band in a short time by preserving both the spectral and spatial properties of the product.

The benefit of band reconstruction arises from the subsequent use of the thorough multispectral product, as required by the current applications of EO data.

IV. BAND RECONSTRUCTION FOR MULTISPECTRAL IMAGES

The extensive adoption of deep-learning-based methods in numerous fields demonstrated that problems impossible to solve with classical methods became approachable and even solvable using neural networks. One such example is the following situation: due to the lack of correlation of spectral bands' distribution, it was considered unrealistic to reconstruct one band using the others. However, once neural-network-based solutions were

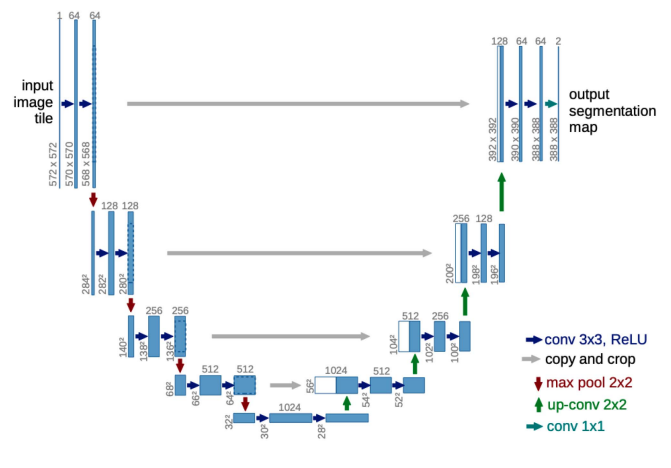


Fig. 3. Original U-Net architecture, an example for 32×32 pixels in the lowest resolution [38].

used, the opposite was demonstrated, the results obtained being more than satisfactory and, in some cases, even more eloquent than the ones obtained using classical methods.

A. Proposed Deep Learning Architecture

U-Net [38] is a state-of-the-art CNN build upon on the “fully convolutional network” introduced by Long et al. [39]. The main characteristic of this network is the U-shape architecture (see Fig. 3) containing a contracting path and an expanding path. The contracting operation is obtained through pooling operators, while the expanding is achieved through upsampling operators. The two branches, down and up, are interconnected through concatenation operations in order to pass spatial and spectral information. Therefore, the symmetry between the two parts of the network is almost perfect.

While the original U-net architecture has been implemented for the segmentation of neuronal structures in electron microscopic stacks, in this article, we introduce a modified U-Net architecture called UNetBRec, presented in Fig. 4, to address

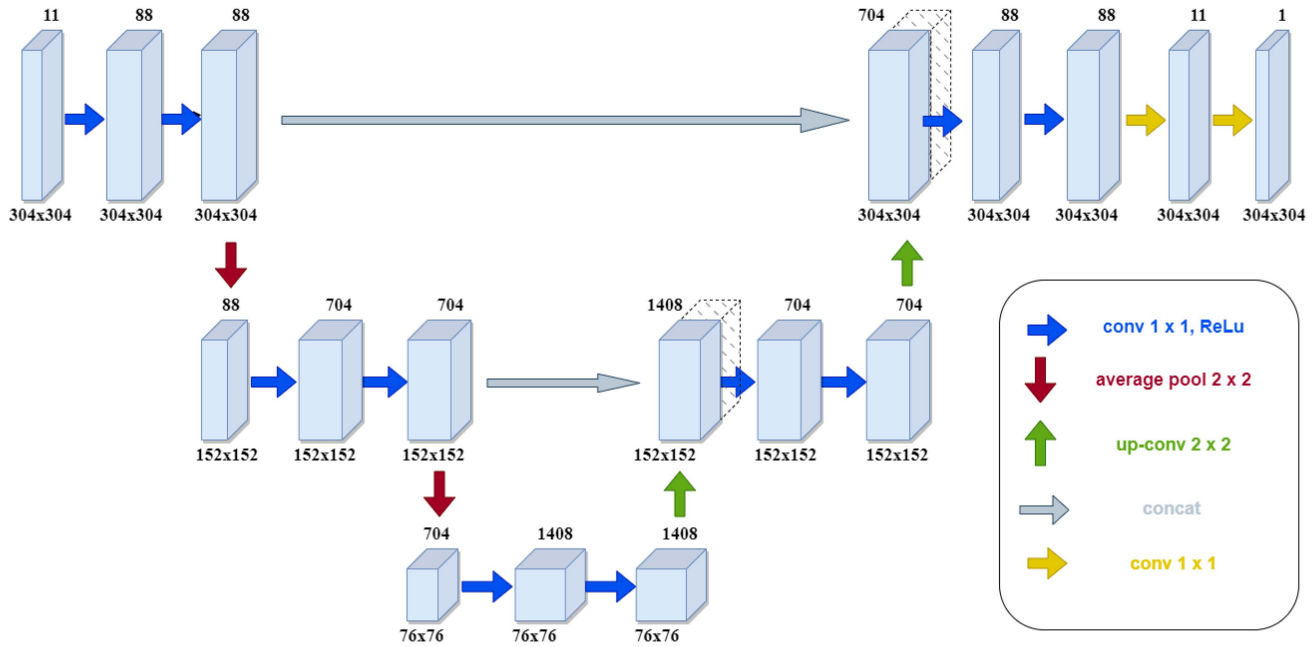


Fig. 4. UNetBRec architecture contains only 14 layers instead of 23 in the original U-Net, meaning that the number of levels on downsizing has decreased from four to two. The dimensions of 304×304 have been chosen experimentally, considering the fact that, placing a smaller dimension as network's input, one can afford to diminish the number of layers in the training stage.

the missing band reconstruction challenge. UNetBRec has the same U-shape structure; the dimensions of the parameters have changed: the number of layers decreased from 23 to 14, meaning that the levels on downsizing has decreased from four to two, to reduce the number of network's parameters. The input patch dimensions decreased from 572×572 single band to $304 \times 304 \times 11$ bands. The main adjustment is related to sequential convolutions that have been settled to 1×1 so that the output will be a single band, with the same dimensionality as the input ones. The typical architecture of a CNN is followed in the down branch. The sequence of two 1×1 unpadded convolutions, each followed by a rectified linear unit (ReLU) and a 2×2 average pooling, is repeated during the contracting path, while the expanding path follows a sequence of an upsampling operation, a 1×1 convolution, a concatenation with the correspondent feature map from the contracting path, and two 1×1 convolutions, each followed by the ReLU. The upsampling operation is, in fact, a deconvolution, which may be seen as a transposed convolution. The last layer is a 1×1 convolutional layer used to get to the desired dimension of the resulting reconstructed band, which is equal to the input dimension. The strategy of this architecture is to successfully learn local and global features relevant for band reconstruction.

In Fig. 4, each box represents a multichannel feature map, with the number of channels on top and the dimension of feature map on the bottom. The arrows define different operations, explained by the legend. The multiplication of the feature channels has been chosen as a result of multiple experiments.

B. Physics-Aware Multispectral Image Band Reconstruction

Several satellite sensors record images with multiple spectral bands at different spatial resolutions, the main advantages being:

- simultaneity of spectral band recording;
- illumination similarity between bands;
- atmospheric conditions similarity between bands;
- very precise coregistration;
- acquisition distance similarity between bands.

The reasons behind recording at varying spatial resolution may be transmission bandwidth restrictions, band designation, or storage.

S2 is one of the products of Copernicus Sentinels mission, which uses satellites that record multispectral images with bands at different spatial resolutions. Bands of S2 product have 10-, 20-, and 60-m resolution. The previously mentioned advantages apply in this case, thus making it possible for the proposed solution to rebuild a spectral band. However, it must be taken into account that the exclusive use of information from the concurrent spectral bands can generate, in addition to band reconstruction, an improvement in terms of resolution. Thus, for 60-m-resolution bands, a resolution improvement may occur, while for 10-m-resolution bands, the quality would be preserved due to the presence of multiple bands with the same resolution.

Another important characteristic of a band is its spectral signature. In the process of band reconstruction, the preservation of the signature demonstrates the effectiveness of the applied method.

The following subsections illustrate and emphasize the improvements which the proposed method brings to the results. Both spectral signature and resolution are analyzed.

1) *Reconstruction of 60-m-Spatial-Resolution Bands*: S2 has three bands at 60-m resolution, but as we use the Level 2 products, band 10 is not included due to the fact that it does not contain surface information. The remaining two bands of 60-m resolution are bands 1 and 9.

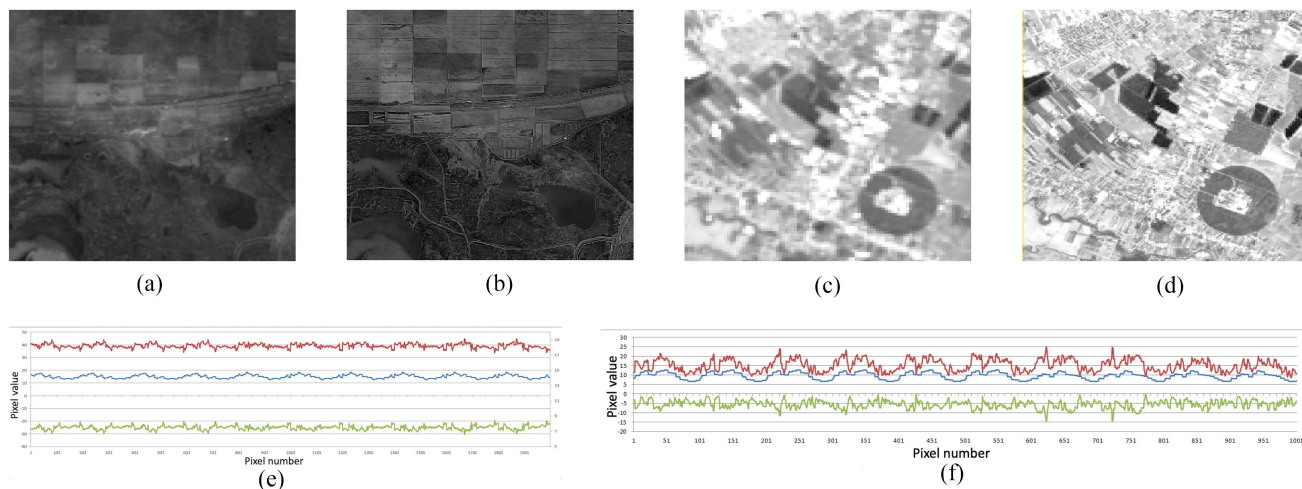


Fig. 5. (a)–(f) Two examples of band 1 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

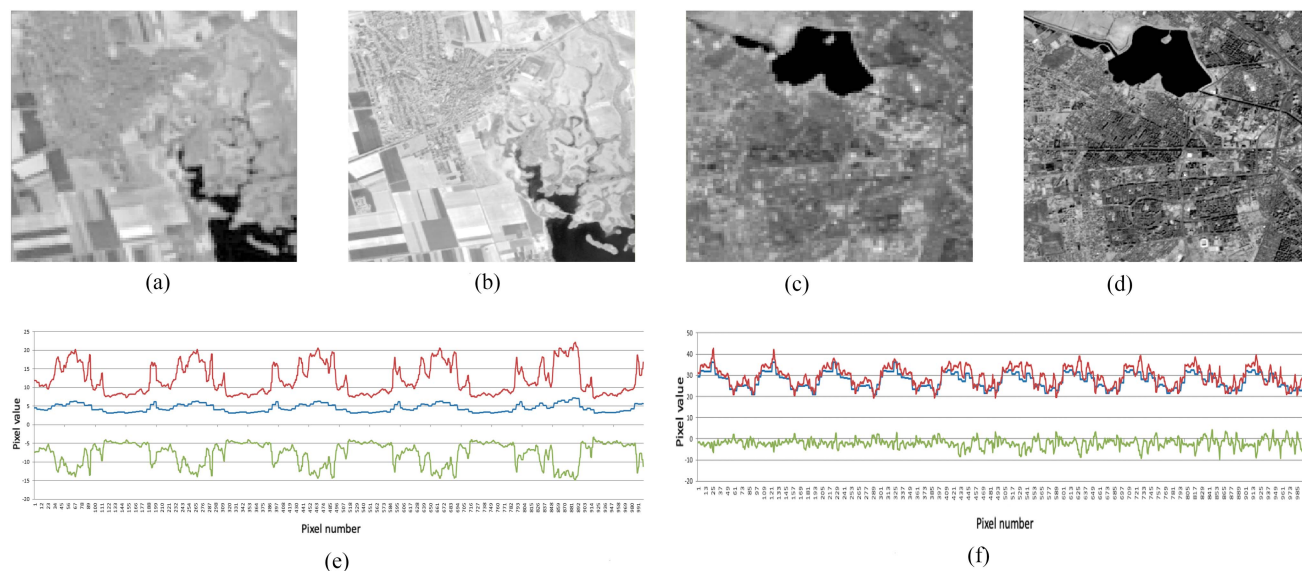


Fig. 6. (a)–(f) Two examples of band 9 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

Fig. 5 presents two examples of band 1 reconstruction, emphasizing the resolution improvement as all the contours are more accurate. Moreover, for each example, there is a graph showing the spectral signature of both initial (blue line) and reconstructed (red line) bands; besides, the green line highlights the pixel differences between the two.

Although the spectral signatures are not identical and do not follow exactly the same pattern, the effect could be explained by the super-resolution itself. Many similar pixels from the initial band belonging to a not clearly defined area in the reconstructed band may have contrasting values in order to define a delineate clear-cut contour of the objects from the earth's surface.

The examples for band 9 prediction (see Fig. 6) also demonstrate that although the main reason was not the super-resolution,

its effect is present and boosts the improvement of the reconstructed band. The presence of more details in the reconstructed scene is also illustrated in the graphs of the spectral signatures, which highlight both the preservation of the patterns and a slight modification due to the increase in contrasts between the neighboring pixels.

2) *Reconstruction of 20-m-Spatial-Resolution Bands:* The number of bands with 20-m resolution in S2 products is equal to six, namely the bands 5, 6, 7, 8A, 11, and 12. Being the resolution that most of the bands are acquired at and being a very small difference compared to the highest resolution, 10 m, the reconstruction is made more accurately. The minor difference between the initial band and the reconstruction is visible in terms of brightness, both in spectral signature pattern and band visualization.

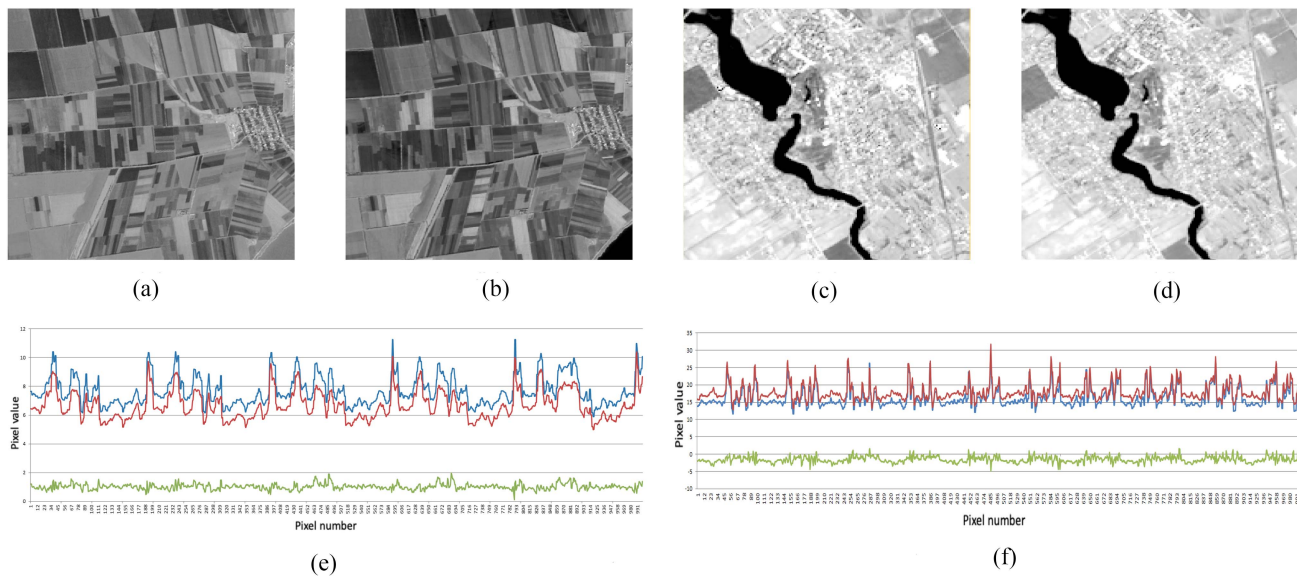


Fig. 7. (a)–(f) Two examples of band 5 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

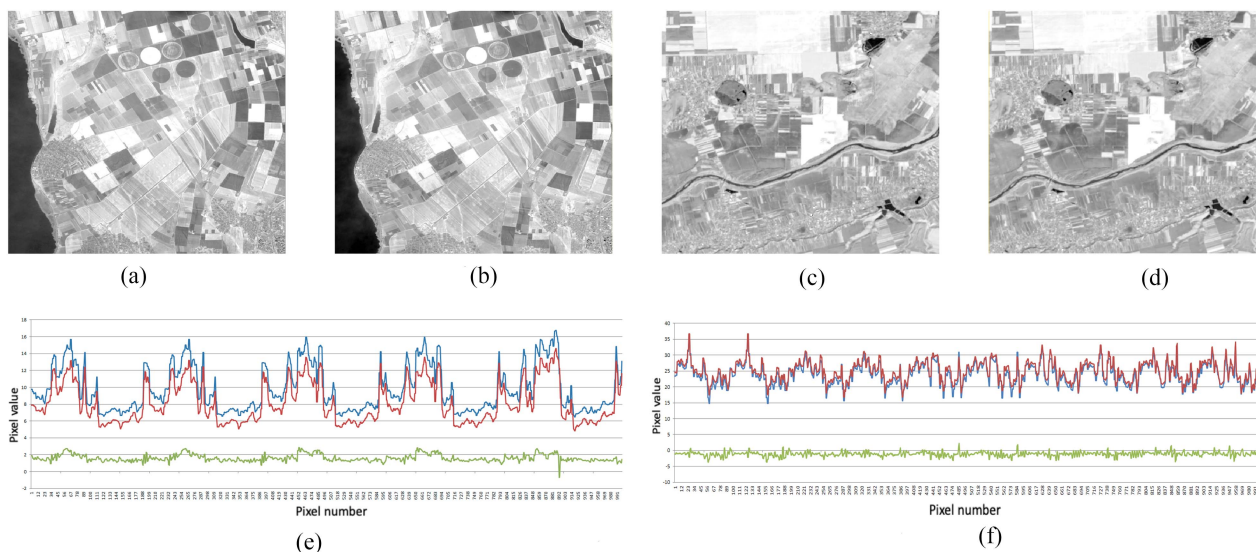


Fig. 8. (a)–(f) Two examples of band 6 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

Figs. 7–12 illustrate two examples for each band recovery.

Unlike the 60-m-resolution bands, in the case of 20-m-resolution bands, from a visual point of view, pixelation differences are no longer noticed, and the spectral signatures keep exactly the same pattern; in some cases, they are even overlapped, which means that they are identical from the point of view of the amplitude of pixel values.

Examples of similar and overlapping spectral signatures of the two bands can be observed in Figs. 7–10, 11(f), and 12(e).

Although from a visual point of view, it is not noticeable, there are small differences in the amplitude of pixel values, but when displayed, the images are calibrated so that the contrasts can

be observed, and small differences remain hidden. Instead, the spectral signature graphs highlight even the smallest difference between the value of a pixel of the initial band and that of the corresponding pixel in the reconstructed one.

Fig. 12(e) and (f) illustrates the greater difference between initial and recovered bands registered in the spectral signature graphs but still preserves the pattern. Both the examples imply higher values of initial band pixels. Taking into account that the wavelengths of these two bands are the longer ones being from the SWIR part of the EM spectrum, it may be deduced that in some cases, although the band will be reconstructed with success, the amplitude of the pixels may not be as higher as in

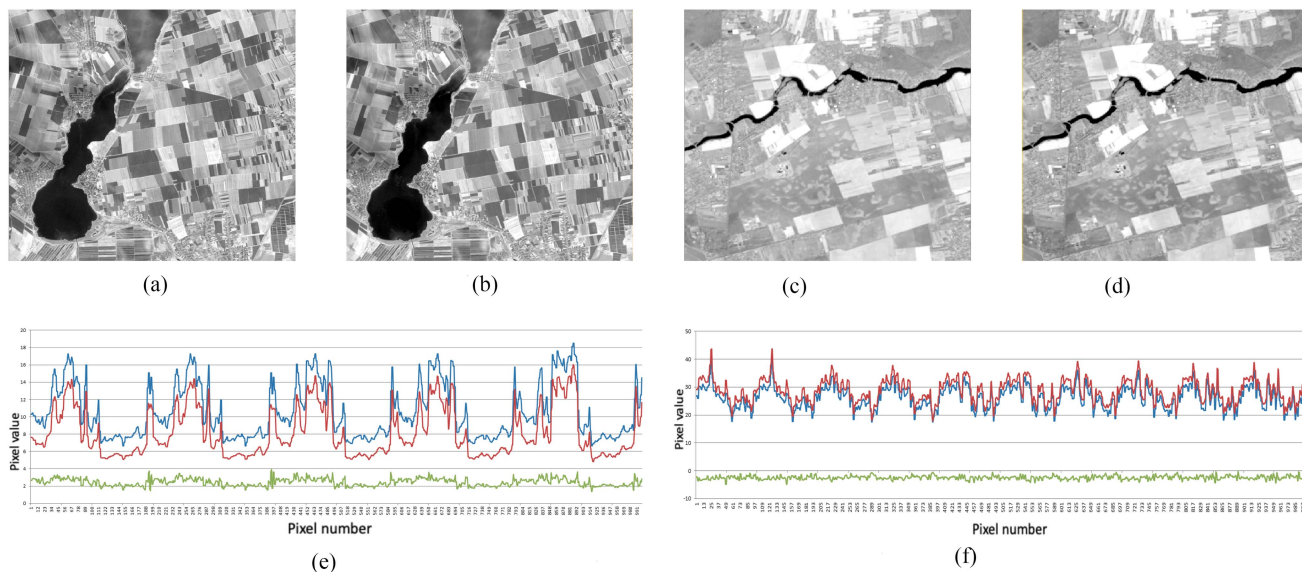


Fig. 9. (a)–(f) Two examples of band 7 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

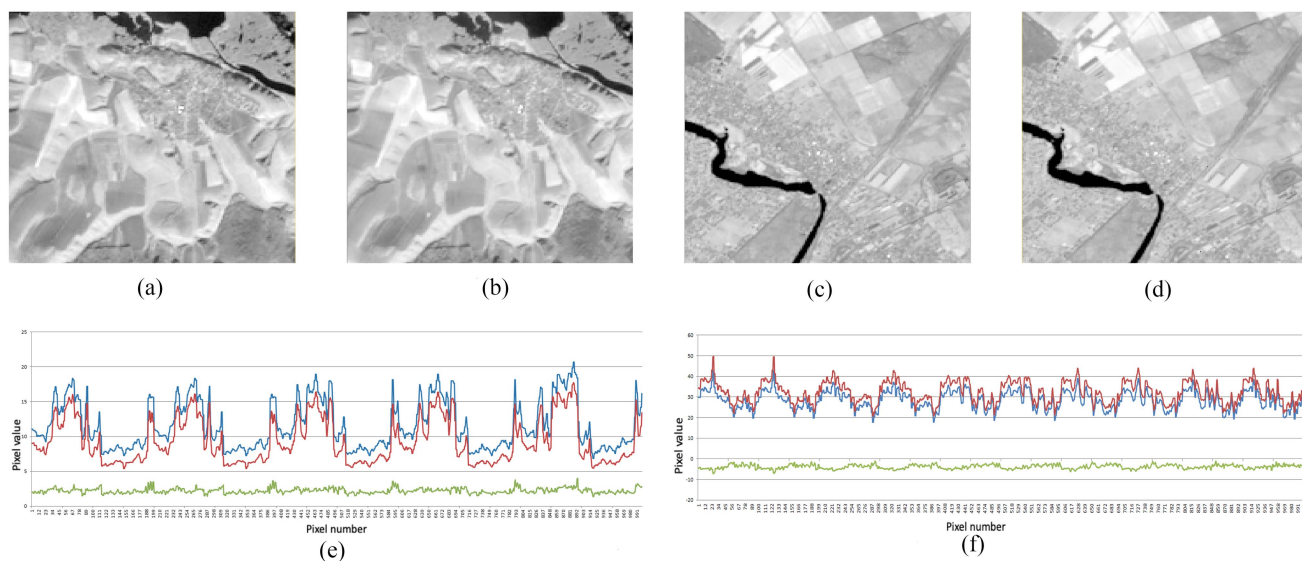


Fig. 10. (a)–(f) Two examples of band 8A reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

the original. The visual analysis of the reconstruction does not unravel the difference.

As highlighted by the examples, the reconstruction of the bands with 20-m resolution proves to be a success, demonstrating the preservation of the spectral signature of the band in question as well as the visual similarity to the original band.

3) *Reconstruction of 10-m-Spatial-Resolution Bands*: Four bands of an S2 product have 10-m resolution, namely bands 2, 3, 4, and 8. First three of these bands have wavelengths from the visible part of the spectrum and are highly used for the true color representation of the image. The fourth one, band 8, has its wavelength in the NIR part of the EM spectrum and is

approximately centered positioned in the range of wavelengths available in an S2 product.

Figs. 13–16 illustrate two examples, each of a certain band recovery. The examples are from two different S2 products over disparate parts of Romania that illustrate both urban areas and agricultural or vegetation areas in order to present as different contexts as possible.

As regards resolution, it was neither improved nor worsened, so that the original quality of each band was preserved. Also, the spectral signature graphs highlight the maintenance of the pattern, registering small amplitude differences between the original and the reconstructed.

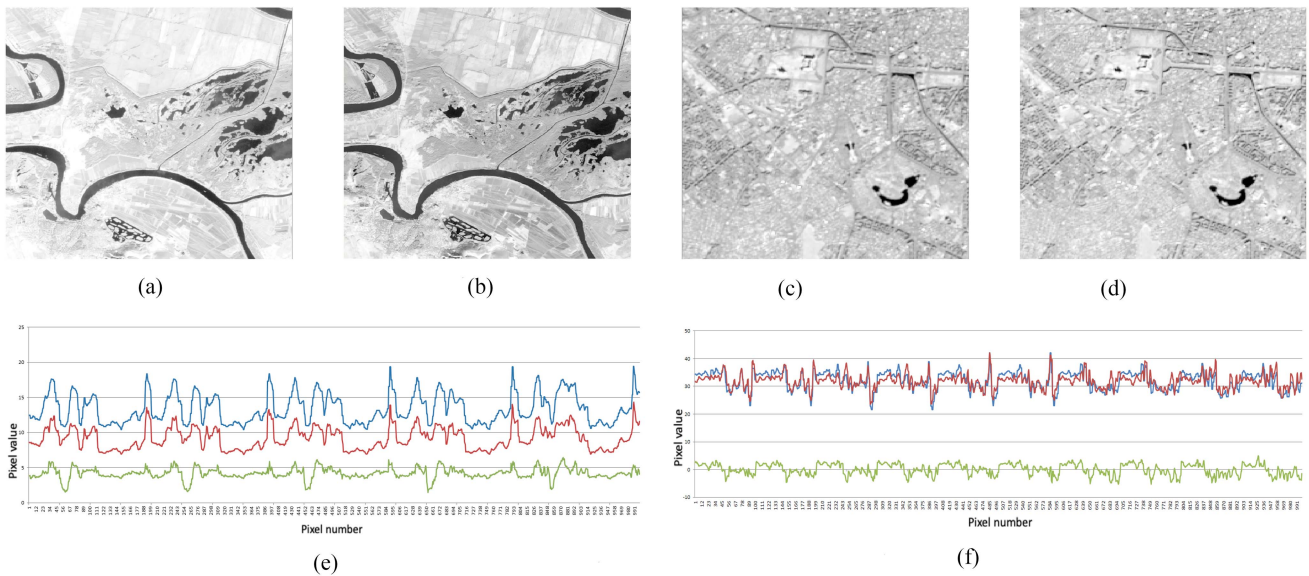


Fig. 11. (a)–(f) Two examples of band 11 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

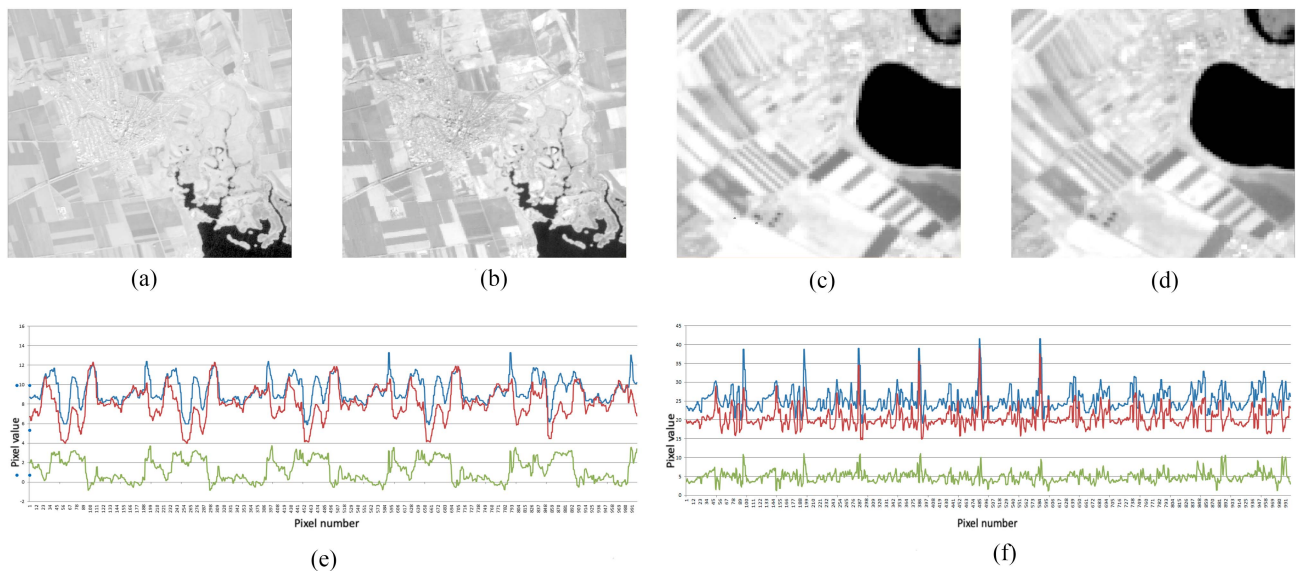


Fig. 12. (a)–(f) Two examples of band 12 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

The quality of 10-m band resolution reconstruction is very similar to that of the ones with a resolution of 20 m, so that the observations regarding the amplitude difference between the pixels are preserved. At the level of details and pixelation, the visual inspection distinguishes neither differences nor the spectral signature graphs, as the patterns are preserved. Spectral signatures only highlight the differences in amplitude, as can be seen in Figs. 13(e) and 16(e).

Regarding the difference compared to the bands with a resolution of 60 m, it can be mentioned that in the case of those, the improvement can be seen as a benefit for the subsequent analysis.

Super-resolution was taken into account only considering the physical characteristics that define a multispectral product, but it was not an objective in itself. The fact that such a result was obtained can only encourage further research on this aspect.

V. EXPERIMENTAL RESULTS AND EVALUATION

This section details the implementation aspects: the dataset selection for training and testing, the stages of experimental fulfillment, and metrics used to evaluate the proposed method. Finally, experimental results are presented and analyzed.

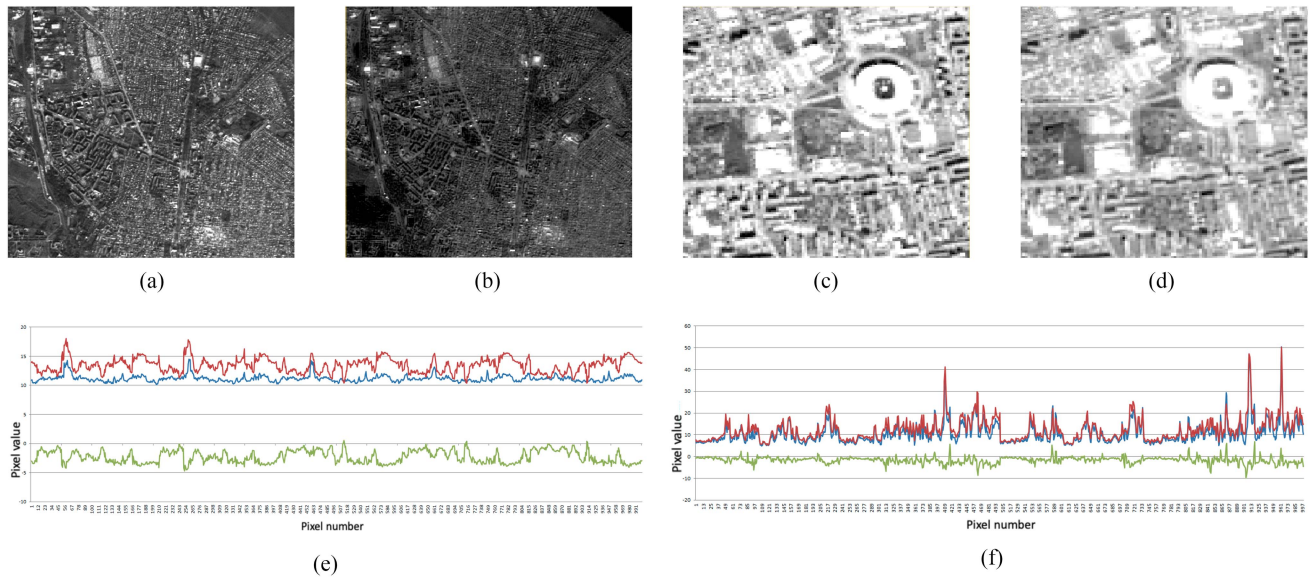


Fig. 13. (a)–(f) Two examples of band 2 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

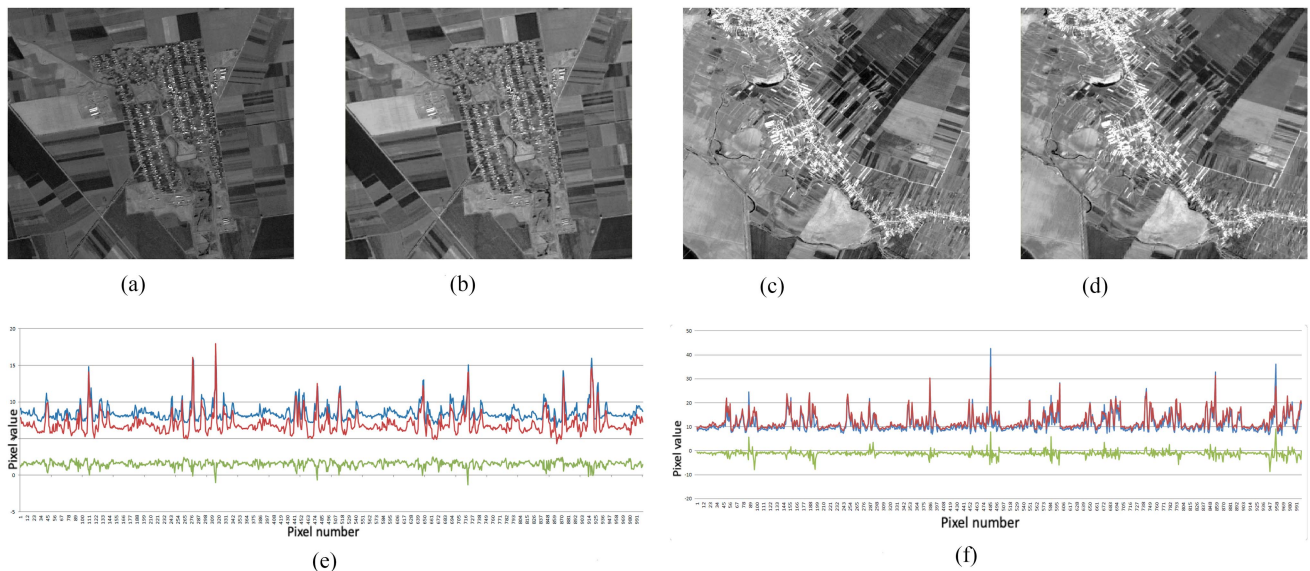


Fig. 14. (a)–(f) Two examples of band 3 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

A. Train and Test Datasets

Deep learning algorithms usually assume the existence or the need to create a dataset for training and testing operations. However, the method proposed in this article does not involve such a need, the use of the products obtained by the sensors of the S2 mission being sufficient. As the method takes into account the worst case scenario, that assumes the nonexistence of the band to be reconstructed, the only condition is that the complementary spectral bands should be available. Also, in the training stage, it is important to have the

initial band so that the network can learn how to reconstruct it.

S2 mission involves two satellites, 2A and 2B, placed on the same orbit, which fly with a phase difference of 180° [40]. The images resulted from the sensing activity of the two MSI identical sensors may be available for usage at different levels of processing. This work has used Level 2A processed products acquired by both the sensors for the experimental results. The locations of the images have been randomly picked, the main objective being the variety in climate zone and distribution across the globe. Various patterns of urban areas and agricultural

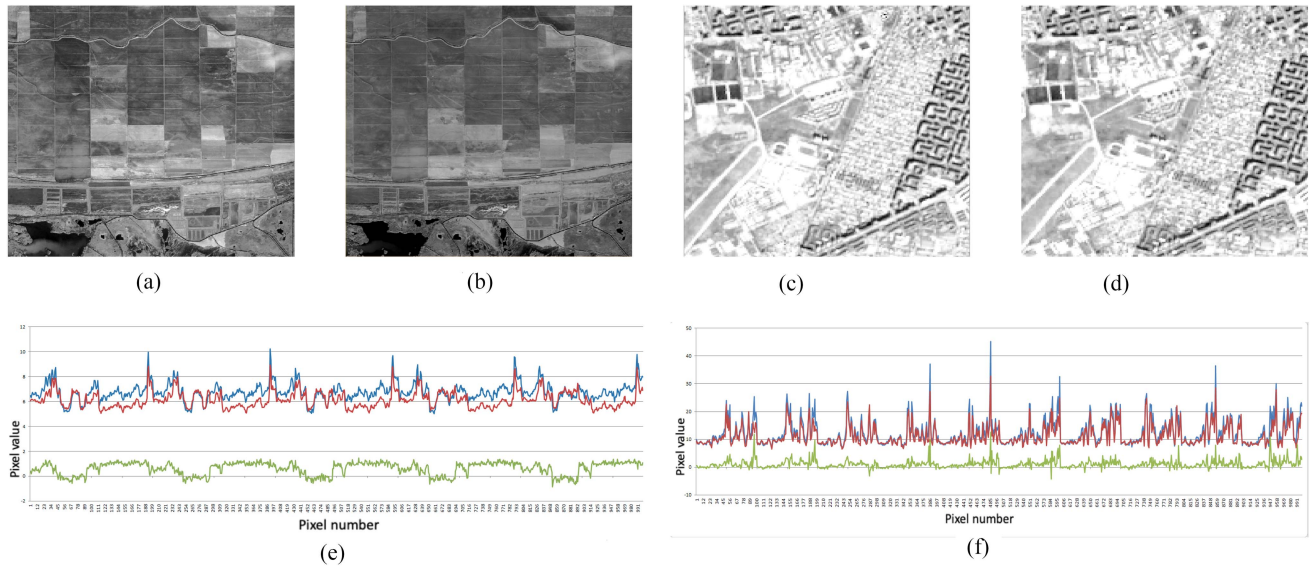


Fig. 15. (a)–(f) Two examples of band 4 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

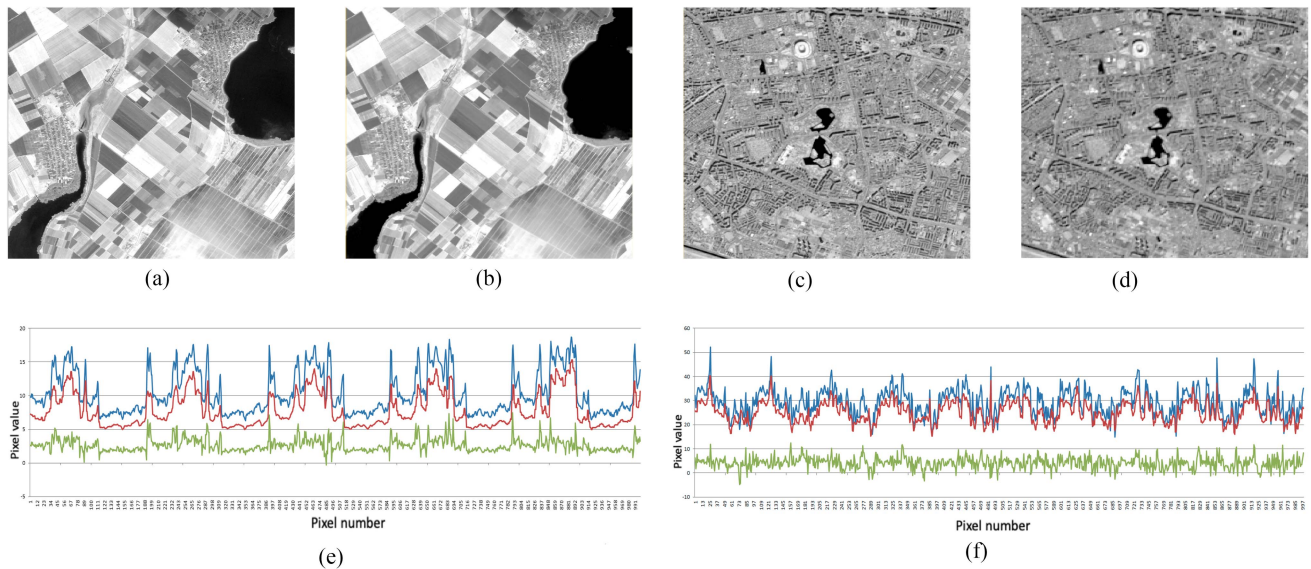


Fig. 16. (a)–(f) Two examples of band 8 reconstruction—visual comparison to emphasize resolution improvement and spectral signature graph to highlight the pattern preservation and development. In (e) and (f), the initial band is shown in blue, the reconstructed band is shown in red, and the difference between them is in green.

regions in North Pole or Kyiv, Ukraine (see Table I) are included in our dataset to prove the efficiency of band reconstruction regardless of land-cover diversity.

An S2 Level 2A product name is defined by a naming convention and has the following form [41]:

MMM_MSIXXX_YYYYMMDDHHMMSS_ Nxyy_ ROOO_ Txxxx_ProductDiscriminator.SAFE,

where:

- 1) *MMM* represents the mission id, naming S2A or S2B;
- 2) *MSIXXX* stands for the level or processing (MSIL1C—Level 1C, MSIL2A—Level 2A);

- 3) *YYYYMMDDHHMMSS* represents the sensing date and time (e.g., 20220323T195051);
- 4) *Nxyy* stands for PDGS processing baseline number (e.g., N0301);
- 5) *ROOO* represents the relative orbit number, which may take values between R001 and R143;
- 6) *Txxxx* is used for the tile number;
- 7) *product discriminator* represents also a date field and is used to determine the difference between end user products sensed at the same date;
- 8) *SAFE* is the product format.

TABLE I
S2 PRODUCTS USED FOR EXPERIMENTAL RESULTS

Mission	Level	PDSG	Relative Orbit Number	Tile number	Acquisition date	City	Country
Training							
S2A	L2A	N0213	R030	T56HKG	December 31, 2019	Wollongong	Australia
S2B	L2A	N0300	R022	T32TQQ	April 8, 2021	Ravenna	Italy
Testing							
S2B	L2A	N0214	R050	T35TMK	March 16, 2020	Bucharest	Romania
S2A	L1C	N0206	R007	T35TPL	December 9, 2017	Tulcea	Romania
S2A	L2A	N0214	R007	T35UQS	April 17, 2020	Kyiv	Ukraine
S2A	L2A	N0301	R010	T26WMD	August 30, 2021	Sermersooq	Greenland

TABLE II
CUSTOM NETWORK SETTINGS FOR EACH MODEL TRAINED

	B1	B2	B3	B4	B5	B6	B7	B8	B8A	B9	B11	B12
Batch dim	18	36	36	36	36	36	36	18	36	36	36	18
Epochs	10	10	10	10	10	5	5	10	10	10	10	5

Taking into account this naming convention, Table I shows the specific products used for training and testing in the proposed method. The first six columns define the principal fields of the naming convention, which help to uniquely identify the products, and the last two columns position the products in the earth's geographical space by specifying the city and country. S2 data can be downloaded for free from the Copernicus Open Access Hub [42].

Each S2 image has 12 spectral bands, being processed at level 2A. The 12 bands initially have different spatial resolution; therefore, in order to have a uniform resolution, we performed an upsampling operation using the nearest neighbor method to bring all the bands to a 10-m GSD. The size of each resulting image is $10\,980 \times 10\,980 \times 12$.

To create the training dataset, a subset of $10\,944 \times 10\,944 \times 12$ was taken from each of the two images shown in Table I, as intended for this purpose. Next, the subset was divided into patches of $304 \times 304 \times 12$, resulting in a number of 1269 patches from each image. Finally, the sets of patches obtained from images were concatenated, thus creating a training dataset of 2592 patches with a size of $304 \times 304 \times 12$.

For testing, each of the images went through the same process as for training, upsampling, subset selection, and patching process, except for the final concatenation. In that way, each image could be passed through the testing process sequentially.

B. Implementation Details

The proposed method implementation was achieved using Python 3.6.13 and TensorFlow 2.3.1 for GPU. Training step was performed on a distributed system containing an Intel(R) Xeon(R) E5-2620v4@2.10 GHz CPU and eight PCIe-connected Tesla K80 GPUs, with 12 GB of RAM each.

As the number of bands of an S2 image is 12, we trained 12 models, one for the reconstruction of each band. The networks were trained having different batch sizes and number of epochs. Table II displays the corresponding batch number and epochs for each model. These numbers were chosen experimentally, after many tries, choosing the ones with better accuracy and lower loss. The duration of training for one model was about 45 min.

The others parameters were set identically for all the trained models. The weights were initialized using *he_normal* [43], and stochastic gradient descent, through the Adam method [44], was chosen for optimization. Learning rate was set to $1e-3$. The loss function used was mean square error (MSE) and was computed between the real band and the reconstructed one.

The filters for the convolutional layers were set in the following way: first two had 88 filters, next two 704, following another two had 1408, next three had 704, another three had 88, one had 11, and the last one had one filter.

For numerical stability, the pixel reflectance values were scaled so that the resulting interval be $[0, \dots, 255]$, according to the following formula:

$$\hat{p} = \frac{p - \min x}{\max x - \min x}, \quad (1)$$

where p represents the radiance value of a pixel from a band and x is the input image. The scaling operation was performed imagewise, before transforming it into a set of $304 \times 304 \times 12$ patches.

C. Evaluation Metrics

In order to quantitatively evaluate the results obtained by the proposed method, different state-of-the-art indexes that measure the accuracy of spatial and spectral profiles' preservation were used. As a consequence, the reconstruction accuracy can be determined and studied. Most of these metrics have demonstrated their use in computer science research and were endorsed also in the remote sensing domain.

1) *Root-Mean-Square Error (RMSE)*: RMSE is a very commonly metric used to measure differences between true values and the ones obtained by an estimator or a model. The formula that defines RMSE is

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (\hat{x} - x)^2}, \quad (2)$$

where \hat{x} is the reconstructed band, x is the initial band, and n is the number of pixels in a band.

2) *Structural Similarity Index Measure (SSIM)*: SSIM measures the similarity between two images, x and \hat{x} , by taking into consideration three comparison measurements: luminance (l), contrast (c) and structure (s), being generally defined as [45]

$$\text{SSIM}(x, \hat{x}) = l(x, \hat{x}) \cdot c(x, \hat{x}) \cdot s(x, \hat{x}). \quad (3)$$

Luminance (l) comparison is defined by the formula

$$l(x, \hat{x}) = \frac{2\mu_x\mu_{\hat{x}} + c_1}{\mu_x^2 + \mu_{\hat{x}}^2 + c_1}, \quad (4)$$

where μ_x is the average of pixel values in x , $\mu_{\hat{x}}$ is the average of pixel values in \hat{x} , and c_1 represents a constant defined by $(K_1 \sim L)^2$. K_1 is a constant $\ll 1$ (0.01) and L is the dynamic range and usually equals 2^{Nr-1} , with Nr being the number of bits per pixel.

Contrast (c) comparison is defined by the following formula:

$$c(x, \hat{x}) = \frac{2\sigma_x\sigma_{\hat{x}} + c_2}{\sigma_x^2 + \sigma_{\hat{x}}^2 + c_2}, \quad (5)$$

where σ_x and $\sigma_{\hat{x}}$ are the variances of x and \hat{x} , respectively. c_2 represents a constant defined by $(K_2 \sim L)^2$. K_2 is a constant equal to 0.03.

The following formula stands for the structure (s) comparison:

$$c(x, \hat{x}) = \frac{\sigma_{x\hat{x}} + c_3}{\sigma_x \sigma_{\hat{x}} + c_3}, \quad (6)$$

where $c_3 = \frac{c_2}{2}$.

As a result, the final formula that defines SSIM is

$$\text{SSIM}(x, \hat{x}) = \frac{(2\mu_x \mu_{\hat{x}} + c_1)(2\sigma_{x\hat{x}} + c_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + c_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + c_2)}. \quad (7)$$

SSIM has a predecessor, which is called ‘‘universal quality index’’ (UQI) defined in [46] and [47]. UQI stands for the special case of SSIM, where $c_1 = c_2 = 0$, but returns unstable results when the sum of averages or variances is very close to 0.

3) *Signal-to-Reconstruction Error (SRE)*: SRE [34] is a useful metric in the case of images that contain clouds, fog, or other phenomena that determine high reflectance values. These high values also generate large absolute reflectance errors. Considering that this metric measures the error relative to the mean image intensity, it is an optimal solution to evaluate the differences between the reconstructed band and the original one. SRE has the property of compensating the effect generated by the high reflectance values.

The formula that defines SRE computation between two bands, x , and its reconstruction, \hat{x} , is illustrated in the following equation:

$$\text{SRE} = 10 \log_{10} \frac{\mu_x^2}{\|\hat{x} - x\|^2/n}, \quad (8)$$

where μ_x^2 is the average of pixel values in x . Resulting values of SRE are expressed in decibels (dB).

4) *Peak-Signal-to-Noise Ratio (PSNR)*: PSNR is delivered as a logarithmic quantity and is expressed in dB. It is a very commonly used metric that expresses a global measurement of image quality. In comparison with SRE, it is not very suited to computing errors between images with different brightness because the peak intensity remains constant [34].

PSNR is defined by the following formula [48]:

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}(x, \hat{x})} \right), \quad (9)$$

where $\text{MSE}(x, \hat{x})$ is defined as

$$\text{MSE}(X, Y) = \frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2 \quad (10)$$

standing for mean squared error, being an often used measure to compute the error function in neural network models.

Optimal PSNR values differ depending on the data type used to represent the pair of images between which the metric is computed [49], [50], [51], [52].

5) *Spectral Angle Mapper (SAM)*: The SAM [53] represents the angular deviation between initial and reconstructed spectral signatures. SAM values are given in degrees. This metric is very useful in the case of spectral images because it ignores brightness, measuring how precise the spectral distribution of a

pixel is respected in the reconstructed band, by comparison with the initial band.

The formula that defines the computation of SAM pixelwise is

$$\text{SAM}(x, \hat{x}) = \cos^{-1} \left(\frac{\sum_{i=1}^{nb} x_i \hat{x}_i}{\sqrt{\sum_{i=1}^{nb} x_i^2} \sqrt{\sum_{i=1}^{nb} \hat{x}_i^2}} \right), \quad (11)$$

where nb is the number of bands.

As the proposed method compares images with only one band, the SAM is computed for each pixel, and the final value represents an average over the whole image.

The metric implementation used in this article was the one proposed by Müller [56]. The code, implementation details, and instructions for usage are available on GitHub [57].

D. Results and Discussion

This article introduces a method to restrain the information contained by a multispectral image in order to reconstruct a missing band. The performances of UNetBRec are quantitatively and qualitatively evaluated. For quantitative analysis, state-of-the-art image reconstruction assessment metrics are computed for UNetBRec and other latest generation methods in order to critically study their achievements. The qualitative analysis is performed by visually comparing the results obtained with the ground truth.

1) *Quantitative Analysis*: The main evaluation metrics of our quantitative comparison are presented in Section V-C, namely RMSE, SSIM, SRE, PSNR, and SAM. UNetBRec has 12 versions, one for each band reconstruction. An overall comparison, as regards used metrics, shows that depending on the resolution of the band, the image for which it is tested, and the differences between the reflectance values, the evaluation metrics lie within an acceptable range in terms of performance. For example, as can be seen from Table III, in the case of band B9, the RMSE registers a higher value compared to that of other bands, which means a bigger error, but the SRE has a value that is among the best values. This is mainly due to the fact that the SRE measures error relative to signal strength, which RMSE does not. For the same reason, the PSNR does not have such high values, compared to other bands. The SAM is, however, the metric according to which the spectral signature is preserved and highlights the good accuracy of the reconstruction.

For each band, the minimum, maximum, and average values were calculated on the entire test dataset. The networks trained for bands B11 and B12 have been evaluated to obtain poorer metric values, but their accuracy remains high, due to the fact that these two bands are part of SWIR and have wavelengths quite far from the bands used for their reconstruction. B12, being right on the edge of this range, records the weakest results at the average values, but if the minimum values are also observed, it can be deduced that for some test data, the accuracy is among the best.

Moreover, the 60-m band reconstruction also encounters lower metric evaluation, but it also can be explained. Amid

TABLE III
QUANTITATIVE EVALUATION OF UNETBREC USING RMSE, SSIM, PSNR, SRE, AND SAM METRICS

UNetBRec	Metrics	RMSE	SSIM (%)	PSNR (dB)	SRE (dB)	SAM (degree)
B1	min	11.10	97.62	55.36	44.56	1.02
	max	17.06	99.73	59.09	67.56	2.01
	average	13.94	99.01	57.25	57.65	1.60
B2	min	4.27	99.47	53.65	52.66	0.26
	max	20.78	99.89	67.40	65.79	1.52
	average	12.05	99.64	60.04	59.58	0.90
B3	min	3.69	99.91	66.78	55.55	0.22
	max	4.58	99.96	68.67	72.49	0.63
	average	4.13	99.93	67.71	64.22	0.40
B4	min	4.43	99.80	55.55	55.49	0.38
	max	16.69	99.89	67.07	66.83	1.08
	average	8.95	99.86	62.49	61.83	0.79
B5	min	4.63	99.84	65.38	57.85	0.40
	max	5.38	99.96	66.69	72.01	1.31
	average	4.98	99.91	66.08	65.01	0.78
B6	min	4.63	99.61	52.33	63.12	0.47
	max	24.19	99.91	66.68	65.18	1.03
	average	11.79	99.80	60.89	64.43	0.81
B7	min	7.09	99.40	51.53	62.00	0.26
	max	26.52	99.83	62.99	65.12	1.05
	average	13.92	99.67	58.77	63.81	0.77
B8	min	11.22	99.21	56.85	60.53	0.54
	max	14.37	99.86	59.00	67.99	1.67
	average	13.31	99.47	57.57	63.77	0.99
B8A	min	8.04	99.58	58.62	61.06	0.16
	max	11.72	99.78	61.89	68.59	0.91
	average	9.58	99.70	60.48	64.93	0.50
B9	min	10.11	99.03	49.47	61.51	0.40
	max	33.60	99.79	59.91	63.85	0.97
	average	18.26	99.51	56.17	62.81	0.65
B11	min	5.17	94.07	50.73	51.42	0.16
	max	29.09	99.91	65.74	64.82	1.07
	average	15.05	97.90	58.57	60.01	0.56
B12	min	13.36	92.10	47.84	52.94	0.23
	max	40.54	98.98	57.48	57.22	2.20
	average	29.93	94.85	51.41	54.62	0.98

Bold values represent the average values band-wise.

reconstruction, these bands benefit from a resolution improvement, so their comparison with the initial bands determines the differences in metric evaluations.

As baseline methods, the ones proposed by Lanaras et al. [34], i.e., DSen2 and VDSen2, and by Brodu [33], i.e., Superres, are used. Although their purpose is different, that of obtaining super-resolution, the main reasons for the comparison are the use of the same type of data, S2 images, and the similarity to a certain extent in terms of approach. The main difference is defined by the lack of bands to reconstruct in training in the case of UNetBRec, while the two methods use all the bands to obtain an improved version from the point of view of resolution. In conclusion, UNetBRec obtains the desired band using the complementary ones, while the competition use all the bands to retrieve super-resolution ones.

Average results over all the test images and all the bands are displayed in Table IV. B2, B3, B4, and B8 are not included in this average quantization as the other methods use the 10-m-resolution bands as grounds to obtain the super-resolution in the others. The state-of-the-art Superres yields rather poor results, while DSen2 and VDSen2 perform similarly, with VDSen2 being slightly better in most of the error metrics. UNetBRec

TABLE IV
AVERAGE COMPUTED METRICS COMPARED BETWEEN STATE-OF-THE-ART METHODS AND UNETBREC

	RMSE	SRE	SAM
Superres	68.95	46.01	1.02
DSen2	25.14	51.10	0.78
VDSen2	23.31	51.73	0.76
UNetBRec	14.68	61.66	0.81

Bold values represent the average values.

TABLE V
QUANTITATIVE EVALUATION BETWEEN UNETBREC AND STATE-OF-THE-ART METHODS SUPERRES, DSEN2, AND VDSEN2 OVER RMSE AND SRE METRICS, BANDWISE

	B1	B5	B6	B7	B8A	B9	B11	B12
RMSE								
Superres	89.78	47.61	67.21	76.85	67.01	86.24	68.92	47.97
DSen2	26.42	19.43	35.12	35.24	14.27	21.59	22.74	26.30
VDSen2	19.65	15.82	33.74	35.79	13.89	20.45	22.07	25.10
UNetBRec	13.94	4.98	11.79	13.92	9.58	18.26	15.05	29.93
SRE								
Superres	47.08	46.10	44.79	44.11	44.47	43.02	49.88	48.63
DSen2	52.08	51.20	49.89	49.21	49.57	48.12	54.98	53.73
VDSen2	52.69	51.42	50.96	49.87	50.05	48.79	55.84	54.24
UNetBRec	57.65	65.01	64.43	63.81	64.93	62.81	60.01	54.62

Bold values represent the average values band-wise.

reduces the RMSE with 37% and increases the SRE with more than 10 dB, thus demonstrating superiority. The SAM error metric is the only exception, encountering a decrease compared with VDSen2 by 4%, and 1.2% referring to DSen2, but an increase by 22.5% relating to Superres. The difference in SAM values is generated by the resolution improvements and comparison with the initial band. Taking this into account, UNetBRec performs and generates very satisfactorily regarding the purposed scope, approaching and even surpassing state-of-the-art methods that use information of all the bands.

On a bandwise analysis, UNetBRec proves its superiority, as shown in Table V. Bands B2, B3, and B4 are not included in this comparison because in the other methods, they remain unchanged and are used as references to obtain the super-resolution of the complementary bands.

All methods' performances vary depending on the band, the differences fitting in the range of what can be called a satisfactory performance. Band B5 turns out to register promising error values along all the methods. This could be explained by the small wavelength difference between it and bands B2, B3, and B4, which, having the best resolution, 10 m, contain a lot of information in both the spectral and spatial domains.

Along all the bands, the superiority of UNetBRec is emphasized through error metric results over state-of-the-art methods.

2) *Qualitative Analysis*: The qualitative analysis of the results obtained with UNetBRec involves comparing the reconstructed band with the original one, also visualizing the difference computed pixelwise between them.

The following examples will highlight two important aspects: the accuracy of the method on scenes from areas of the globe for which training was not carried out, such as the North Pole, and the accuracy on scenes containing atmospheric phenomena, such as thin clouds. The two examples are representative because they show niche cases as regards the large amount of high reflectance values, which cause an increase in the obtained range of pixel

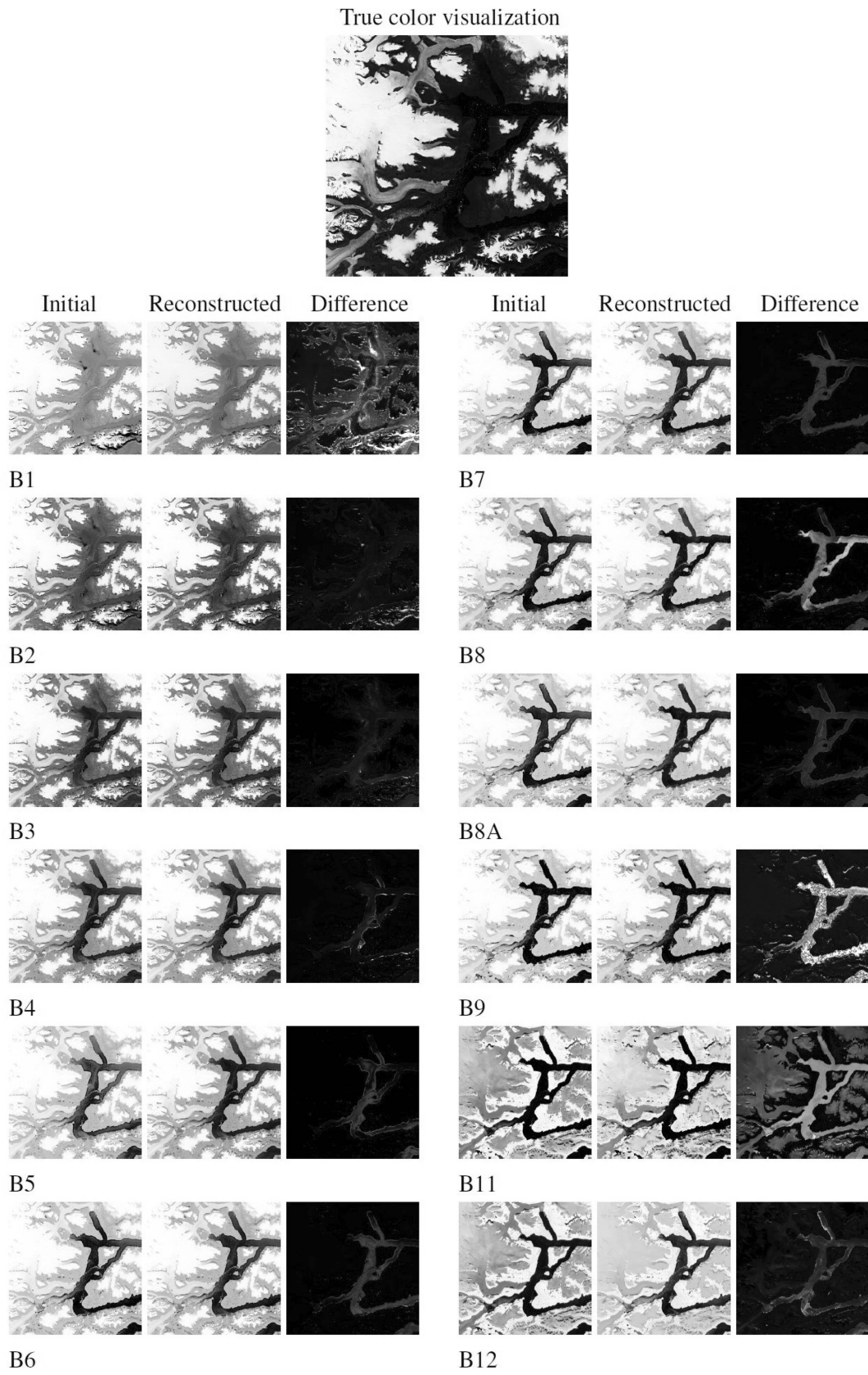


Fig. 17. Bandwise qualitative evaluation of UNetBRec over a scene from the North Pole.

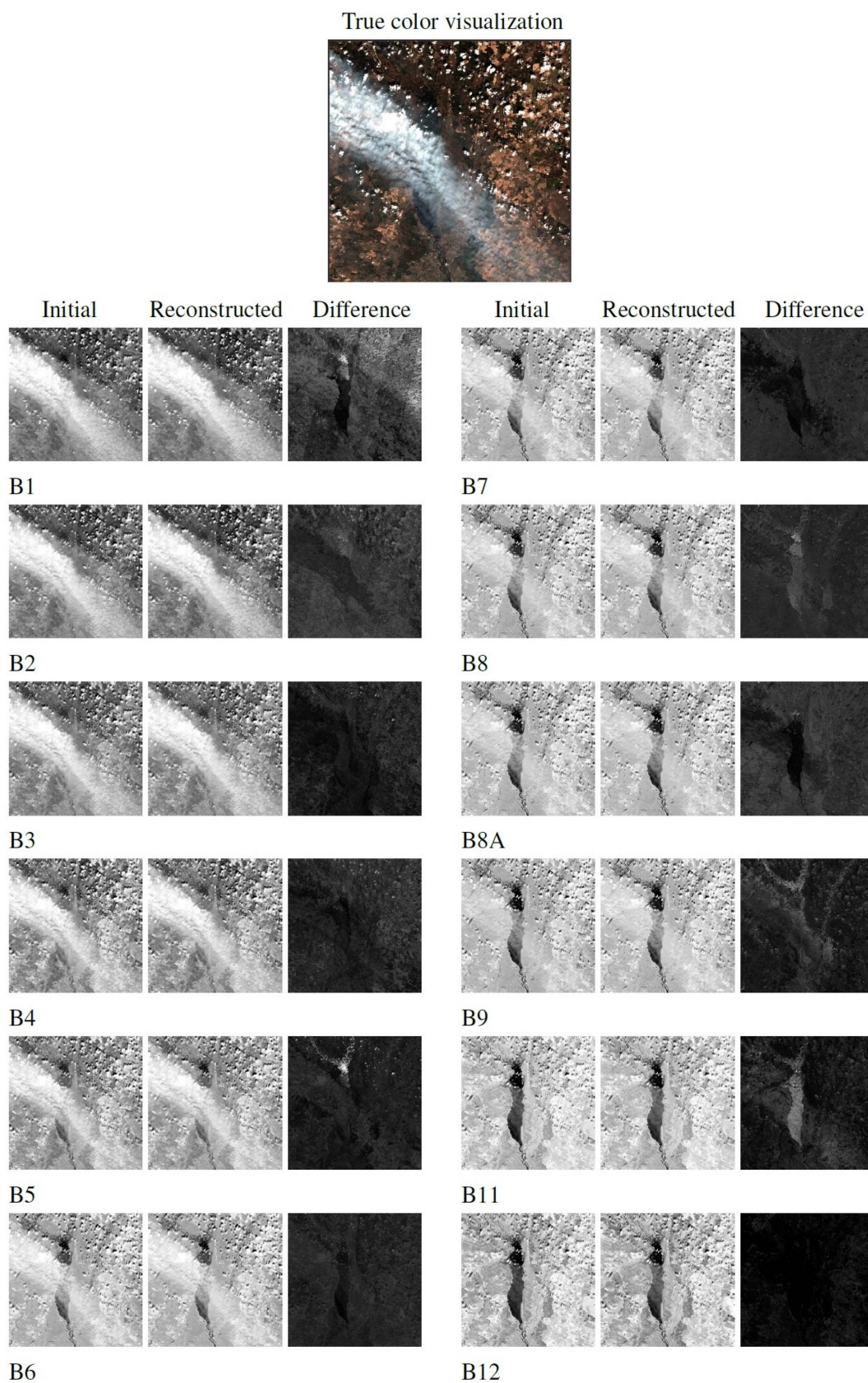


Fig. 18. Bandwise qualitative evaluation of UNetBRec over a scene covered by fog and thin clouds in Kyiv, Ukraine.

values. Thus, these examples represent a strong performance demonstration of UNetBRec regarding the recovery of a band.

Fig. 17 illustrates a scene from the North Pole, which is analyzed on each band by comparing ground truth, reconstruction, and difference bands. Taking into account the discussion on the quantitative analysis and the calculated errors, the observations made remain valid as follows.

- 1) 60-m-resolution bands highlight differences in terms of the contours of the objects, with the difference image recording higher values (marked in white) in those areas.
- 2) SWIR area bands also show a greater difference due to their isolation in an area of the spectrum where there are not so many bands to help the reconstruction with a sufficiently high accuracy.
- 3) Visual spectrum bands benefit from a reconstruction with high accuracy, which makes the visualization of the difference contain many small values (marked in black).

The area most often marked as different is water. It generally registers very low reflectance values, which in the present case represents a very large contrast to the values associated with the considerable area of snow, which surrounds it. These errors along high-contrast edges usually are generated in cases of blurred edges or contrast intensification.

Fig. 18 illustrates one example of reconstruction, which is not influenced by high contrast between the edges of the reflectance values range. The reconstruction of any of the bands is carried out with high accuracy, which can be seen from the visualization of the difference. The method exhibits only slight traces of a behavior involving large errors along high-contrast edges.

The bands from the NIR and SWIR part of the EM spectrum, as they are acquired at a longer wavelength, succeed in capturing information about the earth's surface even through thin clouds, the information content being enriched from one band to another. The rendering of the reconstruction error highlighted by the difference exhibits low values, highlighting the ability of the proposed method to recover the missing band from the concurrent ones. The difference in content and the high reconstruction accuracy are the main factors that demonstrate the power and efficacy of the proposed method.

VI. CONCLUSION

This article provided an efficient and rapid neural-network-based method to recover a missing spectral band of an S2 product. The data needed to perform the reconstruction were obtained from the complementary spectral bands. The starting point of this method was a CNN, namely U-Net. Modifications consisting in reducing the number of convolutional levels and changing the dimensions of the parameters used were essential for its adaptation to our purpose, U-Net being initially used for segmentation. Each band of a multispectral S2 image involved training a model intended to reconstruct it. The adapted network, UNetBRec, demonstrated both quantitatively and qualitatively its improved effectiveness among baseline methods. Moreover, this method proved the efficiency of band reconstruction regardless of land-cover diversity, our datasets' scenes being randomly distributed across the world.

The main advantages of this method are as follows.

- 1) No labeled datasets are required.
- 2) It does not require additional information from other sensors or from the same sensor, but at a short time interval.
- 3) It is an unsupervised method.
- 4) The corrupt band is not used based on the assumption that it does not exist.

We stated that it is an unsupervised method that does not require any labeled datasets due to the fact that the reconstructed band is learned from the other, existing, noise-free bands.

The following disadvantages may be considered: the existence necessity of the complementary bands to the one for which the reconstruction is carried out and the fact that the method it was not trained and tested multisensor .

Considering the balance between pros and cons of the above method, future research will focus on the generalization for multiple sensors and the possibility to derive a super-resolution method.

REFERENCES

- [1] H. Shen et al., "Missing information reconstruction of remote sensing data: A technical review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 61–85, Sep. 2015.
- [2] C. Guillemot and O. Le Meur, "Image inpainting: Overview and recent advances," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 127–144, Jan. 2014.
- [3] C. Zhang, W. Li, and D. Travis, "Gaps-fill of SLC-off landsat ETM satellite image using a geostatistical approach," *Int. J. Remote Sens.*, vol. 28, no. 22, pp. 5103–5122, 2007.
- [4] C. Yu, L. Chen, L. Su, M. Fan, and S. Li, "Kriging interpolation method and its application in retrieval of modis aerosol optical depth," in *Proc. 19th Int. Conf. Geoinf.*, 2011, pp. 1–6.
- [5] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE Trans. Image Process.*, vol. 10, no. 8, pp. 1200–1211, Aug. 2001.
- [6] A. Maalouf, P. Carré, B. Augereau, and C. Fernandez-Maloigne, "A Bandelet-based inpainting technique for clouds removal from remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2363–2371, Jul. 2009.
- [7] T. F. Chan and J. Shen, "Nontexture inpainting by curvature-driven diffusions," *J. Vis. Commun. Image Represent.*, vol. 12, no. 4, pp. 436–449, 2001.
- [8] H. Shen and L. Zhang, "A MAP-based algorithm for destriping and inpainting of remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, May 2009.
- [9] Q. Cheng, H. Shen, L. Zhang, and P. Li, "Inpainting for remotely sensed images with a multichannel nonlocal total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 175–187, Jan. 2009.
- [10] L. Wang, J. J. Qu, X. Xiong, X. Hao, Y. Xie, and N. Che, "A new method for retrieving band 6 of aqua MODIS," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 2, pp. 267–270, Apr. 2006.
- [11] P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Restoration of aqua modis band 6 using histogram matching and local least squares fitting," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 613–627, Feb. 2009.
- [12] H. Shen, C. Zeng, and L. Zhang, "Recovering reflectance of AQUA MODIS band 6 based on within-class local fitting," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 1, pp. 185–192, Mar. 2011.
- [13] I. Gladkova, M. D. Grossberg, F. Shahriar, G. Bonev, and P. Romanov, "Quantitative restoration for MODIS band 6 on AQUA," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2409–2416, Jun. 2012.
- [14] X. Li, H. Shen, L. Zhang, H. Zhang, and Q. Yuan, "Dead pixel completion of AQUA MODIS band 6 using a robust m-estimator multiregression," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 4, pp. 768–772, Apr. 2014.
- [15] L. Rout, "Alert: Adversarial learning with expert regularization using Tikhonov operator for missing band reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 4395–4405, Jun. 2020.

- [16] L. Rout, Y. Bhatija, A. Garg, I. Mishra, S. M. Moorthi, and D. Dhar, "DeepSWIR: A deep learning based approach for the synthesis of short-wave infrared band using multi-sensor concurrent datasets," 2019, *arXiv:1905.02749*.
- [17] C. Zeng, H. Shen, and L. Zhang, "Recovering missing pixels for landsat ETM SLC-off imagery using multi-temporal regression analysis and a regularization method," *Remote Sens. Environ.*, vol. 131, pp. 182–194, 2013.
- [18] C.-H. Lin, K.-H. Lai, Z.-B. Chen, and J.-Y. Chen, "Patch-based information reconstruction of cloud-contaminated multitemporal images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 163–174, Jan. 2014.
- [19] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 442–455, Feb. 2006.
- [20] J. Zhang, M. K. Clayton, and P. A. Townsend, "Missing data and regression models for spatial images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1574–1582, Mar. 2015.
- [21] C. Zeng, H. Shen, M. Zhong, L. Zhang, and P. Wu, "Reconstructing modis LST based on multitemporal classification and robust regression," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 512–516, Mar. 2015.
- [22] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh, "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter," *Remote Sens. Environ.*, vol. 91, nos. 3/4, pp. 332–344, 2004.
- [23] P. Jönsson and L. Eklundh, "TIMESAT—A program for analyzing time-series of satellite sensor data," *Comput. Geosci.*, vol. 30, no. 8, pp. 833–845, 2004.
- [24] W. Zhu, Y. Pan, H. He, L. Wang, M. Mou, and J. Liu, "A changing-weight filter method for reconstructing a high-quality NDVI time series to preserve the integrity of vegetation phenology," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1085–1094, Apr. 2012.
- [25] L. Lorenzi, F. Melgani, and G. Mercier, "Missing-area reconstruction in multispectral images under a compressive sensing perspective," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 3998–4008, Jul. 2013.
- [26] X. Li, H. Shen, L. Zhang, H. Zhang, Q. Yuan, and G. Yang, "Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 7086–7098, Nov. 2014.
- [27] Q. Cheng, H. Shen, L. Zhang, Q. Yuan, and C. Zeng, "Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model," *ISPRS J. Photogrammetry Remote Sens.*, vol. 92, pp. 54–68, 2014.
- [28] S. Benabdelkader and F. Melgani, "Contextual spatio-spectral postreconstruction of cloud-contaminated images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 204–208, Apr. 2008.
- [29] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, Aug. 2018.
- [30] Z. Xiong, Q. Guo, M. Liu, and A. Li, "Pan-sharpening based on convolutional neural network by using the loss function with no-reference," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 897–906, 2021.
- [31] O. Sidorov and J. Y. Hardeberg, "Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops*, 2019, pp. 3844–3851.
- [32] S. Lou, Q. Fan, F. Chen, C. Wang, and J. Li, "Preliminary investigation on single remote sensing image inpainting through a modified GAN," in *Proc. 10th IAPR Workshop Pattern Recognit. Remote Sens.*, 2018, pp. 1–6.
- [33] N. Brodu, "Super-resolving multiresolution images with band-independent geometry of multispectral pixels," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4610–4617, Aug. 2017.
- [34] 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 J. Photogrammetry Remote Sens.*, vol. 146, pp. 305–319, 2018.
- [35] A. Rangnekar, N. Mokashi, E. Ientilucci, C. Kanan, and M. Hoffman, "Aerial spectral super-resolution using conditional adversarial networks," 2017, *arXiv:1712.08690*.
- [36] B. Rodríguez-Suárez, P. Quesada-Barriuso, and F. Argüello, "Design of CGAN models for multispectral reconstruction in remote sensing," *Remote Sens.*, vol. 14, no. 4, 2022, Art. no. 816.
- [37] G. Tsagkatakis, A. Aidini, K. Fotiadou, M. Giannopoulos, A. Pentari, and P. Tsakalides, "Survey of deep-learning approaches for remote sensing observation enhancement," *Sensors*, vol. 19, no. 18, 2019, Art. no. 3929.
- [38] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2015, pp. 234–241.
- [39] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440.
- [40] ESA, "Sentinel-2 Mission," Accessed: Aug. 20, 2022. [Online]. Available: <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-22022>.
- [41] ESA, "Sentinel-2 MSI—Naming Convention," 2022. Accessed: Aug. 20, 2022. [Online]. Available: <http://https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/naming-convention/>
- [42] ESA, "Copernicus Open Access Hub," 2022. Accessed: Aug. 19, 2022. [Online]. Available: <https://scihub.copernicus.eu/dhus/#/home>
- [43] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1026–1034.
- [44] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, Dec. 2014.
- [45] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [46] Z. Wang, "Rate-scalable foveated image and video communications," doctoral dissertation, Dept. of ECE, Univ. Texas, Austin, TX, USA, 2001.
- [47] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
- [48] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. 20th Int. Conf. Pattern Recognit.*, 2010, pp. 2366–2369.
- [49] O. S. Faragallah et al., "A comprehensive survey analysis for present solutions of medical image fusion and future directions," *IEEE Access*, vol. 9, pp. 11358–11371, 2021.
- [50] N. Chervyakov, P. Lyakhov, and N. Nagornov, "Analysis of the quantization noise in discrete wavelet transform filters for 3D medical imaging," *Appl. Sci.*, vol. 10, no. 4, 2020, Art. no. 1223.
- [51] S. T. Welstead, *Fractal and Wavelet Image Compression Techniques*, vol. 40. Bellingham, WA, USA: SPIE, 1999.
- [52] H. Raouf, S. Dietmar, and M. Barni, *Fractal Image Compression Document and Image Compression*. Boca Raton, FL, USA: CRC Press, 2011.
- [53] R. H. Yuhas, A. F. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," in *Proc. 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, pp. 147–149.
- [54] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, p. 594, 2016.
- [55] J. Yang, X. Fu, Y. Hu, Y. Huang, X. Ding, and J. Paisley, "Pannet: A deep network architecture for pan-sharpening," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5449–5457.
- [56] M. Müller, N. Ekhtiari, R. Almeida, and C. Rieke, "Super-resolution of multispectral satellite images using convolutional neural networks," *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. V-1- 2020, pp. 33–40, 2020.
- [57] UP42, "Image Similarity Measures Package," 2022. Accessed: Jul. 7, 2022. [Online]. Available: <https://github.com/up42/image-similarity-measures>



Iulia Coca Neagoe received the B.S. degree in computer science and information technology from the "Ferdinand I" Military Technical Academy, Bucharest, Romania, the M.S. degree in computer science from the University of Bucharest, Romania in 2017, and the Ph.D. degree in electronics, telecommunications, and information technology from the Politehnica University of Bucharest, Romania, in 2023.

She joined the CEOSpaceTech Center, Politehnica University of Bucharest, as a Research Assistant, in 2016. Since then, she has taken part in several research projects. Since 2015, she has been working as an IT Engineer within the Ministry of National Defense and covers subjects such as integrated systems or information technologies and web programming technologies. In 2018, she participated on a three-month internship with the German Aerospace Center, Weßling, Germany, and also on BigSkyEarth Training School, which took place at the Vicomtech Research Center, San Sebastian, Spain. Her research interests include geospatial content-based information retrieval, sensor deep learning, and analytics in multispectral earth observation images.



Daniela Faur (Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the Politehnica University of Bucharest (UPB), Bucharest, Romania, in 2002 and 2006, respectively.

Since 2007, she has been a Professor in information theory with the Department of Applied Electronics and Information Engineering, Faculty of Electronics, Telecommunications and Information Technology, UPB. She is involved in national and European research grants in the field of earth observation data processing and visualization targeting applications in agriculture, disaster and humanitarian crisis management, and biodiversity monitoring. In 2010, she cofounded CEOSpaceTech, the Research Center for Spatial Information that promotes technological and scientific areas of research in earth observation and related fields. She acted as a Project Manager for “GEODIM-Platform for Geo-Information in Support of Disaster Management,” “LEOSITS-Long Term Data Exploitation for Satellite Image Time Series-Extraction of Classes for Scene Dynamic,” and “VATEO—Visual Analytics Tool for Earth Observation Images,” and as a Principal Investigator for “eVADE-ESA ITT-Interactive Visual Analysis Tool for Earth Observation Data.” As a result of winning Copernicus Incubator in 2019, she cofounded OGOR, an agritech startup that builds a live journal based on Copernicus satellite images. She serves as a Europe Principal Investigator in the frame of ESA NASSC Dragon 5 Cooperation, coordinating the project on “Large-scale spatial-temporal analysis for dense satellite image series with deep learning” for the term 2020–2024. From 2019 to 2021, she represented UPB in H2020 SPACE-END: ENDEAVOUR Space HUB. She currently leads, as UPB representative, the H2020 CENTURION: COPERNICUS datacube/AI data cube services for society, industry, and new market generation H2020-SPACE-2018-2021 for the term 2020–2024.

Dr. Faur is a Member of the IEEE Geoscience and Remote Sensing Society.



Corina Vaduva received the B.S. and Ph.D. degrees in electronics and telecommunications from the Politehnica University of Bucharest (UPB), Bucharest, Romania, in 2007 and 2011, respectively.

In 2010, she performed a six-month internship with the German Aerospace Center (DLR), Oberpfaffenhofen, Germany, as a Visiting Ph.D. Student with the Remote Sensing Technology Institute. In 2012, she returned to DLR as a Visiting Scientist for two months. She has been a Research Engineer with the Research Centre for Spatial Information, UPB, since 2007. Her research interests include earth observation image processing and data mining based on information theory, information retrieval, data fusion, change detection, temporal analysis in image time series, and semantic annotation for very high resolution image understanding and interpretation.

Dr. Vaduva was a recipient of a competition prize in the Digital Globe 8 Bands Research Challenge in 2010. She was the Chair and Program Committee Member of the ESA Conference on Big Data from Space in 2016, 2017, and 2019.



Mihai Datcu (Fellow, IEEE) received the M.S. and Ph.D. degrees in electronics and telecommunications from the Politehnica University of Bucharest (UPB), Bucharest, Romania, in 1978 and 1986, respectively, and the Habilitation *a Diriger Des Recherches* degree in computer science from Louis Pasteur University, Strasbourg, France, in 1999.

From 1992 to 2002, he was an Invited Professor with the Swiss Federal Institute of Technology, Zurich, Switzerland. He was a Visiting Professor with the University of Oviedo, Oviedo, Spain; Louis Pasteur University; International Space University, Strasbourg; the University of Siegen, Siegen, Germany; the University of Innsbruck, Innsbruck, Austria; the University of Alcalá, Alcalá de Henares, Spain; the University of Rome Tor Vergata, Rome, Italy; the Universidad Pontificia de Salamanca, Madrid, Spain; the University of Camerino, Camerino, Italy; and the Swiss Center for Scientific Computing, Manno, Switzerland. Since 1981, he has been a Professor with the Department of Applied Electronics and Information Engineering, Faculty of Electronics, Telecommunications and Information Technology, UPB. Since 1993, he has been a Scientist with the German Aerospace Center (DLR), Weßling, Germany, where he is currently a Senior Scientist and the Data Intelligence and Knowledge Discovery Research Group Leader with the Remote Sensing Technology Institute, and involved in the DLRONERA Joint Virtual Center for AI in Aerospace. He is developing algorithms for model-based information retrieval from high-complexity signals and methods for scene understanding from very high resolution synthetic aperture radar (SAR) and interferometric SAR data. Since 2011, he has been leading the Immersive Visual Information Mining Research Laboratory, Munich Aerospace Faculty. He is the Director of the Research Center for Spatial Information, UPB. Since 2001, he had been initiating and leading the Competence Center on Information Extraction and Image Understanding for Earth Observation, ParisTech, Paris Institute of Technology, Telecom Paris, Paris, France, a collaboration of DLR with the French Space Agency (CNES). He was a Professor with the DLR-CNES Chair, ParisTech, Paris Institute of Technology, Telecom Paris. He has initiated the European frame of projects for image information mining (IIM) and is involved in research programs for information extraction, data mining and knowledge discovery, and data understanding with the European Space Agency (ESA), NASA, and in a variety of national and European projects. He and his team have developed and are developing the operational IIM processor in the Payload Ground Segment systems for the German missions, TerraSAR-X, TanDEM-X, and the ESA Sentinel-1 and Sentinel-2. He is a Member of the ESA Working Group Big Data from Space and Visiting Professor with the ESA's Φ -Lab. His research interests include explainable and physics-aware artificial intelligence, smart sensor design, and quantum machine learning with applications in earth observation.

Dr. Datcu was the recipient of the Best Paper Award and the IEEE Geoscience and Remote Sensing Society Prize in 2006, the National Order of Merit with the rank of Knight, for outstanding international research results, awarded by the President of Romania, in 2008, and the Romanian Academy Prize Traian Vuia for the development of the SAADI image analysis system and his activity in image processing in 1987. He was also the recipient of the Chaire d'excellence internationale Blaise Pascal 2017 for international recognition in the field of data science in earth observation. He has served as a co-organizer for international conferences and workshops and as Guest Editor for a special issue on AI and big data of the IEEE and other journals. He is a representative of Romanian in the Earth Observation Program Board. In 2022, he received the IEEE GRSS David Landgrebe Award in recognition of outstanding contributions to earth observation analysis using innovative concepts for big data analysis, image mining, machine learning, smart sensors, and quantum resources.