Using Super-Resolution Algorithms for Small Satellite Imagery: A Systematic Review

Kinga Karwowska and Damian Wierzbicki

***Abstract*—In recent years, we have witnessed significant development in the space sector, in particular regarding Earth imaging. Small satellites, whose size and construction make their production much cheaper, are becoming increasingly popular. As a result, a larger number of satellites may be placed in space, and thus, they may perform more frequent observations of selected spots on Earth. Unfortunately, the construction of these satellites also affects their observation capacity as they have a weaker spatial resolution. Scientists have been dealing with the problem of improving the spatial resolution of satellite imaging for many years. Numerous methods were developed that allow for the best possible representation of high-resolution images based on low-resolution images. However, the application of traditional solutions to improve the resolution of digital images requires an additional high-resolution image. As far as images obtained by small satellites (e.g., nano, micro, or mini) are concerned, the difference between the spatial resolution of panchromatic and multispectral images is small (e.g., for SkySat-3 – SkySat-15 satellites, it is only 0.16 m). The need to increase the spatial resolution of an image that does not have a corresponding higher resolution image (e.g., a panchromatic image or a sequence of images) causes additional problems. This article presents a review of the methods to improve the spatial resolution of small-satellite imaging. The authors analyze the interpolation, pansharpening, and digital image processing methods. Additionally, the article focuses on presenting solutions based on deep learning that enables the enhancement of the spatial resolution of images obtained from small satellites. The methodology of creating databases used for network training is described. Finally, the authors present the main limitations of the analyzed solutions and future development trends that will enable to improve the spatial resolution with the use of a single image.**

***Index Terms*—Convolutional neural networks, deep learning, neural networks, single image super-resolution (SISR), super-resolution.**

I. Introduction

In RECENT years, we have witnessed dynamic progress in technology, which, in turn, has contributed to a significant development of the space sector. According to the data provided by the Union of Concerned Scientists, as of the 1st of January 2021, over 3000 operational satellites were orbiting the Earth. Every sixth of them performs operational tasks [1]. Over the past five years, over 380 new imaging satellites have been placed in orbit (of which approx. 300 weighed less than 500 kg). This number is more than 4.5× higher than in 2011–2015. High-resolution (HR) satellite imaging may be applied in numerous fields of research, such as land-use and landcover mapping [2], [3], urban mapping [4], detection and tracking of objects [5], [6], maritime monitoring [7], and automatic building classification [8], [9].

Contemporary satellites of very HR can record images with a spatial resolution that may reach, in nadir, even 0.30 m. However, nano-, micro-, and minisatellite systems are becoming increasingly popular [10]. Although they are both easier to manufacture and cheaper than large satellites, they have one major disadvantage (as far as the possibility of Earth observation is concerned): their resolution is significantly lower due to their construction. Their small size and weight (nanosatellites: 1–10 kg, microsatellites: 10–100 kg, and minisatellites: 100–500 kg) result in equipment limitations. The main element of an imaging satellite is its optic system, which, in the case of nano- or microsatellites, differs significantly from those installed in large observation satellites, such as WorldView-3 or QuickBird. Nano- and microsatellites are equipped with small matrices, e.g., complementary metal–oxide–semiconductor [11] or CMOSIS CMV [12], which are characterized by low quantum efficiency. Another limitation of small satellites is the inadequacy of the telescope caused by the much smaller number of applied lenses, which leads to a deteriorated quality of the obtained images, e.g., through blurring. The resolution obtained with the use of low-efficiency matrices and imperfect telescopes is significantly lower, even if low earth orbit satellites are employed. An example may be nSight-1, a nanosatellite designed and manufactured by the privately owned South African company SCS Space. The nanosatellite was placed in a 400 km orbit in 2017. It is equipped with the SCS Gecko imager that consists of the sensor unit (SU), control unit, optics, and mechanical support structure. The 2-Mega Pixel SU and Bayer filter allow it to take RGB images of a bandwidth of 64 km and a spatial resolution of 32 m [13] (see Fig. 1).

Scientists have been attempting to improve the spatial resolution of satellite imaging for many years. Numerous methods were developed that allow for the best possible representation of HR images based on low-resolution (LR) images. These methods may be divided into groups based on the number of images necessary to estimate an HR image. This allows us to distinguish between methods that use information obtained from several images and those that require only one image. Another classification is based on the operation of algorithms. These

Manuscript received January 31, 2022; revised March 21, 2022; accepted April 8, 2022. Date of publication April 14, 2022; date of current version May 6, 2022. This work was supported in part by the Military University of Technology, Faculty of Civil Engineering and Geodesy under Grant 531-4000-22-786. *(Corresponding author: Damian Wierzbicki.)*

The authors are with the Faculty of Civil Engineering and Geodesy, Department of Imagery Intelligence, Military University of Technology, 00-908 Warszawa, Poland (e-mail: kinga.karwowska@wat.edu.pl; damian.wierzbicki@wat.edu.pl).

Digital Object Identifier 10.1109/JSTARS.2022.3167646

Fig. 1. Sample photo taken by nSight-1, GSD= 32 m [13].

TABLE I
SUMMARY OF THE NUMBER OF PUBLICATIONS ON THE ISSUE OF IMPROVING THE RESOLUTION OF DIGITAL IMAGES (INCLUDING SATELLITE IMAGERY) WITH THE USE OF SELECTED INTERPOLATION, PANSHARPENING, AND DEEP LEARNING ALGORITHMS

METHOD		Scopus	Web of Science
INTERPOLATION	Nearest Neighbour	1395	1113
	Bilinear	2710	1911
	Cubic	7995	5052
PANSHARPENING	Intensity-hue-saturation (IHS)	135	141
	Principal component analysis (PCA)	128	138
	Brovey	73	53
	Ehlers	28	33
	Gramm-Schmidt	49	30
	Multiplicative	53	60
CNN	SRGAN	185	106
	SRResNet	21	13
	EDSR	63	107
	ESRGAN	46	35
	FSRCNN	43	20
	SRCNN	195	129
	SAN	204	330
	ESPCN	26	13
	EnhanceNet	8	7

methods are based on interpolation, pansharpening, and deep learning algorithms. The problem of the breadth of increasing digital image resolution has often been discussed in the literature. The most popular and oldest ones are the image interpolation methods (see Table I). Additionally, attention should be paid to the methods based on deep learning algorithms, which were described in a greater number of publications than the most popular and most frequently used (in the case of satellite imagery) pansharpening methods.

The least demanding group of solutions are interpolation methods. They only require one image and employ simple mathematical operations to calculate a new pixel value. Unfortunately, the simplicity of calculations significantly affects the interpretation possibilities of the generated images. Despite the increased spatial resolution, the improvement in interpolation possibilities is low. In the 1970s, space imagery became more common, which led to the necessity to improve satellite imagery resolution. Scientists saw the possibility of combining HR panchromatic (PAN) images with spectral information of multispectral images (MS) with a lower resolution very quickly. This method was called pansharpening. Over the years, many solutions based on this dependence have been developed. The great possibilities of these methods are best evidenced by the fact that they are widely used today. On the other hand, their main limitation is the need to have both HR images and the corresponding LR images.

However, as far as images obtained by small satellites (e.g., nano, micro, or mini) are concerned, the difference between the spatial resolution of PAN and multispectral images is small (e.g., for SkySat-3 – SkySat-15 satellites, it is only 0.16 m). The need to increase the spatial resolution of an image when a corresponding higher resolution image (e.g., a PAN image or a sequence of images) is unavailable causes additional problems. Research on improving spatial resolution based on a single image has been conducted for many years, initially without any noteworthy effects. Finally, a breakthrough was driven by the increased computing power of workstations that enabled deep neural networks to process digital images. They allowed for automatic classification of images [14]–[16], detection and tracking of objects [17]–[20], detection of changes [21]–[23], and segmentation [24]–[28]. They were also used to create algorithms to increase the resolution, e.g., SRCNN [29]. Taking the values of the peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) metrics into consideration, it may be concluded that the use of convolutional neural networks may help improve the resolution of digital images significantly. Another milestone in processing digital images was generative adversarial nets (GAN) [30]—as described in the presentation by Goodfellow. The characteristic feature of this solution is network training. During this training, the generator's results (convolutional network) are assessed by another convolutional network (discriminator). This solution provides many possibilities, including generating new (false images) [30] or performing image translation [32]–[34]. Having these results in mind, scientists used GAN to develop new methods to improve the resolution of digital images based on a single LR image. At the same time, considering the construction of satellite images, which may be treated as classic digital images, all solutions using GAN may improve the spatial resolution of satellite images because it is the structure of the image that is important, not what is shown in it. In 2016, the team led by Ledig *et al.* [35] in the publication photorealistic single image super-resolution (SISR) using a generative adversarial network presented the super-resolution generative adversarial networks (SRGAN)

model—the super-resolution using a generative adversarial network that improves the spatial resolution of images, which received great acclaim. Several other authors introduced various modifications to this solution, including enhanced super-resolution generative adversarial networks (ESRGAN) [36]. However, they do not fully solve the problem of improving the resolution of satellite imaging.

This article reviews the methods to improve spatial resolution, starting from commonly used solutions based on interpolation. Then, the pansharpening methods are discussed, which are some of the most widely used solutions. The subsequent sections discuss digital image processing methods that are used, among others, in the SPOT and Pleiades satellites. Finally, the authors focus on solutions based on deep learning, which, as research has demonstrated, offers a variety of new possibilities. They will improve the spatial resolution of images obtained from nano- and microsatellite systems.

The article provides the answers to the following research questions.

1) What are the main problems related to improving the resolution of satellite imaging?
2) How are resolution improvement algorithms developed?
3) What is the difference between resolution improvement methods of small-satellite imaging?
4) What methodology should be applied to build training and validation data to improve the resolution of imaging from mini and nanosatellites using deep learning algorithms?
5) What methods should be utilized to assess the correctness of the operation of the algorithms that are responsible for the enhancement of spatial resolution?

To evaluate individual methods, quantitative analyses were performed in terms of spectral and spatial quality using traditional quality assessment metrics and visual analysis.

The rest of this article is organized as follows. Section II contains a review of the methods to improve the spatial resolution of satellite imaging and a description of the metrics used to assess the correctness of the operation of the analyzed algorithms. Section III includes results and discussion. Finally, Section IV concludes this article.

II. Related Works

The issue of improving the resolution of images has been known for many years, and the dynamic development of the space sector revealed the urgency of this problem in the context of improving the resolution of satellite imaging. This section presents a review of the methods to improve spatial resolution, starting from the most traditional ones, i.e., those based on interpolation, to the dynamically developing SISR methods that employ deep learning.

A. Interpolation

The most popular methods that enable to assign values to newly created elements of the image matrix are image interpolation algorithms. Thanks to its simplicity and computation speed, interpolation is often used to solve the problem of improving image resolution. It is applied in the SISR methods that use only a single LR image to render a super-resolution image (SR). There are numerous methods of interpolation that allow for creating a new super-resolution image with a higher resolution. Table II presents examples of some popular methods along with their advantages and disadvantages.

TABLE II
Comparison of Sample Interpolation Methods That are Used to Enhance the Spatial Resolution of Satellite Images

Method	Advantages	Disadvantages
Nearest Neighbor , Bilinear, cubic interpolation	- They do not require large computing power - Smooth, visible edges, other methods	- They do not improve the interpretation ability - Lack of clearly visible edges – Nearest Neighbor interpolation
Super-Resolution Reconstruction of Satellite Video Images Based on Interpolation Method [45]	- Visual enhancement of Satellite Video Images - Smooth, visible edges	- No significant improvement in interpretation ability
Integration of the bilinear and the bi-cubic interpolation methods [46]	- Significant visual enhancement of the image, smoothened edges - They do not require large computing power - Possibility of modifing the weight at which interpolation results are combined	- No significant improvement in interpretation ability
Splines interpolation [47]	- The method does not require large computing power. It is faster than traditional interpolation methods - Significant visual enhancement of the image, smoothened edges	- No significant improvement in interpretation ability

The objective of linear interpolation methods is to find u function that will meet the condition specified in the following and will then be used for sampling the input image $I^{LR}_{n,m}$ of the dimensions $n \times m$, being integers [37]

$$I^{LR}_{n,m} = u(m,n) \wedge m,n \in \mathbb{Z} \tag{1}$$

where $I^{\mathrm{LR}}_{n,m}$ is the input image with $n \times m$ size and u denotes the image resampling function.

The main methods include, among others, the nearest neighbor, bilinear, and bicubic interpolation. The most used one is the nearest neighbor method (also called pixel duplication), where the pixel of the resulting image takes the value of the pixel of the original image that is situated nearest to the analyzed point. The u function is presented in (2), and the kernel of the nearest neighbor interpolation is shown as follows:

$$u\ (x,\ y) = I^{LR}_{[x],[y]} \tag{2}$$

$$K\,(x,y) = K_1\,(x)\,K_1\,(y)\,, K_1\,(t) = \begin{cases} 1 \text{ for } -0,5 \le t < 0,5 \\ 0 \text{ for others} \end{cases} \tag{3}$$

where $I^{\mathrm{LR}}_{n,m}$ is the input image with $n \times m$ size, u denotes the image resampling function, $K(x,y)$ is the kernel, $K_1(x)K_1(y)$ denotes the tensor product kernel $K(x,y)$, and t is the processed pixel.

A slightly more elaborate method is the bilinear interpolation, where the pixel of the resulting image takes the average value of four neighboring pixels of the original image. Formula (4) presents u function of the bilinear method, where $\langle x \rangle$ means the fraction part equal to $x - [x]$, while $[\cdot]$ implies the function of rounding real numbers down to integers. Finally, (5) presents the interpolation kernel, where $(\cdot)+$ means the positive part

$$\begin{aligned} u\ (x,\ y) = &\ (1 - \langle x \rangle)\,(1 - \langle y \rangle)\, I^{\mathrm{LR}}_{[x],[y]} \\ &+ x - (1 - \ \langle y \rangle)\, I^{\mathrm{LR}}_{[x]+1,[y]} \\ &+ y\,(1 - \ \langle x \rangle)\, I^{\mathrm{LR}}_{[x],[y]+1} + \langle x \rangle\, \langle y \rangle\, I^{\mathrm{LR}}_{[x]+1,[y]+1} \end{aligned} \tag{4}$$

$$K\,(x,y) = K_1\,(x)\,K_1\,(y)\,, \quad K_1\,(t) = (1 - |t|)^+ \tag{5}$$

where $u\,(x,y)$ denotes the image resampling function, $\langle x \rangle$ denotes the fractional part equal to $x - [x]$, $[x]$ and $[y]$ denote the real numbers that were rounded down to integers function, $K\,(x,y)$ denotes the kernel, $K_1(x)K_1(y)$ denotes the tensor product kernel $K(x,y)$, t denotes the processed pixel, and $(.)^+$ denotes the positive part of a real or extended real-valued function.

Bilinear Interpolation creates affine functions, and it is the simplest method that delivers a continuous function. Thanks to its efficiency and simplicity, it is commonly used in numerous algorithms.

Another method is a bicubic interpolation, where the value is calculated based on the values of all pixels that are adjacent to the analyzed pixel. The interpolation kernel formula (6) contains an additional parameter α, which generates an additional degree of freedom of the system

$$K_1\,(t) = \begin{cases} (\alpha+2)\,|t|^3 - (\alpha+3)\,|t|^2 + 1 \text{ if } \leq 1 \\ \alpha|t|^3 - 5\alpha|t|^2 + 8\alpha|t| - 4\alpha \text{ if } 1 < |t| < 2 \\ \quad 0 \text{ dla pozostałych wartości} \end{cases} \tag{6}$$

where α denotes the parameter creating an additional degree of freedom of the system, $K\,(x,y)$ is the kernel, and t denotes the processed pixel.

The methods presented above enable to assign a brightness value to the newly emerging pixels, but this increases only the size of the image instead of the amount of information that may be read. Interpolation methods are used by numerous algorithms that improve the spatial resolution of images. Apart from the interpolation stage, these algorithms also remove blurring and minimize noise, which allows changing the size of the matrix that represents the output image and improves its quality.

One of such methods is partial volume diffusion, developed by Salvado *et al.* [38]. It enables image interpolation while at the same time recovering the blurred edges. Another solution is the curvature-based technique that is based on bilinear interpolation [39], [40] or the new edge detection interpolation algorithm that uses the concept of self-similarity and provides satisfactory super-resolution quality. Still another solution is the projection onto convex sets (POCS) [41] that is used for interpolation in the fusion process. In this method, pixel values of the SR image are determined as the points of intersection of closed convex sets. The most crucial property of POCS consists in the fact that the values of the newly generated image may be included in the model to enable finding the searched point. Another popular method is iteration backprojection (IBP). In this approach, the SR image is estimated by representing the difference between simulated LR images and observed LR images by blurring the image. The literature on remote sensing features the IBP algorithm, modified by Li *et al.* [42], where different blur kernels are applied for each sensor. Additional elements that allow enhancing the resolution of the estimated image with the use of the discussed methods are blurring and noise reduction. Gilman *et al.* [43] found that optimal kernels are similar for different images, so it is possible to use one kernel for multiple images, and the obtained results are satisfactory. The authors developed two image models that enable capturing the important properties of the image and using them to develop the optimum kernels. The introduction of optimal interpolators allowed combining linear and cubic interpolation. Certain other methods use more than one image to enhance resolution. An example of such a solution is the combination of interpolation and gradient methods. It consists in combining LR images with the target HR image. In this solution, the gradient method iterations are performed in a HR grid of spectral and spatial coordinates [44].

Authors of numerous studies have attempted to use these relations to enhance the spatial resolution of satellite imaging. An example is an approach that uses a series of LR images to generate a single HR image by using the information from moving pixels representing the same object [45]. Another solution consists in integrating the bilinear and the bicubic interpolation methods, where the results of the application of interpolation are combined, being assigned the relevant weights [46]. On the other hand, Malpica [47] presents splines interpolation. This method enhances the radiometric resolution of satellite images using less computational power than classic interpolation methods, such as the nearest neighbor, linear, and cubic interpolation, which makes this method faster, and its results are comparable to those of linear interpolation (see Fig. 2).

B. Pansharpening

The optic systems installed on Earth-imaging satellites enable them to take photographs of various spatial and spectral resolutions. PAN images are characterized by very HR that may reach even up to 0.31 m (civilian satellites). In contrast, multispectral images provide plenty of valuable information, thanks to imaging in multiple spectral bands, but their spatial resolution is approximately four times lower. Fortunately, a method exists that enables to combine HR PAN images with the HR of a multispectral (or hyperspectral) image [48]–[51]. It is referred to as pansharpening. This operation results in generating an HR multispectral image. The main pansharpening methods include intensity-hue-saturation (IHS) [52], [53], principal component analysis (PCA), Brovey color standardization, Ehlers fusion, image multiplication, and Gram–Schmidt orthogonalization, but there are also numerous algorithms [54]–[56] (see Table III).

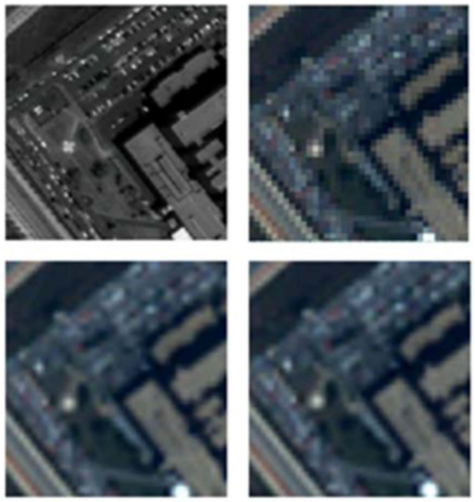

Fig. 2. Detail of Ikonos image scene from Alcala University Campus. From top left, in clockwise direction, Pan image (GSD = 1 m), nearest-neighbour, splines, linear interpolation [47].

The IHS method uses three selected LR spectral bands to generate a color composition. It consists of four stages: The first one is the transformation of the selected MS bands (M_i) to the IHS space [(7) and (8)], where *N* is the number of selected bands [57]–[59]. Then the color composite image is sampled to the resolution of the PAN image. At this stage, each image band is also normalized to the [0, 1] range, and the histogram of the HR image is adjusted. The correctness of this operation is verified by observation of the mean and standard deviation values of the PAN image and the multispectral image (9). In the third phase, the intensity component is replaced with the PAN image (10). The final fourth stage is returning to the RGB color scheme [60]

$$I = \sum_{i=1}^{N} \alpha_i M_i \tag{7}$$

$$\alpha_i = \frac{1}{N} \tag{8}$$

$$P = \frac{\sigma I}{\sigma P} (P - \mu P) + \mu I \tag{9}$$

$$F_i = M_i + (P - I) \tag{10}$$

where *N* represents the number of channels, M_i represents the selected multispectral Image channels, F_i represents pan-sharpened image, P_H represents the histogram-matched high-spatial-resolution PAN data, P represents PAN data, I represents up-sampled intensity component, σ represents deviation, and μ represents mean.

However, some authors noticed that the resulting image contains several spectral distortions despite excellent efficiency and high spatial resolution [61]–[63]. These findings inspired numerous modifications to the algorithm. One of them is the adaptive IHS method. Rehmani *et al.* [64] noted that spectral distortions were smaller and the intensity component was more similar to the PAN image. Another solution to the problem of spectral distortions was the edge-adaptive IHS. In this method, first, the edges are detected in the PAN image, and then the IHS method is applied to the location of the edges. The other pixels of the newly generated image (*F*) (those that do not belong to edges) are represented by the multispectral image (formula 11), where *h(x)* is the edge detection function that takes the value of 1 if an edge occurs, and 0 in other cases

$$F_i = M_i + h(x)(P - I) \tag{11}$$

where F_i denotes the pan-sharpened image, P_H denotes the histogram-matched high-spatial-resolution PAN data, M_i denotes the selected multispectral image channels, and $h(x)$ denotes the edge detection function.

TABLE III
COMPARISON OF SAMPLE INTERPOLATION METHODS THAT ARE USED TO ENHANCE THE SPATIAL RESOLUTION OF SATELLITE IMAGES

Method	Advantages	Disadvantages
Intensity-hue-saturation	- Enhanced spatial resolution - Improved interpretation ability - Excellent performance	- The resolution is enhanced only in three selected LR bands - Presence of spectral distortions
Principal Component Analysis	- Any number of spectral ranges are used - Enhanced spatial resolution - Improved interpretation ability - Excellent performance	- Presence of spectral distortions
Brovey color standardization	- Enhanced spatial resolution - Improved interpretation ability - Excellent performance - Limited spectral distortions in comparison to the other methods	- The resolution is enhanced only in three selected LR bands - Presence of spectral distortions
Ehlers fusion	- Enhanced spatial resolution - Improved interpretation ability	- The resolution is enhanced only in three selected LR bands - Presence of spectral distortions (transformation to IHS color space)
Image multiplication	- Slight color distortions	
Gram-Schmidt orthogonalisation	- Enhanced spatial resolution - Improved interpretation ability - Slight colour distortions - Excellent performance	

Another method that enables the reduction of spectral distortions is the use of discrete ripple transform and compressed sensing [65]. In this solution, the transform is implemented for the intensity component and the PAN image to obtain large-scale subimages. The resulting HR subimages are subjected to local variance, and compressed sensing is applied to LR subimages to reconstruct the intensity component. This enables the integration of local information from the intensity component and the PAN image. The final stage consists in applying inverse ripplet transform and reverse IHS transform.

A similar principle is employed in the Brovey color standardization method that uses any three MS bands. In this method, the selected bands are normalized by dividing these values by the sum of the triplet of bands. The next stage consists in multiplying the obtained results by the PAN image containing spatial information, which may be presented in the form of the following equation, where *i* denotes the three selected bands [66]–[68]:

$$\mathrm{DN}\ (i) = \left[\frac{\mathrm{DN}\,(i)}{\sum \mathrm{DN}\,(i)}\right] \cdot \mathrm{DN}_{\mathrm{pan}} \tag{12}$$

where DN(*i*) represents the digital number, I–3 are the selected channels, and $\mathrm{DN}_{\mathrm{pan}}$ is the digital number of PAN images.

This method may also be modified in certain ways. One of them is the combination of the Brovey method with high-pass filters, such as the Laplacian filter [68], [69]. This operation enhances the contract between the edges that are present in the image and the background, which significantly improves the visibility of details.

Another solution is the PCA pansharpening method. It uses any number of spectral ranges, transforming generally correlated bands into a new set of noncorrelated image data. The transformation matrix (*V*) (13) consists of the eigenvectors of the matrix of covariance or the matrix of correlation of the dataset for the nonstandardized PCA and normalized PCA, respectively. The new HR image is generated with the use of the three first spectral ranges of the PCA, and the PC1 component (which mainly contains spatial information) is replaced with the PAN image (PAN^s) after the histogram is adjusted to PC1. The final stage consists in performing a reverse transformation of the PCA, which creates a new HR multispectral image (*F*) [70]

$$\begin{bmatrix} F^1 \\ F^2 \\ \dots \\ F^n \end{bmatrix} = \begin{bmatrix} v^{11} & v^{12} & \dots & v^{1n} \\ v^{21} & v^{22} & \dots & v^{2n} \\ \dots & \dots & \dots & \dots \\ v^{n1} & v^{n2} & \dots & v^{nn} \end{bmatrix} \begin{bmatrix} \mathrm{PAN}^s \\ \mathrm{PC}^2 \\ \dots \\ \mathrm{PC}^n \end{bmatrix} = V \begin{bmatrix} \mathrm{PAN}^s \\ \mathrm{PC}^2 \\ \dots \\ \mathrm{PC}^n \end{bmatrix} \tag{13}$$

where PAN denotes the panchromatic image, PC^n denotes the principal component, F^n denotes *n* component of the *F* image, where $F = [F^1,\ F^2, \dots,\ F^n]^T$, and *V* denotes the transformation matrix.

The PCA pansharpening method may be modified in several ways [71]–[73]. One of them is the additive wavelet principal component (AWPC). In this method, the PAN image's histogram is adjusted to the histogram of the first principal component PC1, and then the "à trous" algorithm is applied to distinguish wavelet planes. The last stage before applying the reverse transformation of the PCA consists in adding wavelet planes to the principal component [74]. This method was improved by Ghadjati *et al.* In their publication; these authors presented an approach where the spatial information content obtained from the PAN image is automatically adjusted until the highest quality is obtained. This dependence is presented in the following equation, where the HPF is the high-pass filtration of the PAN image (see Fig. 3) [75]:

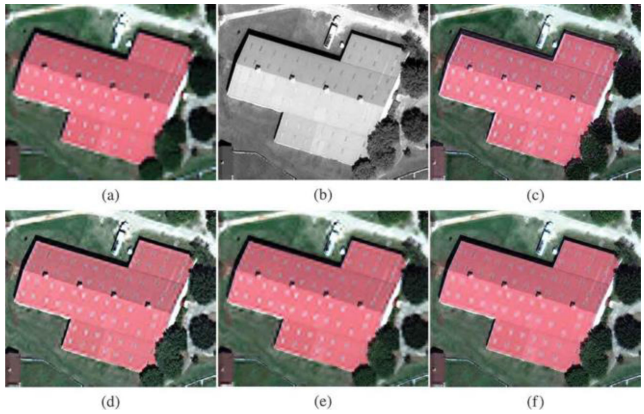

Fig. 3. Subsets from the study area in true color combination of (a) LMS image ($\mathrm{GSD}_{\mathrm{MS}} = 1.24$ m), (b) PAN image ($\mathrm{GSD}_{\mathrm{PAN}} = 0.31$ m), (c) PCA fused image, (d) AWPC fused image, (e) IPCA fused image (one iteration), and (f) IPCA fused image (four iterations) [75].

$$\begin{bmatrix} F_0^1 \\ F_0^2 \\ \dots \\ F_0^n \end{bmatrix} = V \begin{bmatrix} \mathrm{PC}_0^1 + \mathrm{HPF}\ [\mathrm{PAN}] \\ \mathrm{PC}_0^2 \\ \dots \\ \mathrm{PC}_0^n \end{bmatrix} \tag{14}$$

where F_0^n represents the *n* component of the *F* image, where $F = [F_0^1,\ F_0^2, \dots,\ F_0^n]^T$, *V* is the transformation matrix, PC^n is the principal component, PAN is the panchromatic image, and HPF is the high-pass filtering of a PAN image.

Another method that uses the selected bands and the transformation to the IHS color space is the Ehler method. In this solution, the PAN image and the intensity component (*I*) are transformed to the frequency level with the fast Fourier transform. After that, details of the image are separated from the PAN image with the use of high-pass filtering, and component *I* is subjected to low-pass filtering in order to blur the image. Images prepared in this way are adjusted to the original histogram based on intensity and transformed into an HR RGB image with the use of the reverse Fourier transform. This method utilizes several iterations for the combinations of all available bands [76], [77]. The application of this method was discussed in numerous publications on the enhancement of the spatial resolution of multispectral satellite images [78], [79], hyperspectral images (HSI) [80], and radar images [81].

The method with the smallest degree of color distortion, as a result of the presence of the intensity component, is the multiplicative method. The new pixel value is the product of the multiplication of the DN value of the PAN image pixel and the pixel of each of the multispectral image bands [82]. This solution is frequently used [78], [83]–[86].

TABLE IV
QUANTITATIVE EVALUATION WITH REFERENCE IMAGES: QUALITY METRICS RESULTS FOR DIFFERENT PANSHARPENING METHODS USING DIFFERENT SATELLITE IMAGES (MEAN VALUES)

Test Images	Metrics Method	CC	ERGAS	RASE	RMSE	SAM	SID	SCC	UIQI	SSIM
Deimos-2 images	GIHS [62]	0.0285	5.6958	22.5444	30.1408	1.7161	0.0311	0.9916	0.9280	0.8886
	AIHS [92]	0.0043	5.7104	22.4819	29.7967	2.1875	0.0443	0.9903	0.9315	0.8961
	AWLP [92]	0.0213	4.3249	17.4812	23.4077	2.33E-7	0.0013	0.9509	0.9627	0.9581
	PCA [93]	0.2683	7.7880	31.1688	41.2881	4.0265	0.0515	0.9471	0.8141	0.8007
	PRACS [94]	0.0204	4.2847	17.4401	23.1224	1.2166	0.0677	0.9688	0.9643	0.9483
	FRSQ [95]	0.0317	4.7057	18.9366	25.3195	1.1819	0.0620	0.8451	0.9560	0.9128
	Wady et al. Method [61]	0.0132	3.4207	13.5333	18.0741	0.7302	0.0082	0.9832	0.9755	0.9763
WV-2 images	GIHS	0.0089	4.5698	17.2795	50.7240	1.4450	0.0076	0.9910	0.9555	0.8103
	AIHS	0.0010	4.3303	16.7635	49.1225	1.2774	0.0133	0.9896	0.9635	0.8301
	AWLP	0.0086	4.5449	18.2989	53.7066	2.43E-7	0.0004	0.9481	0.9579	0.8338
	PCA	0.1783	8.6275	32.9548	97.1180	4.6856	0.0329	0.9168	0.7952	0.7236
	PRACS	0.0056	4.4537	18.1805	53.2766	1.0780	0.0032	0.9820	0.9638	0.8391
	FRSQ	0.0119	5.5075	21.1254	61.8930	1.4337	0.0692	0.8298	0.9495	0.7992
	Wady et al. Method [61]	0.0055	3.7033	14.0013	41.0648	0.8565	0.0110	0.9874	0.9717	0.8579
WV-3 images	GIHS	0.0406	5.7943	18.8313	60.6717	2.2267	0.0512	0.9818	0.9309	0.5300
	AIHS	0.0111	6.0724	19.7170	63.4380	2.0189	0.0478	0.9849	0.9246	0.5421
	AWLP	0.0514	4.4063	17.0148	54.9083	0.0306	0.0721	0.9373	0.9534	0.7377
	PCA	0.3438	7.4928	26.6934	85.4653	3.4903	0.0487	0.9416	0.8712	0.4411
	PRACS	0.0171	3.7532	12.8878	41.4724	1.3918	0.0016	0.9293	0.9721	0.7387
	FRSQ	0.0151	5.1888	15.1875	48.9939	2.3728	0.1263	0.8228	0.9531	0.6688
	Wady et al. Method [61]	0.0089	3.6405	11.8389	38.0540	1.2723	0.0397	0.9676	0.9728	0.7017

The best values of the analyzed metrics are highlighted in green and the worst ones in red [61].

Gram–Schmidt orthogonalization is another method that replaces the intensity component with the PAN band. In this solution, the algorithm takes nonorthogonal vectors and then rotates them so that they become orthogonal. For images, each band is treated like a multidimensional vector [77], [87]. This method is very popular and brings very good results, which was proven by a number of publications that emphasized the high quality of the obtained images [63], [88]–[91].

Table IV presents an example of testing the correctness of the method proposed by the authors, and Fig. 4 shows a visual comparison of the application of several methods to an exemplary image. Based on the metrics, it can be concluded that the method proposed by the Wada team [61] gives the best results, which is confirmed by visual assessment [see Fig. 4(i)]. For this method, the relative dimensionless global error in synthesis (ERGAS), relative average spectral error (RASE), and root-mean-squared error (RMSE) metrics take the lowest values among the compared methods, while universal image quality index (UIQI) and SSIM take the highest values (close to $+1$). The AWLP method [92] presents slightly weaker values of the assessment metrics, although they do not differ significantly from the winners of this comparison in the visual assessment. The values of the examined metrics confirm the small differences: ERGAS, RASE, and RMSE are much lower than other compared methods, whereas the value of spectral angle mapper (SAM) takes a value close to zero, which means the high quality of mapping of the spectral components. Due to the highest spectral distortions, the worst results were achieved by the PCA method, which is confirmed by the tested metrics, e.g., very high RMSE value (for WV-2 images RMSE $= 97.1180$) or low SSIM value (for WV-3 images SSIM $= 0.44$) [93]. In Table IV, the best values of the analyzed metrics are highlighted in green, while the worst ones are in red.

C. Other Digital Image Processing Methods to Improve Image Resolution

Other solutions are methods that utilize numerous LR images in order to estimate an HR image. One of them is the algorithm that enhances the resolution of the PAN SPOT5 scene, a quincunx sampling mode named Très Haute Résolution (THR), proposed by the CNES company. This innovative model of acquiring and processing images allows obtaining PAN images with a resolution of approx. 2.5 m is based on two images of the resolution of 5 m that are taken simultaneously. This is enabled by the fact that the images are obtained with the use of two CCD detector matrices that are shifted in the focal plane so that they meet the Shannon requirement at first order [96] (see Fig. 5). This means that the acquisition system is well optimized, thanks to

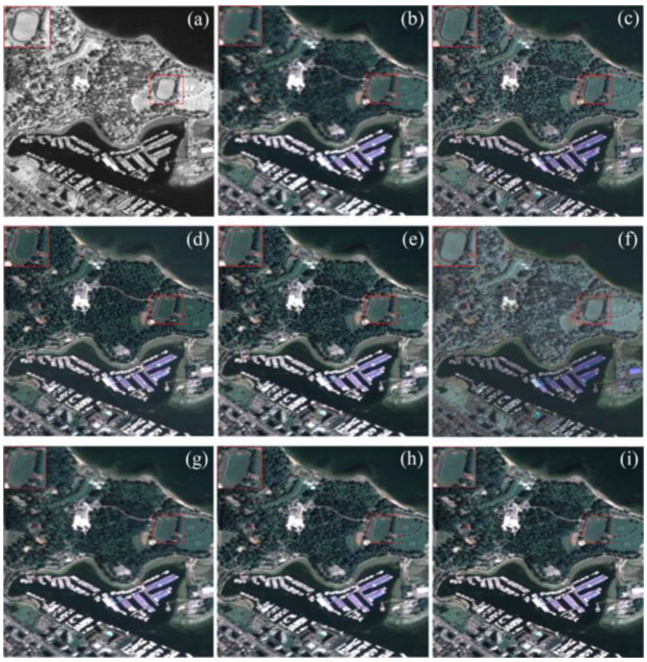

Fig. 4. Fusion results of the second Deimos-2 image. (a) PAN image (GSD_{PAN}= 1 m). (b) Up-sampled MS image (GSD_{MS}= 4 m). (c) GIHS. (d) AIHS. (e) AWLP. (f) PCA. (g) PRACS. (h) FPSQ. (i) Wady *et al.* Method [61].

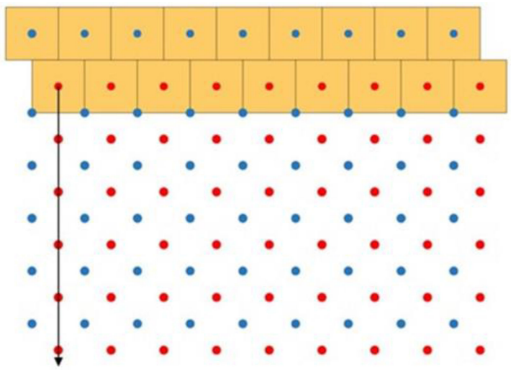

Fig. 5. SPOT5 THR quincunx sampling grid. A double CCD linear array generates two classical square sampling grids, shifted by half a sampling interval in both row and column directions. Interleaving the two shifted grids yields a quincunx sampling [98].

the correct match of the sampling and the modulation transfer function (MTF).

The THR process consists of three stages: interpolation, deconvolution, and noise reduction. In the first stage, the obtained LR images are interleaved, creating an image of a radiometric resolution of 2.5 m. However, only half of the pixels originate from the images obtained by the sensor ($\frac{2 \cdot x \cdot y}{2 \cdot x \cdot 2 \cdot y} = \frac{1}{2}$, where x and y denote the number of rows and columns of the image, respectively). As a result, the values of the newly created pixels are interpolated based on the values of the adjacent pixels. A characteristic blur is visible in the generated image, which is caused by low MTF values for high frequencies. Due to that, the second stage involves removing the blur by applying a filter that represents the instrument's inverse transfer function. This operation results not only in removing the blur but also in enhancing high-frequency noise, which is then removed at the last stage of the algorithm's operation. This phase utilizes the fixed chosen noise restoration algorithm [97], whose task is to remove noise in a controlled way in the most sensitive places. This is enabled by a nonlinear reduction of the wavelet packet coefficients, employing joint-adaptive space and frequency wavelet packet decomposition. This means that the noise removal process consists of thresholding noisy wavelet coefficients in the image at different image restoration levels [98], [99]. The application of this algorithm enables to enhance the spatial resolution of the image, which, in turn, improves the level of detail, and, as a consequence, the interpretation ability of the image, which was ultimately reflected in numerous studies on monitoring the environment [100]–[102], detecting changes [103]–[105], and mapping [106]–[108].

Another solution to estimate the value of pixels of HR images based on LR ones is the super-resolution variable-pixel linear reconstruction (SRVPLR) algorithm. It is based on the VPLR algorithm (also referred to as the Drizzle algorithm) [109], which combines astronomical dithered undersampled images. In this algorithm, pixels in LR images are represented in the newly created HR image, taking into account the shift, rotation, and distortion of the imaging sensor. At the same time, the transferred pixel's size is decreased to limit the number of artifacts in the image. Such decreased pixel i_{xy} of the weight w_{xy} is then added to the pixel of image I_{xy} with the weight W_{xy} and a fraction, overlaying pixel $0 < a_{xy} < 1$. As a result, the new values of the image I'_{xy} and weight W'_{xy} equal

$$W'_{xy} = a_{xy} w_{xy} + W_{xy} \tag{15}$$

$$I'_{xy} = \frac{a_{xy} i_{xy} w_{xy} + I_{xy} W_{xy}}{W'_{xy}} \tag{16}$$

where x, y denote pixel position, i_{xy} denotes the DN value of the reduced pixel, i_{xy} denotes the reduced pixel weight value, W'_{xy} is the final pixel weight value, I_{xy} is the DN value of pixel before modification, I_{xy} is the DN value of pixel after modification, a_{xy} denotes the fractional pixel overlap parameter, where a_{xy} ϵ(0, 1), and DN(i) is the digital number.

However, satellite images, as opposed to astronomical images that present stars, asteroids, nebulae, or galaxies, contain multiple elements that differ in terms of shape, structure, or texture. Due to that, the SRVPLR algorithms adjust the histogram of LR images, taking into account the selected values of pixels of the original LR image as reference values (i.e., their value should remain unchanged after changing the resolution) and using tools to mask the areas that are not subject to the resolution enhancement process.

Other authors have also used the multiframe super-resolution technique to enhance the resolution of electro-optic images [110] and radar images. Ito [111] proposed an algorithm to enhance the resolution of SAR images by performing local arithmetic operations [112] that take into account the orbit deviations of a satellite and temporal changes. The presented solution enables us to estimate a HR image based on n-frames of LR images.

TABLE V
COMPARISON OF SAMPLE DEEP LEARNING METHODS

	Method	Advantages	Disadvantages
CNN	Classic solutions	- improved interpretation ability of SR images - only a LR image is required to enhance resolution	- access to large databases is required, as they are the basis for network training - correct functioning only for specific image parameters - high computational capacity of the workstation is required, particularly at the network training stage
	Networks using residual connections	- improved interpretation ability of SR images - only a LR image is required to enhance resolution - solution for the problem of gradient fading and representation bottlenecks - accelerated network training process - improved interpretation ability of SR images	- access to large databases is required, as they are the basis for network training - correct functioning only for specific image parameters - high computational capacity of the workstation is required, particularly at the network training stage
	Multi-column generator	- the algorithm may accept images of various spatial resolutions Additionally, for solutions that utilize residual connections: - solution for the problem of gradient fading and representation bottlenecks - accelerated network training process - improved interpretation ability of SR images	- access to large databases is required, as they are the basis for network training - high computational capacity of the workstation is required, particularly at the network training stage
GAN	Classic solutions	- significant improvement in the interpretation ability of SR images compared to SR images obtained with use of algorithms that employ classic deep convolutional networks	- unstable training - modifications of network training parameters have a strong influence on the generated HR images - at the training stage, LR images are required, along with corresponding images of high (target) resolution - high computational capacity of the workstation is required, particularly at the network training stage
	Classic solutions	- significant improvement in the interpretation ability of SR images compared to SR images obtained with use of algorithms that employ classic deep convolutional networks - solution for the problem of gradient fading and representation bottlenecks - accelerated network training process	- unstable training - modifications of network training parameters have a strong influence on the generated HR images - at the training stage, LR images are required, along with corresponding images of high (target) resolution - high computational capacity of the workstation is required, particularly at the network training stage

As far as HSI are concerned, tensor-based methods are becoming increasingly popular. An example of the application of such a method is the coupled tensor factorization approach [113], which utilizes the multidimensional structure of the HSI and a multispectral image (MSI) based on the coupled tensor factorization framework. This solution enables solving the problem of the loss of structural information of the HSI and MSI data "cubes," and allows the identifiability of the super-resolution image (SRI) under realistic assumptions. Another solution combines the LR HSI with an HR multispectral image to obtain an HR HSI [114]. This method utilizes nonlocal similar patches to form a nonlocal patch tensor, which is modeled with the use of the method based on the tensor-tensor product (t-product). Another approach to solving this problem exploits ideas from the field of tensor completion to directly impose a low-rank property spatially and spectrally while avoiding the computationally complex patch clustering and dictionary learning common to competing for fusion techniques [115].

D. Deep Learning

Apart from the interpolation methods that utilize HR PAN images or digital image processing, there is also a group of methods that employ deep learning algorithms. These solutions recreate a HR image based on the knowledge about the links between LR images and the corresponding HR ones, which were obtained during model learning. Over the last decade, we have been witnessing a significant improvement in the efficiency of deep learning algorithms that enhance spatial resolution. Even more, they now enable a much more accurate recreation of HR images based on LR ones. Another advantage of this group of solutions over the pansharpening methods, which are a very popular means of enhancing satellite images' spatial resolution, is the fact that deep learning solutions do not require another corresponding HR image. As a result, the interpretation ability of digital images is improved, provided that the neural network has been prepared correctly.

The methods presented in the literature may be divided into two groups: solutions that utilize convolutional neural networks [116]–[121] and methods based on generative adversarial networks [122]–[128]. However, these solutions are used in a slightly different way than those presented above. It may be divided into two stages: training of the network on which the algorithm is based and using the algorithm (see Table V). At the same time, one should bear in mind that the method's capability depends on the training stage. Apart from the adequate architecture and network parameters, network training requires a database. The most common method of creating databases dedicated to improving spatial resolution is reducing the resolution of an HR image to the resolution of the LR image and then conducting training based on LR images that are compared to the actual HR images. As a result, any digital image or any database of digital images (e.g., ImageNet [129], LSUN [130], MC COCO [131]) may be used to train the network. Still, a database dedicated to the issue of improving resolution exists. It contains LR images and the corresponding HR images [132]. However, to use an existing database for network training, it should be selected so that the image parameters are similar to those the algorithm will work on in the future.

TABLE VI
COMPARISON OF THE PSNR [DB]/SSIM METRICS OF THE MOST POPULAR MODELS OF ENHANCING THE RESOLUTION OF DIGITAL IMAGES BASED ON DATABASES: SET5-4X, SET14-4X, BSD100- 4X, URBAN100- 4X, FFHQ 256X256- 4X, FFHQ 512X512- 4X, AND FFHQ 1024X1024- 4X

Model	Set5-4x [139]	Set14-4x [140]	BSD100- 4x [141]	URBAN100- 4x [142]	FFHQ 256x256- 4x [143]	FFHQ 512x512- 4x [143]	FFHQ 1024x1024- 4x [143]
SRGAN [35]	29.40/0.8472	26.02/0.7397	25.16/0.6688	-	17.57/0.415	27.494/0.735	21.49/0.515
SRResNet [35]	32.05/0.9019	28.49/0.8184	27.58/0.7620	-	-	-	-
Bicubic [35]	28.43/0.8211	25.99/0.7486	25.94/0.6935	-	-	-	-
EDSR [144]	32.46/0.8968	28.80/0.7876	27.71/0.7420	26.64/0.8033	22.47/0.706	30.188/0.824	28.34/0.827
ESRGAN [36]	32.73/0.9011	28.99/0.7917	27.85/0.7455	27.03/0.8153	15.43/0.267	27.134/0.741	19.84/0.353
FSRCNN [138]	30.70/0.8657	27.59/0.7535	26.98/0.7398 (BSD200)	-	-	-	-
SRCNN [29]	30.49/0.8628	27.50/0.7513	26.90/0.7101	24.52/0.7221	23.12/0.688	-	27.40/0.801
SAN [145]	32.70/0.9013	29.05/0.7921	27.86/0.7457	27.23/0.8169	-	-	-
HAN+ [146]	32.75/0.9016	28.99/0.7907	27.85/0.7454	27.02/0.8131	-	-	-
WSRGAN [147], [148]	28.24/0.8779	25.08/0.7396	-	-	-	-	-
ESPCN [149]	28.16/0.8001	24.76/0.6962	-	-	-	-	-
EnhanceNet [149]–[151]	31.74/0.8869	28.42/0.7774	27.50/0.7326	25.66/0.7703	23.64/0.701	-	29.42/0.832
SRFBN [152]	32.47/0.8983	28.81/0.7868	27.72/0.7409	26.60/0.8015	21.96/0.693	29.58/0.8270	27.90/0.822

For several years now, participants in many image processing and machine learning competitions have attempted to solve the problem of improving the resolution of digital images. One of them is the NTIRE Challenge organized as part of the Conference on Computer Vision and Pattern Recognition. The NTIRE Challenge deals with numerous issues related to digital image processing, including the improvement of digital image resolution [133]–[136] and video [137].

Most of the research on improving resolution is based on classic digital images (Set5-4x, Set14-4x, BSD100- 4x, URBAN100- 4x, FFHQ 256x256- 4x, FFHQ 512x512- 4x, FFHQ 1024x1024- 4x). These images are characterized by HR and, thus, a great level of detail. In the case of improving the resolution of vertical images acquired from the aerial level, especially for satellite images, the same objects are characterized by different features, and the image resolution is much lower. Therefore, to use the solutions proposed by scientists to improve the resolution of digital images (obtained mainly from the Earth's ceiling), it is necessary to train models whose weights have been estimated based on the images mentioned above (or conduct training from the beginning—which is more time-consuming and requires Big Data).

The application of deep learning neural networks to enhance the spatial resolution of satellite images was inspired by the results of works on improving the resolution of digital images. The authors used the PSNR and the SSIM to assess the correct operation of the applied algorithms (see Table VI). The first indicator carries information about the ratio between the maximum signal strength and the power of noise interfering with the signal, while the other one takes into account distortions of luminance, contrast, and structure. One of the most popular solutions is the SRCNN, which consists of three parts: patch extraction and representation, nonlinear mapping, and reconstruction. Its modification that enables slightly improve the PSNR coefficient (by 0.21 dB for the Set5-4x dataset) is the FSRCNN, where the deconvolution layer was introduced at the end of the network, so the mapping is learned directly from the original LR image (without interpolation) to the HR one, and the mapping layer was modified by changing the dimensions of the element before and after mapping. Additionally, the size of filters was reduced while at the same time increasing the number of mapping layers (see Fig. 6).

Another solution is the ESPCNN, which is a convolutional neural network that allows playing a video recording in real-time, at the resolution of 1080 p, with the use of the K2 graphic processor. The property maps are separated in the LR space in the proposed model architecture. Additionally, the subpixel convolution layer was introduced to enable improved reconstruction of the HR image. An example of a solution that employs generative adversarial networks is WSRGAN, which was created based on the analysis of MSE and VGG losses of the SRGAN model. The authors modified the range of weight values from 0 to 1 and the sampling interval to 0.1, which improved the PSNR and SSIM values of the generated images. However, numerous solutions present higher values of the coefficients. An example is the EDSR model, where the redundant modules of convolutional neural networks (Conv + ReLU + Conv) were removed, and the model's size was increased (32 residual blocks for EDSR and 80 residual blocks for MDSR). Another solution is a second-order attention network (SAN) that focuses on improved learning of correlations between properties, which, as a consequence, results in a better representation of the image. The authors

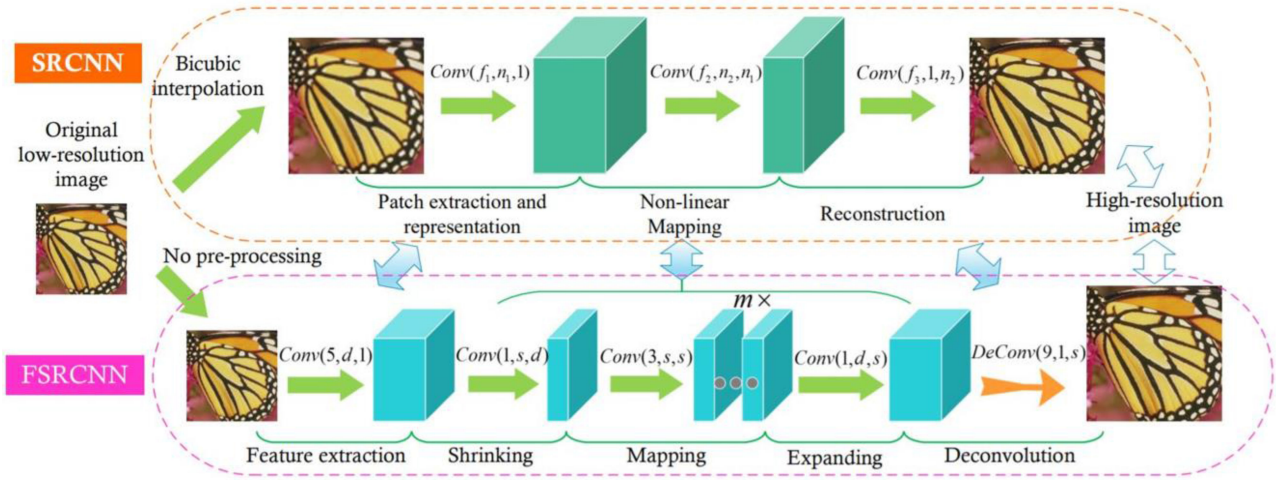

Fig. 6. Comparison of the structures of SRCNN and FSRCNN [138].

have developed an innovative second-order channel attention module to adaptively rescale the channelwise features by using second-order feature statistics for more discriminative representations. The super-resolution feedback network (SRFBN) that utilizes feedback also brings very good results. On the other hand, the authors of the EnhanceNet model focused on the correct representation of the textures of generated images. They proposed the application of automated texture synthesis in combination with perceptual loss. The application of feed-forward fully convolutional neural networks in the discriminator model resulted in an improved quality of the generated SR images. The model that presents the best PSNR and SSIM results for the Set5-4x dataset is the holistic attention network (HAN+) which consists of a layer attention module and a channel-spatial attention module that enables holistic modeling of the correlations between layers, channels, and positions. The comparison of the PSNR and SSIM metrics of the most commonly used resolution enhancement methods revealed that the ESRGAN method [36] brought the best results for the Set-4x database, whereas the best solution for the Set14-4x, BSD100- 4x, and URBAN100-4x was SAN, developed by Dai *et al.* [145]. On the other hand, the worst results were obtained with the use of the ESPCNN algorithm [149]. What is interesting, this solution was characterized by poorer metrics than bicubic interpolation for the Set5-4x and Set14-4x databases (no data were available for the other analyzed databases).

1) Solutions Based on Classic Deep Neural Networks: An example of the application of deep neural networks is the VDSR network [153] which enables the enhancement of the spatial resolution of multispectral images. This solution utilizes blocks that consist of the convolutional layer and an activation unit (ReLU). At the last stage of operation of this model, the input image is added to the obtained results. This combination results in an SR image [154]. Several solutions that also use residual connections have been introduced in recent years.

An example is an algorithm used to enhance the resolution of space photographs presented by Feng *et al.* [155], whose task is to remove noise from the image and then enhance its resolution. The proposed network architecture consists of nine residual blocks containing two convolutional layers separated by the activation unit.

Some of the other solutions presented in the literature are methods that utilize multicolumn networks. In these solutions, each branch is dedicated to one of the resolutions of the input image [156], [157]. As a result of the use of many branches, input data may take different dimensions, which allows adjusting the size of the image to the dimensions of the object whose resolution is to be enhanced. The authors of these publications assume that observing the given object in various spatial resolutions allows us to see more. This translates into practical applications, as is shown in Fig. 7. New SR images are characterized by a much better representation of the shape of the contours of objects in the image, which is additionally confirmed by the PSNR coefficient.

The methods presented in the previous sections allow for enhancing the resolution of images that consist of several bands. However, deep learning solutions also enable enhancing the resolution of HSI. This task is much more difficult than improving the resolution of multispectral images. The reason is the problem with obtaining HR images that are necessary at the training stage of convolutional neural networks.

However, numerous teams have attempted to solve this issue [158]–[163]. One of them was Li *et al.* [164]. The authors presented a new network consisting of two streams: spatial and spectral. The 1-D spectral convolution encodes minor changes in the spectrum, while the 2-D convolution, combined with the attention mechanism, encodes spatial information. Additionally, hierarchical lateral connections are used to enable combining spectral and spatial information. Moreover, attention blocks have been applied. They are implemented between blocks in the spatial network and enable considering the relations between nonlocal regions.

2) GAN-Based Solutions*:* The second group of solutions that utilize deep learning methods is those based on generative adversarial networks. The original GAN model consists of the generator G and discriminator D. The task of the G generator is to accept a random vector on the input and decode it in

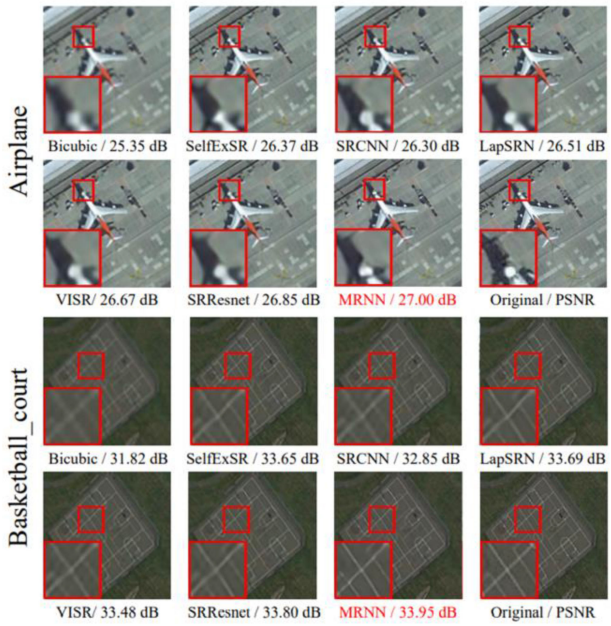

Fig. 7. Comparison of the results of the operation of the multicolumn network MRNN with Bicubic, SelfExSR, SRCNN, LapSRN, VISR, SRResnet, and the original image [157].

order to generate a synthetic image, while the discriminator is responsible for distinguishing between actual and decoded data. The generator network is trained so as to enable it to "cheat" discriminator D. Thus, with the progress in training, it will generate more and more realistic images, which enables, among others, to enhance their resolution. The original GAN model may be expressed as [165]

$$\min_G \max_D V(D, G) = E_{x\sim p_{\text{data}}(x)}[\log D(x)] + E_{z\sim p_{\text{data}}(z)}[\log(1 - D(G(z)))] \tag{17}$$

where x denotes the input image $x \sim p_{\text{data}}(x)$, z represents the random noise from a probability distribution of the *a priori* distribution $z \sim pz(z)$, and E is the empirical approximation of the expected probability. The above formula means that the process is iterative in a way that allows it to maximize the probability that the discriminator will correctly distinguish between real and synthetic images. At the same time, the generator learns how to minimize the probability of detection.

Based on this dependence, Ledig *et al.* [35] proposed the SRGAN algorithm (see Fig. 8). In this solution, LR images are generated based on HR ones. This is achieved by blurring the image with the use of the Gaussian filter, followed by decimation of the image at the sampling coefficient r. The training of the generator network is conducted in the form of feedback, while the weights and bias of the l-layer deep neural network may be calculated as

$$\theta_G = \{W_{1:L}; b_{1:L}\} = \text{argmin}\frac{1}{N}\sum_{n=1}^{N} l^{\text{SR}}\left(G_{\theta_G}\left(I_n^{\text{LR}}\right), I_n^{\text{HR}}\right) \tag{18}$$

where $\theta_G = \{W_{1:L}; b_{1:L}\}$ denotes the weights and biases of an L layer deep network and is obtained by optimizing an SR specific loss function l^{SR}, N denotes the number of HR images in the database I^{HR} and the corresponding LR images I^{LR}, and G_{θ_G} is the generator model (feed-forward CNN).

The network training takes place with the use of the method of maximizing the minimum gain, which is based on a theoretical game scenario, where images estimated by the generator compete with original HR images in the discriminator model. The training of generative adversarial networks utilizes a zero-sum game, where the reward of the generating network is determined (19). This solution motivates the discriminator to classify the received images (samples) correctly as true or false. At the same time, the generator is trying to cheat the adversary by saying that the provided samples are true

$$\min_{\theta_G}\max_{\theta_D} E_{I^{\text{HR}}\sim p_{\text{train}}(I^{\text{HR}})}\left[\log D_{\theta_D}\left(I^{\text{HR}}\right)\right] + E_{I^{\text{LR}}\sim p_G(I^{\text{LR}})}\left[\log\left(1 - D_{\theta_D}\left(G_{\theta_G}\left(I^{\text{HR}}\right)\right)\right)\right] \tag{19}$$

where E is the empirical approximation of the expected probability, D_{θ_D} is the discriminator network, I^{HR} is the HR image, G_{θ_G} is the generator network, and I^{LR} is the LR image.

The authors of SRGAN noted that for images with high PSNR, calculating the content loss based on the MSE of pixels results in the emergence of smooth textures (which is caused by the absence of high frequencies). As a result, VGG loss was applied based on the ReLU activation layers of the pretrained VGG19 network

$$l^{\text{SR}}_{\text{VGG}/\,i,j} = \frac{1}{W_{i,\,j}H_{i,j}} \times \sum_{x=1}^{W_{i,j}}\sum_{y=1}^{H_{i,j}}\left(\phi_{i,j}\left(I^{\text{HR}}\right)_{x,y} - \phi_{i,j}\left(G_{\theta_G}\left(I^{\text{LR}}\right)\right)_{x,y}\right)^2. \tag{20}$$

In this formula, $\phi_{i,j}$ denotes a map of properties of the dimensions $W_{i,j}$ and $H_{i,j}$, obtained by the jth convolution (after activation) before the ith max-pooling layer of the VGG19 network. Then, the VGG loss is defined as the Euclidean distance between the representations of properties of the recreated image $G_{\theta_G}(I^{\text{LR}})$ and the reference image I^{HR}.

The SRGAN model inspired numerous modifications [166], [167], including HSI [168]. However, the most popular modification of this model is ESRGAN [36]. The batch normalization (BN) layers were removed from its generator (BN), and the basic block was replaced with a residual-in-residual dense block (RRDB), being a combination of a multilevel residual block and dense connections.

The removal of the BN layers resulted in stable training and improved network capacity (the time required for training became significantly shorter), which resulted from reduced computational complexity. The authors of ESRGAN also modified the discriminator by replacing it with a relativistic discriminator. As opposed to the standard discriminator used in SRGAN (21), which estimates the probability of whether the assessed image belongs to the set of HR images, the relativistic discriminator attempts to predict the probability that the real image I^{HR} is

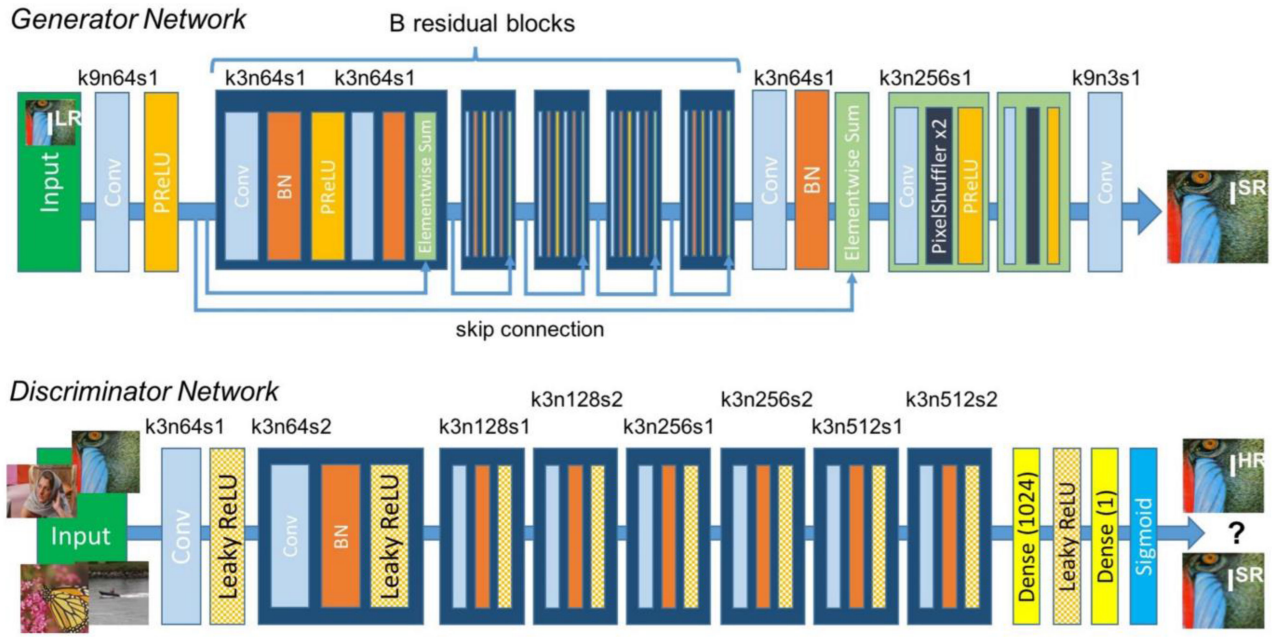

Fig. 8. Architecture of the SRGAN network [35].

relatively more realistic than the fake image I^{SR} (22)

$$D(x) = \sigma(C(x)) \tag{21}$$

$$D_{Ra}\left(I^{\mathrm{HR}}, I^{\mathrm{SR}}\right) = \sigma\left(C\left(I^{\mathrm{HR}}\right) - E_{x_f}\left[C\left(I^{\mathrm{SR}}\right)\right]\right) \tag{22}$$

where σ denotes the sigmoid function, $C(x)$ represents the output data of the generator before the application of the final activation function, and $E[\cdot]$ is the average of all fake data in the minibatch.

Another modification of the SRGAN model is the fact that perceptual loss is applied before the activation layers (not after them). This allows for increasing the number of features taken for calculating $l^{\mathrm{SR}}_{\mathrm{VGG}/\,i,j}$, which improves network efficiency. Additionally, it enables significantly better reconstruction of the brightness of the SR images. The ESRGAN network is trained based on network interpolation, which removes the noise on the estimated SR images. It consists in training the network G_{PSNR} focused on PSNR, and then the network is adjusted to obtain the G_{GAN} network. As it was in the previous case, this solution is also used to enhance the resolution of satellite images [169]–[171]. Two datasets, a satellite video sequence, were used to verify the possibilities and differences between the methods mentioned above. The first set of data contains video sequences collected by SkySat-1 on March 25, 2014 from Las Vegas in the USA, on April 9, 2014 from Burj Khalifa in Dubai, and on February 1, 2019 from Burj al-Arab in Dubai. The resolution was approximately 1.5 m, and the frame size was 1280 × 720 pixels at 30 frames per second [172].

The second dataset included video sequences from the Jilin-1 mission video collected on October 6, 2017 from Beirut in Lebanon and on September 10, 2018 from Florence in Italy. For the Beirut (Lebanon) set, the resolution was approximately 1.3 (or 1.12 m), and the frame size was 1920 × 1080 pixels at 25 frames per second with a video duration of 32 s. For the Florence (Italy) set, the resolution was approximately 0.9 m, and the frame size was 3840 × 2160 pixels at ten frames per second with a video duration of 31 s [Chang Guang Satellite Technology, Company, Ltd.]. First, the authors used the pretrained generator G of ESRGAN, originally referenced on ImageNet and distributed as RRDB_ESRGAN_x4.pth. Then the networks (SRGAN and ESRGAN) were trained based on the above-described datasets, where the LR image was obtained by downsampling the HR image. Fig. 11 shows examples of improvement of spatial resolution of video frame fragments obtained by Jilin-1 Smart Video Satellite using SRGAN and ESRGAN.

Based on the visual analysis, it may be noticed that the introduced modification, i.e., the ESRGAN network, reproduces the structure and texture of the image much better. Moreover, by sharpening small objects, the image interpretation possibilities increase.

The application of residual connections allowed for a significant improvement of results while at the same time shortening the time required for network training [156], [157], [164], [169], [173]–[176] (see Figs. 9 and 10). This solution is used in numerous algorithms. One of them is the solution proposed by Courtrai *et al.* [177], who developed a method to improve the spatial resolution of small objects in both aerial and satellite images (see Fig. 11). The generator model utilizes residual connections and integration with the cycle model. Additionally, the authors used the Wasserstein GAN version [178] with the addition of a gradient penalty, which is the last component of the following loss function of the discriminator (or critic). Such a structure of the model resulted in significant improvement.

The possibilities and reliability of images recreated using deep learning methods are proven by the fact that they are utilized for the improvement of the resolution of medical imaging [179]–[181]. An example is a solution proposed by Zamzmi

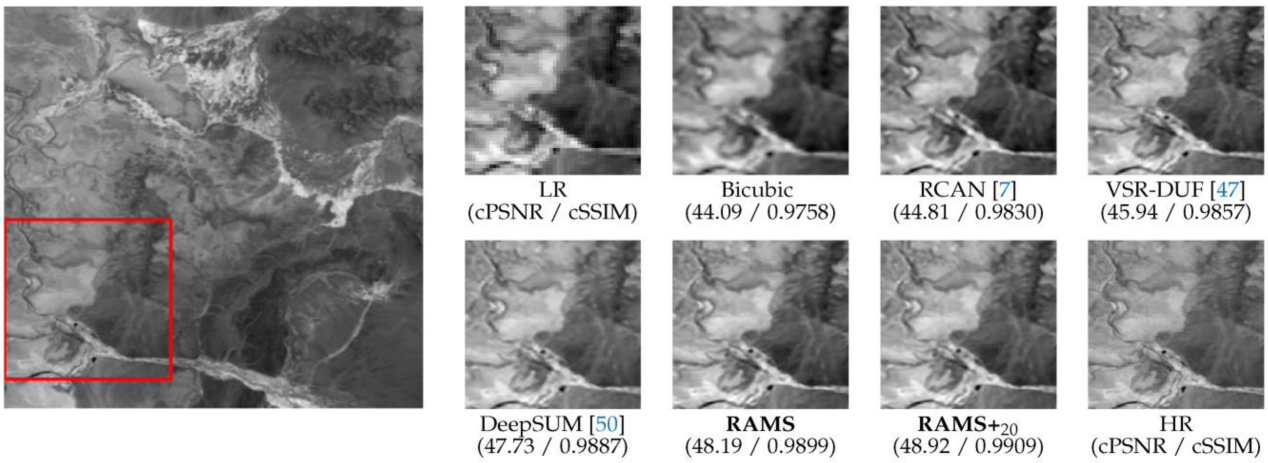

Fig. 9. Visual effects of the application of residual connections on the improved interpretation ability of PAN satellite images (PROBA-V, GSD= 1 km) [174].

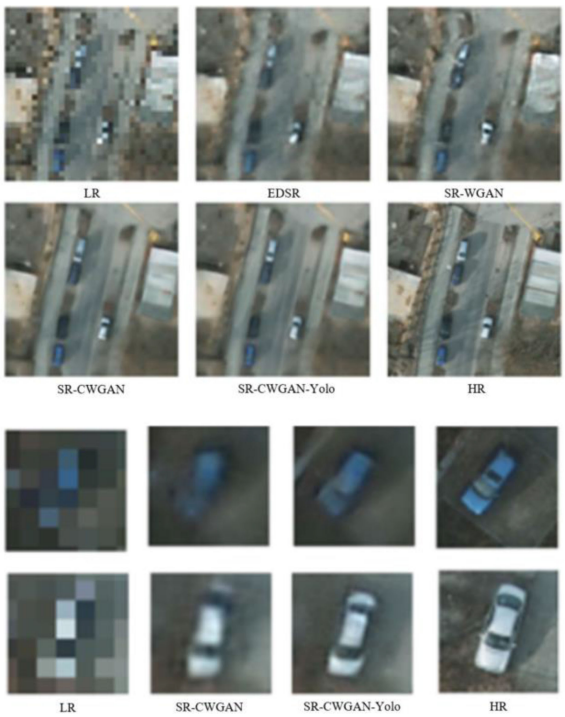

Fig. 10. Possibilities of enhancing the resolution of small objects. (a) Super-resolved images compared to LR (GSD = 1 m) and HR images (GSD = 12.5 cm). (b) Zoom on two vehicles [177].

et al. [182] that enables to enhance the resolution of X-ray images. The first stage of this algorithm consists in performing bicubic interpolation. Then the image is sent to the BN layer. The subsequent element of the architecture is a system of blocks consisting of three layers each (convolutional, zero padding, and activation (ReLU) layers). These blocks differ in terms of the number of convolutional layer filters. The image after bicubic interpolation is added to the resulting image at the last stage. The network analyses conducted by the authors demonstrated that it is characterized by much fewer training parameters than the VDSR [153] so that the training duration is significantly shortened with a simultaneous improvement of results.

As the above examples demonstrate, methods of enhancing the resolution of digital images with the use of deep neural networks are widely applied in issues related to the enhancement of satellite imaging resolution and for all digital images from space photography to medical imaging.

E. Evaluation Metrics

In order to assess the correctness of the operation of the methods, several metrics are used. One of them is the very popular RMSE, defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{m \cdot n} \sum_{i=1}^{m} \sum_{j=1}^{n} |\mathrm{MS}(i,\ j) - \mathrm{PS}(i,j)|^2} \quad (23)$$

where MS and PS are original multispectral images and fused multispectral images, and m x n represents the size of the image.

RMSE error is the basis of RASE [183], which is used to evaluate the average performance of image fusion methods for each spectral band and is calculated using

$$\mathrm{RASE} = \frac{100}{\mu} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \mathrm{RMSE}^2(B_i)} \quad (24)$$

where μ is the mean, N is the number of bands, and B_i represents an ith band of the input MS image [184]. Also, ERGAS [185] determines the number of spectral distortions of the image

$$\mathrm{ERGAS} = 100 \cdot \frac{h}{l} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{\mathrm{RMSE}(i)}{\mu(i)} \right)^2} \quad (25)$$

where $\frac{h}{l}$ is the ratio of pixel sizes of the input PAN and MS images, $\mu(i)$ is the mean of the ith band, while N is the number of bands [184].

Another solution for assessing the correctness of the method is the correlation coefficient (CC) [186], which determines the correlation measure between the PAN and multispectral (MS) images determined according to (26) shown at the bottom of this page. Additionally, modification of the spatial correlation

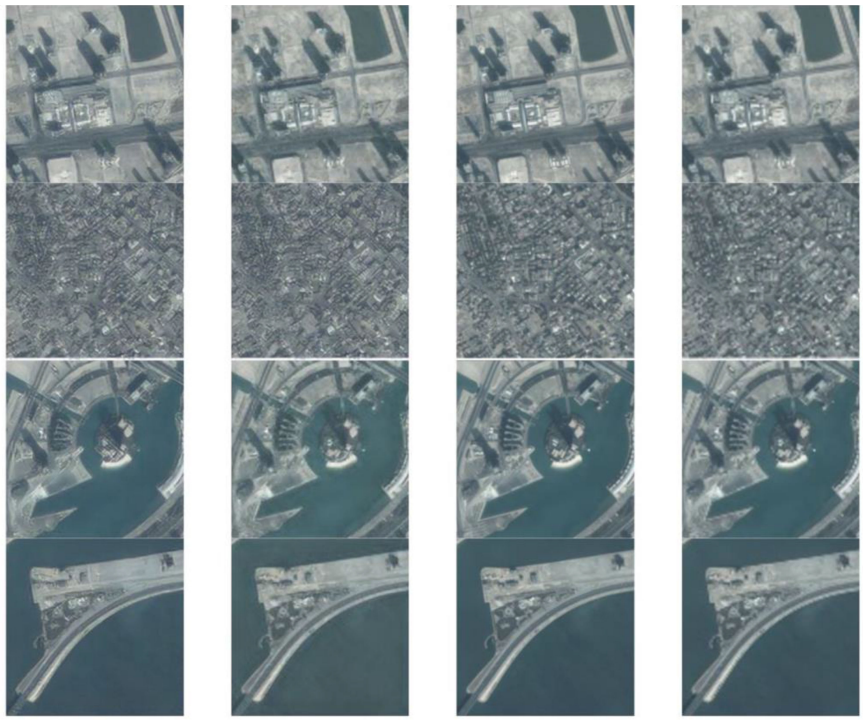

Fig. 11. Comparison of the results of spatial resolution enhancement with the use of SRGAN and ESRGAN on fragments of a video frame captured by microsatellite (95 kg) Jilin-1 Smart Video Satellite. (a) ESRGAN. (b) SRGAN. (c) HR (GSD = 1.3 m). (d) LR (GSD = 5.2 m).

coefficient (SCC) metric [187] can be used, (26) shown at the bottom of this page where PAN represents the panchromatic image, MS the multispectral image, $\overline{\text{MS}}$ the mean value of the MS images, $\overline{\text{PAN}}$ the mean value of the PAN images, and *n,m* the image dimension [184].

Another example of a metric is the SAM [188], which defines the average change of all angles in the spectral component

$$\text{SAM}\ (v,w) = \cos^{-1}\left(\frac{\sum_{i=1}^{L} v_i w_i}{\sqrt{\sum_{i=1}^{L} v_i^2}\sqrt{\sum_{i=1}^{L} w_i^2}}\right) \tag{27}$$

where L represents the number of layers and $v,\ w$ the test spectrum and reference spectrum (each has L components).

The spectral information divergence (SID) [189] method computes spectral similarity based on the divergence between the probability distributions of the two spectra as

$$\text{SID}\ (x,y) = \sum_{i=1}^{L} p_i\ \log\left(\frac{p_i}{q_i}\right) + \sum_{i\ =\ 1}^{L} q_i \log\left(\frac{q_i}{p_i}\right) \tag{28}$$

$$q_i = \frac{r_i}{\sum_{i=1}^{N} r_i} \tag{29}$$

$$p_i = \frac{t_i}{\sum_{i=1}^{N} t_i} \tag{30}$$

where r represents the reference spectrum, t the test spectrum, and L the band number of multispectral images.

Another group of solutions for assessing the correctness of the resolution improvement methods are metrics that take into account three types of distortions: luminance, contrast, and image structure after pansharpening (y) in relation to the reference image (x): UIQI [190] (31) and SSIM (32). Their numerical values are determined based on the following formulas:

$$\text{UIQI}\ (x,y) = \frac{4\mu_x\mu_y\mu_{xy}}{\left(\mu_x^2+\mu_y^2\right)\left(\sigma_x^2+\sigma_y^2\right)} \tag{31}$$

$$\text{SSIM}\ (x,y) = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy}+C_2\right)}{\left(\mu_x^2+\mu_y^2+C_1\right)\left(\sigma_x^2+\sigma_y^2+C_2\right)} \tag{32}$$

where $\mu_x,\ \mu_y, \sigma_x,\ \sigma_y,\ \sigma_{xy}$ are the local means, standard deviations, and cross covariance for images x and y, x is the reference image, and y is the pansharpening image.

Another metric to evaluate the processed image is PSNR. It defines the ratio of the maximum signal power (the maximum image value) to the power disturbing this signal, i.e., mean square

$$\text{CC}\ (\text{MS},\text{PAN}) = = \frac{\sum_{i\ =\ 1}^{M}\sum_{j\ =\ 1}^{N}\left(\text{MS}_{ij}-\overline{\text{MS}}\right)\left(\text{PAN}_{ij}-\overline{\text{PAN}}\right)}{\sqrt{\left(\sum_{i=1}^{M}\sum_{j\ =\ 1}^{N}\left(\text{MS}_{ij}-\overline{\text{MS}}\right)^2\right)\left(\sum_{i\ =\ 1}^{M}\sum_{j\ =\ 1}^{N}\left(\text{PAN}_{ij}-\overline{\text{PAN}}\right)^2\right)}} \tag{26}$$

error (33). PSNR values are expressed in decibels

$$\text{PSNR} = 10 \cdot \log_{10} \frac{[\max(\text{HR}(n,m))]^2}{\text{MSE}} \tag{33}$$

where MSE represents the mean square error and max(HR(*n,m*)) represents the maximum reference image value with *n x m* size.

The evaluation metrics presented above are commonly used in remote sensing tasks. The application of those metrics enables to compare the correctness of the operation of various algorithms, which is given in Tables IV and V. Tasks related to the enhancement of the spatial resolution of digital images involve attempts to achieve the lowest possible values of the CC, ERGAS, RASE, RMSE, SAM, and SID errors and the highest possible values of the PSNR, SCC, UIQI, and SSIM metrics. As far as the last three methods of assessment are concerned, the values of errors approach 1 if the SR images are represented with high quality.

III. Discussion

A. Review of the Methods

This article presents a review of the methods to enhance the spatial resolution of satellite images with the use of classic image sharpening methods, i.e., interpolation and pansharpening, and new solutions that employ deep learning: CNN and GAN. Until now, no such detailed review of the solutions for sharpening satellite images from the point of view of remote sensing has been performed.

Tables II and III present the advantages and disadvantages of classic solutions. The main limitation of the interpolation methods is the lack of a significant increase in interpretative abilities despite reducing the pixel size or the visual improvement of the image quality. At the same time, these solutions do not require the high computing power of working units and use only one image. On the other hand, the use of a PAN image to improve the resolution of multispectral images (MS) with lower resolution allows for a significant improvement in the interpretability of MS images, which provides a significant advantage over solutions that use interpolation. Unfortunately, this method does not prove effective in enhancing the resolution of images that depict moving objects. In such cases, SR images contain artifacts, which emerge as a consequence of the shift between the position of the moving object in the HR PAN image and its position in the multispectral image. This is one of the reasons why these algorithms cannot be applied to enhance the resolution of image sequences. Additionally, considering the methodology of operation of these solutions, they can only be used after improving the resolution of multispectral images that have their HR counterpart.

Table VI presents a summary and comparison of various neural network methods to enhance the resolution of digital images based on the same datasets. The comparison was based on the PSNR and SSIM quality indicators. The highest values of these indicators were obtained for ESRGAN [36] and SRResNet [35]. This method is characterized by significantly more stable training, higher efficiency, and improved reconstruction of details in the image. The obtained results confirm that this architecture may be applied to enhance the resolution of images from nanosatellite systems even up to four times. Besides, the ESRGAN method offers an excellent ability to represent detail, which is confirmed by a vast number of studies in the field of computer vision [36], [191].

In comparison to other methods, architectures based on GANs enable to obtain better results of resolution enhancement than classic methods. Due to the great potential and wide scope of applications of GAN, scientists are developing their applications in remote sensing analysis. However, the problems related to the application of GAN to sharpen whole satellite scenes, not only their fragments, still remain valid. Some other issues connected to the use of deep learning methods include the instability of data training or the vanishing gradient issue. Moreover, the conducted literature review allows us to state that GAN networks have a problem with nonconvergence and with the sensitivity to the selection of hyperparameters. Numerous methods, such as SRCNN (presented in Table VI), cannot be employed to enhance the spatial resolution of satellite images due to relatively poor representation of the texture, which may lead to misinterpretation of the images. The best architecture models were proposed by Wang [161] and Courtrai [177]. These methods use residual blocks, which is particularly important in the event of the application of generative adversarial networks, as they stabilize and significantly accelerate network training. Other important aspects that have been discussed in this literature review are the PSNR, SSIM, and other quality indicators that allow for an objective assessment of the usefulness of the given method for the sharpening of satellite images.

B. Future Possibilities of SISR Architecture

Currently used resolution enhancement methods that employ deep neural networks, including GAN, require very large datasets and high computational capacity of the equipment that is used to train the networks. This is caused by a very high number of layers, which, in turn, results in a large number of parameters that have to be adjusted. They are still based on large volumes of training data so that the trained models are universal. However, the model may become overtrained if the database is too small. In such an event, even if the model achieves good results during network training, it does not function correctly for other data. Although creating databases is not a problem for tasks that enhance the resolution of LR images generated based on HR images, it may become a major limitation in the enhancement of images from nanosatellite systems, where no corresponding HR images are available. Due to that, the SISR methods require further improvement, particularly with sharpening nanosatellite images. The success of deep neural networks is attributed precisely to the vast databases that enable training on powerful representations. Unfortunately, scientists are still unable to recreate the learning process, so deep neural networks are considered a “black box.” Another challenge in the enhancement of the resolution of nanosatellite images is the need to define unambiguous evaluation metrics. Currently, the used MSE, SSIM, or PSNR metrics may often take excellent values, although the interpretation ability of the SR image does not improve, e.g., due to blurring. In addition, in publications, the results of image resolution improvement algorithms using

DL are presented in small images. Often, LR images consist of less than 10 000 pixels (e.g., Set5-4x, Set14-4x, BSD100-4x, URBAN100- 4x, FFHQ 256x256- 4x, FFHQ 512x512- 4x, FFHQ 1024x1024- 4x). This results, to a large extent, from the computational requirements of the algorithms that use GAN.

IV. Conclusion

This literature review presents state of the art. SISR has the potential to be applied to small satellite images. Apart from presenting the advantages and limitations of classic methods, the authors conducted an extensive review of contemporary super-resolution methods based on deep learning, particularly on generative adversarial networks. Deep learning methods show great potential in super-resolution satellite imagery. Despite their advanced sophisticated structure, some types of architecture based on convolutional networks cannot be successfully applied in SISR satellite imaging because they do not improve the interpretation abilities despite the enhanced resolution.

The authors have also demonstrated that numerous research projects so far have focused on enhancing the resolution of small images and the issues of applying SiSR methods in remote sensing, particularly in satellite imagery, to whole satellite scenes, whose dimensions are bigger than the dimension of images presented in publications. Furthermore, the authors presented future development directions with respect to sharpening whole scenes, optimizing parameters, and improving the architecture based on GAN to make them resistant to training instability.

References

[1] Satellite Database | Union of Concerned Scientists, Accessed: Sep. 25, 2022. [Online]. Available: https://www.ucsusa.org/resources/satellite-database

[2] D. Amarsaikhan, H. Blotevogel, J. van Genderen, M. Ganzorig, R. Gantuya, and B. Nergui, "Fusing high-resolution SAR and optical imagery for improved urban land cover study and classification," *Int. J. Image Data Fusion*, vol. 1, pp. 83–97, Mar. 2010, doi: 10.1080/19479830903562041.

[3] H. Luo, L. Li, H. Zhu, X. Kuai, Z. Zhang, and Y. Liu, "Land cover extraction from high resolution ZY-3 satellite imagery using ontology-based method," *ISPRS Int. J. Geo-Inf.*, vol. 5, no. 3, Mar. 2016, Art. no. 31, doi: 10.3390/ijgi5030031.

[4] P. Han *et al.*, "Monitoring trends in light pollution in China based on nighttime satellite imagery," *Remote Sens.*, vol. 6, no. 6, pp. 5541–5558, Jun. 2014, doi: 10.3390/rs6065541.

[5] W. Guo, W. Yang, H. Zhang, and G. Hua, "Geospatial object detection in high resolution satellite images based on multi-scale convolutional neural network," *Remote Sens.*, vol. 10, no. 1, Jan. 2018, Art. no. 131, doi: 10.3390/rs10010131.

[6] C. Witharana and H. J. Lynch, "An object-based image analysis approach for detecting penguin guano in very high spatial resolution satellite images," *Remote Sens.*, vol. 8, no. 5, May 2016, Art. no. 375, doi: 10.3390/rs8050375.

[7] H. Greidanus, "Satellite imaging for maritime surveillance of the european seas," in *Remote Sensing of the European Seas*, Berlin, Germany: Springer, 2008, pp. 343–358, doi: 10.1007/978-1-4020-6772-3_26.

[8] N. Gavankar and S. Ghosh, "Automatic building footprint extraction from high-resolution satellite image using mathematical morphology," *Eur. J. Remote Sens.*, vol. 51, pp. 182–193, Jan. 2018, doi: 10.1080/22797254.2017.1416676.

[9] J. Lee, H. Jang, J. Yang, and K. Yu, "Machine learning classification of buildings for map generalization," *ISPRS Int. J. Geo-Inf.*, vol. 6, no. 10, Oct. 2017, Art. no. 309, doi: 10.3390/ijgi6100309.

[10] G. W. Nagel, E. M. L. de M Novo, and M. Kampel, "Nanosatellites applied to optical earth observation: A review," *Rev. Ambiente Água*, vol. 15, pp. 1–19, Jun. 2020, doi: 10.4136/ambi-agua.2513.

[11] H. Kuuste *et al.*, "Imaging system for nanosatellite proximity operations," *Proc. Estonian Acad. Sci.*, vol. 63, no. 2S, 2014, Art. no. 250, doi: 10.3176/proc.2014.2S.06.

[12] J. Blommaert *et al.*, "CHIEM: A new compact camera for hyperspectral imaging," in *Proc. Small Satell. Conf.*, 2017, Paper 103.

[13] P. Mhangara, W. Mapurisa, and N. Mudau, "Image interpretability of nSight-1 nanosatellite imagery for remote sensing applications," *Aerospace*, vol. 7, Feb. 2020, Art. no. 19, doi: 10.3390/aerospace7020019.

[14] C. Yao, X. Luo, Y. Zhao, W. Zeng, and X. Chen, "A review on image classification of remote sensing using deep learning," in *Proc. 3rd IEEE Int. Conf. Comput. Commun.*, Dec. 2017, pp. 1947–1955, doi: 10.1109/CompComm.2017.8322878.

[15] J. Zhang, H. Zhao, and J. Li, "TRS: Transformers for remote sensing scene classification," *Remote Sens.*, vol. 13, no. 20, Jan. 2021, Art. no. 4143, doi: 10.3390/rs13204143.

[16] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao, "Review of image classification algorithms based on convolutional neural networks," *Remote Sens.*, vol. 13, no. 22, Jan. 2021, Art. no. 4712, doi: 10.3390/rs13224712.

[17] G. Tochon, J. Chanussot, M. D. Mura, and A. L. Bertozzi, "Object tracking by hierarchical decomposition of hyperspectral video sequences: Application to chemical gas plume tracking," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4567–4585, Aug. 2017, doi: 10.1109/TGRS.2017.2694159.

[18] D. Yu and S. Ji, "A new spatial-oriented object detection framework for remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2021, doi: 10.1109/TGRS.2021.3127232.

[19] Z. Dong, M. Wang, Y. Wang, Y. Zhu, and Z. Zhang, "Object detection in high resolution remote sensing imagery based on convolutional neural networks with suitable object scale features," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 2104–2114, Mar. 2020, doi: 10.1109/TGRS.2019.2953119.

[20] X. Li, J. Deng, and Y. Fang, "Few-shot object detection on remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 14–14, 2022, doi: 10.1109/TGRS.2021.3051383.

[21] Y. Li, C. Peng, Y. Chen, L. Jiao, L. Zhou, and R. Shang, "A deep learning method for change detection in synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5751–5763, Aug. 2019, doi: 10.1109/TGRS.2019.2901945.

[22] W. Zhao, L. Mou, J. Chen, Y. Bo, and W. J. Emery, "Incorporating metric learning and adversarial network for seasonal invariant change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2720–2731, Apr. 2020, doi: 10.1109/TGRS.2019.2953879.

[23] J. Zhang *et al.*, "Water body detection in high-resolution SAR images with cascaded fully-convolutional network and variable focal loss," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 316–332, Jan. 2021, doi: 10.1109/TGRS.2020.2999405.

[24] H. Liu, L. He, and J. Li, "Remote sensing image classification based on convolutional neural networks with two-fold sparse regularization," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2017, pp. 992–995, doi: 10.1109/IGARSS.2017.8127121.

[25] Z. Li, X. Cui, L. Wang, H. Zhang, X. Zhu, and Y. Zhang, "Spectral and spatial global context attention for hyperspectral image classification," *Remote Sens.*, vol. 13, no. 4, Jan. 2021, Art. no. 771, doi: 10.3390/rs13040771.

[26] J. M. Ramirez and H. Arguello, "Multiresolution compressive feature fusion for spectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9900–9911, Dec. 2019, doi: 10.1109/TGRS.2019.2930093.

[27] Z. Xue, J. Li, L. Cheng, and P. Du, "Spectral–spatial classification of hyperspectral data via morphological component analysis-based image separation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 70–84, Jan. 2015, doi: 10.1109/TGRS.2014.2318332.

[28] J.-Y. Rau, J.-P. Jhan, and Y.-C. Hsu, "Analysis of oblique aerial images for land cover and point cloud classification in an urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1304–1319, Mar. 2015, doi: 10.1109/TGRS.2014.2337658.

[29] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," Jul. 2015, Accessed: Jul. 15, 2021. [Online]. Available: http://arxiv.org/abs/1501.00092

[30] I. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2014, Paper 27, Accessed: Dec. 8, 2021. [Online]. Available: https://proceedings.neurips.cc/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html

[31] B. Zhao, S. Zhang, C. Xu, Y. Sun, and C. Deng, "Deep fake geography? When geospatial data encounter artificial intelligence," *Cartogr. Geographic Inf. Sci.*, vol. 48, no. 4, pp. 338–352, Jul. 2021, doi: 10.1080/15230406.2021.1910075.

[32] Z. Tan, M. Gao, X. Li, and L. Jiang, "A flexible reference-insensitive spatiotemporal fusion model for remote sensing images using conditional generative adversarial network," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, Jan. 2022, doi: 10.1109/TGRS.2021.3050551.

[33] X. Li, Z. Du, Y. Huang, and Z. Tan, "A deep translation (GAN) based change detection network for optical and SAR remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 179, pp. 14–34, Sep. 2021, doi: 10.1016/j.isprsjprs.2021.07.007.

[34] Y. Liu *et al.*, "CscGAN: Conditional scale-consistent generation network for multi-level remote sensing image to map translation," *Remote Sens.*, vol. 13, no. 10, Jan. 2021, Art. no. 1936, doi: 10.3390/rs13101936.

[35] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," May 2017, Accessed: Jun. 14, 2021. [Online]. Available: http://arxiv.org/abs/1609.04802

[36] X. Wang *et al.*, "ESRGAN: Enhanced super-resolution generative adversarial networks," Sep. 2018, Accessed: Jun. 15, 2021. [Online]. Available: http://arxiv.org/abs/1809.00219

[37] P. Getreuer, "Linear methods for image interpolation," *Image Process. Line*, vol. 1, pp. 238–259, Sep. 2011, doi: 10.5201/ipol.2011.g_lmii.

[38] O. Salvado, C. M. Hillenbrand, and D. L. Wilson, "Partial volume reduction by interpolation with reverse diffusion," *Int. J. Biomed. Imag.*, vol. 2006, 2006, Art. no. 92092, doi: 10.1155/IJBI/2006/92092.

[39] G. Chaudhary, K. Das, and M. Gopi, "Curvature minimizing depth interpolation for intuitive and interactive space curve sketching," [Online]. Available: https://www.researchgate.net/profile/MGopi/publication/228964270_Curvature_Minimizing_Depth_Interpolation_for_Intuitive_and_Interactive_Space_Curve_Sketching

[40] B. S. Morse and D. Schwartzwald, "Isophote-based interpolation," in *Proc. Int. Conf. Image Process.*, Oct. 1998, pp. 227–231, doi: 10.1109/ICIP.1998.999013.

[41] P. Oskoui-Fard and H. Stark, "Tomographic image reconstruction using the theory of convex projections," *IEEE Trans. Med. Imag.*, vol. 7, no. 1, pp. 45–58, Mar. 1988, doi: 10.1109/42.3928.

[42] F. Li, D. Fraser, and X. Jia, "Improved IBP for super-resolving remote sensing images," *Ann. Gis Geographic Inf. Sci.*, vol. 12, pp. 106–111, Dec. 2006, doi: 10.1080/10824000609480624.

[43] A. Gilman, D. G. Bailey, and S. R. Marsland, "Interpolation models for image super-resolution," in *Proc. 4th IEEE Int. Symp. Electron. Des. Test Appl.*, Jan. 2008, pp. 55–60, doi: 10.1109/DELTA.2008.104.

[44] A. Belov and A. Denisova, "Spatial interpolation methods for spectral-spatial remote sensing image super-resolution algorithm based on gradient descent approach," *J. Phys., Conf. Ser.*, vol. 1368, Nov. 2019, Art. no. 032006, doi: 10.1088/1742-6596/1368/3/032006.

[45] X. Qifang, Y. Guoqing, and L. Pin, "Super-resolution reconstruction of satellite video images based on interpolation method," *Procedia Comput. Sci.*, vol. 107, pp. 454–459, Jan. 2017, doi: 10.1016/j.procs.2017.03.089.

[46] W. Shi, Y. Tian, and K. Liu, "An integrated method for satellite image interpolation," *Int. J. Remote Sens.*, vol. 28, no. 6, pp. 1355–1371, Mar. 2007, doi: 10.1080/01431160600851876.

[47] J. A. Malpica, "Splines interpolation in high resolution satellite imagery," in *Advances in Visual Computing*. Berlin, Germany: Springer, 2005, pp. 562–570, doi: 10.1007/11595755_68.

[48] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 2, pp. 29–56, Jun. 2017, doi: 10.1109/MGRS.2016.2637824.

[49] J. Marcello, E. Ibarrola-Ulzurrun, C. Gonzalo-Martín, J. Chanussot, and G. Vivone, "Assessment of hyperspectral sharpening methods for the monitoring of natural areas using multiplatform remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 8208–8222, Oct. 2019, doi: 10.1109/TGRS.2019.2918932.

[50] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, *Hyperspectral Image Fusion*. Berlin, Germany: Springer, 2021.

[51] A. Mookambiga and V. Gomathi, "Comprehensive review on fusion techniques for spatial information enhancement in hyperspectral imagery," *Multidimensional Syst. Signal Process.*, vol. 27, no. 4, pp. 863–889, Oct. 2016, doi: 10.1007/s11045-016-0415-2.

[52] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, "The use of intensity-hue-saturation transformations for merging SPOT panchromatic and ∼ultispectralImage data," *Photocramm. Eng. Remote Sens.*, vol. 56, Apr. 1990, Art. no. 459467.

[53] A. D. Philip, "The use of intensity-hue-saturation transformation for producing color shaded relief images," *Photogramm. Eng. Remote Sens.*, vol. 60, pp. 1369–1374, 1994.

[54] K. Kpalma, M. Chikr El-Mezouar, and N. Taleb, "Recent trends in satellite image pan-sharpening techniques," Vrniacka Banja, Serbia, Jun. 2014, Accessed: Jun. 17, 2021. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01075703

[55] G. Vivone *et al.*, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015, doi: 10.1109/TGRS.2014.2361734.

[56] Shruti and S. Budhiraja, "Multiscale image fusion for pan-sharpening of multispectral images using saliency detection," in *Proc. 9th Int. Conf. Contemporary Comput.*, Aug. 2016, pp. 1–6. doi: 10.1109/IC3.2016.7880253.

[57] Y. Zhang, "(PDF) understanding image fusion," 2008, Accessed: May 31, 2021. [Online]. Available: https://www.researchgate.net/publication/342318144_Understanding_Image_Fusion

[58] M. Choi, "A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1672–1682, Jul. 2006, doi: 10.1109/TGRS.2006.869923.

[59] M.-J. Choi, H.-C. Kim, N. Cho, and H. Kim, "An improved intensity-hue-saturation method for IKONOS image fusion," *Int. J. Remote Sens.*, vol. 13, pp. 1–5, Jan. 2006.

[60] R. Welch and M. Ehlers, "Merging multiresolution SPOT HRV and landsat TM data," *Photogramm.Eng. Remote Sens.*, vol. 53, pp. 301–303, Mar. 1987.

[61] S. M. A. Wady, Y. Bentoutou, A. Bengermikh, A. Bounoua, and N. Taleb, "A new IHS and wavelet based pansharpening algorithm for high spatial resolution satellite imagery," *Adv. Space Res.*, vol. 66, no. 7, pp. 1507–1521, Oct. 2020, doi: 10.1016/j.asr.2020.06.001.

[62] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, Sep. 2001, doi: 10.1016/S1566-2535(01)00036-7.

[63] J. Pushparaj and A. V. Hegde, "Evaluation of pan-sharpening methods for spatial and spectral quality," *Appl. Geomat.*, vol. 9, no. 1, pp. 1–12, Mar. 2017, doi: 10.1007/s12518-016-0179-2.

[64] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 746–750, Oct. 2010, doi: 10.1109/LGRS.2010.2046715.

[65] C. Yang, Q. Zhan, H. Liu, and R. Ma, "An IHS-based pan-sharpening method for spectral fidelity improvement using ripplet transform and compressed sensing," *Sensors*, vol. 18, no. 11, Oct. 2018, Art. no. 3624, doi: 10.3390/s18113624.

[66] W. Hallada and S. Cox, "Image sharpening for mixed spatial and spectral resolution satellite systems," in *Proc. Int. Symp. Remote Sens. Environ.*, 1983, Accessed: May 31, 2021. [Online]. Available: /paper/Image-sharpening-for-mixed-spatial-and-spectral-Hallada-Cox/42fdced145103292dd3ea3c5441145abf6a835e5

[67] H. R. Shahdoosti, "MS and PAN image fusion by combining brovey and wavelet methods," p. 8, doi: 10.48550/arXiv.1701.01996.

[68] H. R. Shahdoosti, "Improved adaptive brovey as a new method for image fusion," Jul. 2018, Accessed: May 24, 2021. [Online]. Available: http://arxiv.org/abs/1807.09610

[69] S. S. Khan, Q. Ran, M. Khan, and Z. Ji, "Pan-sharpening framework based on laplacian sharpening with brovey," in *Proc. IEEE Int. Conf. Signal Inf. Data Process.*, Dec. 2019, pp. 1–5, doi: 10.1109/ICSIDP47821.2019.9173129.

[70] J. Chavez Pat, S. Sides, and J. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *Photogramm. Eng. Remote Sens.*, vol. 57, pp. 265–303, Mar. 1991.

[71] V. P. Shah, N. H. Younan, and R. L. King, "An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1323–1335, May 2008, doi: 10.1109/TGRS.2008.916211.

[72] Z. Zhou, N. Ma, Y. Li, P. Yang, P. Zhang, and Y. Li, "Variational PCA fusion for pan-sharpening very high resolution imagery," *Sci. China Inf. Sci.*, vol. 57, no. 11, pp. 1–10, Nov. 2014, doi: 10.1007/s11432-014-5108-6.

[73] V. Shah, N. H. Younan, and R. King, "An adaptive PCA-based approach to pan-sharpening," in *Proc. SPIE, Int. Soc. Opt. Eng.*, Oct. 2007, Paper 678402, doi: 10.1117/12.736674.

[74] A. M. Gonzalez, X. Otazu, O. Fors, R. García, and J. Nunez, "Fusion of different spatial and spectral resolution images: Development, application and comparison of new methods based on wavelets," in *Proc. RAQRS*, Jan. 2002, pp. 228–237.

[75] M. Ghadjati, A. Moussaoui, and A. Boukharouba, "A novel iterative PCA–based pansharpening method," *Remote Sens. Lett.*, vol. 10, no. 3, pp. 264–273, Mar. 2019, doi: 10.1080/2150704X.2018.1547443.

[76] M. Ehlers, "Multisensor image fusion techniques in remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 46, no. 1, pp. 19–30, Feb. 1991, doi: 10.1016/0924-2716(91)90003-E.

[77] C. A. Laben and V. B. Brower, "Process for enhancing the spatial resolution of multispectral imagery using Pan-sharpening," U.S. Patent 6 011 875 A, Apr. 29, 1998. [Online]. Available: https://lens.org/135-660-046-023-136

[78] H. Sunuprapto, P. Danoedoro, and S. Ritohardoyo, "Evaluation of Pansharpening method: Applied to artisanal gold mining monitoring in gunung pani forest area," *Procedia Environ. Sci.*, vol. 33, pp. 230–238, 2016, doi: 10.1016/j.proenv.2016.03.074.

[79] S. Klonus and M. Ehlers, "Image fusion using the Ehlers spectral characteristics preservation algorithm," *GIScience Remote Sens.*, vol. 44, no. 2, pp. 93–116, Jun. 2007, doi: 10.2747/1548-1603.44.2.93.

[80] S. Xu and M. Ehlers, "Hyperspectral image sharpening based on ehlers fusion," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLII-2/W7, pp. 941–947, Sep. 2017, doi: 10.5194/isprs-archives-XLII-2-W7-941-2017.

[81] S. Klonus and M. Ehlers, "Pansharpening with Terrasar-X and optical data," *3rd TerraSAR-X Sci. Team Meeting*, pp. 25–26, 2008.

[82] R. E. Crippen, "A simple spatial filtering routine for the cosmetic removal of scan-line noise from landsat TM P-Tape imagery," *Photogramm. Eng. Remote Sens.*, vol. 55, pp. 327–331, 1989.

[83] A. Garzelli, B. Aiazzi, L. Alparone, S. Lolli, and G. Vivone, "Multispectral pansharpening with radiative transfer-based detail-injection modeling for preserving changes in vegetation cover," *Remote Sens*, vol. 10, no. 8, Aug. 2018, Art. no. 1308, doi: 10.3390/rs10081308.

[84] S. Kahraman and A. Ertürk, "A comprehensive review of pansharpening algorithms for Göktürk-2 satellite images," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. IV-4/W4, pp. 263–270, Nov. 2017, doi: 10.5194/isprs-annals-IV-4-W4-263-2017.

[85] M. S. Karoui, K. Djerriri, and I. Boukerch, "Pansharpening multispectral remote sensing data by multiplicative joint nonnegative matrix factorization," *Int. J. Remote Sens.*, vol. 37, pp. 805–818, Feb. 2016, doi: 10.1080/01431161.2015.1137650.

[86] G. Palubinskas, "Model-based image adjustment for a successful pansharpening," Mar. 2021, Accessed: Jun. 08, 2021. [Online]. Available: http://arxiv.org/abs/2103.03062

[87] P. F. Fougere, "A defense of the Gram-Schmidt orthogonalization procedure applied to spherical harmonic analysis," *J. Geophys. Res. 1896–1977*, vol. 71, no. 21, pp. 5171–5174, 1966, doi: 10.1029/JZ071i021p05171.

[88] T. Maurer, "How to pan-sharpen images using the Gram-Schmidt pan-sharpen method—A recipe," *ISPRS—Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XL-1/W1, pp. 239–244, May 2013, doi: 10.5194/isprsarchives-XL-1-W1-239-2013.

[89] Z. Huang, Q. Chen, Q. Chen, and X. Liu, "Variational pansharpening for hyperspectral imagery constrained by spectral shape and Gram–Schmidt transformation," *Sensors*, vol. 18, no. 12, Dec. 2018, Art. no. 4330, doi: 10.3390/s18124330.

[90] H. Li, Y. Zhang, Y. Gao, and S. Yue, "Using guided filtering to improve gram-schmidt based pansharpening method for geoeye-1 satellite images," in *Proc. 4th Int. Conf. Inf. Syst. Comput. Technol.*, 2016, pp. 33–37, doi: 10.2991/isct-16.2016.6.

[91] G. Sarp, "Spectral and spatial quality analysis of pan-sharpening algorithms: A case study in Istanbul," *Eur. J. Remote Sens.*, vol. 47, no. 1, pp. 19–28, Jan. 2014, doi: 10.5721/EuJRS20144702.

[92] X. Otazu, M. Gonzalez-Audicana, O. Fors, and J. Nunez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, Oct. 2005, doi: 10.1109/TGRS.2005.856106.

[93] M. Gonzalez-Audicana, J. L. Saleta, R. G. Catalan, and R. Garcia, "Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1291–1299, Jun. 2004, doi: 10.1109/TGRS.2004.825593.

[94] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, Jan. 2011, doi: 10.1109/TGRS.2010.2051674.

[95] H. R. Shahdoosti and H. Ghassemian, "Fusion of MS and PAN images preserving spectral quality," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 611–615, Mar. 2015, doi: 10.1109/LGRS.2014.2353135.

[96] C. E. Shannon, "A mathematical theory of communication," *Bell Labs Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948, doi: 10.1002/j.1538-7305.1948.tb01338.x.

[97] P. Dherete and B. Rouge, "Image de-blurring and application to SPOT5 THR satellite imaging," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2003, vol. 1, pp. 318–320, doi: 10.1109/IGARSS.2003.1293762.

[98] C. Latry and B. Rouge, "Super resolution: Quincunx sampling and fusion processing," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2003, vol. 1, pp. 315–317, doi: 10.1109/IGARSS.2003.1293761.

[99] C. Latry and B. Rougé, "In-flight commissioning of SPOT5 THR Quincunx sampling mode," in *Proc. Sensors, Syst. Next-Gener. Satellites VI*, 2003, Paper 199, doi: 10.1117/12.462632.

[100] A. Navarro *et al.*, "Crop monitoring based on SPOT-5 take-5 and Sentinel-1A data for the estimation of crop water requirements," *Remote Sens.*, vol. 8, no. 6, Jun. 2016, Art. no. 525, doi: 10.3390/rs8060525.

[101] D. C. Wehlage, J. A. Gamon, D. Thayer, and D. V. Hildebrand, "Interannual variability in dry mixed-grass prairie yield: A comparison of MODIS, SPOT, and field measurements," *Remote Sens.*, vol. 8, no. 10, Oct. 2016, Art. no. 872, doi: 10.3390/rs8100872.

[102] M. A. Castillo-Santiago, M. Ricker, and B. H. J. de Jong, "Estimation of tropical forest structure from SPOT-5 satellite images," *Int. J. Remote Sens.*, vol. 31, no. 10, pp. 2767–2782, May 2010, doi: 10.1080/01431160903095460.

[103] M. Hao, W. Shi, H. Zhang, Q. Wang, and K. Deng, "A scale-driven change detection method incorporating uncertainty analysis for remote sensing images," *Remote Sens.*, vol. 8, no. 9, Sep. 2016, Art. no. 745, doi: 10.3390/rs8090745.

[104] N.-B. Chang, M. Han, W. Yao, L.-C. Chen, and S. Xu, "Change detection of land use and land cover in an urban region with SPOT-5 images and partial Lanczos extreme learning machine," *J. Appl. Remote Sens.*, vol. 4, no. 1, Nov. 2010, Art. no. 043551, doi: 10.1117/1.3518096.

[105] D. Lu, M. Batistella, and E. Moran, "Integration of landsat TM and SPOT HRG images for vegetation change detection in the Brazilian Amazon," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 4, pp. 421–430, Apr. 2008, doi: 10.14358/PERS.74.4.421.

[106] E. Sertel and S. S. Akay, "High resolution mapping of urban areas using SPOT-5 images and ancillary data," *Int. J. Environ. Geoinform.*, vol. 2, no. 2, pp. 63–76, Aug. 2015, doi: 10.30897/ijegeo.303545.

[107] V. Pasqualini *et al.*, "Use of SPOT 5 for mapping seagrasses: An application to Posidonia oceanica," *Remote Sens. Environ.*, vol. 94, no. 1, pp. 39–45, Jan. 2005, doi: 10.1016/j.rse.2004.09.010.

[108] A. M. Borghuis, K. Chang, and H. Y. Lee, "Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery," *Int. J. Remote Sens.*, vol. 28, no. 8, pp. 1843–1856, Apr. 2007, doi: 10.1080/01431160600935638.

[109] A. Fruchter and R. N. Hook, "Novel image reconstruction method applied to deep hubble space telescope images," in *Proc. Appl. Digit. Image Process. XX*, Oct. 1997, vol. 3164, pp. 120–125, doi: 10.1117/12.292751.

[110] M. T. Merino and J. Nunez, "Super-resolution of remotely sensed images with variable-pixel linear reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1446–1457, May 2007, doi: 10.1109/TGRS.2007.893271.

[111] Y. Ito, "Resolution enhancement of SAR image using a multiframe super resolution technique," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2009, vol. 4, pp. 446–449, doi: 10.1109/IGARSS.2009.5417409.

[112] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP Graph. Models Image Process.*, vol. 53, no. 3, pp. 231–239, May 1991, doi: 10.1016/1049-9652(91)90045-L.

[113] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Hyperspectral super-resolution: A coupled tensor factorization approach," *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6503–6517, Dec. 2018, doi: 10.1109/TSP.2018.2876362.

[114] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Nonlocal patch tensor sparse representation for hyperspectral image super-resolution," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 3034–3047, Jun. 2019, doi: 10.1109/TIP.2019.2893530.

[115] N. Liu, L. Li, W. Li, R. Tao, J. E. Fowler, and J. Chanussot, "Hyperspectral restoration and fusion with multispectral imagery via low-rank tensor-approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7817–7830, Sep. 2021, doi: 10.1109/TGRS.2020.3049014.

[116] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, "A new pan-sharpening method with deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1037–1041, May 2015, doi: 10.1109/LGRS.2014.2376034.

[117] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, Jul. 2016, Art. no. 594, doi: 10.3390/rs8070594.
[118] G. Scarpa, S. Vitale, and D. Cozzolino, "Target-adaptive CNN-based pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5443–5457, Sep. 2018, doi: 10.1109/TGRS.2018.2817393.
[119] J. Wang, Z. Shao, X. Huang, T. Lu, and R. Zhang, "A dual-path fusion network for pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jun. 2022, Art. no. 5403214, doi: 10.1109/TGRS.2021.3090585.
[120] Z. Shao, L. Wang, Z. Wang, and J. Deng, "Remote sensing image super-resolution using sparse representation and coupled sparse autoencoder," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 2663–2674, Aug. 2019, doi: 10.1109/JSTARS.2019.2925456.
[121] Z. Shao and J. Cai, "Remote sensing image fusion with deep convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1656–1669, May 2018, doi: 10.1109/JSTARS.2018.2805923.
[122] C. Shang *et al.*, "Spatiotemporal reflectance fusion using a generative adversarial network," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Mar. 2021, Art. no. 5400915, doi: 10.1109/TGRS.2021.3065418.
[123] H. Zhang, Y. Song, C. Han, and L. Zhang, "Remote sensing image spatiotemporal fusion using a generative adversarial network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4273–4286, May 2021, doi: 10.1109/TGRS.2020.3010530.
[124] D. Feng, D. An, L. Chen, and X. Huang, "Holographic SAR tomography 3-D reconstruction based on iterative adaptive approach and generalized likelihood ratio test," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 305–315, Jan. 2021, doi: 10.1109/TGRS.2020.2994201.
[125] J. Leinonen, D. Nerini, and A. Berne, "Stochastic super-resolution for downscaling time-evolving atmospheric fields with a generative adversarial network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 1–13, Sep. 2020, doi: 10.1109/TGRS.2020.3032790.
[126] F. Han, H. Zhang, S. Chatterjee, Q. Guo, and S. Wan, "A modified generative adversarial nets integrated with stochastic approach for realizing super-resolution reservoir simulation," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1325–1336, Feb. 2020, doi: 10.1109/TGRS.2019.2945946.
[127] J. Wang, Z. Shao, X. Huang, T. Lu, R. Zhang, and J. Ma, "Enhanced image prior for unsupervised remoting sensing super-resolution," *Neural Netw.*, vol. 143, no. C, pp. 400–412, Nov. 2021, doi: 10.1016/j.neunet.2021.06.005.
[128] Z. Shao, J. Cai, P. Fu, L. Hu, and T. Liu, "Deep learning-based fusion of Landsat-8 and Sentinel-2 images for a harmonized surface reflectance product," *Remote Sens. Environ.*, vol. 235, Dec. 2019, Art. no. 111425, doi: 10.1016/j.rse.2019.111425.
[129] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255, doi: 10.1109/CVPR.2009.5206848.
[130] F. Yu, A. Seff, Y. Zhang, S. Song, T. Funkhouser, and J. Xiao, "LSUN: Construction of a large-scale image dataset using deep learning with humans in the loop," Jun. 2016, Accessed: Dec. 28, 2021. [Online]. Available: http://arxiv.org/abs/1506.03365
[131] T.-Y. Lin *et al.*, "Microsoft COCO: Common objects in context," Feb. 2015, Accessed: Dec. 28, 2021. [Online]. Available: http://arxiv.org/abs/1405.0312
[132] E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Honolulu, HI, USA, Jul. 2017, pp. 1122–1131, doi: 10.1109/CVPRW.2017.150.
[133] H. Ren, A. Kheradmand, M. El-Khamy, S. Wang, D. Bai, and J. Lee, "Real-world super-resolution using generative adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Seattle, WA, USA, Jun. 2020, pp. 1760–1768, doi: 10.1109/CVPRW50498.2020.00226.
[134] Z.-S. Liu, W.-C. Siu, L.-W. Wang, C.-T. Li, M.-P. Cani, and Y.-L. Chan, "Unsupervised real image super-resolution via generative variational autoencoder," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Seattle, WA, USA, Jun. 2020, pp. 1788–1797, doi: 10.1109/CVPRW50498.2020.00229.
[135] R. M. Umer, G. L. Foresti, and C. Micheloni, "Deep generative adversarial residual convolutional networks for real-world super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Seattle, WA, USA, Jun. 2020, pp. 1769–1777, doi: 10.1109/CVPRW50498.2020.00227.
[136] A. Lugmayr *et al.*, "NTIRE 2021 learning the super-resolution space challenge," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2021, pp. 596–612, doi: 10.1109/CVPRW53098.2021.00072.
[137] S. Son *et al.*, "NTIRE 2021 challenge on video super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2021, pp. 166–181, doi: 10.1109/CVPRW53098.2021.00026.
[138] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," Aug. 2016, Accessed: Jul. 15, 2021. [Online]. Available: http://arxiv.org/abs/1608.00367
[139] M. Bevilacqua, A. Roumy, C. Guillemot, and M. A. Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in *Proc. Brit. Mach. Vis. Conf.*, 2012, pp. 135.1–135.10, doi: 10.5244/C.26.135.
[140] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Proc. Int. Conf. Curves Surfaces*, 2010, vol. 6920, pp. 711–730, doi: 10.1007/978-3-642-27413-8_47.
[141] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th IEEE Int. Conf. Comput. Vis.*, Jul. 2001, vol. 2, pp. 416–423, doi: 10.1109/ICCV.2001.937655.
[142] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 5197–5206, doi: 10.1109/CVPR.2015.7299156.
[143] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks,", Mar. 2019, Accessed: Sep. 17, 2021. [Online]. Available: http://arxiv.org/abs/1812.04948
[144] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," Jul. 2017, Accessed: Jul. 15, 2021. [Online]. Available: http://arxiv.org/abs/1707.02921
[145] T. Dai, J. Cai, Y. Zhang, S.-T. Xia, and L. Zhang, "Second-order attention network for single image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 11057–11066, doi: 10.1109/CVPR.2019.01132.
[146] B. Niu *et al.*, "Single image super-resolution via a holistic attention network," Aug. 2020, Accessed: Jul. 15, 2021. [Online]. Available: http://arxiv.org/abs/2008.08767
[147] H. Cao and S. Mi, "Weighted SRGAN and reconstruction loss analysis for accurate image super resolution," *J. Phys., Conf. Ser.*, vol. 1903, no. 1, Apr. 2021, Art. no. 012050, doi: 10.1088/1742-6596/1903/1/012050.
[148] J.-T. Jiang, H. L. Huang, and L. Hu, "Single image super-resolution: Depthwise separable convolution super-resolution generative adversarial network," *Appl. Sci.*, vol. 10, Jan. 2020, Art. no. 375, doi: 10.3390/app10010375.
[149] W. Shi *et al.*, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," Sep. 2016, Accessed: Jul. 19, 2021. [Online]. Available: http://arxiv.org/abs/1609.05158
[150] M. S. M. Sajjadi, B. Schölkopf, and M. Hirsch, "EnhanceNet: Single image super-resolution through automated texture synthesis," Jul. 2017, Accessed: Jul. 15, 2021. [Online]. Available: http://arxiv.org/abs/1612.07919
[151] R. Kalarot, T. Li, and F. Porikli, "Component attention guided face super-resolution network: CAGFace," Oct. 2019, Accessed: Jul. 15, 2021. [Online]. Available: http://arxiv.org/abs/1910.08761
[152] Z. Li, J. Yang, Z. Liu, X. Yang, G. Jeon, and W. Wu, "Feedback network for image super-resolution," Jun. 2019, Accessed: Jul. 15, 2021. [Online]. Available: http://arxiv.org/abs/1903.09814
[153] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1646–1654, doi: 10.1109/CVPR.2016.182.
[154] L. Wagner, L. Liebel, and M. Körner, "Deep residual learning for single-image super-resolution of multi-spectral satellite imagery," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. IV-2/W7, pp. 189–196, Sep. 2019, doi: 10.5194/isprs-annals-IV-2-W7-189-2019.
[155] X. Feng, X. Su, J. Shen, and H. Jin, "Single space object image denoising and super-resolution reconstructing using deep convolutional networks," *Remote Sens.*, vol. 11, no. 16, Jan. 2019, Art. no. 1910, doi: 10.3390/rs11161910.
[156] N. Latte and P. Lejeune, "PlanetScope radiometric normalization and sentinel-2 super-resolution (2.5m): A straightforward spectral-spatial fusion of multi-satellite multi-sensor images using residual convolutional neural networks," *Remote Sens.*, vol. 12, no. 15, Jan. 2020, Art. no. 2366, doi: 10.3390/rs12152366.

[157] T. Lu, J. Wang, Y. Zhang, Z. Wang, and J. Jiang, "Satellite image super-resolution via multi-scale residual deep neural network," *Remote Sens.*, vol. 11, no. 13, Jan. 2019, Art. no. 1588, doi: 10.3390/rs11131588.
[158] W. Dong, Y. Yang, J. Qu, W. Xie, and Y. Li, "Fusion of hyperspectral and panchromatic images using generative adversarial network and image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, May 2021, Art. no. 5508413, doi: 10.1109/TGRS.2021.3078711.
[159] T. Li and Y. Gu, "Progressive spatial-spectral joint network for hyperspectral image reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, May 2021, Art. no. 5507414, doi: 10.1109/TGRS.2021.3079969.
[160] J. Ying, H.-L. Shen, and S.-Y. Cao, "Unaligned hyperspectral image fusion via registration and interpolation modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, May 2021, Art. no. 5511114, doi: 10.1109/TGRS.2021.3081136.
[161] X. Wang, J. Ma, and J. Jiang, "Hyperspectral image super-resolution via recurrent feedback embedding and spatial-spectral consistency regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Mar. 2021, Art. no. 5503113, doi: 10.1109/TGRS.2021.3064450.
[162] D. Liu, J. Li, and Q. Yuan, "A spectral grouping and attention-driven residual dense network for hyperspectral image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7711–7725, Sep. 2021, doi: 10.1109/TGRS.2021.3049875.
[163] R. Dian, S. Li, L. Fang, and Q. Wei, "Multispectral and hyperspectral image fusion with spatial-spectral sparse representation," *Inf. Fusion*, vol. 49, pp. 262–270, Sep. 2019, doi: 10.1016/j.inffus.2018.11.012.
[164] J. Li, R. Cui, B. Li, R. Song, Y. Li, and Q. Du, "Hyperspectral image super-resolution with 1D–2D attentional convolutional neural network," *Remote Sens.*, vol. 11, no. 23, Jan. 2019, Art. no. 2859, doi: 10.3390/rs11232859.
[165] I. J. Goodfellow *et al.*, "Generative adversarial networks," Jun. 2014, Accessed: Sep. 13, 2021. [Online]. Available: http://arxiv.org/abs/1406.2661
[166] L. Liebel and M. Körner, "Single-image super resolution for multispectral remote sensing data using convolutional neural networks," *ISPRS—Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 41B3, pp. 883–890, Jun. 2016, doi: 10.5194/isprs-archives-XLI-B3-883-2016.
[167] Y. Xiong *et al.*, "Improved SRGAN for remote sensing image super-resolution across locations and sensors," *Remote Sens.*, vol. 12, no. 8, Jan. 2020, Art. no. 1263, doi: 10.3390/rs12081263.
[168] X. Dou, C. Li, Q. Shi, and M. Liu, "Super-resolution for hyperspectral remote sensing images based on the 3D attention-SRGAN network," *Remote Sens.,* vol. 12, no. 7, Jan. 2020, Art. no. 1204, doi: 10.3390/rs12071204.
[169] L. S. Romero, J. Marcello, and V. Vilaplana, "Super-resolution of Sentinel-2 imagery using generative adversarial networks," *Remote Sens.*, vol. 12, no. 15, Jan. 2020, Art. no. 2424, doi: 10.3390/rs12152424.
[170] M. Pashaei, M. J. Starek, H. Kamangir, and J. Berryhill, "Deep learning-based single image super-resolution: An investigation for dense scene reconstruction with UAS photogrammetry," *Remote Sens.*, vol. 12, no. 11, Jan. 2020, Art. no. 1757, doi: 10.3390/rs12111757.
[171] J. Rabbi, N. Ray, M. Schubert, S. Chowdhury, and D. Chao, "Small-object detection in remote sensing images with end-to-end edge-enhanced GAN and object detector network," *Remote Sens.*, vol. 12, no. 9, Jan. 2020, Art. no. 1432, doi: 10.3390/rs12091432.
[172] Planet | Homepage, *Planet*. Accessed: Dec. 10, 2021. [Online]. Available: https://www.planet.com/
[173] M. Galar, R. Sesma, C. Ayala, L. Albizua, and C. Aranda, "Super-resolution of Sentinel-2 images using convolutional neural networks and real ground truth data," *Remote Sens.*, vol. 12, no. 18, Jan. 2020, Art. no. 2941, doi: 10.3390/rs12182941.
[174] F. Salvetti, V. Mazzia, A. Khaliq, and M. Chiaberge, "Multi-image super resolution of remotely sensed images using residual attention deep neural networks," *Remote Sens.*, vol. 12, no. 14, Jan. 2020, Art. no. 2207, doi: 10.3390/rs12142207.
[175] M. Qin *et al.*, "Achieving higher resolution lake area from remote sensing images through an unsupervised deep learning super-resolution method," *Remote Sens.*, vol. 12, no. 12, Jan. 2020, Art. no. 1937, doi: 10.3390/rs12121937.
[176] H. Ahn and C. Yim, "Convolutional neural networks using skip connections with layer groups for super-resolution image reconstruction based on deep learning," *Appl. Sci.*, vol. 10, no. 6, Jan. 2020, Art. no. 1959, doi: 10.3390/app10061959.
[177] L. Courtrai, M.-T. Pham, and S. Lefèvre, "Small object detection in remote sensing images based on super-resolution with auxiliary generative adversarial networks," *Remote Sens.*, vol. 12, no. 19, Jan. 2020, Art. no. 3152, doi: 10.3390/rs12193152.
[178] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of wasserstein GANs," Dec. 2017, Accessed: Jun. 16, 2021. [Online]. Available: http://arxiv.org/abs/1704.00028
[179] Y. Gao, H. Li, J. Dong, and G. Feng, "A deep convolutional network for medical image super-resolution," in *Proc. Chin. Automat. Congr.*, Oct. 2017, pp. 5310–5315, doi: 10.1109/CAC.2017.8243724.
[180] D. Mahapatra and B. Bozorgtabar, "Progressive generative adversarial networks for medical image super resolution," Feb. 2019, Accessed: May 28, 2021. [Online]. Available: http://arxiv.org/abs/1902.02144
[181] T. Gulati, S. Sengupta, and V. Lakshminarayanan, "Application of an enhanced deep super-resolution network in retinal image analysis," in *Proc. Ophthalmic Technol. XXX*, Feb. 2020, vol. 11218, Paper 112181K, doi: 10.1117/12.2543791.
[182] G. Zamzmi, S. Rajaraman, and S. Antani, "Accelerating super-resolution and visual task analysis in medical images," *Appl. Sci.*, vol. 10, no. 12, Jan. 2020, Art. no. 4282, doi: 10.3390/app10124282.
[183] "A pan-sharpening method based on evolutionary optimization and IHS transformation," Accessed: Dec. 10, 2021. [Online]. Available: https://www.hindawi.com/journals/mpe/2017/8269078/
[184] S. Panchal and R. Thakker, "Implementation and comparative quantitative assessment of different multispectral image pansharpening approaches," *Signal Image Process., Int. J.*, vol. 6, no. 5, pp. 35–48, Jun. 2015, doi: 10.5121/sipij.2015.6503.
[185] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion ?," in *Proc. 3rd Int. Conf. Fusion Earth Data, Merging Point Meas. Raster Maps Remotely Sensed Images*, Sophia Antipolis, France, Jan. 2000, pp. 99–103, Accessed: Sep. 17, 2021. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00395027
[186] Q. Du, O. Gungor, and J. Shan, "Performance evaluation for pansharpening techniques," in *Proc. Geosci. Remote Sens. Symp.*, 2005, Art. no. 4266, doi: 10.1109/IGARSS.2005.1525860.
[187] Y. Chen, "A new methodology of spatial cross-correlation analysis," *PLoS One*, vol. 10, no. 5, May 2015, Art. no. e0126158, doi: 10.1371/journal.pone.0126158.
[188] A. Goetz, W. Boardman, and R. Yunas, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," *Proc. 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, pp. 147–149.
[189] C.-I. Chang, "Spectral information divergence for hyperspectral image analysis," in *Proc. Int. Geosci. Remote Sens. Symp.*, Jun. 1999, vol. 1, pp. 509–511, doi: 10.1109/IGARSS.1999.773549.
[190] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002, doi: 10.1109/97.995823.
[191] X. Wang, L. Xie, C. Dong, and Y. Shan, "Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data," Aug. 2021, Accessed: Mar. 16, 2022. [Online]. Available: http://arxiv.org/abs/2107.10833

Kinga Karwowska received the M.Sc. degree in geoinformatics in 2020 from the Military University of Technology, Warsaw, Poland, where she is currently working toward the Ph.D. degree with the Doctoral School.

Damian Wierzbicki received the Ph.D. degree in photogrammetry and remote sensing from the Military University of Technology, Warsaw, Poland, in 2015.

He is currently an Associate Professor with the Department of Imagery Intelligence, Faculty of Civil Engineering and Geodesy, Military University of Technology, where he teaches photogrammetry and remote sensing and image processing. His research interests include UAV navigation and image processing, deep learning in remote sensing, the development of new algorithms for object detection, and the classification of image sequences from UAVs.