



## Article

# Full-Resolution Quality Assessment for Pansharpening

Giuseppe Scarpa and Matteo Ciotola \*

Department of Electrical Engineering and Information Technology (DIETI), University Federico II, 80125 Naples, Italy; giscarpa@unina.it

\* Correspondence: matteo.ciotola@unina.it

**Abstract:** A reliable quality assessment procedure for pansharpening methods is of critical importance for the development of the related solutions. Unfortunately, the lack of ground truths to be used as guidance for an objective evaluation has pushed the community to resort to two approaches, which can also be jointly applied. Hence, two kinds of indexes can be found in the literature: (i) reference-based reduced-resolution indexes aimed to assess the synthesis ability; (ii) no-reference subjective quality indexes for full-resolution datasets aimed to assess spectral and spatial consistency. Both reference-based and no-reference indexes present critical shortcomings, which motivate the community to explore new solutions. In this work, we propose an alternative no-reference full-resolution assessment framework. On one side, we introduce a protocol, namely the reprojection protocol, to take care of the spectral consistency issue. On the other side, a new index of the spatial consistency between the pansharpened image and the panchromatic band at full resolution is also proposed. Experimental results carried out on different datasets/sensors demonstrate the effectiveness of the proposed approach.

**Keywords:** super-resolution; image enhancement; convolutional neural networks; data fusion; multispectral images



**Citation:** Scarpa, G.; Ciotola, M.

Full-Resolution Quality Assessment for Pansharpening. *Remote Sens.* **2022**, *14*, 1808. <https://doi.org/10.3390/rs14081808>

Academic Editor: Geert Verhoeven

Received: 8 March 2022

Accepted: 6 April 2022

Published: 8 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Image pansharpening is the process of merging two observations of the same scene, a low-resolution multispectral (MS) component and a high-resolution panchromatic (PAN) component, to generate a new multispectral image that displays both the rich spectral content of the MS and the high resolution of the PAN. By following the taxonomy proposed in [1], pansharpening methods can be roughly grouped into four main categories: component substitution (CS) [2], multiresolution analysis (MRA) [3], variational optimization (VO) [4,5], and machine/deep learning (ML) [6,7].

In the CS approach, the multispectral image is transformed in a suitable domain, where one of its components is replaced with the PAN. In the particular case of three spectral bands, the Intensity–Hue–Saturation (IHS) transform is an option where the intensity component can be replaced with the PAN band [8]. This method has been generalized in [9] (GIHS) to handle a larger number of bands. Other useful transforms to implement a CS solution include the principal component analysis [10], the Brovey transform [11], and the Gram–Schmidt (GS) decomposition [12]. More recently, adaptive CS methods have also been proposed, such as the advanced versions of GIHS and GS [13], the partial replacement CS method (PRACS) [14], or the band-dependent spatial detail (BDSD) injection method and its variants [15–17].

In MRA approaches [3], the pansharpening task is addressed from the perspective of a pyramidal decomposition to separate low-frequency content from detail components. The high-frequency spatial details are extracted by means of multi-resolution decomposition, such as decimated or undecimated wavelet transforms [3,18–20], Laplacian pyramids [21–25], or other nonseparable transforms, e.g., contourlet [26]. Extracted details are then properly injected into the upscaled MS component.

A further set of methods address the pansharpening problem through the variational optimization of suitable models of the fused image. In [4], the optimization target involves the degradation filters mapping high-resolution to low-resolution images, while [27] leverage sparse representations for detail injection. Palsson et al. proposed several methods of this class. A total variation regularized least square formulation is provided in [28]. The same research team has framed the pansharpening as a maximum *a posteriori* problem in [29] and, more recently, explored the use of low-rank representations of the joint PAN-MS [5]. Other methods do not fit the above categories and can be roughly classified as statistical [30–34], dictionary-based [35–40], or matrix factorization approaches [41–43]. The reader is referred to [1,44] for a more comprehensive review.

In recent years, a paradigm shift from model-based to data-driven approaches has revolutionized all fields of image processing, from computer vision [45–49] to remote sensing [7,50–52]. In pansharpening, the first method based on convolutional neural networks (CNN) was proposed by Masi et al. in 2016 [6], after which many more followed in the span of a few years [7,53–64]. It seems safe to say that deep learning is currently the most popular approach for pansharpening. Nonetheless, it suffers from a major problem: the lack of ground truth data for supervised training. In fact, multi-resolution sensors can only provide the original PAN-MS data, downgraded in space or spectrum, and never their high-resolution versions, which remain to be estimated.

Based on this brief and certainly not exhaustive overview, it appears that pansharpening is a very active research field, with many new methods proposed every year. Reliable quality assessment procedures are of critical importance for correctly advancing the state of the art, and an incorrect evaluation paradigm may negatively impact the design or tuning of any new solution. Unfortunately, by the very nature of pansharpening, no ground truth (GT) data are available to perform a reference-based assessment. As a consequence, two kinds of quality assessments are usually employed:

- (i) reference-based reduced-resolution assessment (synthesis check);
- (ii) no-reference full-resolution assessment (consistency check).

Lacking GTs, the synthesis capabilities of any pansharpening method can only be assessed on “synthetic” data. In particular, Wald’s protocol [65] suggests taking the real PAN-MS data and applying a proper resolution downgrade process for scale reduction. The scaled PAN-MS pair is the synthesized image whose GT is now given by the original MS component. The way in which the resolution downgrade has to be performed has been the object of intense research in the last two decades and has a non-negligible impact on the reliability of the consequent quality assessment. However, no matter how a GT is obtained, there exist plenty of reference-based image quality indicators. The spectral angle mapper was introduced in [66] and assesses the balance among the spectral bands. On the contrary, the spatial correlation coefficient [67] computes the correlation coefficient across the high-pass-filtered bands of the pansharpened image and of the GT. It is therefore oriented to the assessment of the spatial quality. The *Erreur Relative Globale Adimensionnelle de Synthèse* [68] generalizes the root mean squared error, introducing band-wise correction weights. The universal image quality index [69] compares image statistics to take into account local correlation, intensity, and contrast. Based on this general-purpose image quality index, domain-specific variants suitably adapted to the pansharpening have been proposed in [70,71]. In addition to these indexes, other popular general-purpose options are sometimes considered for pansharpening, e.g., the peak signal-to-noise ratio or the structural similarity (SSIM) index.

Furthermore, spectral or spatial consistency indexes suited for full-resolution images that do not require any GT have also been developed. A first spectral distortion index proposed by Zhou et al. [67] compares a rescaled version of the pansharpened image with the MS image. Other spectral distortion indexes follow a similar procedure [72,73]. In [74], the Quality for Low Resolution (QLR) index based on SSIM was proposed and later updated, exchanging SSIM with the composite image quality measure CMSC based on means, standard deviations, and correlation coefficient [75]. Basically, most of the known

spectral distortion indexes follow this protocol: degradation of the pansharpened image and comparison with the MS image. The different behavior stems from the degradation model and the error measure. For example, [74] is based on the structural similarity indexes, while [75] leverages a composite image quality measure based on statistics such as mean, variance, and correlation coefficient. On the other hand, the spatial consistency can be checked through the cross-scale invariance of some statistics [72]. A similar approach is proposed in [73], where only high-frequency components are concerned. In [74], it is proposed a Quality for High Resolution (QHR) index, which assesses the SSIM index between the panchromatic band and a projection of the pansharpened image in the PAN domain. A variant of this approach with a different error term was proposed in [75]. Another similar approach is proposed in [76], where the coefficient of determination is used to compare the PAN image with its projection from the pansharpened image. It is also worth mentioning other no-reference quality indexes, such as the Natural Image Quality Evaluator (NIQE) [77,78], which seeks to assess the image quality by itself, rather than checking consistency.

Finally, often, the spectral and the spatial consistency indexes are somehow combined, e.g., through geometric means, to provide a single hybrid consistency index representing the overall quality no-reference assessment of the fused images [72–75,78–83].

Both reference-based and no-reference indexes present inherent limitations, which are analyzed in the next section.

In this work, we propose new full-resolution quality indexes that overcome some of these problems and provide more reliable guidance for the development of ever more accurate pansharpening methods. To this purpose, the developed indexes have been made available to the community through a web repository at <https://github.com/matciotola/fr-pansh-eval-tool/> (accessed on 5 April 2022).

The remainder of the paper is organized as follows. Section 2 provides a brief critical survey on pansharpening assessment. Section 3 introduces the proposed approach. Section 4 discusses the experimental results and, finally, Section 5 draws conclusions.

## 2. A Review of Pansharpening Indexes

Due to the lack of ground truths, visual inspection by human experts is the ultimate benchmark for pansharpening methods. However, this is a lengthy and tedious task, and numerical indexes are essential for a viable development process. Reduced-resolution (RR) indexes are computed, according to Wald's protocol [65], by reducing the scale of both MS and PAN components. Then, pansharpening is performed on these rescaled data, and the original MS is used as GT to compute reference-based indexes. Instead, full-resolution (FR) indexes are computed on the original data and try to assess their spectral and/or spatial consistency with the pansharpened output. In a comprehensive evaluation, following Wald's protocol [65], usually, both types of indexes are considered together, in addition to visual inspection [1]. In order to highlight the critical points of the current approaches to the quality assessment, we provide a brief overview of some of the most representative indexes in the following.

### 2.1. Reduced-Resolution Assessment

Let  $M$  and  $P$  be the original MS and PAN components, respectively, and  $\hat{M}$  be the pansharpened image. The synthesis properties cannot be directly verified on the full-resolution pair  $(M, P)$  due to the lack of GT with which to compare  $\hat{M}$ . Leveraging on a scale invariance hypothesis, it is, however, possible to synthesize datasets with GT by properly downgrading the available  $(M, P)$  pairs by a factor  $R$  (PAN-MS resolution ratio). The resulting reduced-resolution pair  $(M_{\downarrow}, P_{\downarrow})$  can therefore be used as the source input for the pansharpening task, whereas the original MS image  $M$  acts as GT. Once moved to the reduced-resolution space, many different quality indexes can be computed to assess the mismatch between the pansharpened image and its reference GT, such as root mean square

error, structural similarity, and so on. In particular, regarding pansharpening, the most common are the following.

- (a) SAM (Spectral Angle Mapper) [66]. It determines the spectral similarity in terms of the pixel-wise average angle between spectral signatures. If  $\mathbf{v}$  and  $\hat{\mathbf{v}}$  are two corresponding pixel spectral responses to be compared, SAM is obtained by averaging over all image locations the following “angle” among vectors:

$$\text{SAM}(\mathbf{v}, \hat{\mathbf{v}}) = \arccos\left(\frac{\langle \mathbf{v}, \hat{\mathbf{v}} \rangle}{\|\mathbf{v}\|_2 \cdot \|\hat{\mathbf{v}}\|_2}\right). \quad (1)$$

- (b) ERGAS (*Erreur Relative Globale Adimensionnelle de Synthèse*) [68]. This is one of the most popular indexes to assess both spectral and structural fidelity between a synthesized image and a target GT. Such an index presents interesting invariance properties. Indeed, it is insensitive to the radiometric range, number of bands, and resolution ratio. If  $B$  is the number of spectral bands, it is defined as

$$\text{ERGAS} = \frac{100}{R} \sqrt{\frac{1}{B} \sum_{b=1}^B \left(\frac{\text{RMSE}_b}{\mu_b^{\text{GT}}}\right)^2}, \quad (2)$$

where  $\text{RMSE}_b$  is the root mean square error over the  $b$ -th spectral band, and  $\mu_b^{\text{GT}}$  is the average intensity of the  $b$ -th band of the GT image.

- (c)  $Q2^n$  [71]. This is a multiband extension of the universal image quality index (UIQI) [69]. Each pixel of an image with  $B$  spectral bands is accommodated into a hypercomplex (HC) number with one real part and  $B-1$  imaginary parts. Let  $\mathbf{z}$  and  $\hat{\mathbf{z}}$  denote the HC representations of a generic GT pixel and its prediction, respectively, and then  $Q2^n$  can be written as the product of three terms:

$$Q2^n = \frac{|\sigma_{\mathbf{z}\hat{\mathbf{z}}}|}{\sigma_{\mathbf{z}}\sigma_{\hat{\mathbf{z}}}} \cdot \frac{2\sigma_{\mathbf{z}}\sigma_{\hat{\mathbf{z}}}}{\sigma_{\mathbf{z}}^2 + \sigma_{\hat{\mathbf{z}}}^2} \cdot \frac{2\mu_{\mathbf{z}}\mu_{\hat{\mathbf{z}}}}{|\mu_{\mathbf{z}}|^2 + |\mu_{\hat{\mathbf{z}}}|^2}. \quad (3)$$

The first factor provides the modulus of the HC correlation coefficient between  $\mathbf{z}$  and  $\hat{\mathbf{z}}$ . The second and the third terms measure contrast changes and mean bias, respectively, on all bands simultaneously. Statistics are typically computed on  $32 \times 32$  pixel blocks, and an average over the blocks of the whole image provides the global  $Q2^n$  score, which takes values in the  $[0, 1]$  interval, being 1 the optimal value achieved if and only if  $\mathbf{z} = \hat{\mathbf{z}}$  in each location.

Behind its appealing simplicity, regardless of the specific error measurement employed, the reduced-resolution assessment hides some major subtle pitfalls that undermine its usefulness. On one hand, its accuracy depends critically on the method used for scaling, which may not correspond to the actual sensing conditions. More fundamentally, it relies on the arbitrary assumption that a method optimized for the RR scale will keep its good behavior at the FR scale. Experimental evidence shows this not to be the case. Furthermore, optimizing parameters for a given scale may lead to a sort of “scale overfitting”, with the perverse effect of degrading performance on a different scale. For these reasons, many recent studies focus on no-reference full-resolution quality indexes [76,78,80,82]. It is also worth recalling that the assessment of spectral and spatial consistencies is not necessarily linked to a fusion task and is of broader interest, with applications in the medical [84], agricultural [85], and food [86] sciences, to mention a few.

## 2.2. Full-Resolution No-Reference Assessment

As a consequence of the above limitations, many studies have moved toward a *no-reference* approach to the pansharpening quality assessment. In particular, the same Wald’s protocol [65] recognizes the problem of invoking a complementary *consistency* quality check in addition to the synthesis capacity assessment. In practice, for consistency-based

assessment, the input data are pansharpened in their original (full) resolution space. The outcome is then compared with both input components, MS and PAN, for spectral and consistency assessment, respectively. The spectral consistency check requires a spatial resolution degradation of the pansharpened image, which is typically done by means of a low-pass filtering (LPF), followed by decimation, using LPFs that approximate the sensor Modulation Transfer Function (MTF) [87]. Other approaches have also been explored, such as wavelet transforms [88,89], averaging operators [90], and pyramid decomposition [21]. The spatial (or *structural*) consistency check is indeed more controversial, lacking a clear relationship that allows one to project the pansharpened MS in the PAN domain [91]. No-reference indexes are also often referred to as FR indexes, as they do not require any resolution degradation of the input data, contrarily to the reference-based ones, which do need it because of the lack of GTs. For this reason, the latter are also referred to as RR indexes.

Among the several FR quality assessment approaches proposed in the last few years [26,67,72,73,76], we will recall a few of them, which are the most frequently used. In particular, Alparone et al. [72] proposed a spectral distortion index  $D_\lambda$  defined as

$$D_\lambda = \sqrt[p]{\frac{1}{B(B-1)} \sum_{i=1}^B \sum_{j=1, j \neq i}^B |d_{i,j}(M, \hat{M})|^p}, \quad (4)$$

where  $d_{i,j}(M, \hat{M}) = Q(M_i, M_j) - Q(\hat{M}_i, \hat{M}_j)$ , with subscripts indicating selected bands and  $Q(\cdot, \cdot)$  being the UIQI [69]. Such an index was also enclosed in the pansharpening benchmark toolbox [44]. As can be observed,  $D_\lambda$  compares the inter-band relationship through the UIQI, separately for the input MS  $M$  and the pansharpened image  $\hat{M}$ , and then quantifies their divergence over each coupling of spectral bands. As can be assumed, lacking a direct measurement of the distance between  $M$  and  $\hat{M}$ , and leveraging an assumption of cross-scale inter-band correlation invariance only, there can be important spectral aberrations that can remain unseen by  $D_\lambda$ . The recently published revisited version [1] of the same toolbox [44], indeed, makes use of an approximated version of Khan's index proposed in [73] to monitor the spectral consistency. In particular, Khan's index is defined as follows

$$D_\lambda^{(K)} = 1 - Q^{2^n}(\hat{M}_\downarrow^{\text{lp}}, M), \quad (5)$$

where  $\hat{M}_\downarrow^{\text{lp}}$  indicates the MTF-based low-pass-filtered and decimated version of the pansharpened image  $\hat{M}$ , while  $M$  is the original input MS. Its approximation, used in the toolbox [1], named  $\tilde{D}_\lambda^{(K)}$ , avoids the decimation of  $\hat{M}$  by using an upscaled version  $\tilde{M}$  of  $M$ :

$$\tilde{D}_\lambda^{(K)} = 1 - Q^{2^n}(\hat{M}^{\text{lp}}, \tilde{M}), \quad (6)$$

being  $\hat{M}^{\text{lp}}$  the low-pass version of  $\hat{M}$ . These indexes range between 0 (optimal value) and 1, and clearly relate to the spectral consistency, since the resolution downgrade process applied to  $\hat{M}$  eliminates the high-pass content retrieved from the PAN. However, two critical points should be taken into account:

- (a) dependence on the accuracy of the estimated MTF;
- (b) sensitivity to the PAN-MS alignment.

In particular, as a consequence of (b), methods such as several CS approaches [12–14,25,92] that intrinsically address the coregistration problem are penalized by both  $D_\lambda^{(K)}$  and  $\tilde{D}_\lambda^{(K)}$  on misaligned datasets. Hence, in order to fairly use Khan's index, it is usually recommended to keep separate the registration and pansharpening problems by means of a prior coregistration of the datasets to be used for testing [1,44].

Several indexes have also been proposed for FR spatial consistency check [72–74,76]. In particular, the spatial distortion index proposed in [72] is used in the most recent pansharpening assessment toolbox [1] and is defined as:

$$D_S = \left\| \frac{1}{B} \sum_{i=1}^B Q(\hat{M}_i, P) - Q(M_i, P_{\downarrow}) \right\|_q, \quad (7)$$

where  $P_{\downarrow}$  is the original PAN component resized to the MS scale and  $Q(\cdot, \cdot)$  is the UIQI quality index [69]. This index is somehow similar to  $D_{\lambda}$  as it checks the cross-scale invariance of the PAN-MS spectral correlation. Therefore, it inherits similar limitations, which are:

- i. no direct comparison between the pansharpened image  $\hat{M}$  and the PAN  $P$ ;
- ii. a cross-scale invariance assumption for which there are no guarantees.

### 3. Proposed Full-Resolution Indexes

We now propose some new indexes for assessing the quality of pansharpening, aiming to overcome some limitations of traditional ones. Given the weakness of the scale invariance assumption of reduced-resolution quality assessment indexes, we focus on full-resolution ones, moving our perspective from synthesis to consistency. In the following, we will discuss spectral and spatial assessment separately.

#### 3.1. Reprojection Protocol for Spectral Accuracy Assessment

The assessment of the spectral consistency is not as controversial as the assessment of the spatial one. In particular, as far as we know, Khan's approach is the most trusted one for this purpose. Nonetheless, as remarked above, it has some limitations and, in particular, in order to work properly, the datasets must be well aligned.

To cope with this issue without the need to operate any coregistration on the available datasets, let us now consider a variation of Khan's index where the decimation of  $\hat{M}^{lp}$  is performed using a band-wise shift that maximizes the correlation coefficient between the low-pass-filtered PAN  $P^{lp}$  and each ideally interpolated MS band  $\tilde{M}_b$ . In the case of aligned datasets, it returns the usual Khan's index because no shifts are applied. Otherwise, before decimation, we apply to  $\hat{M}$  the same shift that would align  $\hat{M}$  and  $P^{lp}$ . Let us indicate as  $D_{\lambda, \text{align}}^{(K)}$  such an index that is formally given by

$$D_{\lambda, \text{align}}^{(K)} = 1 - Q^{2^n}(\hat{M}_{\downarrow a}^{lp}, M), \quad (8)$$

where  $\downarrow a$  indicates decimation with band-wise alignment.

Let  $\mathcal{M}$  indicate a generic reference-based distortion index, such as  $Q^{2^n}$ , SAM, or ERGAS. We define the reprojection index R- $\mathcal{M}$  as

$$\text{R-}\mathcal{M} = \mathcal{M}(\hat{M}_{\downarrow a}^{lp}, M), \quad (9)$$

which, in the particular case of  $\mathcal{M} = Q^{2^n}$ , reduces to the complement of the  $D_{\lambda, \text{align}}^{(K)}$  variant defined above (Equation(8)), i.e.,

$$\text{R-}Q^{2^n} = Q^{2^n}(\hat{M}_{\downarrow a}^{lp}, M) = 1 - D_{\lambda, \text{align}}^{(K)}. \quad (10)$$

In other words, the pansharpened image  $\hat{M}$  is projected in the original MS space prior to being compared with the MS image  $M$  (reference) using the error index  $\mathcal{M}$ . The projection accounts for both eventual misalignment and sensor MTF using matched low-pass filters. Of course, because of the low-pass filtering, the high-frequency image content (spatial details) is discarded and some further indexes will be necessary to assess spatial quality. Note also that, under the hypothesis that the sensor MTF is correctly simulated, the

reprojection indexes ensure that an ideal pansharpening (GT) would achieve the best score as expected. This will be experimentally analyzed in Section 4.3.

A pictorial representation of the proposed reprojection indexes, which relates them to the synthesis assessment protocol proposed by Wald [65], is given in Figure 1. The common starting point (top-left) is the available  $(P, M)$  pair. Indexes based on Wald's protocol follow the counterclockwise path: decimation–pansharpening–measure. The proposed reprojection protocol, instead, consists in following the clockwise path: pansharpening–decimation–measure, where the actual pansharpened image is first computed, and then reprojected on the low-resolution domain for quality assessment. Eventually, in both cases, the original MS is used as a reference for any error measurement. In our proposal, however, pansharpening precedes the resolution downgrading, thereby exploiting the most valuable source of spatial information, the original PAN. Furthermore, the overall evaluation no longer depends on how the PAN is downscaled. Since the reprojection protocol involves only the low-resolution component of the synthesized image, it has to be considered a spectral consistency rather than a synthesis check. A pseudo-code of the reprojection indexes is given in Algorithm 1.

---

### Algorithm 1 Reprojection error assessment

---

**Require:**  $\tilde{M}$ ,  $M$ ,  $P$ ,  $\mathbf{g}_{\text{MTF}}$  (MTF parameters),  $\mathcal{M}$  (error function)

**Ensure:**  $R\text{-}\mathcal{M}$

1:  $\tilde{M} = \text{upscale}(M, R)$

2:  $P^{\text{lp}} = \text{LPF}(P, \mathbf{g}_{\text{MTF}})$

3: **for**  $b = 1 : B$  **do**

4:  $\mathbf{s}(b) = \arg \max_{(m,n) \in [1,R]^2} \text{corrcoef}(P_{i,j}^{\text{lp}}, \tilde{M}(b)_{i+m,j+n})$

5: **end for**

6:  $\hat{M}^{\text{lp}} = \text{LPF}(\tilde{M}, \mathbf{g}_{\text{MTF}})$

7:  $\hat{M}_{\downarrow a}^{\text{lp}} = \text{decimation}(\hat{M}^{\text{lp}}, R, \mathbf{s})$

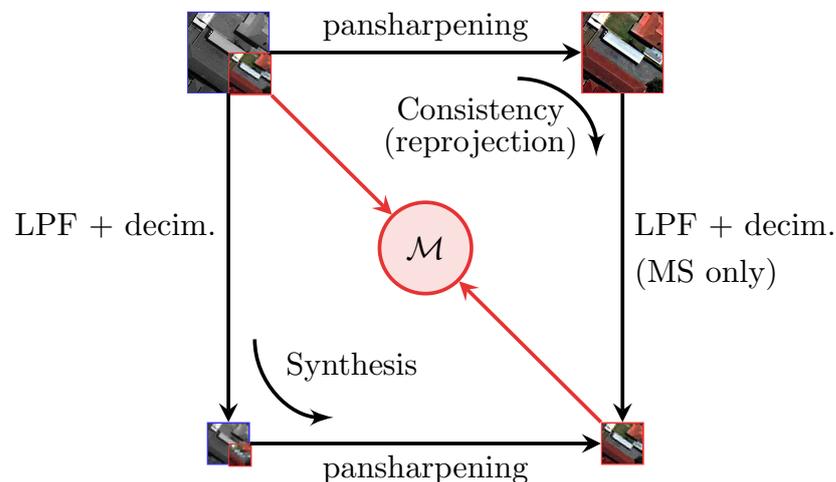
8:  $R\text{-}\mathcal{M} = \mathcal{M}(\hat{M}_{\downarrow a}^{\text{lp}}, M)$

▷ polynomial interpolation  
▷ MTF-tailored low-pass filtering

▷ band-wise shift for alignment

▷ Equation (9)

---



**Figure 1.** Spectral consistency check using reprojection (clockwise) vs. synthesis check according to Wald's protocol (counterclockwise).  $\mathcal{M}$  is any reference-based index. In gray and red are shown the PAN and MS channels, respectively.

### 3.2. Correlation-Based Spatial Consistency Index

In order to complement the reprojection-based assessment, we also propose a new index to evaluate the spatial consistency with the aim of obtaining indications better correlated to human judgment than those provided by  $D_S$ . In particular, to assess the preservation of fine-scale spatial structures, we compute the average local correlation between the pansharpened image and the PAN.

Let  $X_{ij}^\sigma$  indicate a  $\sigma \times \sigma$  patch of  $X$  centered on location  $(i, j)$ . We compute the correlation field  $\rho_{P\hat{M}}^\sigma$  given by the local correlation coefficients between  $P$  and each band  $b$  of  $\hat{M}$ :

$$\rho_{P\hat{M}}^\sigma(i, j, b) = \text{corrcoef}\left(P_{ij}^\sigma, \hat{M}(b)_{ij}^\sigma\right) \quad (11)$$

Then, we reduce it to its average value  $\overline{\rho_{P\hat{M}}^\sigma}$  over space and spectral bands. The final index is then defined as

$$D_\rho \triangleq 1 - \overline{\rho_{P\hat{M}}^\sigma} \quad (12)$$

such that  $D_\rho = 0$  corresponds to perfect correlation. From a different point of view, we are studying to what extent a  $\sigma \times \sigma$  patch of any band of  $\hat{M}$  can be linearly predicted from the corresponding PAN patch. Therefore,  $D_\rho$  is strictly related to the matching between the spatial layouts of  $\hat{M}$  and  $P$  at a fine scale.

The choice of the scale parameter  $\sigma$  is of critical importance, as we are interested in the exploitation of the complementary information discarded by the reprojection indexes, i.e., details appearing only at the PAN scale but not at the MS scale. Therefore, if  $R$  is the resolution ratio, by choosing  $\sigma = R$  (equal to 4 for our datasets), we can place the focus only on that spatial content that is not visible at the MS scale. By doing so, we reduce the dependence between the proposed spatial ( $D_\rho$ ) and spectral ( $R\text{-}\mathcal{M}$ ) distortion indexes, assigning to them clear but well-separated evaluation objectives.

As for  $D_S$  and  $D_\lambda^{(K)}$ , or its variants,  $D_\rho$  computed for  $\hat{M} = \text{GT}$  does not reach the zero value, given the slightly different spatial layout of different bands in natural images and the complex relationship between the high-resolution image and its panchromatic projection [91]. Nonetheless, we expect good-quality pansharpened images to present relatively small values of  $D_\rho$ . A detailed discussion in this regard, supported by our experimental results, will be provided in the following.

#### 4. Experimental Results and Discussion

In this section, we present several experimental analyses. Preliminarily, a summary of the datasets employed and the involved methods is provided. The first experiment focuses on the dependence of the spectral consistency indexes on the alignments of the MS bands with the PAN. Next, we move to the reduced-resolution domain to cross-check reference-based and no-reference indexes thanks to the availability of the ground truth. Then, it follows an experimental analysis focused on the assessment of the spatial consistency, before closing with an overall comparison.

##### 4.1. Datasets and Methods

The experimental validation relies on 25 methods provided in the benchmark toolbox [1] belonging to the four main categories recalled in Section 1, CS (8), MRA (9), VO (3), and ML (4), plus an ideal interpolator (EXP). The dataset is composed of two WorldView-2 (WV2) and two WorldView-3 (WV3) large images, courtesy sample products of DigitalGlobe©. In total, 20  $512 \times 512$  tiles were extracted from the WV2 images (Washington and Stockholm) and 20 from the WV3 images (Adelaide and Mexico City) for the experiments at full resolution. Likewise, 20 + 20  $2048 \times 2048$  tiles were extracted and downsampled to size  $512 \times 512$  for the experiments at reduced resolution. Table 1 summarizes the main spectral and spatial characteristics of the WV-2 and WV-3 sensors.

**Table 1.** Bandwidths of the multispectral channels (left-hand side) and Ground Sample Distance (GSD) [m] at Nadir (rightmost column) for WorldView-2 and WorldView-3 images.

Sensor	Bandwidths of the MS Channels [nm]								GSD [m]
	Coastal	Red	Blue	Red Edge	Green	Near-IR1	Yellow	Near-IR2	
WV2	396-458	624-694	442-515	699-749	506-586	765-901	584-632	856-1043	0.46/1.84
WV3	400-450	630-690	450-510	705-745	510-580	770-895	585-625	860-1040	0.31/1.24

#### 4.2. Spectral Distortion Dependence on PAN-MS Misalignment

The impact of band misregistration on the quality of the fused products has been already recognized in the past [93–95]. Here, we propose an ad hoc experimental analysis to show the robustness of the proposed reprojection indexes to the data misregistration. The starting point is a WV3 dataset composed of ten  $2048 \times 2048$  tiles extracted from a larger image of Adelaide (DigitalGlobe© sample product). This dataset presents band misalignment that we have corrected to produce a companion aligned dataset. In Table 2, we summarize the average spectral distortion scores for each method for both datasets. For a convenient reading of these numbers, in Figure 2, we show the impact of data misregistration on  $D_{\lambda}^{(K)}$  and  $\tilde{D}_{\lambda}^{(K)}$  in differential terms.

**Table 2.** Misregistration impact on  $D_{\lambda}$ ,  $D_{\lambda}^{(K)}$ ,  $\tilde{D}_{\lambda}^{(K)}$  and  $D_{\lambda,align}^{(K)}$ . Detailed information about the methods can be found in [1].

	Aligned Dataset	$D_{\lambda}$		$D_{\lambda}^{(K)}$		$\tilde{D}_{\lambda}^{(K)}$		$D_{\lambda,align}^{(K)}$	
		Yes	No	Yes	No	Yes	No	Yes	No
	EXP (interpolator)	0.0000	0.0000	0.0249	0.0336	0.0394	0.0479	0.1238	0.0670
CS	BT-H [92]	0.0698	0.0823	0.1231	0.0696	0.2310	0.1324	0.0874	0.0697
	BDS [15]	0.0429	0.0377	0.0989	0.0867	0.1834	0.1561	0.1511	0.1064
	C-BDS [16]	0.0514	0.0557	0.1436	0.1195	0.2158	0.1836	0.2346	0.1664
	BDS-PC [17]	0.0141	0.0160	0.0782	0.0896	0.1464	0.1530	0.0705	0.0892
	GS [12]	0.0196	0.0177	0.1494	0.0824	0.2611	0.1414	0.0824	0.0839
	GSA [13]	0.0576	0.0573	0.1182	0.0604	0.2457	0.1334	0.0760	0.0624
	C-GSA [25]	0.0309	0.0333	0.0929	0.0583	0.1925	0.1240	0.0741	0.0625
	PRACS [14]	0.0178	0.0193	0.0728	0.0468	0.1464	0.0889	0.0834	0.0610
MRA	AWLP [96]	0.0332	0.0416	0.0282	0.0273	0.0415	0.0320	0.0920	0.0495
	MTF-GLP [96]	0.0679	0.0759	0.0231	0.0195	0.0434	0.0352	0.0912	0.0428
	MTF-GLP-FS [97]	0.0544	0.0669	0.0222	0.0199	0.0400	0.0347	0.0944	0.0439
	MTF-GLP-HPM [96]	0.0620	0.0722	0.0310	0.0210	0.0508	0.0376	0.0909	0.0411
	MTF-GLP-HPM-H [92]	0.1019	0.1177	0.0270	0.0200	0.0524	0.0401	0.0885	0.0402
	MTF-GLP-HPM-R [98]	0.0497	0.0620	0.0272	0.0210	0.0453	0.0362	0.0928	0.0420
	MTF-GLP-CBD [87]	0.0551	0.0657	0.0222	0.0200	0.0402	0.0346	0.0942	0.0441
	C-MTF-GLP-CBD [25]	0.0089	0.0375	0.0224	0.0225	0.0360	0.0347	0.1129	0.0495
MF [99]	0.0585	0.0711	0.0421	0.0281	0.0681	0.0441	0.0733	0.0371	
VO	FE-HPM [4]	0.0579	0.0666	0.0340	0.0210	0.0529	0.0355	0.0784	0.0406
	SR-D [27]	0.0225	0.0381	0.0043	0.0064	0.0117	0.0190	0.1254	0.0450
	TV [28]	0.0166	0.0238	0.0225	0.0173	0.0459	0.0373	0.0583	0.0252
ML	PNN [6]	0.0589	0.0553	0.0740	0.0629	0.1477	0.1307	0.1815	0.0878
	PNN-IDX [6]	0.0837	0.0848	0.2671	0.0546	0.1570	0.1005	0.3344	0.0889
	A-PNN [100]	0.0527	0.0636	0.0371	0.0312	0.0773	0.0607	0.0963	0.0522
	A-PNN-FT [100]	0.0196	0.0203	0.0414	0.0308	0.0815	0.0579	0.0830	0.0489

Each bar indicates the difference between the values of the given indicator computed on aligned and misaligned datasets, respectively. As can be clearly noticed, traditional CS methods such as BT-H, GS, GSA, C-GSA, PRACS, which, by construction, provide fused

images that are strongly anchored to the PAN geometry, show a considerable spectral loss according to  $D_{\lambda}^{(K)}$  or  $\tilde{D}_{\lambda}^{(K)}$ . These results are not aligned with our expectations, as these CS methods are expected to be more robust to misregistration. The reader can refer to [93] for a theoretical comparison between CS and MRA methods in the presence of misregistration, with the former category being superior to the latter.

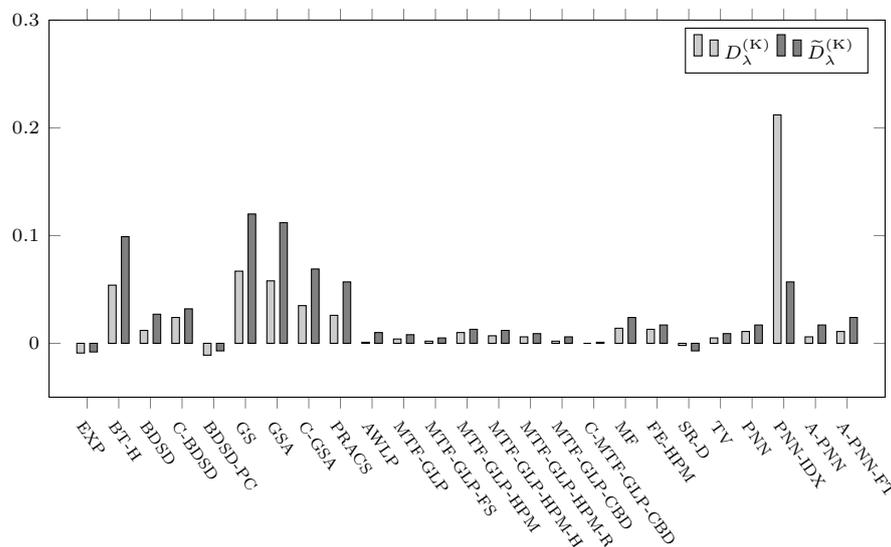


Figure 2. Misregistration impact on  $D_{\lambda}^{(K)}$  and  $\tilde{D}_{\lambda}^{(K)}$ . Bar levels indicate the variation in the indexes due to data misregistration.

In Figure 3, instead, we compare Khan’s index  $D_{\lambda}^{(K)}$  with its variant  $D_{\lambda,align}^{(K)}$  with alignment. Again, bars indicate their variations due to dataset misalignment.

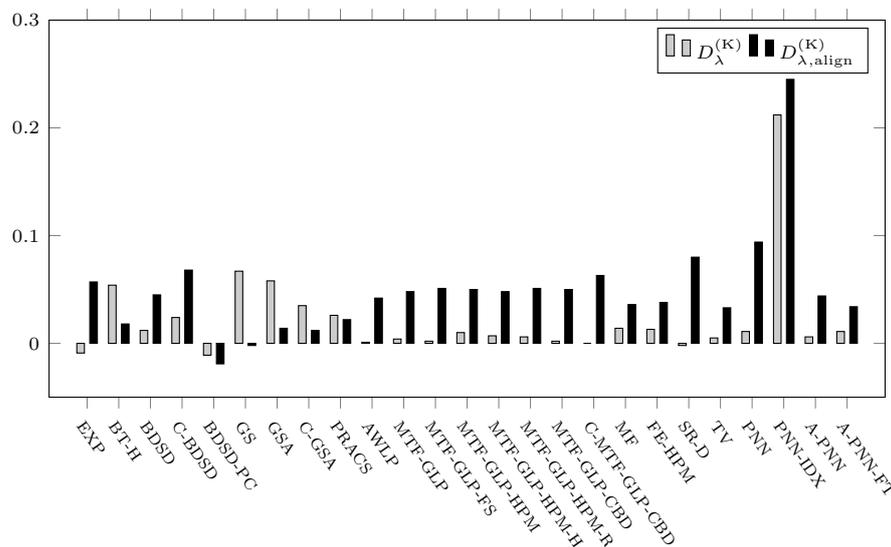


Figure 3. Misregistration impact on  $D_{\lambda}^{(K)}$  and  $D_{\lambda,align}^{(K)}$ . Bar levels indicate the variation in the indexes due to data misregistration.

By inspecting the figure, it can be noticed that misregistered data generally cause an increase in  $D_{\lambda,align}^{(K)}$ , except for some CS methods, such as BT-H, GS, GSA, C-GSA, and PRACS, which tend to preserve the PAN geometry operating an intrinsic alignment. On the contrary, other methods, notably those belonging to the MRA, VO, and ML categories,

which are oriented to the spectral preservation but do not operate alignment, register an increase in the spectral distortion when the metric takes into account the misalignment.

Concluding, the proposed experiment suggests that Khan's index with alignment correction, which is the complement of the proposed reprojection index  $R-Q2^n$ , is more robust to misalignment or, at least, more consistent with the theoretical expectation [93] than the original Khan's index  $D_\lambda^{(K)}$  or its variant  $\tilde{D}_\lambda^{(K)}$ .

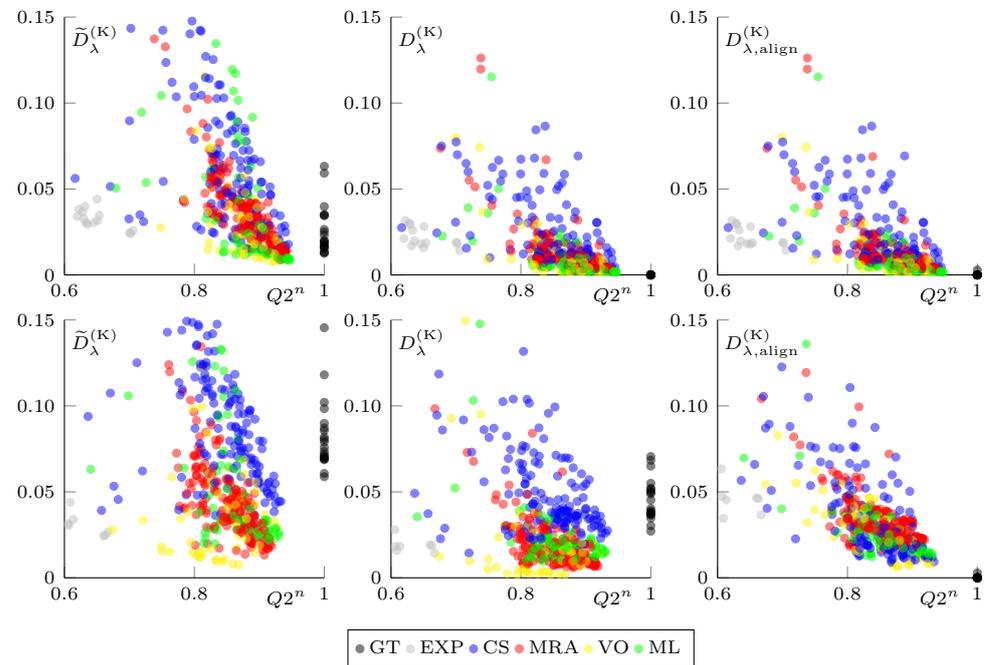
#### 4.3. Reference vs. No-Reference Index Cross-Checking in the Reduced-Resolution Space

In order to assess the consistency between no-reference and reference-based indicators, we have designed a set of experiments in the reduced-resolution space. Although no-reference indexes are conceived to work in the full-resolution framework, their use on (simulated) reduced-resolution datasets allows us to carry out studies on the correlation between them and objective error measurements (reference-based indexes) thanks to the availability of GTs. In particular, for this experiment, we have resorted to our WV2 dataset composed of twenty  $2048 \times 2048$  images at full resolution (for the sake of brevity, analogous results obtained on WV3 images will not be presented). These images come from two larger images of Washington (13) and Stockholm (7), respectively, courtesy samples of DigitalGlobe©. Each such tile was resized to  $512 \times 512$  pixels using the usual Wald's downgrading protocol. This dataset was already well coregistered; therefore, we have proceeded to create a misregistered counterpart operating a simple modification of the downgrading process: a 1-pixel shift (in both directions) has been introduced in the decimation (after LPF) of the MS bands.

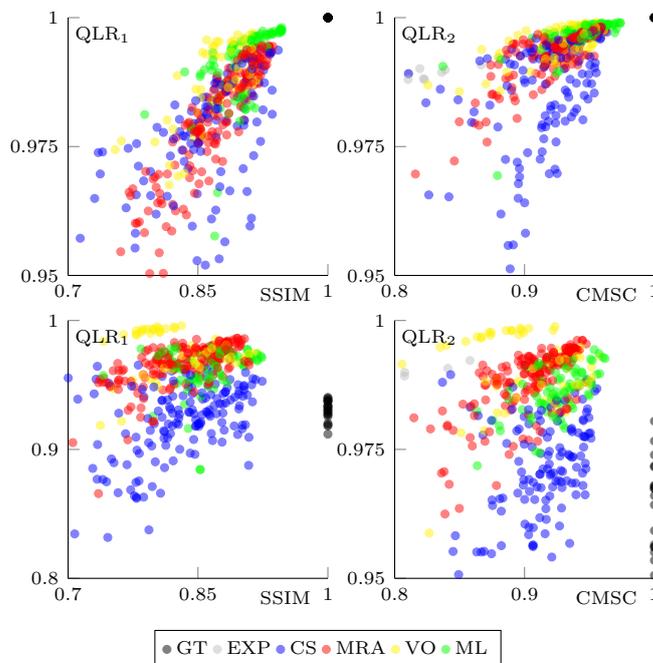
Then, 25 pansharpening algorithms provided by the toolbox [1] were run on each RR sample image, generating 500 results, for each of the two datasets (registered and not), for which all the indexes of interest have been computed. Eventually, we have obtained hundreds of points in a multi-dimensional evaluation space, which enable plenty of analyses, with some dimensions corresponding to the reference-based indexes (SAM, ERGAS and  $Q2^n$ , SSIM, CMSC), and the remaining ones associated with the proposed indexes ( $R-SAM$ ,  $R-ERGAS$ ,  $R-Q2^n = 1 - D_{\lambda,align}^{(K)}$ ,  $D_\rho$ ) and other state-of-the-art no-reference indexes ( $D_\lambda$ ,  $D_\lambda^{(K)}$ ,  $\tilde{D}_\lambda^{(K)}$ ,  $D_S$ ,  $QLR_1$ ,  $QLR_2$ ). In particular, by construction, we expect to observe a good correlation between Khan's index variants and  $Q2^n$ . Therefore, we start from the scatter plots shown in Figure 4 with such variants, in turn, vs.  $Q2^n$ .

Both the "aligned" (top) and the "misaligned" (bottom) datasets are considered. For the aligned case, by visual inspection, we can appreciate that  $D_\lambda^{(K)}$  correlates better than  $\tilde{D}_\lambda^{(K)}$  with  $Q2^n$ . Moreover,  $D_{\lambda,align}^{(K)}$  behaves similarly to  $D_\lambda^{(K)}$  because no alignment is operated. For both  $D_\lambda^{(K)}$  and  $D_{\lambda,align}^{(K)}$ , the GTs (black dots), for which  $Q2^n$  is always 1, obtain the ideal value (0) of spectral distortion. Rarely, the alignment process (based on the maximization of the correlation) embedded in the computation of  $D_{\lambda,align}^{(K)}$  fails for some spectral band, giving rise to non-zero results for the GT even if the test image is aligned. Moreover,  $\tilde{D}_\lambda^{(K)}$  is not minimized by the GT, coherently with its definition (Equation (6)) based on the comparison between the smoothed GT and the upscaled MS. It is also worth noticing that the correlation degree between  $Q2^n$  and the different variants of spectral distortion grows when assessed by category (by colors), supporting the idea that no-reference indexes are more reliable for intra-class assessment [1]. Moving to the misaligned dataset (bottom scatters), it can be clearly recognized a bias for both  $\tilde{D}_\lambda^{(K)}$  and  $D_\lambda^{(K)}$ , which does not occur for  $D_{\lambda,align}^{(K)}$ , represented by the shift of the GT scores, which no longer achieve the ideal value. In general, the distortion scores in terms of  $\tilde{D}_\lambda^{(K)}$  and  $D_\lambda^{(K)}$  register a degradation, with CS methods (blue) much more penalized than others. This last observation gives further support to the interpretation provided above about the results of Figures 2 and 3. A similar behavior is also registered for other state-of-the-art indexes.

For example, in Figure 5, the score scatters in the QLR<sub>1</sub>-SSIM (left) and QLR<sub>2</sub>-CMSC (right) planes, with registered (top) and misaligned (bottom) datasets.

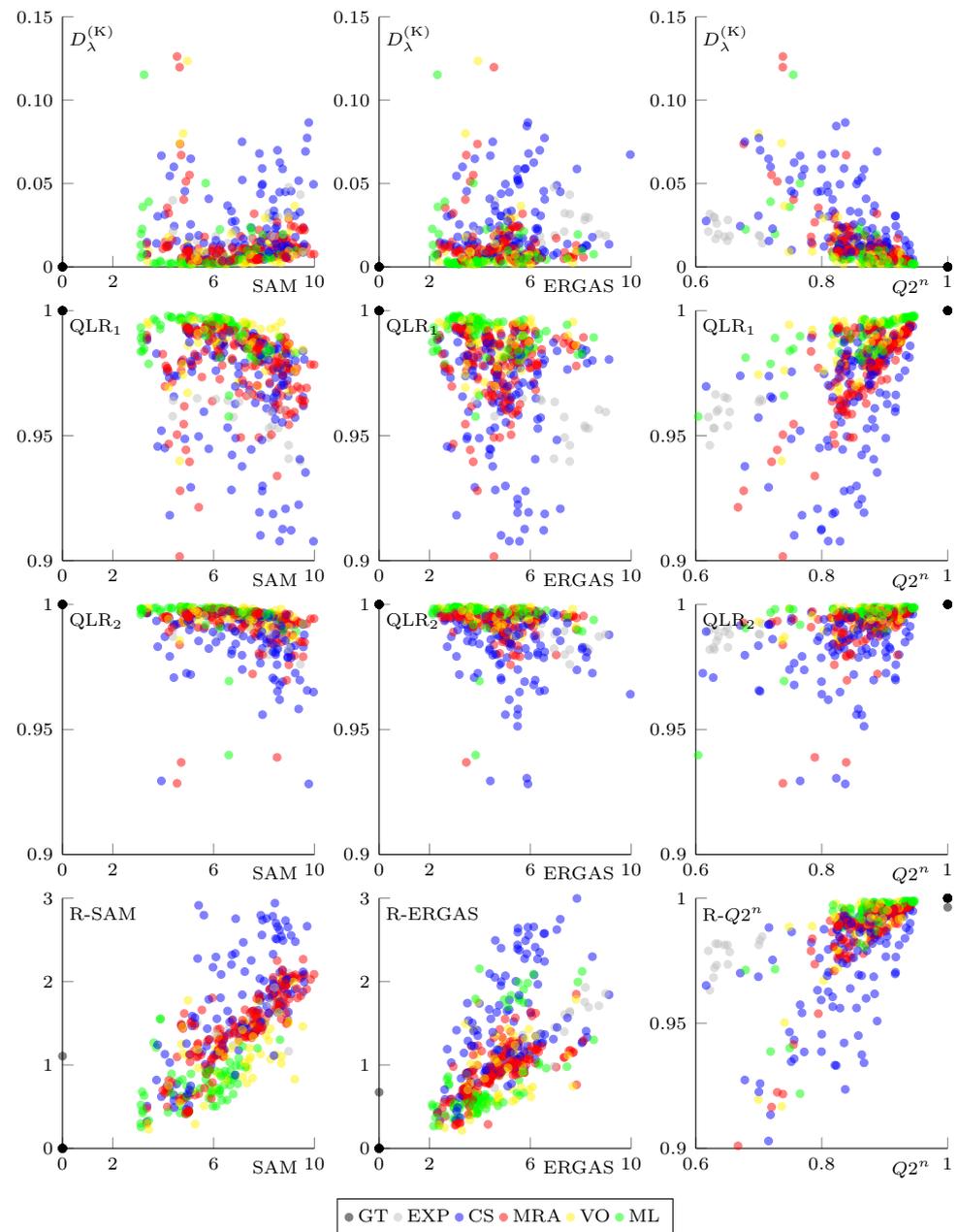


**Figure 4.** From left to right,  $\tilde{D}_\lambda^{(K)}$ ,  $D_\lambda^{(K)}$ , and  $D_{\lambda,align}^{(K)} = 1 - R - Q2^n$  vs.  $Q2^n$  on the WV2 RR dataset. Both the “aligned” (top) and “misaligned” (bottom) dataset cases are shown. Each marker is associated with a single pansharpening result on a  $512 \times 512$  tile. Black markers correspond to the ground truth. Light gray ones (EXP) correspond to a simple interpolation. The remaining clusters correspond to the different algorithms grouped by categories.



**Figure 5.** From left to right, QLR<sub>1</sub> vs. SSIM and QLR<sub>2</sub> vs. CMSC on the WorldView-2 RR dataset. Both the “aligned” (top) and “misaligned” (bottom) dataset cases are shown. Each marker is associated with a single pansharpening result on a  $512 \times 512$  tile. Black markers correspond to the ground truth. Light gray ones (EXP) correspond to a simple interpolation. The remaining clusters correspond to the different algorithms grouped by categories.

Let us now leave the registration problem out, considering aligned datasets only and looking at the relationship between the spectral consistency indexes and the three most used reference-based indexes, i.e., SAM, ERGAS, and  $Q2^n$ , with the help of Figure 6. From top to bottom, we can see how different compared spectral consistency indexes agree with the three (column-wise distributed) reference-based indexes.



**Figure 6.** No-reference vs. reference-based indexes' scores on the aligned WV2 RR dataset. From left to right, SAM, ERGAS, and  $Q2^n$  are selected as reference-based indexes (x-axis). From top to bottom,  $D_\lambda^{(K)}$ ,  $QLR_1$ ,  $QLR_2$ , and  $R-M$  are chosen as no-reference indexes (y-axis).

From the top row scatters, it clearly appears that  $D_\lambda^{(K)}$  agrees with SAM and ERGAS much less than with  $Q2^n$ . This is because  $D_\lambda^{(K)}$  is based on the same  $Q2^n$  index but also because SAM and ERGAS encode a different concept of quality. Similar considerations apply for  $QLR_1$  and  $QLR_2$ , which, as well as  $Q2^n$ , are both based on the comparison among local statistics. For these reasons, we believe that, in addition to  $R-Q2^n$ , it makes sense to also provide  $R-SAM$  and  $R-ERGAS$  for a more comprehensive evaluation of pansharpening

methods. On the bottom row, we show all three ( $\mathcal{M}$ , R- $\mathcal{M}$ ) scatter plots, which speak in favor of a good level of agreement between each objective index  $\mathcal{M}$  and the reprojected counterpart R- $\mathcal{M}$ .

Besides the qualitative interpretation of the score scatters, we can quantify the level of agreement among the reference-based indexes and the compared no-reference indexes in terms of correlation coefficients. These are shown in the usual matrix form, for both the aligned (Table 3) and misaligned (Table 4) datasets, respectively. As expected, the reprojected indexes show a relatively high correlation with the non-reprojected counterpart, both for aligned and misaligned datasets.  $D_\lambda^{(K)}$  and  $\tilde{D}_\lambda^{(K)}$  also correlate well with  $Q2^n$ , but only in the aligned case. Likewise,  $QLR_1$  and  $QLR_2$  also correlate well with different reference-based indexes in ideal conditions but register a drop on misaligned data. It is also worth remarking that GT scores have been discarded so as not to penalize excessively the competitors on the misaligned dataset (see GT score distribution in Figures 4 and 5).

**Table 3.** Correlation coefficient among different indexes (some have been complemented for a uniform polarization) obtained on the aligned RR WV2 datasets.

	R-SAM	R-ERGAS	R- $Q2^n$	$D_\lambda$	$D_\lambda^{(K)}$	$\tilde{D}_\lambda^{(K)}$	$D_{\lambda,align}^{(K)}$	$QLR_1$	$QLR_2$
SAM	0.595	0.353	0.096	0.372	0.101	−0.012	0.096	0.409	0.348
ERGAS	0.141	0.743	0.169	0.050	0.169	0.078	0.169	0.173	0.493
$Q2^n$	0.121	0.306	0.384	−0.326	0.382	0.417	0.384	0.486	0.356
SSIM	0.204	0.338	0.199	−0.106	0.199	0.179	0.199	0.603	0.359
CMSC	0.144	0.390	0.318	−0.095	0.320	0.211	0.318	0.466	0.556

**Table 4.** Correlation coefficient among different indexes (some have been complemented for a uniform polarization) obtained on the misaligned RR WV2 datasets.

	R-SAM	R-ERGAS	R- $Q2^n$	$D_\lambda$	$D_\lambda^{(K)}$	$\tilde{D}_\lambda^{(K)}$	$D_{\lambda,align}^{(K)}$	$QLR_1$	$QLR_2$
SAM	0.745	0.200	0.119	0.336	0.054	−0.066	0.119	0.238	0.253
ERGAS	0.022	0.932	0.221	0.041	0.171	0.062	0.221	0.028	0.361
$Q2^n$	0.023	0.152	0.439	−0.329	0.287	0.293	0.439	0.226	0.110
SSIM	0.164	0.168	0.255	−0.142	0.066	0.035	0.255	0.218	0.046
CMSC	0.151	0.298	0.429	−0.106	0.238	0.079	0.429	0.159	0.294

#### 4.4. A Qualitative Assessment of the Proposed Spatial Distortion Index

While objective synthesis quality indexes such as  $Q2^n$  and ERGAS account for both spectral and spatial inaccuracies, their reprojected versions clearly limit their scope to the spectral component as they are computed on low-pass versions of the fused products, ignoring image “details”. It is therefore necessary to complement these indexes with some spatial quality indicator.

The assessment of the spatial quality of a pansharpened image in the full-resolution framework is very difficult and somewhat controversial, lacking a degradation model to describe the relationship between the full-resolution spatial–spectral data cube (fused image) and the panchromatic image. In fact, while the spectral degradation can be reasonably modeled through the sensor MTF, allowing for a spectral consistency check, the spatial degradation cannot be modeled through a simple global weighted band average [91]. Indeed, in addition to the obvious spatial dependency of the spectral mixing process that provides the PAN image, there is also a mismatch between the PAN spectral coverage and the MS coverage [91]. This means that there could be details “seen” by the PAN but not by the virtual full-resolution MS counterpart, and vice versa. This is the origin of the ill-posedness of the pansharpening problem, mostly residing in the spatial reconstruction, which makes the spatial consistency assessment subjective to some extent. On the basis of this observation, we provide here an interpretation of some sample results through visual inspection, comparing the proposed index  $D_\rho$  with the state-of-the-art spatial distortion index  $D_S$ . Of course, given the subjective nature of this kind of analysis, we

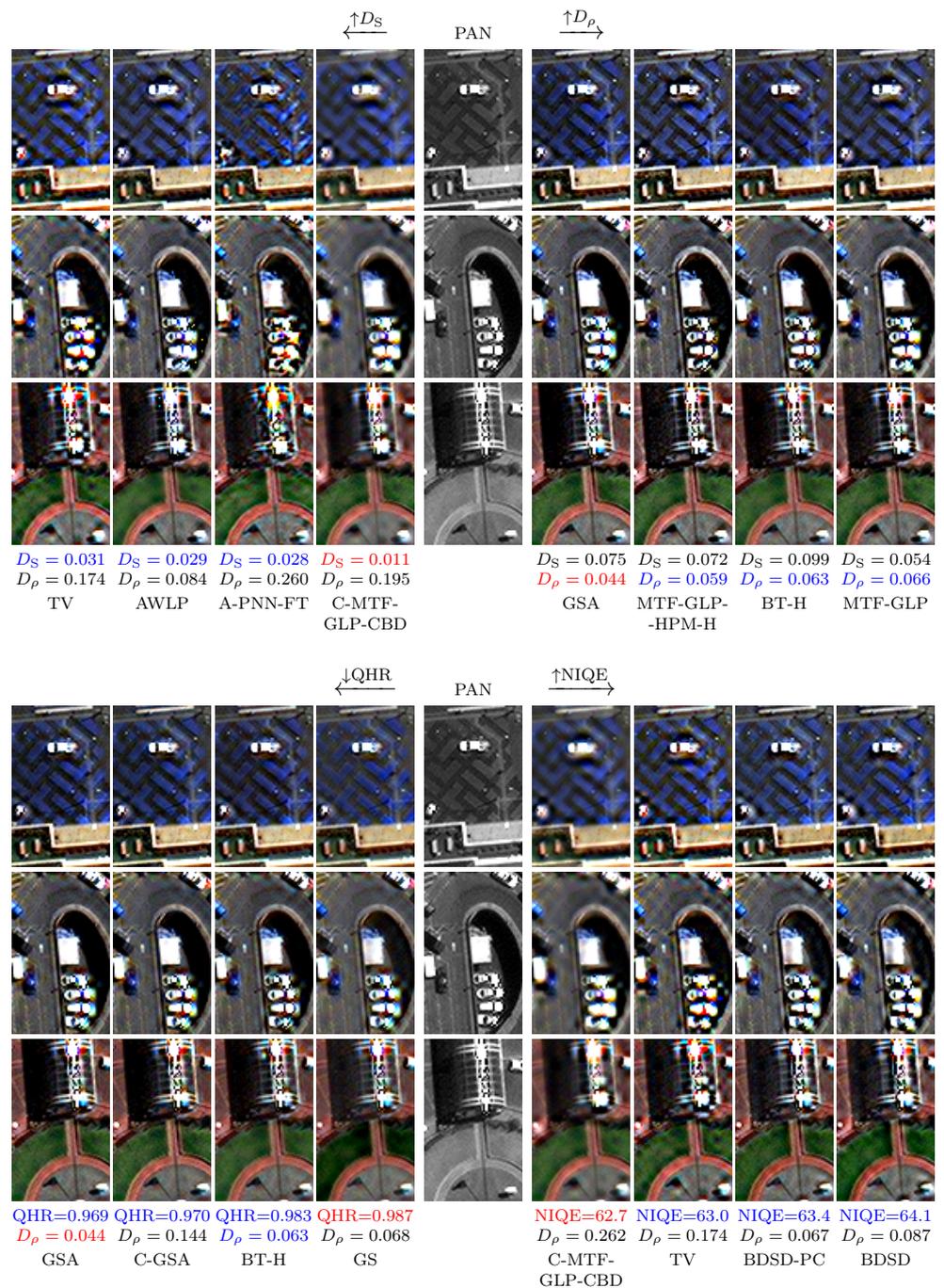
will focus on clearly visible phenomena, leaving to the reader the final say based on his/her visual perception of more subtle patterns, if any. In particular, we propose two experiments: a “horizontal” comparison among the pansharpening toolbox methods and a “vertical” comparison where a single machine learning method is optimized using the proposed  $D_\rho$  (varying the related scale parameter  $\sigma$ ) jointly with a term controlling the spectral consistency.

Let us start with the first experiment. Figure 7 shows some FR pansharpening results for crops extracted from a single tile of the WV3 Adelaide image (For visualization purposes, hereinafter, all displayed crops extracted from multispectral images are obtained combining the red, green, and blue channels (see Table 1)). The PAN component, used as a reference, is shown in the middle of different groups of results. In the top half of the figure, the top four solutions according to  $D_S$  (left) and the top four according to  $D_\rho$  (right) are shown with  $D_S$  and  $D_\rho$  computed on the whole tile. Similarly, in the bottom half, the top four results according to QHR and NIQE are shown with related scores. It appears clearly that images selected according to the  $D_\rho$  index ensure better agreement with the reference in terms of spatial layout and, in general, better quality, with sharp contours, accurate textures, and a lack of annoying patterns such as those present in some top  $D_S$  and top NIQE images. It is interesting to observe that, among the top four  $D_S$  results, the most convincing one from visual inspection corresponds also to the relatively lowest  $D_\rho$  (AWLP). It is also worth observing a certain degree of agreement between  $D_\rho$  and QHR. Similar phenomena are observed on all other tiles.

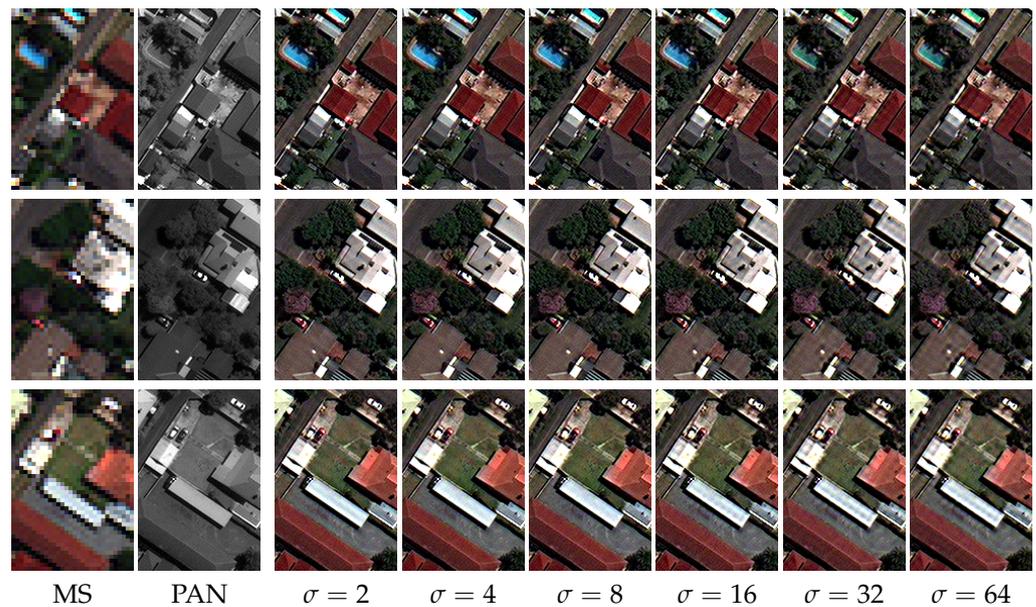
In the previous experiment, we set the scale parameter  $\sigma = R = 4$  according to the above-discussed theoretical motivations (Section 3.2). To gain insight into the effectiveness of this choice, we have designed an additional ad hoc experiment to provide pansharpening results with controlled spectral and spatial qualities, with the latter quantified through  $D_\rho$ , configurable using different scale settings. This is achieved by leveraging a CNN model working in target-adaptive modality [101] and using a combination of the wanted consistency measures as a loss. In this context, the choice of the specific CNN model is not critical as we work in adaptive modality, by running unsupervised tuning iterations on each test image until the loss terms reach the desired level (overfitting). In particular, the optimized loss is defined as

$$\mathcal{L} = \max(\|\hat{M}_{\downarrow a}^{\text{lp}} - M\|_1, \mathcal{L}_\lambda^*) + \beta \max(D_\rho^{(\sigma)}, D_\rho^*), \quad (13)$$

where  $\mathcal{L}_\lambda^*$  and  $D_\rho^*$  are two prefixed target levels for the spectral ( $\|\hat{M}_{\downarrow a}^{\text{lp}} - M\|_1$ ) and spatial ( $D_\rho^{(\sigma)}$ ) quality indicators, respectively. In practice, the network parameters are pushed to overfit the test image so that the corresponding loss reaches the target quality  $(\mathcal{L}_\lambda^*, D_\rho^*)$ . These two threshold values could be ideally set to zero, but this would lead to extremely long tuning and may also generate instability because of a conflicting interaction between the two loss terms. In particular, we have set  $\mathcal{L}_\lambda^* = 0.015$ , quite a low value considering the dynamic range of the pixel values, and  $D_\rho^* = 0.1$ , which is also quite low according to our experiments. By doing so, each test requires a different number of tuning iterations. We have therefore oversized the number (experimentally set to 5000) of iterations to ensure convergence for all. This process is repeated for several choices of the scale parameter  $\sigma$  ranging from 2 to 64. In Figure 8, we show some crops from the WV3 dataset. These sample results reflect a general behavior that we have observed in a wide range of experiments, which is a relatively good response in the range of  $\sigma \in [4, 8]$ . For smaller ( $\sigma = 2$ ) or larger ( $\sigma > 8$ ) scale values, noisy patterns arise. These are particularly noticeable on roofs and roads for the selected samples. The above observations provide experimental validation of the choice  $\sigma = R = 4$  proposed in this work on the basis of the theoretical motivations discussed in Section 3.2.



**Figure 7.** Pansharpening results on crops from a full-resolution WV3 Adelaide tile. The reference PAN is in the middle column. In the top half, the four best results in terms of  $D_S$  and  $D_\rho$  are on its left and right, respectively, with best scores (in red) closer to the reference. In the bottom half are shown the top results according to QHR (leftward) and NIQE (rightward). Best scores are highlighted in red; second, third, and fourth best results are in blue.



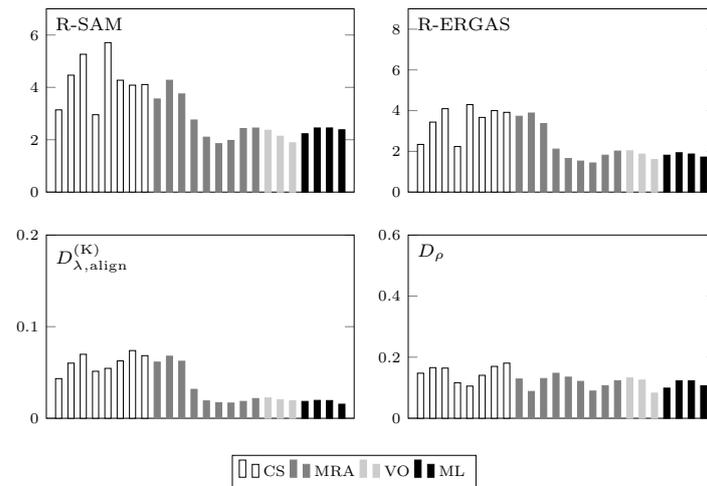
**Figure 8.** Pansharpening results with controlled spectral ( $\ell_1$ -norm) and spatial ( $D_\rho$ ) distortions. From left to right: MS and PAN, followed by the outputs optimized using  $\sigma \in \{2, 4, 8, 16, 32, 64\}$ .

The present experiment gives us the opportunity to also make some considerations about the computational load related to the proposed indexes. In general, in the context of quality assessment, the computation time is not a critical issue as it is carried out offline. However, with the advent of deep learning, researchers have started to use such quality indexes as loss functions for training purposes. In this regard, it is worth distinguishing pixel-based indexes (e.g., SAM, ERGAS,  $\ell_\alpha$ -norms), which are relatively light to compute, from other indexes that involve local statistics computed at a certain scale for each location (e.g.,  $Q2^n$ ,  $D_\rho$ ,  $D_\lambda^{(K)}$ ). When using the latter as loss, the training may slow down. In the particular case of  $D_\rho$ , which has been included in the loss function to fine-tune the sample CNN employed in the experiment, we have actually registered a moderate impact on the training time. In fact, using a training batch composed of a single  $2048 \times 2048$  image, the time consumption per iteration, without or with the additional  $D_\rho$  loss component (second term in Equation (13)), shifts from 1.27 s to 1.6 s on an NVIDIA P6000 GPU.

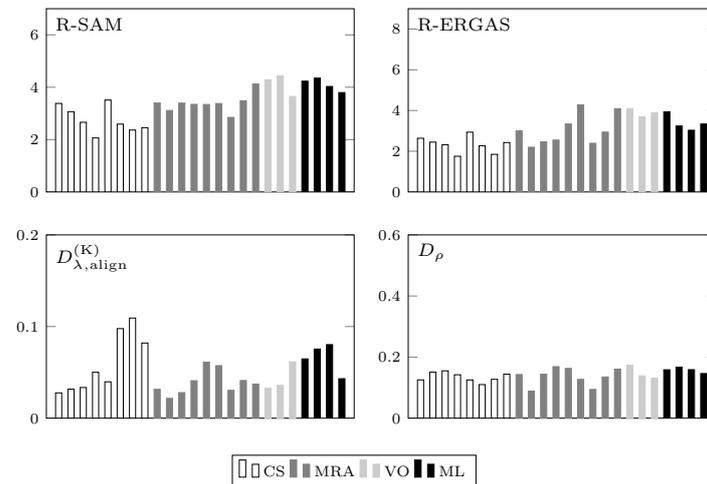
#### 4.5. Comparative Results

To conclude, we present here an overall comparison among the available pansharpening methods provided by the toolbox [1], using the proposed quality indicators and supported by the visual inspection of some sample results. In Figures 9 and 10, we gather the average numerical results obtained on our WV3 and WV2 datasets, respectively.

As expected, the numerical results clearly show a good level of agreement among the three reprojection error indexes, all essentially linked to the spectral consistency. However, some exceptions can be observed, particularly for the WV2 dataset (Figure 10), where some CS methods register a performance loss in terms of  $D_{\lambda, \text{align}}^{(K)}$  not observed on R-SAM and R-ERGAS. We also recognize good agreement with some literature findings, particularly on WV3 (Figure 9), such as the spectral accuracy gap between MRA and CS methods [44] or the competitiveness of some MRA, VO, and ML solutions. Moreover, for the ML solutions, it is worth noticing a performance gap moving from one dataset (WV3) to another (WV2). Indeed, such variability is not unusual for data-driven approaches, which can suffer generalization limits when the training dataset is not sufficiently representative. In this particular situation, it could likely be the case that the WV2 test images used are too different from the images used for the training of the involved ML methods.



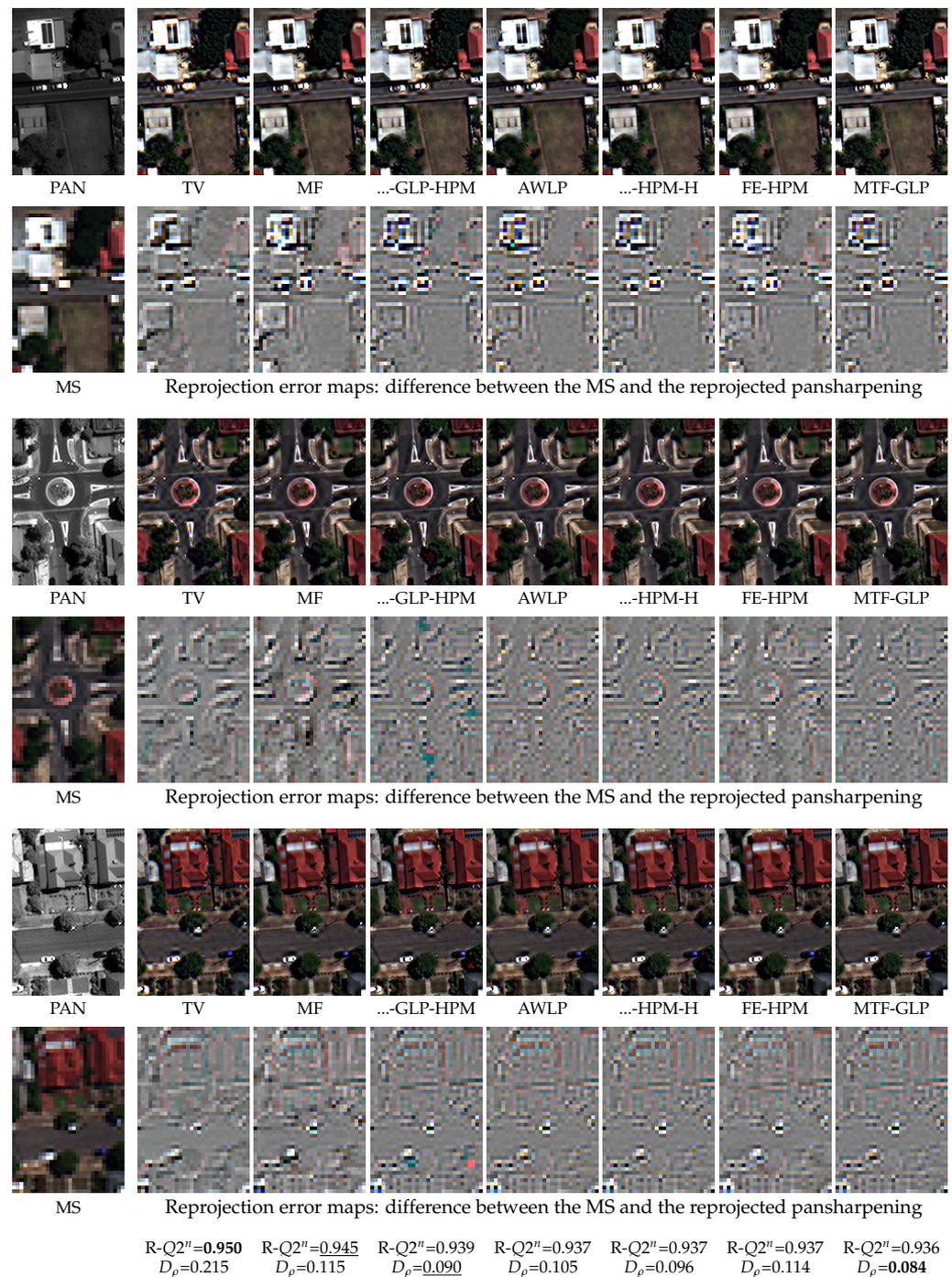
**Figure 9.** Full-resolution assessment on the WV3 dataset.  $R-Q2^n$  is replaced by its complement  $D_{\lambda,align}^{(K)}$  for the sake of readability (0 is the ideal value for all). For details on methods and related categories (CS, MRA, VO, ML), refer to Table 2.



**Figure 10.** Full-resolution assessment on the WV2 dataset.  $R-Q2^n$  is replaced by its complement  $D_{\lambda,align}^{(K)}$  for the sake of readability (0 is the ideal value for all). For details on methods and related categories (CS, MRA, VO, ML), refer to Table 2.

To gain insight into the effectiveness of the proposed indexes, let us take a look at some sample results for a complementary subjective analysis. In Figure 11, we show some clips from a single WV3 tile of the Adelaide dataset. On the leftmost column are gathered the PAN and MS input components on two consecutive rows. Then, the corresponding pansharpening results of the best-performing solutions according to  $R-Q2^n$  are shown in decreasing order from left to right, next to the PAN. Next to the MS, we also show the “reprojection” error map, which is the difference between the input MS and the reprojection (LPF plus decimation) of the output. On the bottom lines are gathered the  $R-Q2^n$  and  $D_{\rho}$  scores. Therefore, the spectral quality of the results decreases left to right. This can also be appreciated through the inspection of the reprojection error maps. Moreover, the best methods from the spatial perspective ( $D_{\rho}$ ) follow a different ordering. This partially reflects a tradeoff between spectral and spatial features. Particularly interesting is the case of the method TV, which scores first in terms of  $R-Q2^n$  (actually, it is the best according to R-SAM and R-ERGAS, as well) but shows the worst  $D_{\rho}$  value, 0.215, corresponding to a 78.5% average correlation between the PAN and the pansharpened bands. The impact of this low correlation level is clearly visible in the pansharpening results of TV, which show underlying noisy patterns. Such patterns disappear with the reprojection, explaining why

the reprojection indexes do not worsen. The other methods achieve much lower values of  $D_\rho$ , ranging from 0.084 to 0.115 (around 90% of correlation), which corresponds to a higher coherence between the spatial features of the results and the PAN, easily appreciable through visual inspection. For completeness, in Figure 12, we show analogous results for a WV2 tile, which basically confirms the same considerations made above for WV3.



**Figure 11.** Sample crops from a WV3 sample image (Adelaide) and related full-resolution pansharpening. For each crop, in addition to the PAN and MS components (left-most column), the best-performing solutions according to R-Q2<sup>n</sup> are shown from left to right in decreasing order, next to the PAN. For each method, both the R-Q2<sup>n</sup> and  $D_\rho$  scores are reported on the bottom lines of the figure. Moreover, next to the MS, for each pansharpening result, the corresponding reprojection error map is shown.



**Figure 12.** Sample crops from a WV2 sample image (Washington) and related full-resolution pansharpening. For each crop, in addition to the PAN and MS components (**left-most** column), the best-performing solutions according to  $R-Q2^n$  are shown from left to right in decreasing order, next to the PAN. For each method, both the  $R-Q2^n$  and  $D_\rho$  scores are reported on the bottom lines of the figure. Moreover, next to the MS, for each pansharpening result, the corresponding reprojection error map is shown.

## 5. Conclusions

To cope with the limitations of the reference-based reduced-resolution procedures for the quality assessment of pansharpening methods, in this work, we propose a new full-resolution no-reference evaluation framework. By following Wald's protocol [65], the full-resolution assessment must be carried out checking the “consistency” of the fused

products with the MS and PAN input components rather than the “synthesis” capacity, which would require the availability of GTs. In particular, inspired by Khan’s index [73], we have proposed the use of reprojection-based indexes with embedded alignment to handle misregistered datasets for the assessment of the spectral consistency between the pansharpened image and the input MS. Moreover, the spatial consistency between the fused image and the PAN is quantified by averaging, spatially and spectrally, the fine-scale local correlation of individual super-resolved bands with the high-resolution PAN.

A key qualifying aspect of the proposed indexes is the absence of any resolution downgrading of the input data, which frees the assessment from the effect of scale-dependent phenomena. Experiments on reduced-resolution datasets show that the reprojection indexes are reliable predictors of image quality as quantified by reference-based indexes, supporting their use in the full-resolution domain. On the other hand, experiments on full-resolution data make clear that the local correlation-based index provides indications on image quality that largely agree with the judgement of human experts. The proposed approach can be readily generalized to fusion tasks other than pansharpening, such as, for example, the combination of low-resolution hyperspectral and high-resolution multispectral images.

On the other hand, the user must also be aware of some limitations. In particular, the reprojection requires the knowledge of an accurate estimation of the sensor MTF for a correct low-pass filtering. A wrong estimation would lead to inaccurate spectral consistency assessment, a problem shared with preexisting solutions. Moreover, the proposed correlation distortion index is based on the assumption that the PAN correlates well with all spectral bands. Such a hypothesis is globally acceptable but there can be rare cases, spatially and spectrally well localized, for which it may be too strong.

In order to ensure reproducible research, the code for the proposed indexes can be found at <https://github.com/matciotola/fr-pansh-eval-tool/> (accessed on 5 April 2022).

**Author Contributions:** Conceptualization, G.S. and M.C.; methodology, G.S. and M.C.; software, M.C.; validation, G.S. and M.C.; formal analysis, G.S. and M.C.; investigation, G.S. and M.C.; resources, G.S. and M.C.; data curation, M.C.; writing—original draft preparation, G.S.; writing—review and editing, G.S. and M.C.; visualization, M.C.; supervision, G.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank DigitalGlobe© and ESA for providing sample images for research purposes.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

MS	Multispectral
PAN	Panchromatic
GT	Ground Truth
CS	Component Substitution
MRA	Multiresolution Analysis
VO	Variational Optimization
ML	Machine Learning
IHS	Intensity–Hue–Saturation
GIHS	Generalized IHS
GS	Gram–Schmidt
PRACS	Partial Replacement Adaptive Component Substitution
BSDS	Band-Dependent Spatial-Detail
RR	Reduced-Resolution
FR	Full-Resolution

SAM	Spectral Angle Mapper
ERGAS	<i>Erreur Relative Globale Adimensionnelle de Synthèse</i>
RMSE	Root Mean Squared Error
HC	HyperComplex
LPF	Low-Pass Filter
UIQI	Universal Image Quality Index
WV2	WorldView-2
WV3	WorldView-3
GSD	Ground Sample Distance
SSIM	Structural SIMilarity
QLR	Quality Low Resolution
QHR	Quality High Resolution
NIQE	Natural Image Quality Evaluator

## References

- Vivone, G.; Dalla Mura, M.; Garzelli, A.; Restaino, R.; Scarpa, G.; Ulfarsson, M.O.; Alparone, L.; Chanussot, J. A New Benchmark Based on Recent Advances in Multispectral Pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods. *IEEE Geosci. Remote Sens. Mag.* **2020**, *9*, 53–81. [[CrossRef](#)]
- Shettigara, V. A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set. *Photogramm. Eng. Remote Sens.* **1992**, *58*, 561–567.
- Ranchin, T.; Wald, L. Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 49–61.
- Vivone, G.; Simões, M.; Dalla Mura, M.; Restaino, R.; Bioucas-Dias, J.M.; Licciardi, G.A.; Chanussot, J. Pansharpening Based on Semiblind Deconvolution. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1997–2010. [[CrossRef](#)]
- Palsson, F.; Ulfarsson, M.O.; Sveinsson, J.R. Model-Based Reduced-Rank Pansharpening. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 656–660. [[CrossRef](#)]
- Masi, G.; Cozzolino, D.; Verdoliva, L.; Scarpa, G. Pansharpening by Convolutional Neural Networks. *Remote Sens.* **2016**, *8*, 594. [[CrossRef](#)]
- Yang, J.; Fu, X.; Hu, Y.; Huang, Y.; Ding, X.; Paisley, J. PanNet: A Deep Network Architecture for Pan-Sharpener. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017. [[CrossRef](#)]
- Tu, T.M.; Su, S.C.; Shyu, H.C.; Huang, P.S. A new look at IHS-like image fusion methods. *Inf. Fusion* **2001**, *2*, 177–186. [[CrossRef](#)]
- Tu, T.M.; Huang, P.S.; Hung, C.L.; Chang, C.P. A fast intensity hue-saturation fusion technique with spectral adjustment for IKONOS imagery. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 309–312. [[CrossRef](#)]
- Chavez, P.; Kwarteng, A. Extracting spectral contrast in Landsat thematic mapper image data using selective principal component analysis. *Photogramm. Eng. Remote Sens.* **1989**, *55*, 339–348.
- Gillespie, A.R.; Kahle, A.B.; Walker, R.E. Color enhancement of highly correlated images. II. Channel ratio and “chromaticity” transformation techniques. *Remote Sens. Environ.* **1987**, *22*, 343–365. [[CrossRef](#)]
- Laben, C.; Brower, B. Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-Sharpener. U.S. Patent 6011875, 4 January 2000.
- Aiazzi, B.; Baronti, S.; Selva, M. Improving component substitution pansharpening through multivariate regression of MS+Pan data. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3230–3239. [[CrossRef](#)]
- Choi, J.; Yu, K.; Kim, Y. A New Adaptive Component-Substitution-Based Satellite Image Fusion by Using Partial Replacement. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 295–309. [[CrossRef](#)]
- Garzelli, A.; Nencini, F.; Capobianco, L. Optimal MMSE pan sharpening of very high resolution multispectral images. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 228–236. [[CrossRef](#)]
- Garzelli, A. Pansharpening of Multispectral Images Based on Nonlocal Parameter Optimization. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2096–2107. [[CrossRef](#)]
- Vivone, G. Robust Band-Dependent Spatial-Detail Approaches for Panchromatic Sharpening. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6421–6433. [[CrossRef](#)]
- Nunez, J.; Otazu, X.; Fors, O.; Prades, A.; Pala, V.; Arbiol, R. Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1204–1211. [[CrossRef](#)]
- Otazu, X.; Gonzalez-Audicana, M.; Fors, O.; Nunez, J. Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2376–2385. [[CrossRef](#)]
- Khan, M.; Chanussot, J.; Condat, L.; Montanvert, A. Indusion: Fusion of Multispectral and Panchromatic Images Using the Induction Scaling Technique. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 98–102. [[CrossRef](#)]
- Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2300–2312. [[CrossRef](#)]

22. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A.; Selva, M. An MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas. In Proceedings of the GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Berlin, Germany, 22–23 May 2003; pp. 90–94.
23. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A.; Selva, M. MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 591–596. [[CrossRef](#)]
24. Lee, J.; Lee, C. Fast and Efficient Panchromatic Sharpening. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 155–163. [[CrossRef](#)]
25. Restaino, R.; Mura, M.D.; Vivone, G.; Chanussot, J. Context-Adaptive Pansharpening Based on Image Segmentation. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 753–766. [[CrossRef](#)]
26. Shah, V.P.; Younan, N.H.; King, R.L. An Efficient Pan-Sharpener Method via a Combined Adaptive PCA Approach and Contourlets. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1323–1335. [[CrossRef](#)]
27. Vicinanza, M.R.; Restaino, R.; Vivone, G.; Mura, M.D.; Chanussot, J. A Pansharpening Method Based on the Sparse Representation of Injected Details. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 180–184. [[CrossRef](#)]
28. Palsson, F.; Sveinsson, J.; Ulfarsson, M. A New Pansharpening Algorithm Based on Total Variation. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 318–322. [[CrossRef](#)]
29. Palsson, F.; Sveinsson, J.R.; Ulfarsson, M.O.; Benediktsson, J.A. Model-Based Fusion of Multi- and Hyperspectral Images Using PCA and Wavelets. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2652–2663. [[CrossRef](#)]
30. Fasbender, D.; Radoux, J.; Bogaert, P. Bayesian Data Fusion for Adaptable Image Pansharpening. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1847–1857. [[CrossRef](#)]
31. Zhang, L.; Shen, H.; Gong, W.; Zhang, H. Adjustable Model-Based Fusion Method for Multispectral and Panchromatic Images. *IEEE Trans. Syst. Man Cybern. B Cybern.* **2012**, *42*, 1693–1704. [[CrossRef](#)]
32. Meng, X.; Shen, H.; Li, H.; Yuan, Q.; Zhang, H.; Zhang, L. Improving the spatial resolution of hyperspectral image using panchromatic and multispectral images: An integrated method. In Proceedings of the 2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Tokyo, Japan, 2–5 June 2015.
33. Shen, H.; Meng, X.; Zhang, L. An Integrated Framework for the Spatio-Temporal-Spectral Fusion of Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 7135–7148. [[CrossRef](#)]
34. Zhong, S.; Zhang, Y.; Chen, Y.; Wu, D. Combining Component Substitution and Multiresolution Analysis: A Novel Generalized BDSF Pansharpening Algorithm. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2867–2875. [[CrossRef](#)]
35. Li, S.; Yang, B. A New Pan-Sharpener Method Using a Compressed Sensing Technique. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 738–746. [[CrossRef](#)]
36. Li, S.; Yin, H.; Fang, L. Remote Sensing Image Fusion via Sparse Representations Over Learned Dictionaries. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 4779–4789. [[CrossRef](#)]
37. Zhu, X.; Bamler, R. A Sparse Image Fusion Algorithm With Application to Pan-Sharpener. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2827–2836. [[CrossRef](#)]
38. Cheng, M.; Wang, C.; Li, J. Sparse representation based pansharpening using trained dictionary. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 293–297. [[CrossRef](#)]
39. Zhu, X.X.; Grohnfeldt, C.; Bamler, R. Exploiting Joint Sparsity for Pansharpening: The J-SparseFI Algorithm. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 2664–2681. [[CrossRef](#)]
40. Hong, D.; Yokoya, N.; Chanussot, J.; Zhu, X.X. An Augmented Linear Mixing Model to Address Spectral Variability for Hyperspectral Unmixing. *IEEE Trans. Image Process.* **2019**, *28*, 1923–1938. [[CrossRef](#)]
41. Yokoya, N.; Yairi, T.; Iwasaki, A. Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 528–537. [[CrossRef](#)]
42. Lanaras, C.; Baltsavias, E.; Schindler, K. Hyperspectral Super-Resolution by Coupled Spectral Unmixing. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 3586–3594.
43. Hong, D.; Yokoya, N.; Chanussot, J.; Zhu, X.X. CoSpace: Common Subspace Learning From Hyperspectral-Multispectral Correspondences. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4349–4359. [[CrossRef](#)]
44. Vivone, G.; Alparone, L.; Chanussot, J.; Mura, M.D.; Garzelli, A.; Licciardi, G.A.; Restaino, R.; Wald, L. A Critical Comparison Among Pansharpening Algorithms. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2565–2586. [[CrossRef](#)]
45. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–8 December 2012; pp. 1106–1114.
46. Dong, C.; Loy, C.; He, K.; Tang, X. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 295–307. [[CrossRef](#)]
47. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2980–2988.
48. Lateef, F.; Ruichek, Y. Survey on semantic segmentation using deep learning techniques. *Neurocomputing* **2019**, *338*, 321–348. [[CrossRef](#)]
49. Zhao, Z.; Zheng, P.; Xu, S.; Wu, X. Object Detection With Deep Learning: A Review. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3212–3232. [[CrossRef](#)] [[PubMed](#)]
50. Scarpa, G.; Gargiulo, M.; Mazza, A.; Gaetano, R. A CNN-Based Fusion Method for Feature Extraction from Sentinel Data. *Remote Sens.* **2018**, *10*, 236. [[CrossRef](#)]

51. Benedetti, P.; Ienco, D.; Gaetano, R.; Ose, K.; Pensa, R.G.; Dupuy, S.  $M^3$ Fusion: A Deep Learning Architecture for Multiscale Multimodal Multitemporal Satellite Data Fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 4939–4949. [[CrossRef](#)]
52. Mazza, A.; Sica, F.; Rizzoli, P.; Scarpa, G. TanDEM-X Forest Mapping Using Convolutional Neural Networks. *Remote Sens.* **2019**, *11*, 2980. [[CrossRef](#)]
53. Wei, Y.; Yuan, Q. Deep residual learning for remote sensed imagery pansharpening. In Proceedings of the 2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP), Shanghai, China, 19–21 May 2017; pp. 1–4.
54. Wei, Y.; Yuan, Q.; Shen, H.; Zhang, L. Boosting the Accuracy of Multispectral Image Pansharpening by Learning a Deep Residual Network. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1795–1799. [[CrossRef](#)]
55. Rao, Y.; He, L.; Zhu, J. A residual convolutional neural network for pan-sharpening. In Proceedings of the 2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP), Shanghai, China, 19–21 May 2017; pp. 1–4.
56. Masi, G.; Cozzolino, D.; Verdoliva, L.; Scarpa, G. CNN-based Pansharpening of Multi-Resolution Remote-Sensing Images. In Proceedings of the Joint Urban Remote Sensing Event 2017, Dubai, United Arab Emirates, 6–8 March 2017.
57. Azarang, A.; Ghassemian, H. A new pansharpening method using multi resolution analysis framework and deep neural networks. In Proceedings of the 2017 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA), Shahrekord, Iran, 19–20 April 2017; pp. 1–6.
58. Yuan, Q.; Wei, Y.; Meng, X.; Shen, H.; Zhang, L. A Multiscale and Multidimensional Convolutional Neural Network for Remote Sensing Imagery Pan-Sharpener. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 978–989. [[CrossRef](#)]
59. Liu, X.; Wang, Y.; Liu, Q. Psgan: A Generative Adversarial Network for Remote Sensing Image Pan-Sharpener. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 873–877.
60. Shao, Z.; Cai, J. Remote Sensing Image Fusion With Deep Convolutional Neural Network. *IEEE J. Sel. Top. Appl. Earth Observ.* **2018**, *11*, 1656–1669. [[CrossRef](#)]
61. Vitale, S.; Ferraioli, G.; Scarpa, G. A CNN-Based Model for Pansharpening of WorldView-3 Images. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 5108–5111. [[CrossRef](#)]
62. Zhang, Y.; Liu, C.; Sun, M.; Ou, Y. Pan-Sharpener Using an Efficient Bidirectional Pyramid Network. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5549–5563. [[CrossRef](#)]
63. Dong, W.; Zhang, T.; Qu, J.; Xiao, S.; Liang, J.; Li, Y. Laplacian Pyramid Dense Network for Hyperspectral Pansharpening. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–13. [[CrossRef](#)]
64. Dong, W.; Hou, S.; Xiao, S.; Qu, J.; Du, Q.; Li, Y. Generative Dual-Adversarial Network With Spectral Fidelity and Spatial Enhancement for Hyperspectral Pansharpening. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, 1–15. [[CrossRef](#)] [[PubMed](#)]
65. Wald, L.; Ranchin, T.; Mangolini, M. Fusion of satellite images of different spatial resolution: Assessing the quality of resulting images. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 691–699.
66. Kruse, F.A.; Lefkoff, A.; Boardman, J.; Heidebrecht, K.; Shapiro, A.; Barloon, P.; Goetz, A. The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.* **1993**, *44*, 145–163. [[CrossRef](#)]
67. Zhou, J.; Civco, D.L.; Silander, J.A. A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *Int. J. Remote Sens.* **1998**, *19*, 743–757. [[CrossRef](#)]
68. Wald, L. Quality of high resolution synthesised images: Is there a simple criterion? In Proceedings of the Third conference “Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images”. SEE/URISCA, Sophia Antipolis, France, 26–28 January 2000; pp. 99–103.
69. Wang, Z.; Bovik, A.C. A universal image quality index. *IEEE Signal Process. Lett.* **2002**, *9*, 81–84. [[CrossRef](#)]
70. Alparone, L.; Baronti, S.; Garzelli, A.; Nencini, F. A global quality measurement of pan-sharpened multispectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 313–317. [[CrossRef](#)]
71. Garzelli, A.; Nencini, F. Hypercomplex Quality Assessment of Multi/Hyperspectral Images. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 662–665. [[CrossRef](#)]
72. Alparone, L.; Aiazzi, B.; Baronti, S.; Garzelli, A.; Nencini, F.; Selva, M. Multispectral and panchromatic data fusion assessment without reference. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 193–200. [[CrossRef](#)]
73. Khan, M.M.; Alparone, L.; Chanussot, J. Pansharpening Quality Assessment Using the Modulation Transfer Functions of Instruments. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3880–3891. [[CrossRef](#)]
74. Palubinskas, G. Quality assessment of pan-sharpening methods. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 2526–2529.
75. Palubinskas, G. Joint quality measure for evaluation of pansharpening accuracy. *Remote Sens.* **2015**, *7*, 9292–9310. [[CrossRef](#)]
76. Alparone, L.; Garzelli, A.; Vivone, G. Spatial consistency for full-scale assessment of pansharpening. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 5132–5134.
77. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [[CrossRef](#)]
78. Kwan, C.; Budavari, B.; Bovik, A.C.; Marchisio, G. Blind Quality Assessment of Fused WorldView-3 Images by Using the Combinations of Pansharpening and Hypersharpening Paradigms. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1835–1839. [[CrossRef](#)]
79. Aiazzi, B.; Alparone, L.; Baronti, S.; Carlà, R.; Garzelli, A.; Santurri, L. Full scale assessment of pansharpening methods and data products. *Proc. SPIE* **2014**, *9244*, 1–12.

80. Meng, X.; Bao, K.; Shu, J.; Zhou, B.; Shao, F.; Sun, W.; Li, S. A Blind Full-Resolution Quality Evaluation Method for Pansharpening. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–16. [[CrossRef](#)]
81. Carla, R.; Santurri, L.; Aiazzi, B.; Baronti, S. Full-scale assessment of pansharpening through polynomial fitting of multiscale measurements. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6344–6355. [[CrossRef](#)]
82. Vivone, G.; Restaino, R.; Chanussot, J. A Bayesian Procedure for Full-Resolution Quality Assessment of Pansharpened Products. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4820–4834. [[CrossRef](#)]
83. Vivone, G.; Addesso, P.; Chanussot, J. A combiner-based full resolution quality assessment index for pansharpening. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 437–441. [[CrossRef](#)]
84. Szczykutowicz, T.P.; Rose, S.D.; Kitt, A. Sub pixel resolution using spectral-spatial encoding in x-ray imaging. *PLoS ONE* **2021**, *16*, e0258481. [[CrossRef](#)]
85. Kordi Ghasrodashti, E.; Sharma, N. Hyperspectral image classification using an extended Auto-Encoder method. *Signal Process. Image Commun.* **2021**, *92*, 116111. [[CrossRef](#)]
86. Zhu, H.; Gowen, A.; Feng, H.; Yu, K.; Xu, J.L. Deep Spectral-Spatial Features of Near Infrared Hyperspectral Images for Pixel-Wise Classification of Food Products. *Sensors* **2020**, *20*, 5322. [[CrossRef](#)]
87. Alparone, L.; Wald, L.; Chanussot, J.; Thomas, C.; Gamba, P.; Bruce, L. Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S Data-Fusion Contest. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3012–3021. [[CrossRef](#)]
88. Daubechies, I. Orthonormal bases of compactly supported wavelets II. Variations on a theme. *SIAM J. Math. Anal.* **1993**, *24*, 499–519. [[CrossRef](#)]
89. Ranchin, T.; Wald, L. The wavelet transform for the analysis of remotely sensed images. *Int. J. Remote Sens.* **1993**, *14*, 615–619. [[CrossRef](#)]
90. Munechika, C.K.; Warnick, J.S.; Salvaggio, C.; Schott, J.R. Resolution enhancement of multispectral image data to improve classification accuracy. *Photogramm. Eng. Remote Sens.* **1993**, *59*, 67–72.
91. Thomas, C.; Ranchin, T.; Wald, L.; Chanussot, J. Synthesis of Multispectral Images to High Spatial Resolution: A Critical Review of Fusion Methods Based on Remote Sensing Physics. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1301–1312. [[CrossRef](#)]
92. Lolli, S.; Alparone, L.; Garzelli, A.; Vivone, G. Haze Correction for Contrast-Based Multispectral Pansharpening. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2255–2259. [[CrossRef](#)]
93. Baronti, S.; Aiazzi, B.; Selva, M.; Garzelli, A.; Alparone, L. A Theoretical Analysis of the Effects of Aliasing and Misregistration on Pansharpened Imagery. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 446–453. [[CrossRef](#)]
94. Jing, L.; Cheng, Q.; Guo, H.; Lin, Q. Image misalignment caused by decimation in image fusion evaluation. *Int. J. Remote Sens.* **2012**, *33*, 4967–4981. [[CrossRef](#)]
95. Xu, Q.; Zhang, Y.; Li, B. Recent advances in pansharpening and key problems in applications. *Int. J. Image Data Fusion* **2014**, *5*, 175–195. [[CrossRef](#)]
96. Alparone, L.; Garzelli, A.; Vivone, G. Intersensor Statistical Matching for Pansharpening: Theoretical Issues and Practical Solutions. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4682–4695. [[CrossRef](#)]
97. Vivone, G.; Restaino, R.; Chanussot, J. Full Scale Regression-Based Injection Coefficients for Panchromatic Sharpening. *IEEE Trans. Image Process.* **2018**, *27*, 3418–3431. [[CrossRef](#)] [[PubMed](#)]
98. Vivone, G.; Restaino, R.; Chanussot, J. A Regression-Based High-Pass Modulation Pansharpening Approach. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 984–996. [[CrossRef](#)]
99. Restaino, R.; Vivone, G.; Dalla Mura, M.; Chanussot, J. Fusion of Multispectral and Panchromatic Images Based on Morphological Operators. *IEEE Trans. Image Process.* **2016**, *25*, 2882–2895. [[CrossRef](#)] [[PubMed](#)]
100. Scarpa, G.; Vitale, S.; Cozzolino, D. Target-Adaptive CNN-Based Pansharpening. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 5443–5457. [[CrossRef](#)]
101. Ciotola, M.; Vitale, S.; Mazza, A.; Poggi, G.; Scarpa, G. Pansharpening by convolutional neural networks in the full resolution framework. *IEEE Trans. Geosci. Remote Sens.* **2022**. [[CrossRef](#)]