

# Image Fusion in Remote Sensing: An Overview and Meta Analysis

Hessah Albanwan, Rongjun Qin, Yang Tang

**Abstract**— Image fusion in Remote Sensing (RS) has been a consistent demand due to its ability to turn raw images of different resolutions, sources, and modalities into accurate, complete, and spatio-temporally coherent images. It greatly facilitates downstream applications such as pan-sharpening, change detection, land-cover classification, etc. Yet, image fusion solutions are highly disparate to various remote sensing problems and thus are often narrowly defined in existing reviews as topical applications, such as pan-sharpening, and spatial-temporal image fusion. Considering that image fusion can be theoretically applied to any gridded data through pixel-level operations, in this paper, we expanded its scope by comprehensively surveying relevant works with a simple taxonomy: 1) many-to-one image fusion; 2) many-to-many image fusion. This simple taxonomy defines image fusion as a mapping problem that turns either a single or a set of images into another single or set of images, depending on the desired coherence, e.g., spectral, spatial/resolution coherence, etc. We show that this simple taxonomy, despite the significant modality difference it covers, can be presented by a conceptually easy framework. In addition, we provide a meta-analysis to review the major papers studying the various types of image fusion and their applications over the years (from the 1980s to date), covering 5,926 peer-reviewed papers. Finally, we discuss the main benefits and emerging challenges to provide open research directions and potential future works.

**Index Terms**— Fusion, Meta-Analysis, Remote sensing images, Learning-based fusion methods

## 1. INTRODUCTION

**I**MAGE fusion can be broadly defined as the process of combining information from single or multiple images into more accurate, complete, and spatiotemporally coherent images. It has received great attention in remote sensing (RS) over the years due to the ever-growing development of sensors and platforms, which provide a nearly unlimited archive of earth observation data. Complementary images from different sensors/platforms enable augmented observations of RS data to enrich and expand the knowledge

The study is partially supported by the Office of Naval Research (Grand No. N000142012141 & N000142312670). (Corresponding author: Rongjun Qin, qin.324@osu.edu).

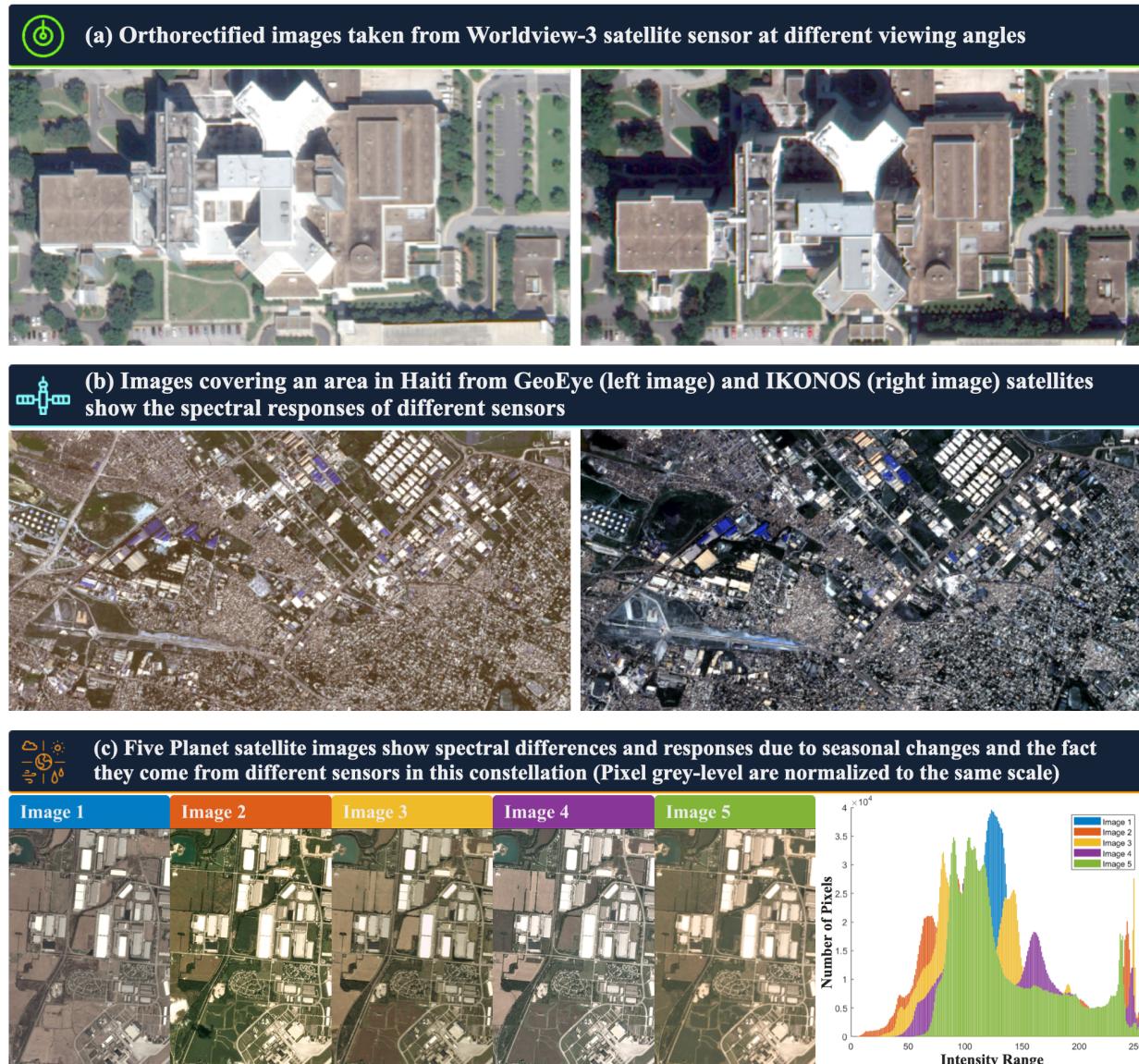
Hessah Albanwan is with the Geospatial Data Analytics Lab and Department of Civil, Environment and Geodetic Engineering, The Ohio State University, Columbus, USA (email: [albanwan.1@osu.edu](mailto:albanwan.1@osu.edu))

Rongjun Qin is with the Geospatial Data Analytics Lab, Department of Civil, Environment and Geodetic Engineering and the Department of Electrical and Computer Engineering, and Translational Data Analytics Institute, The Ohio state University, Columbus, USA (email: [qin.324@osu.edu](mailto:qin.324@osu.edu))

Yang Tang is with the Geospatial Data Analytics Lab and Department of Civil, Environment and Geodetic Engineering, The Ohio State University, Columbus, USA (email: [tang.1693@osu.edu](mailto:tang.1693@osu.edu))

around them [1], [2], [3], [4]. Existing image fusion methods that fundamentally achieve augmented observations were shown to be of great interest in many RS applications [1], [2], [3], [4], [5], [6], [7], [8]. These methods aim at enhancing spatial, temporal, and spectral resolutions at the image level and, at the same time, improving the learnable features from various sources to achieve better image interpretation performances [1], [2], [4]. For instance, by combining low-spatial high-temporal resolution images from MODIS with high-spatial low-temporal resolution images from Landsat, we can generate high-spatial and high-temporal images for more accurate land-cover and biomass monitoring applications [5], [6], [7], [9]. In contrast, high-spectral low-spatial resolution data (e.g., multi- / hyper-spectral images) can be used to enhance the spectral resolution of panchromatic images, i.e., pan-sharpening [6], [8], [10]. Beyond single-modal data (i.e., optical images), researchers have also explored using multi-modality data for image-based feature learning [11], [12]. These multi-modality data include but are not limited to optical images, LiDAR (Light Detection and Ranging), and SAR (Synthetic Aperture Radar) data, which have shown that such a feature-level fusion can significantly enhance the performance of machine learning tasks for remote sensing [13], [14], [15].

The past literature has provided syntheses of the existing image fusion efforts with a great focus on individual topical applications and their specific fusion algorithms, for example, pan-sharpening, super-resolution, or feature-level fusion for classification [1], [16], [17], [18]. Their limited area of investigation restricts comprehensive understanding and views of the image fusion concepts, which can also be applicable to other RS tasks and problems. For instance, with some adaptation, fusion techniques used in pan-sharpening can be applied to super-resolution or spatiotemporal inference problems. In RS, the majority of data is still operating through a raster format defined under a geo-referenced coordinate frame (i.e., projection space) [19], [20]. On the one hand, this made it relatively easy to register data with respect to each other and, on the other hand, made it highly effective to implement algorithms to two-dimensional (2D) grids where multiple data layers can be fused. Hence, fusion applications on 2D image grids can be easily extended through various RS applications and data products, such as optical images, classification maps, thematic maps, or Digital surface models (DSMs) [1], [2], [3], [4], [5], [6], [7], [8]. These products, although being presented as 2D raster, possess completely different content associated with different uncertainties and numerical values, in addition to being used in completely different applications. Therefore, it can be beneficial to comprehensively survey image fusion methods that are



**Figure 1.** Examples of radiometric inconsistencies between RS images due to differences in (a) viewing direction, (b) sensors, and (c) season.

generally applicable to various applications with a broader scope.

This systematic review focuses on providing a comprehensive revision and a meta-analysis on the topic of image fusion in the RS field. Despite the significant difference in the type of image fusion and its applications, we consider that their fusion methods can share similar concepts. Therefore, the framework used in one method can be used to augment other types of images and applications. For example, fusion techniques used for super-resolution can also be applied for pan-sharpening, while methods used to process and fuse time series analysis, such as filtering, can be used for image pre-processing. Moreover, publications informed by the meta-analysis are used as measuring entities to evaluate current scientific activities such as fusion algorithms and applications.

We hereby conduct a search in Scopus for all original peer-reviewed publications on the subject of image fusion in RS published between 1980 and 2023. We reviewed 5,926 articles related to image fusion and its applications in RS, covering specific fusion algorithms and downstream applications in studying urban and environmental issues. Therefore, we expand the scope of image fusion, as practiced in the RS community, by providing a simple yet comprehensive taxonomy that defines fusion as a mapping function operating on different numbers of images to output single or multiple images. We also show that despite the significant “input-output” difference, the fusion problem can be explained by a conceptually easy framework. The meta-analysis plays a significant role in providing insights into the most recent fusion techniques and current applications, specifically the



**Figure 2.** Examples of spatiotemporal inconsistencies are shown in RS images due to (a) natural causes, such as crop cycles, and (b) natural and human causes, like earthquakes and new constructions.

learning-based and deep learning methods, which have become very popular and require a literature update.

The rest of the paper is divided into the following Sections: Section 2 introduces the background information of RS image fusion, which includes the main challenges of image fusion in the RS field, the proposed taxonomy of image fusion, and the commonly used image fusion approaches. Section 3 describes the method and guidelines for this meta-analysis to review image fusion and its evolution in the RS field over the years. Section 4 presents the meta-analysis results and statistical information concerning the number of publications, applications, methods, etc. Finally, Section 5 summarizes this review with an open discussion on the main challenges and potential future works of RS image fusion.

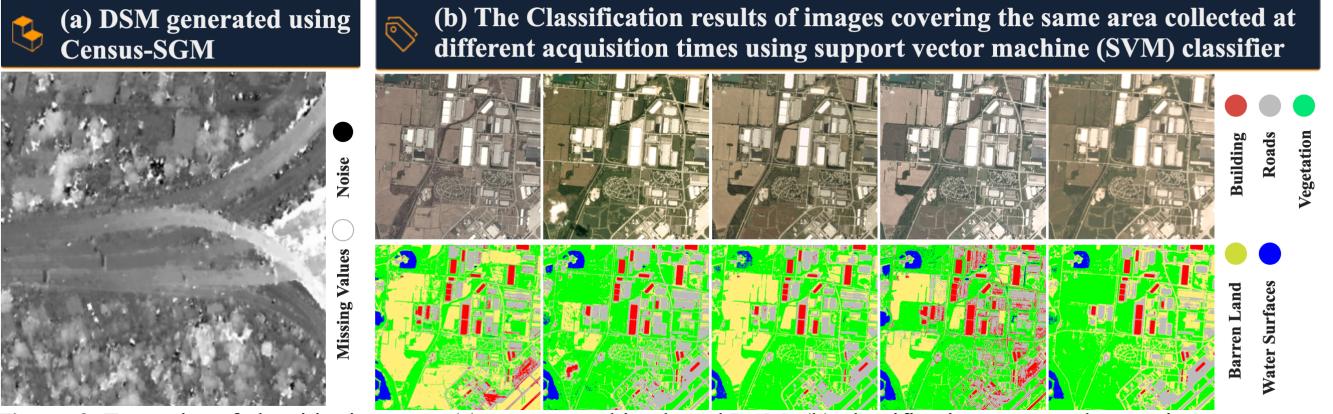
## 2. AN OVERVIEW OF REMOTE SENSING IMAGE FUSION

Despite current RS technologies and sensor advances, image fusion still faces many challenges. Factors of concern such as the image quality (e.g., noisy level, blurriness, radiometric imbalances), the complexity of the scene (e.g., convoluted dense urban and greenspace, homogeneous vegetated regions with high seasonal differences), and processing algorithms (e.g., modeling errors and propagated

errors) play important roles fusion solutions. Therefore, understanding these contributing factors is essential to comprehend the fusion problem and the capabilities of its algorithms. In this section, we discuss the main factors impacting image fusion quality, the proposed taxonomy that categorizes fusion into two groups based on current practices, and finally, the fusion techniques currently in use for various RS applications.

### 2.1. Factors Impacting the Quality of Remote Sensing Image Fusion

As mentioned in Section 1, a typical RS image is a 2D raster data representing physical information about a surface area on earth. The raster “data” can be of any type, such as optical images, elevation maps (e.g., Digital Elevation Model (DEM), Digital Surface Model (DSM), and Digital Terrain Model (DTM)), classification maps, thematic maps, radar images, etc. The pixel information of these images can be directly obtained from sensors (e.g., radiometry) or derived from specific algorithms (e.g., elevation maps from photogrammetry methods). In both cases, inconsistency can be developed due to the varying conditions (e.g., acquisition angle, weather, and season) or errors propagating through the processing algorithms (e.g., DSM errors), leading to



**Figure 3.** Examples of algorithmic errors: (a) stereo matching based DSM; (b) classification maps under varying illumination conditions.

low-quality input prior to fusion. This section summarizes three crucial factors, including radiometric inconsistency, spatiotemporal consistency, and derived/propagated errors from image products such as elevation maps. Details are in the following sections.

#### 2.1.1. Radiometric Inconsistency

Radiometric inconsistency refers to heterogeneous spectral information of objects when observed by different images under varying acquisition conditions [21]. For instance, Figure 1(a) shows the same objects observed by images under different sun angles and sensor's viewing directions, where corresponding pixels show different brightness and chromaticity. The sensors themselves, due to different calibration and spectral responses, may also produce spectral heterogeneities [22], [23], [24]. As an example, Figure 1(b) shows two images for an area in Haiti taken by GeoEye (original Ground Sampling Distance (GSD): 0.5m) and IKONOS (original GSD: 1m), resampled to the same GSD (1m), which still show very different contrast and saturation. Phenological changes, such as those affected by atmospheric conditions and seasons, may also produce spectrally variant images [25], [26]. Figure 1(c) shows an example of five Planet satellite images capturing the same scene within one year of time, where there are significant appearance differences among different collections. Because most fusion algorithms expect spectrally homogenous images, this problem may lead to incorrect use of pixel information data, leading to low-quality fusion results.

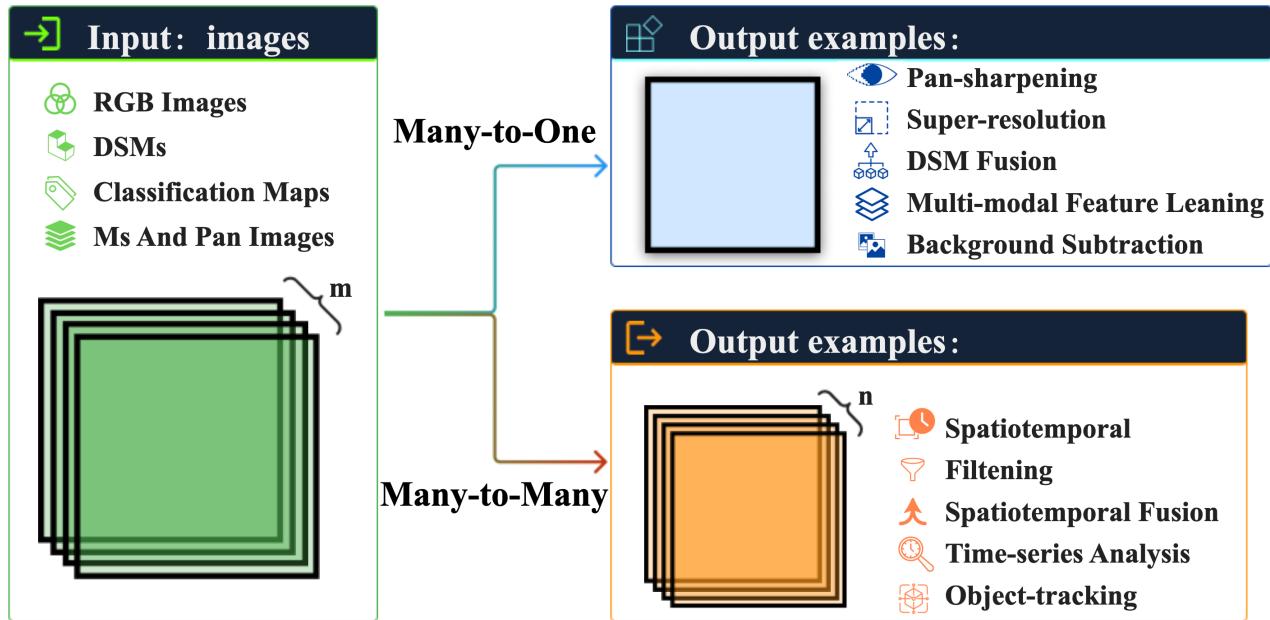
#### 2.1.2. Spatiotemporal Inconsistency

The spatiotemporal inconsistency of RS data refers to images covering the same geographical area but presenting different scene content due to land cover changes over time [27]. This type of inconsistency can sometimes be convoluted by the radiometric inconsistency caused by sensors (explained in 2.1.1.) since images capturing a scene of a different time may come from a different sensor in one or multiple satellite constellations. Typically, spatiotemporal inconsistencies are considered due to natural causes such as seasons, phenological

cycles, disasters, etc., or anthropogenic activities such as urban developments, wars, demolitions, etc. An example of the phenological growth of crops is shown in Figure 2(a), in which we can see the vegetated area in the suburban area in September turned into barren land in December, leading to significant appearance changes. Figure 2(b) shows another example of the spatiotemporal inconsistency, depicting the differences caused by the earthquake (Haiti earthquake in 2010) and its status in 2015. It is obvious that there have been new buildings established to accommodate refugees in houses or shelters (See Figure 2(b)-red box), as well as new buildings (See Figure 2(b) blue and yellow boxes). These inconsistencies, on the one hand, tell the surface dynamics, while on the other hand, when these images are directly used for super-resolution, it will undoubtedly produce haunted effects. Therefore, strategies for dealing with or avoiding these inconsistencies are necessary for successful fusion applications.

#### 2.1.3. Propagated Error

Processed RS data, e.g., classification maps or stereo-based DSM, may contain errors during the data processing [28], [29]. For example, a typical DSM generation pipeline includes steps such as epipolar rectification, stereo matching, and triangulation from point clouds, which may accumulate errors due to numerical and modeling approximation [30], [31] in these steps. An example DSM generated using a semi-global matching (SGM) algorithm [32] is shown in Figure 3(a), where visible noises can be observed on both flat surfaces and building boundaries. Such errors are typical and can be algorithm-dependent, for example, certain stereo-matching algorithms produce sharper boundaries but produce salt-and-pepper noises (such as using normalized cross-correlation algorithms [33], while algorithms that enforce smoothness (e.g., global matching algorithms, such as SGM and deep learning methods (e.g., PSMNet [34])) will produce less salt-and-pepper noise, but may omit important and sharpen structures, such as small and tall buildings. Therefore, understanding such characteristics as priors for fusion is beneficial. Following a similar consideration, when



**Figure 4.** An overview of the types of image fusion in remote sensing. Note: ‘m’ and ‘n’ in the figure denote the numbers of input and output images, respectively.

classification maps are used for fusion, errors from classification maps may yield inconsistent fusion results. For example, Figure 3(b) shows classification maps predicted from images captured at different times with only days of difference, where the illumination differences result in different noise levels, sometimes quite substantial (i.e., a large portion of the Barren land is mixed with the Vegetation class over the short time). The errors introduced by these learning models are still not very well characterized, which can also bring significant challenges for fusion algorithms.

## 2.2. A Generic Taxonomy of Remote Sensing Image Fusion

Image fusion is a highly disparate topic that can be applied to various RS tasks. In the current literature, these different fusion tasks are often narrowly defined within their own methodological domain. For example, pan-sharpening is often considered an independent task other than super-resolution. However, many of these fusion tasks are of a similar nature, and their methodologies share common strategies. The generic fusion problem can be described mathematically: given a set of observations  $X = \{x_1, x_2, x_3, \dots, x_n \mid x_i \in R^w \times h \times c_x\}$ , where  $w, h$ , and  $c_x$  refer to the width, length, and number of channels for raster data. Thus,  $X$  can be represented as a tensor belonging to  $R^{w \times h \times c_x \times n}$ . A simple linear fusion algorithm aims to convolute the tensor  $X$ , through each individual observation via a mapping function  $f(X)$ , and output another tensor  $Y = \{y_1, y_2, y_3, \dots, y_m \mid y_i \in R^{w \times h \times c_y}\}$ , where  $c_y$  refers to the number of channels for each one of out of the  $m$  outputs, thus  $Y$  can be represented as a tensor belonging to  $R^{w \times h \times c_y \times m}$  as follows:

$$Y = f(X) \quad (1)$$

This is easily extendable when  $f(\cdot)$  is a non-linear function involving more complex and multi-step mappings [4], [16], [18]. Here, depending on  $m = 1$  or more, we categorize these strategies into two with the simplest “input-output” fashion, being 1) “many-to-one”, meaning that several input images are fused into a single image output representing enhanced results, where these several inputs often refer to multiple homogenous observations and the fusion of which refers to output with enhanced spatial/spectral resolution or accuracy. 2) “many-to-many”, meaning that several input images are fused into several (often equivalent number of) outputs with enhanced spatiotemporal resolution. Details of these two types will be further explained in the following subsections. As mentioned earlier, we divided the RS image fusion problem into two types: many-to-one (M2O) and many-to-many (M2M). Figure 4 provides an overview of these types of image fusion.

### Many-to-one (M2O) Image Fusion

Many-to-one image fusion refers to the process of combining several images into a single image. This type of problem serves many applications, such as pan-sharpening, super-resolution, DSM fusion, multi-modal feature learning for classification, background image generation for tracking, etc. [7], [12], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45]. These fusion problems have been intensively investigated over the years [1], [2], [4]. The general premise is that the information of different inputs residing in the same pixel location should refer to the same static object with the same or complementary modality. Therefore, the goal of the fusion can be generally categorized into three aspects: 1) estimate a more accurate observation of the same pixel, using the redundant pixel information from these multiple inputs

(e.g., super-resolution, Pan-sharpen, and DSM fusion) [8], [42], [43], [46], [47], [48], [49], [50], [51]; 2) generate a background image to represent the static scene in preparation for object tracking [52], [53], [54], [55], [56], [57]; 3) extract features from different modality (different spectral bands, height, microwave reflectance) to achieve enhanced classification results to generate accurate thematic maps [36], [40], [44], [58], [59], [60], [61], [62], [63].

### Many-to-many (M2M) Image Fusion

Many-to-many image fusion combines the knowledge/information from multiple images by simultaneously or sequentially processing them to yield a group of refined images, often matching the number of inputs. This type of fusion is mostly used for applications that require a bundle of enhanced images, such as multi-image filtering, change detection over time, Multiview 3D reconstruction, spectral coherence between a group of images, and constant monitoring of the earth's surface. [25], [44], [58], [64], [65], [66], [67], [68]. The general premise of this type of image fusion is that the information of different inputs residing in the same pixel location represents a certain state of the objects in that pixel, which are complementary to each other, such that these inputs can be used to augment others and produce refined outputs that have more accurate pixel information per output. For example, spatiotemporal filtering approaches aim to denoise individual images using inputs at the temporal direction, such that pixels polluted by atmospheric conditions can be recovered using other images acquired under a better atmospheric condition. Thus, similar to the M2O category, it exploits spatial and temporal correlations between images, which are the key ingredients to providing accurate and reliable fused images. Examples of applications in this category can be spatiotemporal fusion to achieve an augmented spatial and temporal resolution of an image sequence, satellite video data augmentation, object tracking, and time-series analysis [53], [57], [69], [70], [71], [72], [73], [74].

#### 2.3. General Fusion Approaches in Remote Sensing

Image fusion approaches have been developing rapidly over the years to tackle the fundamental challenges mentioned in Section 2.1 and fuel various RS applications. Methods supporting both M2O and M2M fusion can range from simple arithmetic operations to complex learning-based approaches (i.e., machine learning (ML) and deep learning (DL) methods). The choice of fusion approach mainly depends on the type of integrated images and applications. For example, combining multimodal or multisource images (e.g., RGB image and DSM) to synthesize mixed modality data requires adaptation of differences in image structures, viewing perspectives, resolutions, platforms, etc. Feature-level fusion integrates these diverse modality data into a common feature space prior to fusion. In contrast, unimodal image fusion, due to the modality information, can be directly fused (through linear operations) and often starts from pixel-based operations

(e.g., averaging). The RS applications themselves define the necessary level of accuracy for the fusion algorithm. For example, averaging multitemporal images can be sufficient to achieve moderate accuracy for scene-level information extraction (e.g., crop yields prediction) while performing per-pixel image interpretation (e.g., semantic and panoptic segmentation) for urban scenes with buildings and trees of different scales, weighted averaging or learned-fusion at various levels can be more effective [1], [58], [75], [76]. Understanding these factors is essential to developing an efficient fusion algorithm that can adapt to different images, scenes, and applications.

Generally, fusion methods can be categorized broadly into two groups: classic arithmetic methods (classic method hereafter) and deep learning-based methods (learning-based method hereafter). Classic fusion methods perform basic operations to combine images based on statistical, numerical, and domain transform methods [25], [69], [73], [77], [78], [79], [80], [81]. These methods (mostly based on linear operations) are primarily based on pixel-based and are operated locally (e.g., color transformation from RGB to Intensity-Hue-Saturation (IHS)). Classic fusion methods are not data-driven and do not require complex structures to formulate the fusion problem; thus, they are considered simple to implement and efficient to run. However, the simplicity of the classic methods (such as window-based methods) may have limited performance when processing complex features such as thin structures or repeated textures [25], [82]. Moreover, they assume fixed and known data and noise distribution (e.g., normally distributed) and thus may not adapt well to different data types, scene contents, and distributions. On the other hand, the learning-based methods have recently shown remarkable performances in many image fusion tasks [83], [84]. Examples of these methods include classic statistical ML and DL methods [36], [38], [44], [47], [62], [65], [66], [85], [86], [87], which are found to outperform many of the classic fusion methods due to their ability to extract complex features and patterns from images. Like many existing learning-based approaches, adaptivity and generalization are considered a challenge unless relevant, high-quality input training data is available, which can be very expensive considering the geographically diverse domain of RS images. In this section, we will review both types of techniques under the context of image fusion, and we will focus on describing the most common methods and highlighting their unique advantages and limitations. A summary of Image fusion methods is shown in Table I.

##### 2.3.1. Classic Image Fusion Methods

Classic image fusion methods are defined as mostly non-learning approaches, which are based on statistical, numerical, and domain transformation methods. These methods have a simple structure and do not require complex modeling of the fusion problem or learning from large datasets; thus, they are relatively easy to implement and considered computationally fast, and for a long time, they serve as the mainstream

**Table I. Image fusion methods in remote sensing**

Type of Method	M2O	M2M	Projection and substitution	
<b>Classic Image Fusion Methods (Section 2.3.1)</b>				
<i>Statistical/probability-based methods</i>				
• Statistical filtering				
Basic statistical methods:				
- Average	[37], [88]	[89]	Principal Component Substitution	[122] [123]
- Median	[90]	N/A	Gram-Schmidt	[7] [124]
- Standard deviation	[91]	[92]	• Frequency domain transformation	
- Histogram	[80]	[78]	Fourier transform	[125], [126] [127], [128]
Spatiotemporal filtering:			Discrete Fourier transform	[129], [130] [131]
- Bilateral filter	[93]	[25]	Discrete Cosine transform	[132], [133] [134]
- Spatiotemporal filter	[9]	[58], [69]	• Wavelet transformation	
- Statistical weighting	[43]	[94]	Continuous wavelet transform	[88], [135], [136] [137], [138]
- Image-guided filters	[75], [95]	[96], [97]	Discrete wavelet transform	[79], [139], [140] [141]
• Probabilistic modeling			• Multiresolution analysis (MRA)	
Maximum a-posteriori	[98]	[99]	Curvelet transform	[142] [143], [144]
Bayesian fusion	[100], [101]	[102]	Laplacian pyramid transform	[5], [145] N/A
Dempster-Shafer theory	[54]	[103]		
<i>Prediction-based methods</i>				
• Interpolation/ extrapolation				
Spline	[104]	[105]	Learning-based Image Fusion Methods (Section 2.3.2)	
Bicubic	[106]		<i>Conventional machine learning methods</i>	
Kring	[107]	[108]	Support vector machine	[61], [146] [147]
Regression			Random forest	[148], [149] [150]
Linear regression	[109]	[71], [74], [110]	K-means clustering	[151] [152]
Least squares method	[111]	[74], [112]		
Partial differential equations	[113], [114]	[115]	<i>Deep learning methods</i>	
Optimization			Deep learning for spatial-spectral image fusion	
Iterative methods	[109], [116]	[58]	Convolutional neural networks	[153], [154], [155] [156], [157]
Transformation-based fusion			Residual Network (ResNet)	[47], [158] [159]
• Spatial domain transformation			U-Net	[160], [161] [162]
Color transformation	[117]	[118]	• Deep learning for spatiotemporal image fusion	
- Intensity-Hue-Saturation	[119]	[120], [121]	Recurrent neural networks	[163] [84]
- Brovey transformation			Long short-term memory	[164], [165] [166]
			Gated recurrent unit	[86], [167] [168]
			Attention mechanism	[87], [169] [170]
			Transformers	[171], [172] [173], [174]

methods for practical uses. Most of these methods perform local operations, i.e., informing from neighboring pixels in space or time for noisy removal and purifying the fused results. Figure 5 shows an overview of the classic image fusion, where several DSMs are fused using median filtering and per-pixel operations to update every pixel using information from their neighbors. Depending on how the local pixel information is used, the classic fusion methods are further divided into three categories: 1) Statistical/probability-based methods, which mostly assume linear operations on local pixel values to yield fused results; 2) Prediction-based methods, which use designated functions to explicitly model the mapping between the input pixel values and the fused values. 3) Transformation-based fusion, which fuses information under a transformed space.

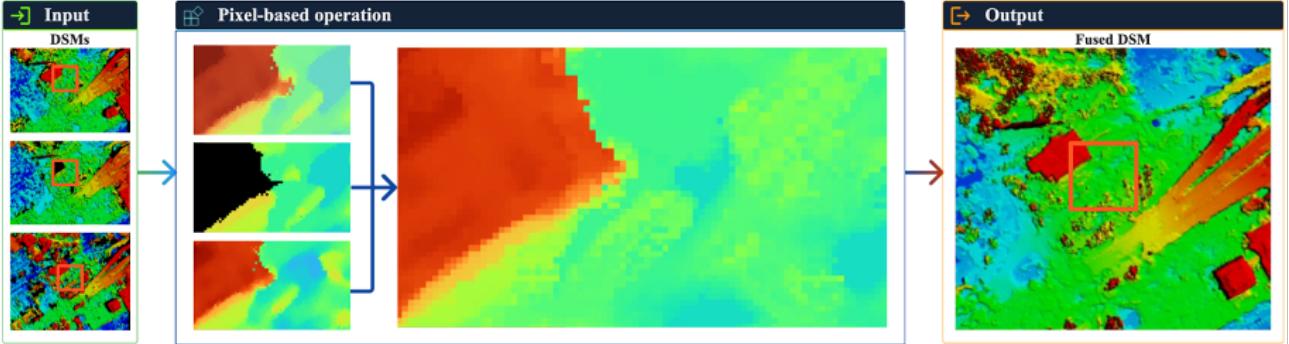
### Statistical/probability-based Methods

Statistical methods explore the statistical properties of pixel values within a local window (spatially and/or temporally). Examples of these methods include using mean, median, standard deviation, and histograms as means to compute the updated pixel values. For a long time, statistical filter-based methods have dominated the fusion approaches [25], [75], [175], [176] due to their simplicity and the ability to easily scale to large-format remote sensing images. Given the input of a stack of images, statistical filters produce smoothed images and suppress unwanted noise during the fusion process, in which the statistical quantities play the role (i.e., mean, median, etc.). The window size was considered a hyperparameter, which may affect the fusion results depending on the noise level of the input. A tangible strategy is to create “soft” weighting for pixels within a window to

reflect the contribution of each pixel in the fusion process. To adaptively determine the weight of each pixel, statistical quantities, such as entropy, can be used as the uncertainty/confidence of the pixels, meaning that the higher the entropy value, the more informative it can be, which should receive a higher weight. Of course, there are also other means of defining the weights, such as the use of an adaptive filtering kernel [90] or through guided filtering [96], [97], as well as filters based on derived statistical hypothesis from the images themselves, such as maximum likelihood, maximum posterior, Dempster-Shafer theory, and Bayesian fusion [2], [54], [98], [100]. Many studies have suggested using image segments to guide the fusion [38], [95], [160], where the window size would change adaptively based on semantic segmentation.

### Prediction-based Methods

Pixel-level prediction considers image fusion as the pixel-level prediction problem, meaning that given a set of input images, the image fusion problem aims to predict pixel values for a new grid as the fused results. Different from the statistical filter method, the pixel-level prediction method allows to predict pixel values of a grid with a size different from the original grid, which extrapolates pixel values at cells where original information is unavailable or at a coarser granularity. Examples of pixel-level prediction methods include super-resolution [73] and spatiotemporal prediction for time-series analysis [72], [105], [108]. For instance, for super-resolution, images are up-sampled to larger sizes, creating a new mesh or gridded format of an image with old values and empty cells, which are to be filled by interpolated values calculated from neighboring pixels using methods such



**Figure 5.** An overview of classic fusion methods workflow performing per-pixel operations to fuse several DSMs into a single DSM (as a case of M2O image fusion), where the value of the central pixel in the orange box is updated using their neighboring pixels in space and time.

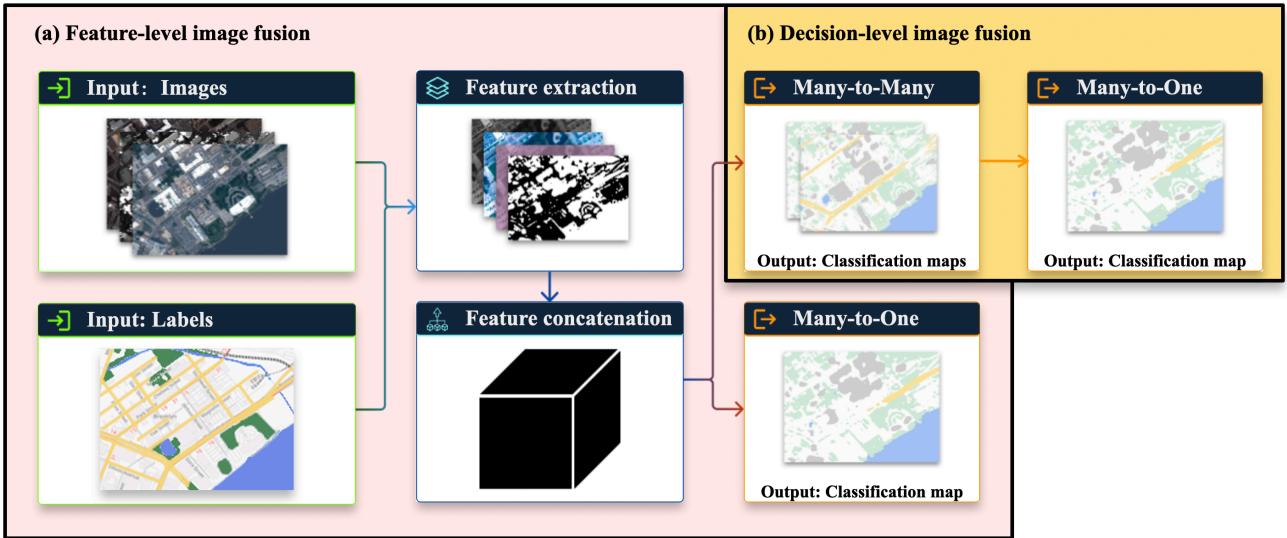
as spline, bicubic, Kring, etc. [46], [68], [73], [177], [178]. Typical interpolation methods operate in local domains, which are fast yet can be easily polluted by data impurities and noise of neighboring pixels. Methods like regression-based approaches aim to utilize a wider data receptive field, which can provide better inference by modeling the relationship between the raw pixel values and the predicted values [78], [179], [110], [180], [109], [74]. Kernel regression serves as one of the key approaches, where the model parameters (e.g., weights and biases) of the regression kernel can be estimated with robust estimation approaches such as robust least squares [111] with sometimes outlier removal algorithms such as random sample consensus (RANSAC) [181]. Postprocessing methods, such as Bayesian regularization [102] and Markov-Random Field [182], can be used to create fusion results under certain priors.

### Transformation-based fusion

Domain-transformation methods decompose images into different domains or representations and perform the data fusion under the transformed domain, such as different color space, frequency, or wavelet domains [79], [79], [183], [184]. These approaches can improve the robustness of image fusion for some tasks. For example, they can be used for spatial resolution enhancement due to their ability to preserve spatial information and fine details in images (e.g., in the frequency domain); at the same time, they are not impacted by spectral distortions. Domain transformation methods can be operated in the spatial, frequency, or other transformed domains, such as wavelets transformation [4], [76], [88], [135], [136], [139], [185]. The spatial domain includes techniques such as color transformation (e.g., Intensity-Hue-Saturation (IHS) and Brovey transformation) or projection and substitution (e.g., Principal Component Substitution (PCS), Gram-Schmidt (GS)) [4], [76], [185]. The frequency domain utilizes techniques related to the Fourier transform and its variants (e.g., Discrete Fourier Transform (DFT) or Cosine transform) [127], [132], whereas wavelet transform includes methods like continuous wavelet transform (CWT), discrete wavelet transform (DWT), as well as their variants [88], [135], [136], [139], [186]. Multi-resolution analysis (MRA) is another

domain transform approach; it is a part of compressive sensing and includes methods such as curvelet transform and Laplacian pyramid transform [4], [76], [185]. These methods follow a similar processing pipeline, including image transformation, coefficient selection, coefficient fusion, and inverse transformation to the image's original domain. For example, pan-sharpening requires converting both the panchromatic and multispectral images to their frequency components using Fourier, wavelet, curvelet, or any multiscale transform-based approach. These coefficients will then be selected and fused prior to an inverse transformation. Finally, the image will be reconstructed to its original form and domain by inverting the transformation. Domain transformation-based fusion has several advantages: 1) identifies surface characteristics such as roughness and repeated patterns or structures [4], [135], [186]; 2) MRA can apply fusion at multiple scales and resolutions and, thus, can effectively extract spatial details, spectral, and textural information; 3) operating in the frequency domain allows more computationally efficient analysis. Nevertheless, these methods can also introduce new artifacts as they are sensitive to noise in the input images, i.e., augmented noises in the frequency domain [135]. They also require domain expert knowledge to determine the suitable approach for a specific RS task, which may not be adaptive to all scene variations and applications.

In sum, most classic fusion methods utilize local operations, which are less effective in utilizing long-range dependencies to handle complex and non-linear inconsistencies of the inputs. In contrast, deep learning-based approaches, due to their ability to learn from examples with complex prediction models, can better address fusing tasks in removing inconsistencies caused by various and very often unknown factors (i.e., uncharacterized errors from derived products or sensor errors). Most classic fusion methods may only address fusion problems whose inconsistencies are due to assumed/relatively simple noise distributions (e.g., Gaussian distribution), whereas they may fall short when processing input images that come with inconsistency due to complex/unknown noise distributions.



**Figure 6.** Examples of the learning-based approaches for classification that apply (a) feature-level and (b) decision-level image fusion to generate single or several classification maps as classic cases of M2M and M2O image fusion, respectively.

### 2.3.2. Learning-based Fusion Methods

Learning how to produce fused results through examples via means of ML and DL has become increasingly attractive for many image fusion tasks. The conventional ML fusion methods utilize algorithms through simpler models such as support vector machine (SVM), random forest, decision trees, etc. [36], [58], [187], which learn either parameters for statistical filtering, or pixel-level prediction by examples in applications such as super-resolution and time-series analysis. In comparison, DL methods apply fusion through more complex models such as convolutional neural networks (CNNs), generative neural networks, recurrent neural networks (RNNs), attention mechanisms, and transformers [169], [188], [86], [87], [153], [158], [189]. Learning-based methods are data-driven, meaning they can extract features and learn patterns directly and automatically from images, enabling image fusion to be performed mostly on the feature and decision levels. The feature-level fusion is where learned features (i.e., feature vectors or feature maps) are extracted separately from images and fused in the following steps, where decision-level fusion combines the predicted images as a postprocessing step (See Figure 6). As is the common problem in ML/DL, learning-based approaches have limited transferability and generalization to new scenes or datasets due to large domain gaps between training and testing images [190], [191], [192], [193], [194]. Oftentimes, there are also limited RS-based training data, which can impact the performance of such models.

#### Conventional machine learning (ML) methods for image fusion

ML methods are prevalent for image fusion tasks [148], [149], [195], [196], which are primarily divided into feature-level and decision-level fusion. The feature-level fusion performs linear or statistical fusion at the feature space (i.e.

hand-crafted feature) in the prediction pipeline [13], [14], [15], [37], [65], [197], the key of which is to learn feature descriptions of multi-modal data into a space where these features can be easily fused without creating conflict of evidence [3], [4], [149], [195]. Feature-level fusion is mostly used in multisource and multimodal data-based fusion, which serves as a means to effectively utilize complementary information for typical tasks such as semantic segmentation, instance segmentation, etc. The decision-level fusion, as what it names, performs fusion on the predicted results, such as from classification/semantic segmentation maps. As compared to feature-level fusion, decision-level fusion is made simpler because the predicted results are already in the same modality with simple fusing strategies such as weight fusion or majority voting, averaging, median, etc. [44], [58], [60], [62], [198]. Of course, the downside of the decision-level fusion is the lack of flexibility to infer earlier processes in feature selection. In sum, both approaches dominated classic fusion literature when handling multi-modal data.

#### Deep learning (DL) methods for image fusion

DL models possess significant advances in learning complex features and pixel-level prediction, as shown in recent literature [84]. While technically, it introduces methodological complexity, it mostly differs from conventional ML by using models that are highly parameterized, whereas the concept of fusion strategies remains, such as the data-level, feature-level, or decision-level fusion. However, unlike ML methods, DL methods can automatically extract low to high-level features without manual interference; however, they may require significant training data and time for learning and processing images. From an application perspective, DL methods are mostly used in two types of applications: spatial-spectral image fusion that focuses on M2O and spatiotemporal fusion that produces predictions through time (M2M). The following subsections

entail relevant works in these two types.

#### a. Deep learning for spatial-spectral image fusion

As mentioned above, DL serves as a more complex mapper in image fusion. One key application in RS image fusion is to create a spatially and spectrally enhanced image from a set of spatially/spectrally low-resolution images (M2O fusion). Earlier success in DL fusion use CNNs and their variants, such as ResNet, U-Net, and AlexNet [47], [153], [154], [158], [199], which can efficiently extract not only spatial details (e.g., 2D shapes), but also valuable contextual information. For example, Pan-sharpen fuses a low-resolution MS image with a high-resolution PAN image into a high-resolution MS image. Typically, the conventional approach explicitly extracts and augments high-frequency details to reserve in the final image while blending spectral information in the low-frequency domain. This process will consequently augment noises from the original image. DL models, on the other hand, can segregate noise from spatial details by learning from examples, such that when fusing the MS and PAN images into the Pan-Sharpened image, unwanted noises are suppressed to minimum. Due to the higher model capacity, DL models can also directly learn, again from examples, predicting pixel-level results directly from MS and PAN images, which learns complex feature representations that best fit the prediction results towards the ground-truth. For instance, in the work of [155], they performed super-resolution on an MS image using super-resolution CNN (SRCNN), then combined the new high-resolution MS with the PAN image using the GS method (a classic fusion method) to get more accurate spatial and spectral representation. A similar line of work can be found in the application of super-resolution [156], [200], [201] and multi-modality fusion based learning as well [202], [203].

#### b. Deep learning for spatiotemporal image fusion

DL also benefits the spatiotemporal image fusion by providing powerful connections for modeling long-range dependencies in both the spatial and the temporal direction, advancing applications such as time-series analysis, change detection, and crop and vegetation dynamics prediction [147], [156], [163]. Well-parameterized models such as recurrent neural networks (RNNs), Long Short-term Memory (LSTM), and Transformers have been widely used in such applications. These types of networks have a built-in structure to store and reuse information with an increasing range of dependencies (from RNN to transformers) [86], [189], [204], [205], [206], [207], and thus they are capable of performing spatiotemporal image fusion considering well-modeled trend for improving the fusion accuracy. For example, typical spatiotemporal filters only consider local neighborhood tensors and are parameterized by simple one-pass linear weights and biases, which lack the capability to model complex phenological patterns. In contrast, DL models with millions of parameters can easily do [84]. The three typical DL models accounting for temporal dependencies, i.e., RNN, LSTM, and Transformer, differ in their structure, mechanism, and memory length due to

their design. Thus, fusion can be applied at different stages depending on the network. For instance, RNNs operate in a loop structure, where important information is stored from the first input image (in what is called a hidden state) and fused with the following image to be processed in an RNN cell, and so forth until the network is set to terminate at a given condition or when all images are consumed. This gives the opportunity to fusion tasks to retain pixel information that is static throughout the time. LSTM and Gated Recurrent Unit (GRU) have similar procedures; however, they have a longer memory to store temporal dependencies and contextual information and use what are called gates to control input, output, and information flow [199]. However, the range of dependencies in the sequential data (e.g., multi-temporal images) is still limited. Recently, researchers explored alternatives that can leverage performance and resources, especially for geographical data, which can be in the form of self-supervision. This led to the adoption of attention mechanisms through transformers, a seminal work that outperforms traditional DL networks in incorporating global dependencies, and thus, it can be naturally used in image fusion tasks [87], [169], [170], [171], [172], [173], [174], [205], [208], [209], [210]. Transformer, in adaptation of the attention mechanism through even longer-range dependencies, shows its ability to exploit global information, which outperforms RNN/LSTM-based structures. Moreover, transformers can also extract features in both the spatial and temporal directions. For example, ViT (Vision Transformer) divides images into multiple smaller patches to allow the attention mechanism (cross-attention) in Transformer to explore global information within the image [170], [171], [172], [173], [174], [205]. Similarly, cross-attention can be applied to divided patches among images captured at different times to allow for temporal information encoding [171]. In sum, the use of transformers for fusion is still growing, and more such methods are expected to be developed soon.

### 3. A META-ANALYSIS OF EXISTING WORKS USING REMOTE SENSING FUSION AS A CORE TECHNIQUE

We aim to review the topic of remote sensing image fusion in a broad sense; thus, we have included a meta-analysis on papers that involve image fusion as a core technique, aiming to analyze the impact of image fusion on the field of RS. The meta-analysis is designed to help inform existing and most recent practices of image fusion methods in the RS field. On the one hand, this helped us to develop our new taxonomy that divides image fusion based on the strategy and the “input-output” formation; on the other hand, it helped us understand the most recent fusion methods that have not been investigated yet in previous review literature. Therefore, for our meta-analysis, we will first present our search method, while in the next Section (Section 4), we will present the results of the statistical analysis showing the development of fusion methods and applications over the years.

#### 3.1. Data Selection and Extraction

In this review, we searched for articles under the remote sensing image fusion topic using the Scopus database (as of 2023), where we set a few queries and searched them in the title, abstract, and keyword sections. The search query includes a few sentences like “Remote sensing” AND “Image fusion”, “Remote sensing” AND “Fusion”, “Fusion” AND “Digital Surface Models” AND “Fusion”, and “Digital Elevation Models” AND “Fusion” (See Table II). We then refine the search by picking only peer-reviewed journal and conference articles and voiding everything else. The initial results led to a total number of 5,926 articles, which we further sorted into journal or conference articles that count for 3,466 and 2,460 articles, respectively.

**Table II. The queries of this review**

Query
“Image fusion” AND “Remote sensing”
“Remote sensing” AND “Fusion”
“Digital Surface Models (DSM)” AND “Fusion”
“Digital Elevation Models (DEM)” AND “Fusion”

### 3.2.Data Information Sources

Our search only includes journal and conference articles that have been peer-reviewed and are well-recognized in the remote sensing society. We summarize these articles in the list shown in Table III. Our list comes from several sources, such as the IEEE Geoscience and Remote Sensing Society (IGARSS), the annals of the International Society for Photogrammetry and Remote Sensing (ISPRS), the Society of Photo-Optical Instrumentation Engineers (SPIE), etc. Table III indicates that the total number of published conference articles is about 2,460, whereas the list of peer-reviewed articles from well-established journals is about 3,466.

**Table III. The titles of the conferences and journals selected for the review with the number of published articles concerning fusion in remote sensing (as of 2023)**

Conference Name	#
International Geoscience and Remote Sensing	1,225
Symposium IGARSS	
Proceedings of SPIE The International Society for Optical Engineering	935
IEEE ISPRS Joint Workshop on Remote Sensing and Data Fusion Over Urban Areas DFUA 2001	68
2nd GRSS ISPRS Joint Workshop on Remote Sensing and Data Fusion Over Urban Areas Urban 2003	64
ISPRS Annals of The Photogrammetry Remote Sensing and Spatial Information Sciences	57
Proceedings International Conference on Image Processing ICIP	27
2011 International Conference on Remote Sensing Environment and Transportation Engineering RSETE 2011 Proceedings	18
Proceedings Applied Imagery Pattern Recognition Workshop	12
IEEE Aerospace Conference Proceedings	12
International Conference on Signal Processing	11

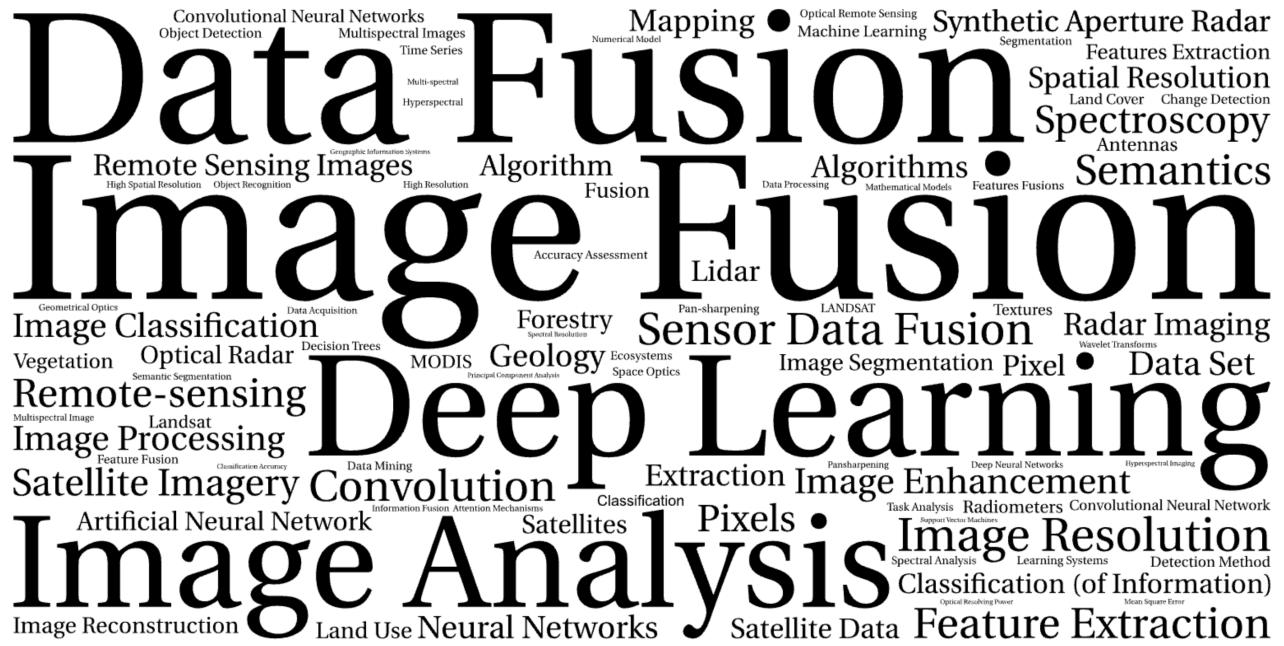
Proceedings ICSP	
IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops	9
Digest International Geoscience and Remote Sensing Symposium IGARSS	8
Conference Record IEEE Instrumentation and Measurement Technology Conference	7
IEEE Conference on Intelligent Transportation Systems Proceedings ITSC	7
<b>Total number of conference publications</b>	<b>2,460</b>
Journal Name	#
Remote Sensing	837
IEEE Transactions in Geoscience and Remote Sensing	619
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing	419
IEEE Geoscience and Remote Sensing Letters	320
International Journal of Remote Sensing	288
Remote Sensing of Environment	166
ISPRS Journal of Photogrammetry and Remote Sensing	142
Journal of Applied Remote Sensing	121
International Journal of Applied Earth Observation and Geoinformation	99
Photogrammetric Engineering and Remote Sensing	62
ISPRS International Journal of Geo-Information	54
Information Fusion	46
Applied Sciences Switzerland	45
International Journal of Image and Data Fusion	34
Journal of Geo-Information Science	31
Remote Sensing Letters	31
Canadian Journal of Remote Sensing	26
IEEE Transactions on Image Processing	22
European Journal of Remote Sensing	18
IEEE Transactions on Instrumentation and Measurement	17
IEEE Transactions on Neural Networks and Learning Systems	17
Computers and Geosciences	13
IEEE Sensors Journal	11
Journal of Environmental Management	10
Photogrammetric Record	9
Pattern Recognition Letters	9
<b>Total number of journal publications</b>	<b>3,466</b>
<b>Total number of conference and journal publications</b>	<b>5,926</b>

## 4. META-ANALYSIS RESULTS

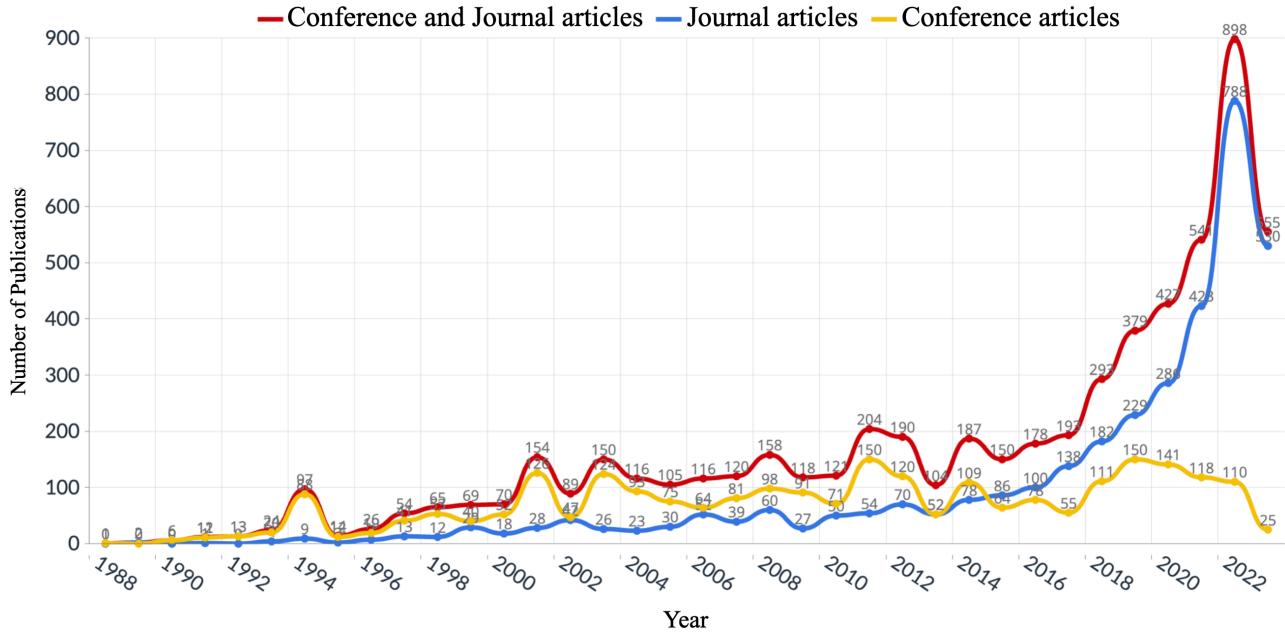
This section presents the meta-analysis results for the selected data from the previous query, as described in Section 3. We will show the word cloud for the keywords from our search, the frequency of annual journal and conference publications, the number of publications per country, and the frequency of publications discussing specific RS applications.

### 4.1. Statistical Analysis for the Keywords in this Review

Figure 7 shows the main keywords as they appear in the selected articles. Words with large fonts indicate their



**Figure 7.** A visualization of the word cloud for the keywords of this review



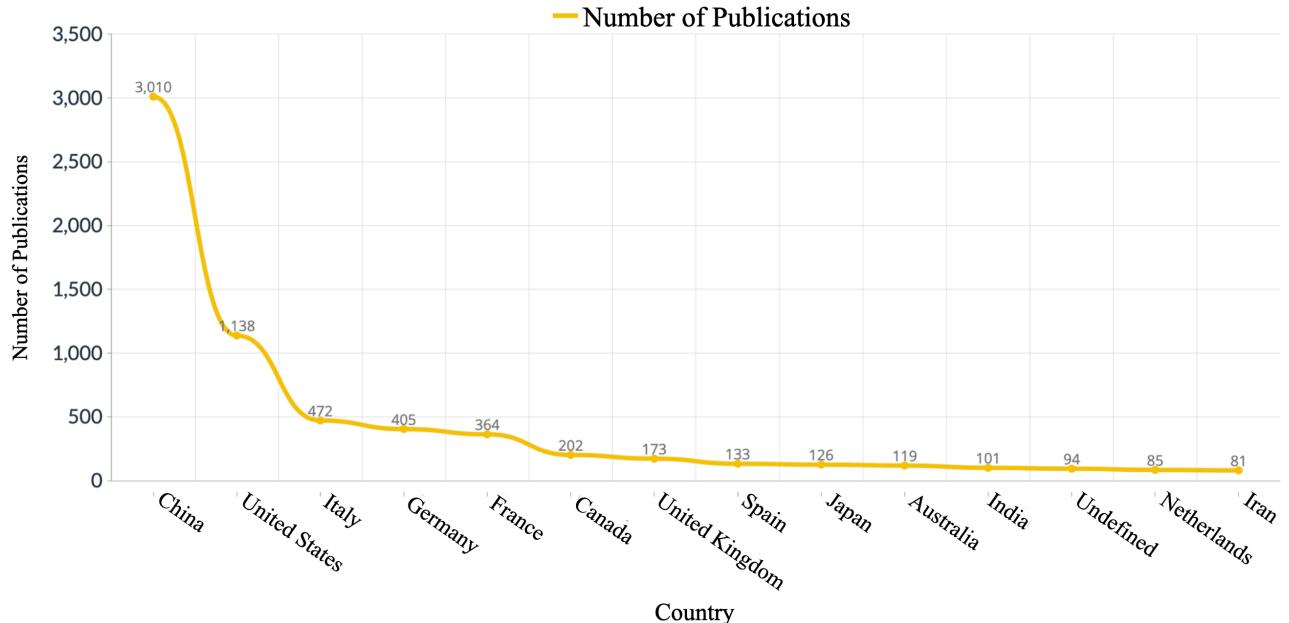
**Figure 8.** The statistical analysis of the annual conference and journal publications.

appearance in high frequency and vice versa for words with smaller fonts. Our results in Figure 7 suggest that the words "Image fusion" and "Data fusion" are the most frequent words to appear. The figure also indicates that the most applications using image fusion are classification, spatial resolution enhancement, and image enhancement, which are applied to images such as satellite images such as MODIS, hyperspectral images, and classification maps.

#### 4.2. Statistical Analysis of the Annual Publications

In Figure 8, we can see the pattern of the number of articles published over the years. We compare the total

number of articles from journals, conferences, and both, as demonstrated by the red, blue, and yellow lines. The figure shows that image fusion has been in the field for decades since the 1980s. It continues to grow gradually over the years from several hundred publications to almost a thousand publications per year, indicating that more research was proposed to enhance the fusion performance in RS (see the red line in Figure 8). The number of conference articles was dominant up to 2015, as seen from the yellow line in Figure 8. After that, the number of journal articles significantly increased over the years, as indicated by the blue line in Figure 8.



**Figure 9.** The statistical analysis of the number of publications per country. Note: the “undefined” in the x-axis refers to unknown origins or countries of the authors.

#### 4.3. Statistical Analysis of the Publications Based on Country

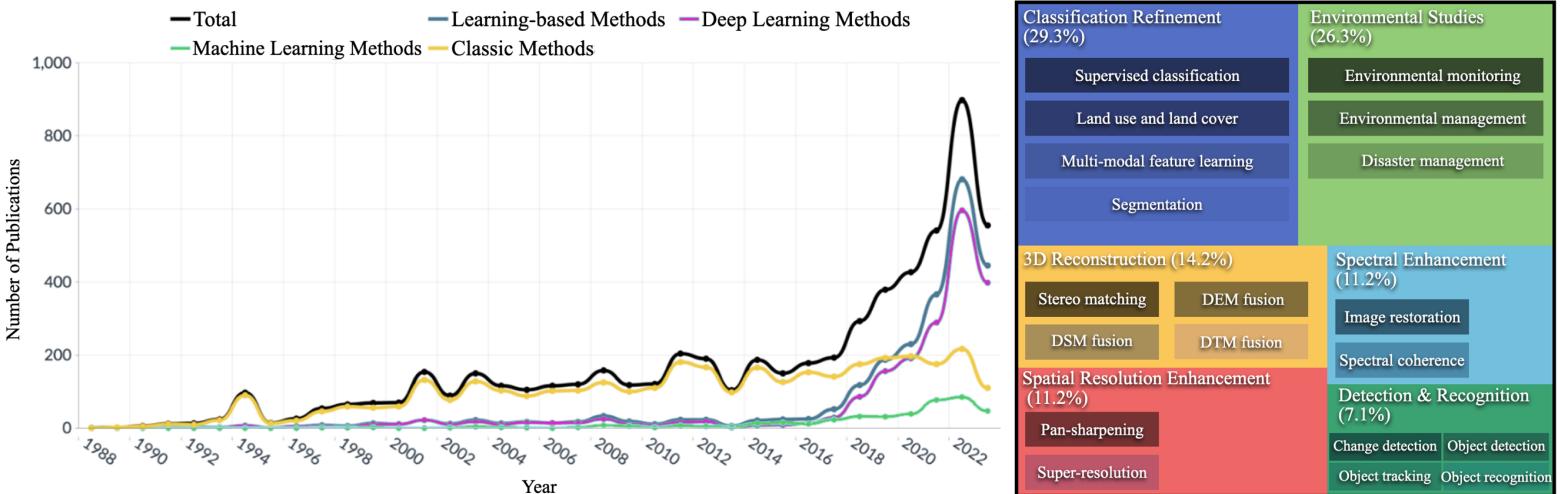
Researchers from 103 countries worldwide have published at least one journal or conference paper related to the image fusion topic. In Figure 9, we provide a plot that summarizes the top 14 countries with the highest number of publications on this subject. The figure shows that China makes at least 50% of the contribution towards this field compared to all other countries, with the number of publications of around ~3,000. The United States of America (USA) also has a high record of publications concerning image fusion in RS, with around ~1,100 articles. Italy, France, Germany, Canada, United Kingdom, Spain, Japan, Australia, and India have a lower publication rate ranging between 100 and 500. The remaining countries still contribute to this field but with a lower number of publications, less than 100.

#### 4.4. Statistical Analysis of the Approaches and Applications of Image Fusion

Image fusion methods, as explored in Section 2.3, are broadly categorized into classic and learning-based methods. Classic methods encompass statistical, numerical, and domain transform techniques [58], [77], [69], [78], [79], [80], [43], while learning-based methods incorporate ML and DL approaches [65], [66], [36], [47], [62], [85], [86], [87]. The progression of the number of publications concerning the development of the classic (yellow line) and learning-based (blue line) image fusion methods over the past 30+ years is shown in Figure 10, which highlights a significant difference. Learning-based methods have shown a remarkable increase, reaching nearly ~1,000 publications in 2022 alone, in contrast to the classic methods, which have maintained a steady output of no more than 200 publications annually. DL methods have demonstrated a dramatic rise in contributions since their establishment in 2015. Two additional lines in Figure 10

illustrate this: the purple line for ML and the green line for DL methods. While both ML and DL methods initially had similar growth in publications, after the year 2015, DL (purple line) publications significantly overtook those of ML (green line) and, by the year 2020, surpassed classic methods as well. This underscores the breakthrough of DL in image fusion post-2015, though classic and ML methods continue to be integral in many applications.

In terms of applications, image fusion is pivotal in Remote Sensing (RS) with diverse uses categorized into six main tasks: 1) Classification refinement, including supervised classification, land use land cover (LULC), and segmentation, and any other application related to refining the accuracy and precision of the classification maps [58], [65], [60], [66], [36], [12], [62], [40], [70], [198], [63]; 2) Detection and recognition, such as building and change detection [64], [179], [56], [45], [211], [212], [123]; 3) Spectral image enhancement for spectral coherence, denoising, filtering, and restoration [25], [213], [116], [22], [214], [69], [215], [23], [72]; 4) Spatial resolution enhancement for tasks like pan-sharpening or super-resolution [6], [47], [86], [216], [68], [217]; 5) 3D reconstruction involving DSM and DEM fusion, and stereo matching [43], [49], [75], [85], [152], [218]; 6) Environmental studies for monitoring and managing natural events or disasters [66], [169], [219], [220], [221], [222]. These applications are summarized in Figure 10. The majority of research focuses on refining classification maps (nearly 30% of total publications), followed by environmental studies (approximately 26.2%), which includes environmental monitoring and disaster management. Spatial resolution enhancement, 3D reconstruction, and spectral image enhancement have comparable publication shares (ranging from 11-14.5%), while detection and recognition hold about 7% of total publications.



**Figure 10.** An overview of image fusion approaches and applications covered in RS shows the publication numbers and percentages over the years. The total number of articles surveyed is 5,926.

## 5. SUMMARY AND FUTURE DIRECTIONS

### 5.1. Summary and Conclusion

In this article, we have reviewed the topic of image fusion for almost all areas in the RS field, i.e., images of different types, fusion techniques, and applications. We focused on providing a broader overview of the fusion problem, where we have introduced a new taxonomy that manifests the two most common categories of image fusion as practiced in the RS field, which includes M2O and M2M image fusion. This taxonomy, although it is simple in notion, is fundamental as it presents fusion as a mapping function that can process different numbers of inputs and outputs. Therefore, the fusion strategies can be extended to augment different applications based on similar concepts and not be limited to specific types of images (i.e., spectral) or applications (i.e., pan-sharpening) as presented in previous literature. Moreover, we have presented a meta-analysis that covered 5,926 peer-reviewed articles to provide a statistical analysis showing the progression of image fusion methods and applications over the years. The meta-analysis indicates a significant growth in image fusion methods, especially with the introduction of DL methods, which has boosted the number of publications from several hundred to almost 1,000 in a year. The statistical results also indicate that classification, environmental studies, and 3D reconstruction are the top three applications that benefit from image fusion, where they account for almost ~15–30% of all publications. On the other hand, applications related to spectral and spatial resolution enhancement account for almost ~11% of all publications. This conclusion leads to rethinking the past literature reviewing RS image fusion, which almost described the fusion problem as an approach for spatial resolution enhancement, specifically for pan-sharpening. Our meta-analysis, on the other hand, indicates that fusion methods can be applied to many types of images and applications. Therefore, we explicitly describe and present the main categories of image fusion methods and their RS

applications to show the unlimited capabilities of such strategies.

### 5.2. Discussion and Future Directions

Image fusion has been a long-investigated topic, yet it has grown to become increasingly complex due to the ever-developing sensory data and diversified inputs (sources, derived products, different modalities). Therefore, existing syntheses focused on their sub-topics due to the divergent applications associated with image fusion, such as image super-resolution, spatiotemporal filtering, and pan-sharpening. In our investigation, we re-synergize these works back to the fundamentals of fusion by categorizing the image fusion through two simple strategies, i.e., many-to-one (M2O) and many-to-many (M2M) image fusion. With an overview, we describe typical strategies and used models related to these general tasks to provide a starting point overview for researchers interested in image fusion approaches and applications, thus fueling future research. Obviously, due to the highly disparate image-fusion tasks, specific considerations should be accounted for when addressing specific image-fusion challenges. At the same time, with the technological developments (both in sensors and algorithms), there is a converging trend of using deeper models for image-fusion tasks. Specific conclusions and recommendations are as follows:

**Knowing the cause of inconsistency:** When preparing images for fusion, i.e., many-to-one (M2O) or many-to-many (M2M) applications, it is essential to understand the tasks and causes of inconsistency for devising the fusion algorithms. Section 2.1 entails various examples of the causes of inconsistency; thus, specific algorithms addressing such challenges in fusion applications should be considered. Moreover, M2O algorithms emphasize information homogeneity among different inputs, while M2M algorithms focus on information compensation and signal-noise separation due to their task of using multiple images to

augment each other. These constitute the core challenges to address, which are being tackled by both classic and various deep learning models (Section 2.2-2.3)

**Multi-modality learning:** The recent advance of learning models suggests a trend towards using multi-modality remote sensing data for fusion, specifically, the use of Radar, optical, and cross-modal input such as GIS (Geographical Information System) data. It is seen that the fundamentals of these tasks are the fusion algorithms in yielding informative remote sensing assets for better interpretation, such as for information retrieval via augmented spatial-temporal resolution, as well as unique 2D/3D signatures (such as thermal, cloud-penetrating Radar, as well as LiDAR) for improved classification, detection and quantification of ground objects. Many of these tasks serve critical environmental and civilian applications.

**Trends and future directions:** Our meta-analysis over 5,926 articles (Section 3-4) yields several conclusions: first, the use of image fusion appears to be highest in publications related to the sub-domain of “remote sensing and geosciences”, where the fusion of images for image interpretation is critical. Second, image fusion is mostly used in applications for image classification and environmental studies, more than 50% of the published works. The rest are shared (roughly equal) among 3D fusion, spatial/spectral enhancement, and detection. Third, there is a converging trend of using deep models for image fusion, while classic methods, although taking a significantly smaller fraction of published works, remain a fair number as compared to previous years, noting that there is still a good use of older methods, as they require much fewer labels. Although with a fairly smaller number, we note that there is increasing work on building foundational models that could interact with various remote sensing tasks, including image fusion and the fusion of cross-modality data, such as the use of images, texts, and audio via using large language models [223], [224], [225].

In sum, our review suggests that image fusion, or the fusion of gridded data, still plays a vital role in this era, where both sensors and models are developed at a rapid pace. In addition to improving classic image fusion algorithms to accommodate newer sensory data and derived data (e.g., DSMs), there has been a paradigm shift from canonical methods to using complex learning and generative models to yield more accurate and comprehensively fused results. Although there has been less focus on already-used algorithms such as pan-sharpening, there is a large room for development when involving multi-modality, non-raster data.

#### ACKNOWLEDGMENT

Rongjun Qin is partially funded by the Office of Naval Research (grant numbers N000142012141 & N000142312670).

#### REFERENCES

- [1] V. Pandit and R. J. Bhiwani, “Image Fusion in Remote Sensing Applications: A Review,” *Int. J. Comput. Appl.*, vol. 120, no. 10, pp. 22–32, Jun. 2015, doi: 10.5120/21263-3846.
- [2] H. Ghassemian, “A Review of Remote Sensing Image Fusion Methods,” *Inf. Fusion*, vol. 32, pp. 75–89, Nov. 2016, doi: 10.1016/j.inffus.2016.03.003.
- [3] P. Ghamsari *et al.*, “Multisource and Multitemporal Data Fusion in Remote Sensing: A Comprehensive Review of the State of the Art,” *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 1, pp. 6–39, Mar. 2019, doi: 10.1109/MGRS.2018.2890023.
- [4] C. Pohl and J. L. Van Genderen, “Review Article Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications,” *Int. J. Remote Sens.*, vol. 19, no. 5, pp. 823–854, Jan. 1998, doi: 10.1080/014311698215748.
- [5] L. Alparone, S. Baronti, B. Aiazzi, and A. Garzelli, “Spatial Methods for Multispectral Pansharpening: Multiresolution Analysis Demystified,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2563–2576, May 2016, doi: 10.1109/TGRS.2015.2503045.
- [6] M. Ehlers, S. Klonus, P. Johan Åstrand, and P. Rosso, “Multisensor image fusion for pansharpening in remote sensing,” *Int. J. Image Data Fusion*, vol. 1, no. 1, pp. 25–45, Mar. 2010, doi: 10.1080/19479830903561985.
- [7] R. Restaino, M. Dalla Mura, G. Vivone, and J. Chanussot, “Context-Adaptive Pansharpening Based on Image Segmentation,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 753–766, Feb. 2017, doi: 10.1109/TGRS.2016.2614367.
- [8] X. Meng, H. Shen, H. Li, L. Zhang, and R. Fu, “Review of the Pansharpening Methods for Remote Sensing Images Based on the Idea of Meta-Analysis: Practical Discussion and Challenges,” *Inf. Fusion*, vol. 46, pp. 102–113, Mar. 2019, doi: 10.1016/j.inffus.2018.05.006.
- [9] F. Gao, J. Masek, M. Schwaller, and F. Hall, “On the Blending of the Landsat and MODIS Surface Reflectance: Predicting Daily Landsat Surface Reflectance,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2207–2218, Aug. 2006, doi: 10.1109/TGRS.2006.872081.
- [10] C. Witharana, D. L. Civco, and T. H. Meyer, “Evaluation of Pansharpening Algorithms in Support of Earth Observation Based Rapid-Mapping Workflows,” *Appl. Geogr.*, vol. 37, pp. 63–87, Feb. 2013, doi: 10.1016/j.apgeog.2012.10.008.
- [11] L. Gómez-Chova, D. Tuia, G. Moser, and G. Camps-Valls, “Multimodal Classification of Remote Sensing Images: A Review and Future Directions,” *Proc. IEEE*, vol. 103, no. 9, pp. 1560–1584, Sep. 2015, doi: 10.1109/JPROC.2015.2449668.
- [12] S. Ji, Z. Zhang, C. Zhang, S. Wei, M. Lu, and Y. Duan, “Learning Discriminative Spatiotemporal Features for Precise Crop Classification from Multi-Temporal Satellite Images,” *Int. J. Remote Sens.*, vol. 41, no. 8, pp. 3162–3174, Apr. 2020, doi: 10.1080/01431161.2019.1699973.
- [13] J. Zhang, “Multi-Source Remote Sensing Data Fusion: Status and Trends,” *Int. J. Image Data Fusion*, vol. 1, no. 1, pp. 5–24, Mar. 2010, doi: 10.1080/19479830903561035.
- [14] Y. Kong, B. Yan, Y. Liu, H. Leung, and X. Peng, “Feature-Level Fusion of Polarized SAR and Optical Images Based on Random Forest and Conditional Random Fields,” *Remote Sens.*, vol. 13, no. 7, p. 1323, Mar. 2021, doi: 10.3390/rs13071323.
- [15] P. Rajah, J. Odindi, and O. Mutanga, “Feature Level Image Fusion of Optical Imagery and Synthetic Aperture Radar (SAR) for Invasive Alien Plant Species Detection and Mapping,” *Remote Sens. Appl. Soc. Environ.*, vol. 10, pp. 198–208, Apr. 2018, doi: 10.1016/j.rsase.2018.04.007.
- [16] H. Shen, X. Meng, and L. Zhang, “An Integrated Framework for the Spatio-Temporal-Spectral Fusion of Remote Sensing Images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7135–7148, Dec. 2016, doi: 10.1109/TGRS.2016.2596290.
- [17] M. Schmitt and X. X. Zhu, “Data Fusion and Remote Sensing: An Ever-Growing Relationship,” *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 4, pp. 6–23, Dec. 2016, doi: 10.1109/MGRS.2016.2561021.
- [18] M. Belguti and A. Stein, “Spatiotemporal Image Fusion in Remote Sensing,” *Remote Sens.*, vol. 11, no. 7, p. 818, Apr. 2019, doi: 10.3390/rs11070818.
- [19] M. Chi, A. Plaza, J. A. Benediktsson, Z. Sun, J. Shen, and Y. Zhu, “Big Data for Remote Sensing: Challenges and Opportunities,” *Proc. IEEE*, vol. 104, no. 11, pp. 2207–2219, Nov. 2016, doi: 10.1109/JPROC.2016.2598228.

- [20] P. Liu, "A Survey of Remote-Sensing Big Data," *Front. Environ. Sci.*, vol. 3, 2015, Accessed: Jan. 14, 2024. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fenvs.2015.00045>
- [21] A. Barsi *et al.*, "Remote Sensing Data Quality Model: From Data Sources to Lifecycle Phases," *Int. J. Image Data Fusion*, vol. 10, no. 4, pp. 280–299, Oct. 2019, doi: 10.1080/19479832.2019.1625977.
- [22] Y. Chen, K. Sun, D. Li, T. Bai, and W. Li, "Improved Relative Radiometric Normalization Method of Remote Sensing Images for Change Detection," *J. Appl. Remote Sens.*, vol. 12, no. 4, p. 045018, Dec. 2018, doi: 10.1117/1.JRS.12.045018.
- [23] M. A. Syariz, B.-Y. Lin, L. G. Denaro, L. M. Jaelani, M. Van Nguyen, and C.-H. Lin, "Spectral-Consistent Relative Radiometric Normalization for Multitemporal Landsat 8 Imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 147, pp. 56–64, Jan. 2019, doi: 10.1016/j.isprsjprs.2018.11.007.
- [24] D. Yuan and C. D. Elvidge, "Comparison of Relative Radiometric Normalization Techniques," *ISPRS J. Photogramm. Remote Sens.*, vol. 51, no. 3, pp. 117–126, Jun. 1996, doi: 10.1016/0924-2716(96)00018-4.
- [25] H. Albanwan and R. Qin, "A Novel Spectrum Enhancement Technique for Multi-Temporal, Multi-Spectral Data Using Spatial-Temporal Filtering," *ISPRS J. Photogramm. Remote Sens.*, vol. 142, pp. 51–63, Aug. 2018, doi: 10.1016/j.isprsjprs.2018.05.020.
- [26] C.-H. Lin, P.-H. Tsai, K.-H. Lai, and J.-Y. Chen, "Cloud Removal from Multitemporal Satellite Images Using Information Cloning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 232–241, Jan. 2013, doi: 10.1109/TGRS.2012.2197682.
- [27] A. Asokan and J. Anitha, "Change Detection Techniques for Remote Sensing Applications: A Survey," *Earth Sci. Inform.*, vol. 12, no. 2, pp. 143–160, Jun. 2019, doi: 10.1007/s12145-019-00380-5.
- [28] M. S. Hamid, N. A. Manap, R. A. Hamzah, and A. F. Kadmin, "Stereo Matching Algorithm Based on Deep Learning: A Survey," *J. King Saud Univ. - Comput. Inf. Sci.*, Aug. 2020, doi: 10.1016/j.jksuci.2020.08.011.
- [29] K. Zhou, X. Meng, and B. Cheng, "Review of Stereo Matching Algorithms Based on Deep Learning," *Comput. Intell. Neurosci.*, vol. 2020, p. e8562323, Mar. 2020, doi: 10.1155/2020/8562323.
- [30] M. Bosch, Z. Kurtz, S. Hagstrom, and M. Brown, "A Multiple View Stereo Benchmark for Satellite Imagery," in *2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, Oct. 2016, pp. 1–9, doi: 10.1109/AIPR.2016.8010543.
- [31] R. Qin, "A Critical Analysis of Satellite Stereo Pairs for Digital Surface Model Generation and a Matching Quality Prediction Model," *ISPRS J. Photogramm. Remote Sens.*, vol. 154, pp. 139–150, Aug. 2019, doi: 10.1016/j.isprsjprs.2019.06.005.
- [32] H. Hirschmüller, "Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Jun. 2005, pp. 807–814 vol. 2, doi: 10.1109/CVPR.2005.56.
- [33] X. Wan, J. G. Liu, S. Li, and H. Yan, "Phase Correlation Decomposition: The Impact of Illumination Variation for Robust Subpixel Remotely Sensed Image Matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6710–6725, Sep. 2019, doi: 10.1109/TGRS.2019.2907933.
- [34] J.-R. Chang and Y.-S. Chen, "Pyramid Stereo Matching Network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5410–5418.
- [35] W. Du, Z. Qin, J. Fan, M. Gao, F. Wang, and B. Abbasi, "An Efficient Approach to Remove Thick Cloud in VNIR Bands of Multi-Temporal Remote Sensing Images," *Remote Sens.*, vol. 11, no. 11, Art. no. 11, Jan. 2019, doi: 10.3390/rs1111284.
- [36] Q. Gao and S. Lim, "A Probabilistic Fusion of a Support Vector Machine and a Joint Sparsity Model for Hyperspectral Imagery Classification," *GIScience Remote Sens.*, vol. 56, no. 8, pp. 1129–1147, Nov. 2019, doi: 10.1080/15481603.2019.1623003.
- [37] X. Kang, S. Li, and J. A. Benediktsson, "Feature Extraction of Hyperspectral Images with Image Fusion and Recursive Filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3742–3752, Jun. 2014, doi: 10.1109/TGRS.2013.2275613.
- [38] X. Ma, S. Hu, S. Liu, J. Fang, and S. Xu, "Remote Sensing Image Fusion Based on Sparse Representation and Guided Filtering," *Electronics*, vol. 8, no. 3, p. 303, Mar. 2019, doi: 10.3390/electronics8030303.
- [39] F. Meng, X. Yang, C. Zhou, and Z. Li, "A Sparse Dictionary Learning-Based Adaptive Patch Inpainting Method for Thick Clouds Removal from High-Spatial Resolution Remote Sensing Imagery," *Sensors*, vol. 17, no. 9, p. 2130, Sep. 2017, doi: 10.3390/s17092130.
- [40] H. T. T. Nguyen, T. M. Doan, E. Tomppo, and R. E. McRoberts, "Land Use/Land Cover Mapping Using Multitemporal Sentinel-2 Imagery and Four Classification Methods—A Case Study from Dak Nong, Vietnam," *Remote Sens.*, vol. 12, no. 9, Art. no. 9, Jan. 2020, doi: 10.3390/rs12091367.
- [41] F. Palsson, J. R. Sveinsson, and M. O. Ulfarsson, "A New Pansharpening Algorithm Based on Total Variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 318–322, Jan. 2014, doi: 10.1109/LGRS.2013.2257669.
- [42] C. Papathanassiou and M. Petrou, "Super Resolution: An Overview," in *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05*, Jul. 2005, pp. 5655–5658, doi: 10.1109/IGARSS.2005.1526059.
- [43] R. Qin, X. Ling, E. M. Farella, and F. Remondino, "Uncertainty-Guided Depth Fusion from Multi-View Satellite Images to Improve the Accuracy in Large-Scale DSM Generation," *Remote Sens.*, vol. 14, no. 6, p. 1309, Mar. 2022, doi: 10.3390/rs14061309.
- [44] G. Vivekananda, R. Swathi, and A. Sujith, "Multi-Temporal Image Analysis for LULC Classification and Change Detection," *Eur. J. Remote Sens.*, vol. 54, no. sup2, pp. 189–199, Mar. 2021, doi: 10.1080/22797254.2020.1771215.
- [45] N. Zerrouki, A. Dairi, F. Harrou, Y. Zerrouki, and Y. Sun, "Efficient Land Desertification Detection Using a Deep Learning-Driven Generative Adversarial Network Approach: A Case Study," *Concurr. Comput. Pract. Exp.*, vol. 34, no. 4, Feb. 2022, doi: 10.1002/cpe.6604.
- [46] A. Belov and A. Y. Denisova, "Remote Sensing Data Fusion Algorithm for Super-Resolution: Multi-Temporal Case," in *Eighth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2020)*, K. Themistocleous, S. Michaelides, V. Ambrosia, D. G. Hadjimitsis, and G. Papadavid, Eds., Paphos, Cyprus: SPIE, Aug. 2020, p. 4, doi: 10.1117/12.2569653.
- [47] M. Kawulok, P. Benecki, S. Piechaczek, K. Hrynczenko, D. Kostrzewa, and J. Nalepa, "Deep Learning for Multiple-Image Super-Resolution," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 6, pp. 1062–1066, Jun. 2020, doi: 10.1109/LGRS.2019.2940483.
- [48] J. Liu, H. Chen, and S. Yang, "Research on Building DSM Fusion Method Based on Adaptive Spline and Target Characteristic Guidance," *Information*, vol. 12, no. 11, p. 467, Nov. 2021, doi: 10.3390/info12110467.
- [49] C. Stucker and K. Schindler, "ResDepth: Learned Residual Stereo Reconstruction," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jun. 2020, pp. 707–716, doi: 10.1109/CVPRW50498.2020.00100.
- [50] G. Vivone, "Robust Band-Dependent Spatial-Detail Approaches for Panchromatic Sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6421–6433, Sep. 2019, doi: 10.1109/TGRS.2019.2906073.
- [51] H. Zhu, X. Tang, J. Xie, W. Song, F. Mo, and X. Gao, "Spatio-Temporal Super-Resolution Reconstruction of Remote-Sensing Images Based on Adaptive Multi-Scale Detail Enhancement," *Sensors*, vol. 18, no. 2, p. 498, Feb. 2018, doi: 10.3390/s18020498.
- [52] O. A. Carvalho Júnior, R. F. Guimarães, A. R. Gillespie, N. C. Silva, and R. A. T. Gomes, "A New Approach to Change Vector Analysis Using Distance and Similarity Measures," *Remote Sens.*, vol. 3, no. 11, Art. no. 11, Nov. 2011, doi: 10.3390/rs3112473.
- [53] R. Chen, V. G. Ferreira, and X. Li, "Detecting Moving Vehicles from Satellite-Based Videos by Tracklet Feature Classification," *Remote Sens.*, vol. 15, no. 1, p. 34, Dec. 2022, doi: 10.3390/rs15010034.
- [54] Y. Lu, J. C. Trinder, and K. Kubik, "Automatic Building Detection Using the Dempster-Shafer Algorithm," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 4, pp. 395–403, Apr. 2006, doi: 10.14358/PERS.72.4.395.
- [55] A. Prakash and R. P. Gupta, "Land-Use Mapping and Change Detection in a Coal Mining Area - a Case Study in the Jharia Coalfield, India," *Int. J. Remote Sens.*, vol. 19, no. 3, pp. 391–410, Jan. 1998, doi: 10.1080/014311698216053.
- [56] C. Wu, L. Zhang, and L. Zhang, "A Scene Change Detection Framework for Multi-Temporal Very High Resolution Remote Sensing Images," *Signal Process.*, vol. 124, pp. 184–197, Jul. 2016, doi: 10.1016/j.sigpro.2015.09.020.
- [57] Z. Zhang, C. Wang, J. Song, and Y. Xu, "Object Tracking Based on Satellite Videos: A Literature Review," *Remote Sens.*, vol. 14, no. 15, p. 3674, Jul. 2022, doi: 10.3390/rs14153674.
- [58] H. Albanwan, R. Qin, X. Lu, M. Li, D. Liu, and J.-M. Guldmann, "3D Iterative Spatiotemporal Filtering for Classification of Multitemporal Satellite Data Sets," *Photogramm. Eng. Remote Sens.*, vol. 86, no. 1, pp. 23–31, Jan. 2020, doi: 10.14358/PERS.86.1.23.

- [59] L. Bruzzone and L. Carlin, "A Multilevel Context-Based System for Classification of Very High Spatial Resolution Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2587–2600, Sep. 2006, doi: 10.1109/TGRS.2006.875360.
- [60] G. Cui, Z. Lv, G. Li, J. Atli Benediktsson, and Y. Lu, "Refining Land Cover Classification Maps Based on Dual-Adaptive Majority Voting Strategy for Very High Resolution Remote Sensing Images," *Remote Sens.*, vol. 10, no. 8, p. 1238, Aug. 2018, doi: 10.3390/rs10081238.
- [61] P. Du, J. Xia, J. Chanussot, and X. He, "Hyperspectral Remote Sensing Image Classification Based on the Integration of Support Vector Machine and Random Forest," in *2012 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2012, pp. 174–177. doi: 10.1109/IGARSS.2012.6351609.
- [62] R. Lucas, A. Rowlands, A. Brown, S. Keyworth, and P. Bunting, "Rule-Based Classification of Multi-Temporal Satellite Imagery for Habitat and Agricultural Land Cover Mapping," *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 3, pp. 165–185, Aug. 2007, doi: 10.1016/j.isprsjprs.2007.03.003.
- [63] W. Zhou, A. Troy, and M. Grove, "Object-Based Land Cover Classification and Change Analysis in the Baltimore Metropolitan Area Using Multitemporal High Resolution Remote Sensing Data," *Sensors*, vol. 8, Mar. 2008, doi: 10.3390/s8031613.
- [64] F. Bovolo and L. Bruzzone, "A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in the Polar Domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007, doi: 10.1109/TGRS.2006.885408.
- [65] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. L. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-Based Framework for Multitemporal and Multisource Remote Sensing Data Classification and Change Detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1822–1835, Jun. 2008, doi: 10.1109/TGRS.2008.916201.
- [66] A. Erkkilä and R. Kalliola, "Patterns and Dynamics of Coastal Waters in Multi-Temporal Satellite Images: Support to Water Quality Monitoring in the Archipelago Sea, Finland," *Estuar. Coast. Shelf Sci.*, vol. 60, no. 2, pp. 165–177, Jun. 2004, doi: 10.1016/j.ecss.2003.11.024.
- [67] A. Gulácsi and F. Kovács, "Drought Monitoring with Spectral Indices Calculated from Modis Satellite Images in Hungary," *J. Environ. Geogr.*, vol. 8, no. 3–4, pp. 11–20, Dec. 2015, doi: 10.1515/jengeo-2015-0008.
- [68] I. Yanovsky and J. Qin, "Spatio-Temporal Super-Resolution Reconstruction of Remote Sensing Data," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, Jul. 2021, pp. 2907–2910. doi: 10.1109/IGARSS47720.2021.9553433.
- [69] W. Gan, H. Albanwan, and R. Qin, "Radiometric Normalization of Multitemporal Landsat and Sentinel-2 Images Using a Reference MODIS Product Through Spatiotemporal Filtering," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 4000–4013, 2021, doi: 10.1109/JSTARS.2021.3069855.
- [70] C. M. Viana, I. Girão, and J. Rocha, "Long-Term Satellite Image Time-Series for Land Use/Land Cover Change Detection Using Refined Open Source Data in a Rural Region," *Remote Sens.*, vol. 11, no. 9, p. 1104, May 2019, doi: 10.3390/rs11091104.
- [71] W. Wu *et al.*, "A Long Time-Series Radiometric Normalization Method for Landsat Images," *Sensors*, vol. 18, no. 12, p. 4505, Dec. 2018, doi: 10.3390/s18124505.
- [72] W. Wu, L. Ge, J. Luo, R. Huan, and Y. Yang, "A Spectral-Temporal Patch-Based Missing Area Reconstruction for Time-Series Images," *Remote Sens.*, vol. 10, no. 10, Art. no. 10, Oct. 2018, doi: 10.3390/rs10101560.
- [73] Q. Xie, G. Yao, and P. Liu, "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.
- [74] C. Zeng, H. Shen, and L. Zhang, "Recovering Missing Pixels for Landsat ETM+ SLC-Off Imagery Using Multi-Temporal Regression Analysis and a Regularization Method," *Remote Sens. Environ.*, vol. 131, pp. 182–194, Apr. 2013, doi: 10.1016/j.rse.2012.12.012.
- [75] H. Albanwan and R. Qin, "Enhancement of Depth Map by Fusion Using Adaptive and Semantic-Guided Spatiotemporal Filtering," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. V-3–2020, pp. 227–232, Aug. 2020, doi: 10.5194/isprs-annals-V-3-2020-227-2020.
- [76] F. Dadrass Javan, F. Samadzadegan, S. Mehravar, A. Toosi, R. Khatami, and A. Stein, "A Review of Image Fusion Techniques for Pan-Sharpening of High-Resolution Satellite Imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 171, pp. 101–117, Jan. 2021, doi: 10.1016/j.isprsjprs.2020.11.001.
- [77] N. Amrani, J. Serra-Sagristà, P. Peter, and J. Weickert, "Diffusion-Based Inpainting for Coding Remote-Sensing Data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1203–1207, Aug. 2017, doi: 10.1109/LGRS.2017.2702106.
- [78] E. H. Helmer and B. Ruefenacht, "Cloud-Free Satellite Image Mosaics with Regression Trees and Histogram Matching," *Photogramm. Eng. Remote Sens. Vol 71 No 9 Sept. 2005 1079–1089*, 2005, Accessed: Feb. 08, 2022. [Online]. Available: <http://www.fs.usda.gov/treesearch/pubs/30086>
- [79] C. W. Lacewell, M. Gebril, R. Buaba, and A. Homaiifar, "Optimization of Image Fusion using Genetic Algorithms and Discrete Wavelet Transform," in *Proceedings of the IEEE 2010 National Aerospace Electronics Conference*, Jul. 2010, pp. 116–121. doi: 10.1109/NAECON.2010.5712933.
- [80] K. Nagamani, K. Divya, K. Sujatha, K. R. Bonagiri, G. B. Kande, and P. S. S. Kumar, "Adaptive Histogram Equalization of Wavelet Sub Bands for the Enhancement of Contrast in Aerial Images," *Mater. Today Proc.*, Nov. 2021, doi: 10.1016/j.matpr.2021.10.297.
- [81] C. Yu, L. Chen, L. Su, M. Fan, and S. Li, "Kriging Interpolation Method and Its Application in Retrieval of MODIS Aerosol Optical Depth," in *2011 19th International Conference on Geoinformatics*, Jun. 2011, pp. 1–6. doi: 10.1109/GeoInformatics.2011.5981052.
- [82] Y. Han, W. Liu, X. Huang, S. Wang, and R. Qin, "Stereo Dense Image Matching by Adaptive Fusion of Multiple-Window Matching Results," *Remote Sens.*, vol. 12, no. 19, p. 3138, Sep. 2020, doi: 10.3390/rs12193138.
- [83] A. E. Maxwell, T. A. Warner, and F. Fang, "Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review," *Int. J. Remote Sens.*, vol. 39, no. 9, pp. 2784–2817, May 2018, doi: 10.1080/01431161.2018.1433343.
- [84] Z. Wang, Y. Ma, and Y. Zhang, "Review of Pixel-Level Remote Sensing Image Fusion Based on Deep Learning," *Inf. Fusion*, vol. 90, pp. 36–58, Feb. 2023, doi: 10.1016/j.inffus.2022.09.008.
- [85] H. Papasaika, E. Koklopoulou, E. Baltsavias, and D. Kressner, "Fusion of Digital Elevation Models Using Sparse Representations," Jan. 2011, pp. 171–184.
- [86] M. R. Arefin *et al.*, "Multi-Image Super-Resolution for Remote Sensing Using Deep Recurrent Networks," presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 206–207. Accessed: Dec. 24, 2023. [Online]. Available: [https://openaccess.thecvf.com/content\\_CVPRW\\_2020/html/w11/Arefin\\_Multi-Image\\_Super-Resolution\\_for\\_Remote\\_Sensing\\_Using\\_Deep\\_Recurrent\\_Networks\\_CVPR\\_W\\_2020\\_paper.html](https://openaccess.thecvf.com/content_CVPRW_2020/html/w11/Arefin_Multi-Image_Super-Resolution_for_Remote_Sensing_Using_Deep_Recurrent_Networks_CVPR_W_2020_paper.html)
- [87] 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, Art. no. 14, 2020, doi: 10.3390/rs12142207.
- [88] J. Wu, H. Huang, Y. Qiu, H. Wu, J. Tian, and J. Liu, "Remote Sensing Image Fusion Based on Average Gradient of Wavelet Transform," in *IEEE International Conference Mechatronics and Automation, 2005*, Jul. 2005, pp. 1817–1821 Vol. 4. doi: 10.1109/ICMA.2005.1626836.
- [89] H. Li, X. Cui, M. Li, J. Deng, and S. Chen, "Adaptive Speckle Filter for Multi-Temporal PolSAR Image with Multi-Dimensional Information Fusion," *Remote Sens.*, vol. 15, no. 14, Art. no. 14, Jan. 2023, doi: 10.3390/rs15143679.
- [90] R. Qin, "Automated 3D Recovery from Very High Resolution Multi-View Satellite Images," presented at the ASPRS Conference (IGTF) 2017, Baltimore, Maryland, USA: ASPRS Conference (IGTF) 2017, 2017.
- [91] D. Yıldırım and O. Güngör, "A Novel Image Fusion Method Using IKONOS Satellite Images," *J. Geod. Geoinformation*, vol. 1, no. 1, pp. 27–34, 2012, doi: 10.9733/jgg.170512.1.
- [92] K. Ezimand, M. Chahardoli, M. Azadbakht, and A. A. Matkan, "Spatiotemporal Analysis of Land Surface Temperature Using Multi-Temporal and Multi-Sensor Image Fusion Techniques," *Sustain. Cities Soc.*, vol. 64, p. 102508, Jan. 2021, doi: 10.1016/j.scs.2020.102508.
- [93] K. Kotwal and S. Chaudhuri, "Visualization of Hyperspectral Images Using Bilateral Filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2308–2316, May 2010, doi: 10.1109/TGRS.2009.2037950.
- [94] T. Lu, S. Li, and W. Fu, "Fusion Based Seamless Mosaic for Remote Sensing Images," *Sens. Imaging*, vol. 15, no. 1, p. 101, Nov. 2014, doi: 10.1007/s11220-014-0101-0.
- [95] H. Albanwan and R. Qin, "An Adaptive and Image-guided Fusion for Stereo Satellite Image Derived Digital Surface Models," *J. Geod.*

- Geoinformation Sci.*, vol. 5, no. 4, pp. 1–9, Dec. 2022, doi: 10.11947/j.JGGS.2022.0401.
- [96] Q. Guo, J. Zhang, and Y. Zhang, “Multitemporal Hyperspectral Images Change Detection Based on Joint Unmixing and Information Coguidance Strategy,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9633–9645, Nov. 2021, doi: 10.1109/TGRS.2020.3045799.
- [97] Z. Huang, L. Fang, and S. Li, “Subpixel-Pixel-Superpixel Guided Fusion for Hyperspectral Anomaly Detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 5998–6007, Sep. 2020, doi: 10.1109/TGRS.2019.2961703.
- [98] H. Shen and L. Zhang, “A MAP-Based Algorithm for Destriping and Inpainting of Remotely Sensed Images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, May 2009, doi: 10.1109/TGRS.2008.2005780.
- [99] D. He, Y. Zhong, and L. Zhang, “Spectral–Spatial–Temporal MAP-Based Sub-Pixel Mapping for Land-Cover Change Detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 1696–1717, Mar. 2020, doi: 10.1109/TGRS.2019.2947708.
- [100] J. Xue, Y. Leung, and T. Fung, “A Bayesian Data Fusion Approach to Spatio-Temporal Fusion of Remotely Sensed Images,” *Remote Sens.*, vol. 9, no. 12, p. 1310, Dec. 2017, doi: 10.3390/rs9121310.
- [101] Y. Zhang, S. De Backer, and P. Scheunders, “Noise-Resistant Wavelet-Based Bayesian Fusion of Multispectral and Hyperspectral Images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3834–3843, Nov. 2009, doi: 10.1109/TGRS.2009.2017737.
- [102] C. He, Z. Zhang, D. Xiong, J. Du, and M. Liao, “Spatio-Temporal Series Remote Sensing Image Prediction Based on Multi-Dictionary Bayesian Fusion,” *ISPRS Int. J. Geo-Inf.*, vol. 6, no. 11, Art. no. 11, Nov. 2017, doi: 10.3390/ijgi6110374.
- [103] L. Cayuela, J. D. Golicher, J. S. Rey, and J. M. R. Benayas, “Classification of a Complex Landscape Using Dempster–Shafer Theory of Evidence,” *Int. J. Remote Sens.*, vol. 27, no. 10, pp. 1951–1971, May 2006, doi: 10.1080/01431160500181788.
- [104] J. Anger, T. Ehret, C. de Franchis, and G. Facciolo, “Fast and Accurate Multi-Frame Super-Resolution of Satellite Images,” *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. V-1–2020, pp. 57–64, Aug. 2020, doi: 10.5194/isprs-annals-V-1-2020-57-2020.
- [105] Z. Pan, Y. Hu, and B. Cao, “Construction of Smooth Daily Remote Sensing Time Series Data: A Higher Spatiotemporal Resolution Perspective,” *Open Geospatial Data Softw. Stand.*, vol. 2, no. 1, p. 25, Oct. 2017, doi: 10.1186/s40965-017-0038-z.
- [106] W. Jing *et al.*, “A Rigorously-Incremental Spatiotemporal Data Fusion Method for Fusing Remote Sensing Images,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 16, pp. 6723–6738, 2023, doi: 10.1109/JSTARS.2023.3296468.
- [107] Y. Li, M. Li, Z. Liu, and C. Li, “Combining Kriging Interpolation to Improve the Accuracy of Forest Aboveground Biomass Estimation Using Remote Sensing Data,” *IEEE Access*, vol. 8, pp. 128124–128139, 2020, doi: 10.1109/ACCESS.2020.3008686.
- [108] D. Hu, H. Shu, H. Hu, and J. Xu, “Spatiotemporal Regression Kriging to Predict Precipitation Using Time-Series MODIS Data,” *Clust. Comput.*, vol. 20, no. 1, pp. 347–357, Mar. 2017, doi: 10.1007/s10586-016-0708-0.
- [109] G. Vivone, R. Restaino, and J. Chanussot, “Full Scale Regression-Based Injection Coefficients for Panchromatic Sharpening,” *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3418–3431, Jul. 2018, doi: 10.1109/TIP.2018.2819501.
- [110] J. Zhang, M. K. Clayton, and P. A. Townsend, “Missing Data and Regression Models for Spatial Images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1574–1582, Mar. 2015, doi: 10.1109/TGRS.2014.2345513.
- [111] D. Huang, P. Yang, J. Li, and C. Ma, “Remote Sensing Image Fusion Using Bidimensional Empirical Mode Decomposition and the Least Squares Theory,” *J. Comput. Commun.*, vol. 05, no. 12, pp. 35–48, 2017, doi: 10.4236/jcc.2017.512004.
- [112] A. A. Nielsen, “Multiset Canonical Correlations Analysis and Multispectral, Truly Multitemporal Remote Sensing Data,” *IEEE Trans. Image Process.*, vol. 11, no. 3, pp. 293–305, Mar. 2002, doi: 10.1109/83.988962.
- [113] D. P. Bavirisetti, G. Xiao, and G. Liu, “Multi-Sensor Image Fusion Based on Fourth Order Partial Differential Equations,” in *2017 20th International Conference on Information Fusion (Fusion)*, Jul. 2017, pp. 1–9, doi: 10.23919/ICIF.2017.8009719.
- [114] P. Liu, F. Huang, G. Li, and Z. Liu, “Remote-Sensing Image Denoising Using Partial Differential Equations and Auxiliary Images as Priors,” *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 3, pp. 358–362, May 2012, doi: 10.1109/LGRS.2011.2168598.
- [115] A. Nazari, A. Zehtabian, M. Gribaldo, and H. Ghassemian, “Remotely sensed image restoration using partial differential equations and watershed transformation,” presented at the *Seventh International Conference on Machine Vision (ICMV 2014)*, A. Verikas, B. Vuksanovic, P. Radeva, and J. Zhou, Eds., Milan, Italy, Feb. 2015, p. 944520. doi: 10.1117/12.2181817.
- [116] M. J. Carty and A. A. Nielsen, “Automatic Radiometric Normalization of Multitemporal Satellite Imagery with the Iteratively Re-Weighted Mad Transformation,” *Remote Sens. Environ.*, vol. 112, no. 3, pp. 1025–1036, Mar. 2008, doi: 10.1016/j.rse.2007.07.013.
- [117] J. Chavez, S. C. Sides, and J. A. Anderson, “Comparison of Three Different Methods to Merge Multiresolution and Multispectral Data: Landsat Tm and Spot Panchromatic,” *AAPG Bull. Am. Assoc. Pet. Geol. USA*, vol. 74:6, Art. no. CONF-900702-, Jun. 1990, Accessed: Jan. 04, 2024. [Online]. Available: <https://www.osti.gov/biblio/6165108>
- [118] Y. Leung, J. Liu, and J. Zhang, “An Improved Adaptive Intensity–Hue–Saturation Method for the Fusion of Remote Sensing Images,” *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 985–989, May 2014, doi: 10.1109/LGRS.2013.2284282.
- [119] R. Gharbia, A. H. El Baz, A. E. Hassanien, and M. F. Tolba, “Remote Sensing Image Fusion Approach Based on Brovey and Wavelets Transforms,” in *Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014*, P. Kömer, A. Abraham, and V. Snásel, Eds., in *Advances in Intelligent Systems and Computing*. Cham: Springer International Publishing, 2014, pp. 311–321. doi: 10.1007/978-3-319-08156-4\_31.
- [120] J. Nichol and M. S. Wong, “Satellite Remote Sensing for Detailed Landslide Inventories Using Change Detection and Image Fusion,” *Int. J. Remote Sens.*, vol. 26, no. 9, pp. 1913–1926, May 2005, doi: 10.1080/01431160512331314047.
- [121] B. Yang, M. Kim, and M. Madden, “Assessing Optimal Image Fusion Methods for Very High Spatial Resolution Satellite Images to Support Coastal Monitoring,” *GIScience Remote Sens.*, vol. 49, no. 5, pp. 687–710, Sep. 2012, doi: 10.2747/1548-1603.49.5.687.
- [122] C. Chen, X. He, B. Guo, X. Zhao, and Y. Chu, “A Pixel-Level Fusion Method for Multi-Source Optical Remote Sensing Image Combining the Principal Component Analysis and Curvelet Transform,” *Earth Sci. Inform.*, vol. 13, no. 4, pp. 1005–1013, Dec. 2020, doi: 10.1007/s12145-020-00472-7.
- [123] S. Achour, M. Chikr Elmezouar, N. Taleb, K. Kpalma, and J. Ronzin, “A PCA-PD Fusion Method for Change Detection in Remote Sensing Multi Temporal Images,” *Geocarto Int.*, vol. 37, no. 1, pp. 196–213, Jan. 2022, doi: 10.1080/10106049.2020.1713228.
- [124] M. Dalla Mura, G. Vivone, R. Restaino, P. Addesso, and J. Chanussot, “Global and Local Gram-Schmidt Methods for Hyperspectral Pansharpening,” in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, Italy: IEEE, Jul. 2015, pp. 37–40. doi: 10.1109/IGARSS.2015.7325691.
- [125] Z. Chen, P. Zhang, Y. Zhang, X. Xu, L. Ji, and H. Tang, “Thick Cloud Removal in Multi-Temporal Remote Sensing Images via Frequency Spectrum-Modulated Tensor Completion,” *Remote Sens.*, vol. 15, no. 5, p. 1230, Feb. 2023, doi: 10.3390/rs15051230.
- [126] Q. Zeng, H. Qin, X. Yan, and T. Yang, “Fourier Domain Anomaly Detection and Spectral Fusion for Stripe Noise Removal of TIR Imagery,” *Remote Sens.*, vol. 12, no. 22, p. 3714, Nov. 2020, doi: 10.3390/rs12223714.
- [127] J. G. Denipote and M. S. V. Paiva, “A Fourier Transform-based Approach to Fusion High Spatial Resolution Remote Sensing Images,” in *2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing*, Dec. 2008, pp. 179–186. doi: 10.1109/ICVGIP.2008.108.
- [128] M. Lillo-Saavedra, C. Gonzalo, A. Arquero, and E. Martinez, “Fusion of Multispectral and Panchromatic Satellite Sensor Imagery Based on Tailored Filtering in the Fourier Domain,” *Int. J. Remote Sens.*, vol. 26, no. 6, pp. 1263–1268, Mar. 2005, doi: 10.1080/01431160412331330239.
- [129] A. Akoguz, B. Kurt, and S. K. Pinar, “Pansharpening of Multispectral Images Using Filtering in Fourier Domain,” presented at the *SPIE Remote Sensing*, L. Bruzzone, J. A. Benediktsson, and F. Bovolo, Eds., Amsterdam, Netherlands, Oct. 2014, p. 92441X. doi: 10.1117/12.2067255.
- [130] C. Zhou, G. Zhang, Z. Yang, and J. Zhou, “A Novel Image Registration Algorithm Using Wavelet Transform and Matrix-Multiply Discrete Fourier Transform,” *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2020.3031335.
- [131] M. Cui, S. Xue, and Y. Huang, “Land Cover Classification Base on Fourier Analysis,” in *2010 Fourth International Conference on Genetic*

- and Evolutionary Computing*, Shenzhen: IEEE, Dec. 2010, pp. 390–393. doi: 10.1109/ICGEC.2010.103.
- [132] E. Alhatami, U. A. Bhatti, M. Huang, and S. L. Feng, “Image Fusion Based on Discrete Cosine Transform with High Compression,” in *2022 7th International Conference on Signal and Image Processing (ICSIP)*, Jul. 2022, pp. 606–610. doi: 10.1109/ICSIP55141.2022.9887300.
- [133] H. Yu, Z. Lin, G. L. Chai, C. H. Yan, and N. Raman, “A DCT-based Change Detection Method for Multi-Temporal SAR Images,” in *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2008, p. III-166-III-169. doi: 10.1109/IGARSS.2008.4779309.
- [134] K. Yang, Y. Luo, M. Li, S. Zhong, Q. Liu, and X. Li, “Reconstruction of Sentinel-2 Image Time Series Using Google Earth Engine,” *Remote Sens.*, vol. 14, no. 17, Art. no. 17, Jan. 2022, doi: 10.3390/rs14174395.
- [135] K. Amolins, Y. Zhang, and P. Dare, “Wavelet Based Image Fusion Techniques — an Introduction, Review and Comparison,” *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 4, pp. 249–263, Sep. 2007, doi: 10.1016/j.isprsjprs.2007.05.009.
- [136] J. Cheng, H. Liu, T. Liu, F. Wang, and H. Li, “Remote Sensing Image Fusion Via Wavelet Transform and Sparse Representation,” *ISPRS J. Photogramm. Remote Sens.*, vol. 104, pp. 158–173, Jun. 2015, doi: 10.1016/j.isprsjprs.2015.02.015.
- [137] S. Muhammad, M. Wachowicz, and L. M. T. de Carvalho, “Evaluation of Wavelet Transform Algorithms for Multi-Resolution Image Fusion,” in *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002. (IEEE Cat.No.02EX5997)*, Jul. 2002, pp. 1573–1580 vol.2. doi: 10.1109/ICIF.2002.1021004.
- [138] J. Zhang *et al.*, “Detection of Canopy Chlorophyll Content of Corn Based on Continuous Wavelet Transform Analysis,” *Remote Sens.*, vol. 12, no. 17, Art. no. 17, Jan. 2020, doi: 10.3390/rs12172741.
- [139] J. Jinju, N. Santhi, K. Ramar, and B. Sathyam Bama, “Spatial Frequency Discrete Wavelet Transform Image Fusion Technique for Remote Sensing Applications,” *Eng. Sci. Technol. Int. J.*, vol. 22, no. 3, pp. 715–726, Jun. 2019, doi: 10.1016/j.estch.2019.01.004.
- [140] H. Zhuang, K. Deng, Y. Yu, and H. Fan, “An Approach Based on Discrete Wavelet Transform to Unsupervised Change Detection in Multispectral Images,” *Int. J. Remote Sens.*, vol. 38, no. 17, pp. 4914–4930, Sep. 2017, doi: 10.1080/01431161.2017.1331475.
- [141] S. Cui and M. Datcu, “Statistical Wavelet Subband Modeling for Multi-Temporal SAR Change Detection,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 5, no. 4, pp. 1095–1109, Aug. 2012, doi: 10.1109/JSTARS.2012.2200655.
- [142] M. Choi, R. Y. Kim, M.-R. Nam, and H. O. Kim, “Fusion of Multispectral and Panchromatic Satellite Images Using the Curvelet Transform,” *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 2, pp. 136–140, Apr. 2005, doi: 10.1109/LGRS.2005.845313.
- [143] F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, “Remote Sensing Image Fusion Using the Curvelet Transform,” *Inf. Fusion*, vol. 8, no. 2, pp. 143–156, Apr. 2007, doi: 10.1016/j.inffus.2006.02.001.
- [144] C. V. Rao, J. M. Rao, A. S. Kumar, D. S. Jain, and V. K. Dadhwal, “Satellite Image Fusion Using Fast Discrete Curvelet Transforms,” presented at the 2014 IEEE International Advance Computing Conference (IACC), Feb. 2014, pp. 952–957. doi: 10.1109/IAdCC.2014.6779451.
- [145] R. Su, L. Sun, and J. Wang, “Comparison of Remote Sensing Image Fusion Rules Based on the Laplacian Pyramid Transformation,” in *2021 IEEE International Conference on Data Science and Computer Application (ICDSCA)*, Oct. 2021, pp. 670–674. doi: 10.1109/ICDSCA53499.2021.9650282.
- [146] S. Zheng, W. Shi, J. Liu, and J. Tian, “Remote Sensing Image Fusion Using Multiscale Mapped LS-SVM,” *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1313–1322, May 2008, doi: 10.1109/TGRS.2007.912737.
- [147] J. Wang, X. Cai, X. Chen, Z. Zhang, and L. Tang, “Classification of Forest Vegetation Type Using Fused NDVI Time Series Data Based on STNLFFM,” in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2019, pp. 6047–6050. doi: 10.1109/IGARSS.2019.8897907.
- [148] X. Nie, R. Gao, R. Wang, and D. Xiang, “Online Multiview Deep Forest for Remote Sensing Image Classification via Data Fusion,” *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 8, pp. 1456–1460, Aug. 2021, doi: 10.1109/LGRS.2020.3002848.
- [149] D. K. Seo, Y. H. Kim, Y. D. Eo, M. H. Lee, and W. Y. Park, “Fusion of SAR and Multispectral Images Using Random Forest Regression for Change Detection,” *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 10, Art. no. 10, Oct. 2018, doi: 10.3390/ijgi7100401.
- [150] H. Farhadi and M. Najafzadeh, “Flood Risk Mapping by Remote Sensing Data and Random Forest Technique,” *Water*, vol. 13, no. 21, Art. no. 21, Jan. 2021, doi: 10.3390/w13213115.
- [151] Z. Lv, T. Liu, C. Shi, J. A. Benediktsson, and H. Du, “Novel Land Cover Change Detection Method Based on k-Means Clustering and Adaptive Majority Voting Using Bitemporal Remote Sensing Images,” *IEEE Access*, vol. 7, pp. 34425–34437, 2019, doi: 10.1109/ACCESS.2019.2892648.
- [152] C. E. Fuss, A. A. Berg, and J. B. Lindsay, “DEM Fusion Using a Modified K-Means Clustering Algorithm,” *Int. J. Digit. Earth*, vol. 9, no. 12, pp. 1242–1255, Dec. 2016, doi: 10.1080/17538947.2016.1208685.
- [153] G. Scarpa, M. Gargiulo, A. Mazza, and R. Gaetano, “A CNN-Based Fusion Method for Feature Extraction from Sentinel Data,” *Remote Sens.*, vol. 10, no. 2, p. 236, Feb. 2018, doi: 10.3390/rs10020236.
- [154] Z. Shao and J. Cai, “Remote Sensing Image Fusion with Deep Convolutional Neural Network,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 5, pp. 1656–1669, May 2018, doi: 10.1109/JSTARS.2018.2805923.
- [155] J. Zhong, B. Yang, G. Huang, F. Zhong, and Z. Chen, “Remote Sensing Image Fusion with Convolutional Neural Network,” *Sens. Imaging*, vol. 17, no. 1, p. 10, Jun. 2016, doi: 10.1007/s11220-016-0135-6.
- [156] D. Kavran, D. Mongus, B. Žalik, and N. Lukac, “Graph Neural Network-Based Method of Spatiotemporal Land Cover Mapping Using Satellite Imagery,” *Sensors*, vol. 23, no. 14, p. 6648, Jul. 2023, doi: 10.3390/s23146648.
- [157] Z. Yin, L. Zou, J. Sun, H. Zhang, W. Zhang, and X. Shen, “A Nonlinear Radiometric Normalization Model for Satellite Images Time Series Based on Artificial Neural Networks and Greedy Algorithm,” *Remote Sens.*, vol. 13, no. 5, Art. no. 5, Jan. 2021, doi: 10.3390/rs13050933.
- [158] Y. Wei, Q. Yuan, H. Shen, and L. Zhang, “Boosting the Accuracy of Multispectral Image Pansharpening by Learning a Deep Residual Network,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1795–1799, Oct. 2017, doi: 10.1109/LGRS.2017.2736020.
- [159] B. Dang and Y. Li, “MSResNet: Multiscale Residual Network via Self-Supervised Learning for Water-Body Detection in Remote Sensing Imagery,” *Remote Sens.*, vol. 13, no. 16, Art. no. 16, Jan. 2021, doi: 10.3390/rs13163122.
- [160] R. Qin, X. Huang, W. Liu, and C. Xiao, “Semantic 3D Reconstruction Using Multi-View High-Resolution Satellite Images Based on U-Net and Image-Guided Depth Fusion,” in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2019, pp. 5057–5060. doi: 10.1109/IGARSS.2019.8900588.
- [161] W. Yao, Z. Zeng, C. Lian, and H. Tang, “Pixel-Wise Regression Using U-Net and Its Application on Pansharpening,” *Neurocomputing*, vol. 312, pp. 364–371, Oct. 2018, doi: 10.1016/j.neucom.2018.05.103.
- [162] J. Fang, J. Yang, A. Khader, and L. Xiao, “A Multiresolution Details Enhanced Attentive Dual-UNet for Hyperspectral and Multispectral Image Fusion,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 16, pp. 638–655, 2023, doi: 10.1109/JSTARS.2022.3228941.
- [163] Z. Yang, C. Diao, and B. Li, “A Robust Hybrid Deep Learning Model for Spatiotemporal Image Fusion,” *Remote Sens.*, vol. 13, no. 24, p. 5005, Dec. 2021, doi: 10.3390/rs13245005.
- [164] Y. Chang and B. Luo, “Bidirectional Convolutional LSTM Neural Network for Remote Sensing Image Super-Resolution,” *Remote Sens.*, vol. 11, no. 20, Art. no. 20, Jan. 2019, doi: 10.3390/rs11202333.
- [165] C. A. Theran, M. A. Álvarez, E. Arzuaga, and H. Sierra, “A Pixel Level Scaled Fusion Model to Provide High Spatial-Spectral Resolution for Satellite Images Using LSTM Networks,” in *2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Sep. 2019, pp. 1–5. doi: 10.1109/WHISPERS.2019.8921269.
- [166] L. Mou, L. Bruzzone, and X. X. Zhu, “Learning Spectral-Spatial-Temporal Features via a Recurrent Convolutional Neural Network for Change Detection in Multispectral Imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 924–935, Feb. 2019, doi: 10.1109/TGRS.2018.2863224.
- [167] T. Zhou, H. He, Y. Wang, and Y. Liao, “Improved Gated Recurrent Units together with Fusion for Semantic Segmentation of Remote Sensing Images based on Parallel Hybrid Network,” In Review, preprint, Oct. 2023. doi: 10.21203/rs.3.rs-3400980/v1.
- [168] Q. Zeng *et al.*, “Prediction of Radar Echo Space-Time Sequence Based on Improving TrajGRU Deep-Learning Model,” *Remote Sens.*, vol. 14, no. 19, Art. no. 19, Jan. 2022, doi: 10.3390/rs14195042.
- [169] J. Liu and Y. Wang, “Water Body Extraction in Remote Sensing Imagery Using Domain Adaptation-Based Network Embedding Selective

- Self-Attention and Multi-Scale Feature Fusion,” *Remote Sens.*, vol. 14, no. 15, p. 3538, Jul. 2022, doi: 10.3390/rs14153538.
- [170] H. Gong *et al.*, “Swin-Transformer-Enabled YOLOv5 with Attention Mechanism for Small Object Detection on Satellite Images,” *Remote Sens.*, vol. 14, no. 12, Art. no. 12, Jan. 2022, doi: 10.3390/rs14122861.
- [171] S. K. Roy, A. Deria, D. Hong, B. Rasti, A. Plaza, and J. Chanussot, “Multimodal Fusion Transformer for Remote Sensing Image Classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–20, 2023, doi: 10.1109/TGRS.2023.3286826.
- [172] L. Scheibenreif, J. Hanna, M. Mommert, and D. Borth, “Self-supervised Vision Transformers for Land-cover Segmentation and Classification,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jun. 2022, pp. 1421–1430, doi: 10.1109/CVPRW56347.2022.00148.
- [173] A. A. Aleissaee *et al.*, “Transformers in Remote Sensing: A Survey,” *Remote Sens.*, vol. 15, no. 7, p. 1860, Mar. 2023, doi: 10.3390/rs15071860.
- [174] Q. Li, Y. Chen, and Y. Zeng, “Transformer with Transfer CNN for Remote-Sensing-Image Object Detection,” *Remote Sens.*, vol. 14, no. 4, Art. no. 4, Jan. 2022, doi: 10.3390/rs14040984.
- [175] Q. Cheng, H. Liu, H. Shen, P. Wu, and L. Zhang, “A Spatial and Temporal Nonlocal Filter-Based Data Fusion Method,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4476–4488, Aug. 2017, doi: 10.1109/TGRS.2017.2692802.
- [176] M. Rothermel, N. Haala, and D. Fritsch, “A Median-Based Depthmap Fusion Strategy for the Generation of Oriented Points,” *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. III-3, pp. 115–122, Jun. 2016, doi: 10.5194/isprs-annals-III-3-115-2016.
- [177] C. Lanaras, J. Bioucas-Dias, E. Baltasvias, and K. Schindler, “Super-Resolution of Multispectral Multiresolution Images from a Single Sensor,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI: IEEE, Jul. 2017, pp. 1505–1513, doi: 10.1109/CVPRW.2017.194.
- [178] K. Nasrollahi and T. B. Moeslund, “Super-Resolution: A Comprehensive Survey,” *Mach. Vis. Appl.*, vol. 25, no. 6, pp. 1423–1468, Aug. 2014, doi: 10.1007/s00138-014-0623-4.
- [179] L. T. Luppino, F. M. Bianchi, G. Moser, and S. N. Anfinsen, “Unsupervised Image Regression for Heterogeneous Change Detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9960–9975, Dec. 2019, doi: 10.1109/TGRS.2019.2930348.
- [180] B. Haack and A. Rafter, “Regression Estimation Techniques with Remote Sensing: A Review and Case Study,” *Geocarto Int.*, vol. 25, no. 1, pp. 71–82, Feb. 2010, doi: 10.1080/10106040802711679.
- [181] M. A. Fischler and R. C. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981, doi: 10.1145/358669.358692.
- [182] R. Touati, M. Mignotte, and M. Dahmane, “Multimodal Change Detection in Remote Sensing Images Using an Unsupervised Pixel Pairwise-Based Markov Random Field Model,” *IEEE Trans. Image Process.*, vol. 29, pp. 757–767, 2020, doi: 10.1109/TIP.2019.2933747.
- [183] T. Guo, H. S. Mousavi, T. H. Vu, and V. Monga, “Deep Wavelet Prediction for Image Super-Resolution,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA: IEEE, Jul. 2017, pp. 1100–1109, doi: 10.1109/CVPRW.2017.148.
- [184] P. Sirguey, R. Mathieu, Y. Arnaud, M. M. Khan, and J. Chanussot, “Improving MODIS Spatial Resolution for Snow Mapping Using Wavelet Fusion and ARSIS Concept,” *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 78–82, Jan. 2008, doi: 10.1109/LGRS.2007.908884.
- [185] I. Amro, J. Mateos, M. Vega, R. Molina, and A. K. Katsaggelos, “A Survey of Classical Methods and New Trends in Pansharpening of Multispectral Images,” *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 1, p. 79, Sep. 2011, doi: 10.1186/1687-6180-2011-79.
- [186] Y. O. Ouma, J. Tetuko, and R. Tateishi, “Analysis of Co-Occurrence and Discrete Wavelet Transform Textures for Differentiation of Forest and Non-Forest Vegetation in Very-High-Resolution Optical-Sensor Imagery,” *Int. J. Remote Sens.*, vol. 29, no. 12, pp. 3417–3456, Jun. 2008, doi: 10.1080/01431160701601782.
- [187] M. Sheykhou, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, “Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 6308–6325, 2020, doi: 10.1109/JSTARS.2020.3026724.
- [188] X. Liu, Y. Wang, and Q. Liu, “Psgan: A Generative Adversarial Network for Remote Sensing Image Pan-Sharpening,” in *2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens: IEEE, Oct. 2018, pp. 873–877, doi: 10.1109/ICIP.2018.8451049.
- [189] C. Chen *et al.*, “Flood Discharge Prediction Based on Remote-Sensed Spatiotemporal Features Fusion and Graph Attention,” *Remote Sens.*, vol. 13, no. 24, p. 5023, Dec. 2021, doi: 10.3390/rs13245023.
- [190] O. Tasar, A. Giros, Y. Tarabalka, P. Alliez, and S. Clerc, “DAugNet: Unsupervised, Multisource, Multitarget, and Life-Long Domain Adaptation for Semantic Segmentation of Satellite Images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1067–1081, Feb. 2021, doi: 10.1109/TGRS.2020.3006161.
- [191] X. Li, L. Zhang, B. Du, L. Zhang, and Q. Shi, “Iterative Reweighting Heterogeneous Transfer Learning Framework for Supervised Remote Sensing Image Classification,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 10, no. 5, pp. 2022–2035, May 2017, doi: 10.1109/JSTARS.2016.2646138.
- [192] S. J. Pan, J. T. Kwok, and Q. Yang, “Transfer Learning Via Dimensionality Reduction,” in *Proceedings of the 23rd national conference on Artificial intelligence - Volume 2*, in AAAI’08. Chicago, Illinois: AAAI Press, Jul. 2008, pp. 677–682.
- [193] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, “Domain Adaptation via Transfer Component Analysis,” *IEEE Trans. Neural Netw.*, vol. 22, no. 2, pp. 199–210, Feb. 2011, doi: 10.1109/TNN.2010.2091281.
- [194] X. Shi, Q. Liu, W. Fan, P. S. Yu, and R. Zhu, “Transfer Learning on Heterogenous Feature Spaces via Spectral Transformation,” in *2010 IEEE International Conference on Data Mining*, Dec. 2010, pp. 1049–1054, doi: 10.1109/ICDM.2010.65.
- [195] M. Pastorino *et al.*, “Multisensor and Multiresolution Remote Sensing Image Classification through a Causal Hierarchical Markov Framework and Decision Tree Ensembles,” *Remote Sens.*, vol. 13, no. 5, Art. no. 5, Jan. 2021, doi: 10.3390/rs13050849.
- [196] P. P. Shingare, P. M. Hemane, and D. S. Dandekar, “Fusion Classification of Multispectral and Panchromatic Image Using Improved Decision Tree Algorithm,” in *2014 International Conference on Signal Propagation and Computer Technology (ICSPCT 2014)*, Jul. 2014, pp. 598–603, doi: 10.1109/ICSPCT.2014.6884944.
- [197] X. Zhu, F. Cai, J. Tian, and T. K.-A. Williams, “Spatiotemporal Fusion of Multisource Remote Sensing Data: Literature Survey, Taxonomy, Principles, Applications, and Future Directions,” *Remote Sens.*, vol. 10, no. 4, Art. no. 4, Apr. 2018, doi: 10.3390/rs10040527.
- [198] F. Yuan, K. E. Sawaya, B. C. Loeffelholz, and M. E. Bauer, “Land Cover Classification and Change Analysis of the Twin Cities (minnesota) Metropolitan Area by Multitemporal Landsat Remote Sensing,” *Remote Sens. Environ.*, vol. 98, no. 2–3, pp. 317–328, Oct. 2005, doi: 10.1016/j.rse.2005.08.006.
- [199] I. H. Sarker, “Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions,” *SN Comput. Sci.*, vol. 2, no. 6, p. 420, Aug. 2021, doi: 10.1007/s42979-021-00815-1.
- [200] P. V. Arun, K. M. Buddhiraju, A. Porwal, and J. Chanussot, “CNN Based Spectral Super-Resolution of Remote Sensing Images,” *Signal Process.*, vol. 169, p. 107394, Apr. 2020, doi: 10.1016/j.sigpro.2019.107394.
- [201] L. Liebel and M. Körner, “SINGLE-IMAGE SUPER RESOLUTION FOR MULTISPECTRAL REMOTE SENSING DATA USING CONVOLUTIONAL NEURAL NETWORKS,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLI-B3, pp. 883–890, Jun. 2016, doi: 10.5194/isprs-archives-XLI-B3-883-2016.
- [202] L. Zeng, Y. Du, H. Lin, J. Wang, J. Yin, and J. Yang, “A Novel Region-Based Image Registration Method for Multisource Remote Sensing Images Via CNN,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 1821–1831, 2021, doi: 10.1109/JSTARS.2020.3047656.
- [203] 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.
- [204] J. R. Bergado, C. Persello, and A. Stein, “Recurrent Multiresolution Convolutional Networks for VHR Image Classification,” *ArXiv180605793 Cs*, Jun. 2018, Accessed: Oct. 14, 2018. [Online]. Available: <http://arxiv.org/abs/1806.05793>
- [205] W. Li, D. Cao, Y. Peng, and C. Yang, “MSNet: A Multi-Stream Fusion Network for Remote Sensing Spatiotemporal Fusion Based on

- Transformer and Convolution,” *Remote Sens.*, vol. 13, no. 18, p. 3724, Sep. 2021, doi: 10.3390/rs13183724.
- [206] E. Maggiori, G. Charpiat, Y. Tarabalka, and P. Alliez, “Recurrent Neural Networks to Correct Satellite Image Classification Maps,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 4962–4971, Sep. 2017, doi: 10.1109/TGRS.2017.2697453.
- [207] T. Zhu, J. Shen, and F. Sun, “Long Short-Term Memory-Based Simulation Study of River Happiness Evaluation – a Case Study of Jiangsu Section of Huaihe River Basin in China,” *Heliyon*, vol. 8, no. 9, p. e10550, Sep. 2022, doi: 10.1016/j.heliyon.2022.e10550.
- [208] Y. Bazi, L. Bashmal, M. M. A. Rahhal, R. A. Dayil, and N. A. Ajlan, “Vision Transformers for Remote Sensing Image Classification,” *Remote Sens.*, vol. 13, no. 3, p. 516, Feb. 2021, doi: 10.3390/rs13030516.
- [209] D. A. Shahajan, M. Varma T., and R. Muthuganapathy, “Point Transformer for Shape Classification and Retrieval of Urban Roof Point Clouds,” *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2021.3061422.
- [210] K. Chen, Z. Zou, and Z. Shi, “Building Extraction from Remote Sensing Images with Sparse Token Transformers,” *Remote Sens.*, vol. 13, no. 21, Art. no. 21, Jan. 2021, doi: 10.3390/rs13214441.
- [211] X. Zhang, P. Xiao, X. Feng, and M. Yuan, “Separate Segmentation of Multi-Temporal High-Resolution Remote Sensing Images for Object-Based Change Detection in Urban Area,” *Remote Sens. Environ.*, vol. 201, pp. 243–255, Nov. 2017, doi: 10.1016/j.rse.2017.09.022.
- [212] U. H. Atasever, P. Civicioglu, E. Besdok, and C. Ozkan, “A New Unsupervised Change Detection Approach Based On DWT Image Fusion And Backtracking Search Optimization Algorithm For Optical Remote Sensing Data,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XL-7, pp. 15–18, Sep. 2014, doi: 10.5194/isprsarchives-XL-7-15-2014.
- [213] M. J. Canty, A. A. Nielsen, and M. Schmidt, “Automatic Radiometric Normalization of Multitemporal Satellite Imagery,” *Remote Sens. Environ.*, vol. 91, no. 3, pp. 441–451, Jun. 2004, doi: 10.1016/j.rse.2003.10.024.
- [214] Y. Du, P. M. Teillet, and J. Cihlar, “Radiometric Normalization of Multitemporal High-Resolution Satellite Images with Quality Control for Land Cover Change Detection,” *Remote Sens. Environ.*, vol. 82, no. 1, pp. 123–134, Sep. 2002, doi: 10.1016/S0034-4257(02)00029-9.
- [215] A. A. Nielsen, “The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data,” *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 463–478, Feb. 2007, doi: 10.1109/TIP.2006.888195.
- [216] H. Shen, M. K. Ng, P. Li, and L. Zhang, “Super-Resolution Reconstruction Algorithm to MODIS Remote Sensing Images,” *Comput. J.*, vol. 52, no. 1, pp. 90–100, Feb. 2008, doi: 10.1093/comjnl/bxm028.
- [217] H. Zhu, W. Song, H. Tan, J. Wang, and D. Jia, “Super Resolution Reconstruction Based on Adaptive Detail Enhancement for Zy-3 Satellite Images,” *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. III-7, pp. 213–217, Jun. 2016, doi: 10.5194/isprsaannals-III-7-213-2016.
- [218] M. Rumpler, A. Wendel, and H. Bischof, “Probabilistic Range Image Integration for DSM and True-Orthophoto Generation,” in *Image Analysis*, J.-K. Kämäärinen and M. Koskela, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2013, pp. 533–544. doi: 10.1007/978-3-642-38886-6\_50.
- [219] J. Ma, H. Shen, P. Wu, J. Wu, M. Gao, and C. Meng, “Generating Gapless Land Surface Temperature with a High Spatio-Temporal Resolution by Fusing Multi-Source Satellite-Observed and Model-Simulated Data,” *Remote Sens. Environ.*, vol. 278, p. 113083, Sep. 2022, doi: 10.1016/j.rse.2022.113083.
- [220] X. Ma, Y. Yao, B. Zhang, and C. He, “Retrieval of High Spatial Resolution Precipitable Water Vapor Maps Using Heterogeneous Earth Observation Data,” *Remote Sens. Environ.*, vol. 278, p. 113100, Sep. 2022, doi: 10.1016/j.rse.2022.113100.
- [221] K. J. Theron, J. S. Pryke, N. Latte, and M. J. Samways, “Mapping an Alien Invasive Shrub Within Conservation Corridors Using Super-Resolution Satellite Imagery,” *J. Environ. Manage.*, vol. 321, p. 116023, Nov. 2022, doi: 10.1016/j.jenvman.2022.116023.
- [222] A. Voutilainen, T. Pyhälähti, K. Y. Kallio, J. Pulliainen, H. Haario, and J. P. Kaipio, “A Filtering Approach for Estimating Lake Water Quality from Remote Sensing Data,” *Int. J. Appl. Earth Obs. Geoinformation*, vol. 9, no. 1, pp. 50–64, Feb. 2007, doi: 10.1016/j.jag.2006.07.001.
- [223] Q. Cheng, Y. Zhou, P. Fu, Y. Xu, and L. Zhang, “A Deep Semantic Alignment Network for the Cross-Modal Image-Text Retrieval in Remote Sensing,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 4284–4297, 2021, doi: 10.1109/JSTARS.2021.3070872.
- [224] M. Guo, C. Zhou, and J. Liu, “Jointly Learning of Visual and Auditory: A New Approach for RS Image and Audio Cross-Modal Retrieval,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 11, pp. 4644–4654, Nov. 2019, doi: 10.1109/JSTARS.2019.2949220.
- [225] G. Mao, Y. Yuan, and L. Xiaoqiang, “Deep Cross-Modal Retrieval for Remote Sensing Image and Audio,” in *2018 10th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS)*, Aug. 2018, pp. 1–7. doi: 10.1109/PRRS.2018.8486338.