

Advances In Satellite Image Processing: Enhancement, Segmentation, Classification, And Change Detection

Ms.Priyanka T. Avhad¹, Dr. Parminder Kaur²

¹Research Scholar, MGM University, Chhatrapati Sambhajanagar, Maharashtra, India

²Professor, MGM University, Chhatrapati Sambhajanagar, Maharashtra, India

Abstract

The accurate classification of the field and territorial boundaries is becoming increasingly important in the face of the dynamic nature of the modern land and water fronts. Historical maps often contain changes in administrative boundaries, vegetation patterns and water bodies, thus reliable tools to detect these changes are important. Satellite imagery has become the main means of observing such changes and therefore a thorough understanding of satellite image processing techniques is needed when dealing with historical cartographic data. The present study gives a thorough assessment of the pros and cons of a wide variety of image processing techniques in this field. There are a number of computational approaches, but their relative effectiveness can vary greatly depending on the nature of the imagery and the needs of the specific application. The application of incorrect technique may produce suboptimal results. This highlights the importance of matching the method to the objectives of the analysis. This work therefore describes the relevant processing techniques and satellite imaging methods suitable for different use cases. Moreover, some comparative evaluations are performed to demonstrate the performances of different approaches under different conditions. The aim of this synthesis is to facilitate the development of informed and innovative solutions to common problems encountered in satellite based image processing applications.

Keywords: remote sensing; change detection; fusion; feature extraction; segmentation; classification.

Introduction

The combination of satellite-based field boundary detection provides a powerful tool for the detection of continuously changing interfaces between terrestrial and marine features. By looking at a series of images taken from a satellite over a period of time, we can see whether particular areas have got bigger, smaller or changed their structure. Such temporal analyses are a direct input for the production of accurate maps, which are used in a broad range of practical activities. Therefore, satellite image processing has become an essential computational tool in many fields like defence, agriculture, disaster mitigation and resource assessment. But the sheer volume and complexity of satellite data poses serious challenges for rapid and accurate processing.

Remote sensing imagery generally contains a lot of information but the usefulness of the imagery is compromised if the images are of poor quality or too few features are selected in the analysis. Traditional visual interpretation uses features like shape, colour, tone and texture but it is often not possible for humans to interpret several images at once. Thus, the time and accuracy of manual classification heavily depend on subjective interpretation and the quality of the images. These limitations have led to a move towards computational methods that can handle large data sets efficiently and include multi-band spectral information that would be unmanageable for human analysts. Therefore, image processing has become an essential part of modern applications of remote sensing, where the accuracy of the processing method has a direct influence on the reliability of the extracted information.

Thus a good knowledge of different image processing techniques is necessary in order to choose the most suitable one for a task. The development of satellite technology has resulted in a demand for timely, high-resolution imagery, particularly for emergency

response operations and public safety. The growing capacities of emerging Earth observation systems to provide continuous monitoring are leading to the development of new strategies to ensure the effective use of available data. Its goal is to provide timely and accurate insights to decision makers, especially during high-stress situations such as natural disasters.

One of the ongoing challenges, despite the wealth of satellite and aerial imagery, is turning the raw data into actionable intelligence. But it is still a difficult task to extract the relevant patterns, features and structures correctly. There are many processing techniques, but traditional methods usually need to be carefully evaluated to know whether they are suitable for a particular use case. The advances in artificial intelligence and machine learning have resulted in more powerful feature extraction, especially from high resolution imagery, and the ability to build multi-layered processing architectures. But there is still a lot of room for improvement in the performance and adaptability of existing systems. A thorough knowledge of traditional approaches is fundamental to design new automated workflows that can address the different needs of satellite image processing. The present work reviews the main image processing techniques and their applications, providing their advantages and disadvantages. The combination of past and current methods is intended to help develop better and new solutions for persistent problems in satellite image processing.

Image Processing in Remote Sensing

The satellite image processing techniques addressed include enhancement, feature extraction, segmentation, fusion, change detection, compression, classification, and feature detection, as illustrated in Figure 1.

Figure 1 Image Processing Methodology

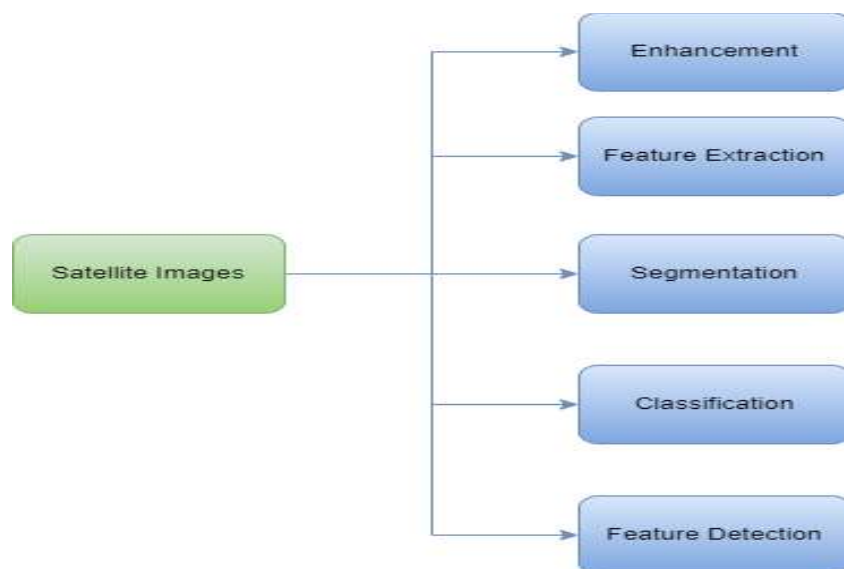


Image Enhancement

Enhancement is an important step for improving the overall visual quality and keeping the important details as the satellite images have low brightness levels. Contrast is one of the perceptual factors that influence the subjective evaluation of an image. Contrast is a critical issue, because it allows us to easily tell the objects apart from the environment. The interpretability of features is primarily determined by the differences in brightness

and colour of the scene elements. The limitation of illumination and contrast has been addressed by a number of enhancement algorithms proposed in the literature. This enhancement is often performed prior to higher level processing such as segmentation and object recognition. The quality and clarity of the input images play a big role in the following analysis. The image enhancement techniques can be broadly classified into two types namely spatial domain and frequency domain [1]. Histogram equalization is still one of the most popular techniques among these that can be used to adjust the intensity distribution either over the entire image or over selected regions . Traditional histogram equalization can greatly improve the global contrast but usually does not preserve the average luminance of the original image. To overcome these limitations, several refined variants such as Bi-Histogram Equalization (BHE), Recursive Mean Separate Histogram Equalization (RMSHE) and the related methods have been proposed to produce more natural enhancement results.

The enhancement methods and their evaluation metrics are shown in **Table 1**.

Table 1. Image Enhancement methods and their performance metrics.

Method	Test Images	Performance Metrics
Histogram Planting [2]	Phobos Images from Satellites	Mean, Average Information
Modified differential evolution [3]	Images from NASA, Satpalda Geospatial Services and Satellite Imaging Corp.	Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE), Universal Quality Index (UQI), Normalized Absolute Error (NAE), Structural Contrast-Quality Index (SC-QI):
Fractional Differential (FD) unsharp masking [4]	Infrared Images	Average brightness (B), Average contrast (V), Average discrete information content (or entropy, E), Sharpness (S), Colorfulness (C), Correlation (R), Energy (E) and Homogeneity (M)

The improved images by RMSHE are effective, but some visual inconsistencies and artifacts are found. Many of these problems can be alleviated by Histogram Planting (HP) [2], a distribution based improvement technique which holds a lot of promises to achieve more balanced contrast adjustments. There are two main distribution strategies of HP: HP-ESD (equality-spaced distribution) and HP-PSD (proportionality-spaced distribution). These approaches often provide a good improvement, but can in some cases lead to small distortions. To overcome such limitations, the researchers have proposed an advanced enhancement technique by combining a modified differential evolution process with Cuckoo Search (CS) optimization algorithm [3]. This hybrid approach, successfully reduces unwanted artifacts while preserving the fine image details and average luminance. The motivation is to find a transformation which can remap intensities of pixels to a more informative range. In this framework, the local feature extraction is performed by means of a sliding window which is computationally expensive. Hence, the MDE algorithm combines global and local extraction processes to balance accuracy and

efficiency.

Further improvements are obtained by using multi-resolution satellite images from different sensors, offering better accuracy with limited reliance on Ground Control Points (GCPs) [5]. Root mean square error values for multi-resolution data sets are calculated using a subset of rational polynomial coefficients with manual selection of tie points. Future developments may automate this tie-point selection process, reducing manual intervention and increasing consistency across large image collections.

Usually, the performance of enhancement techniques is evaluated using several quantitative measures. The quality of the reconstructed image is measured by Peak Signal to noise ratio (PSNR). The Absolute Mean Brightness Error (AMBE) is the difference between the mean brightness of the original and the enhanced images. Entropy is a measure of information content of the image and indicates the improvement of visibility of details. Furthermore, the Mean Square Error (MSE) quantifies the differences squared between the input and the processed images, showing the extent of structural changes that have occurred during the enhancement.

Feature Extraction

Feature extraction is the process of extracting quantifiable features that are used to select and classify parts of an image. This is the basis for subsequent classification tasks. The extracted descriptors have a direct influence on the accuracy and robustness of the final output [6] and, thus, these features are decisive for how well an image classification system performs. The features can generally be divided into two groups: general ones and domain specific ones. There are two types of features: general features which are generic visual properties such as colour, shape or texture, and domain-specific features which are custom properties tailored to a specific application or context.

Among the approaches explored for feature correspondence, a graph-theoretic matching method has been proposed in Reference [7]. This method is robust against nonlinear intensity variations but it is computationally intensive, although effective, since it adopts the Mutual Information (MI) as a similarity measure. Other descriptors have been proposed to overcome such limitations. One of them is the Local Self-Similarity (LSS), which is less complex and more robust to illumination variations. Furthermore, uniform robust (UR) versions of the Scale-Invariant Feature Transform (SIFT) have been proposed to improve the consistency of the feature distribution and hence the stability of matches under scale and orientation changes. The key performance indicators in the evaluation of these methods are recall, the ratio of correctly matched features to all true matches and precision, the ratio of correct matches to all detected matches. Table 2 shows a comparison of these feature extraction methods and their performance metrics.

Table 2 Feature extraction approaches

Method	Test Images	Feature Extracted	Performance Metrics
Uniform Competency Feature Extraction [8]	Synthetic Images: ETM+, World View 2, Real Images: SPOT4, SPOT6, Quick Bird, GeoEye1, ASTER	Rotation and scale invariant Local features	Recall, Precision, Root Mean Square Error
RepTree, Machine Learning and	Envisat images	Continues features such as roads, rivers	Efficiency and processing time

Euclidean distance [9]			
Multi image saliency analysis [10]	SPOT 5 images	ROI extraction such as clouds	Recall, precision, F-measure.
Digital Surface Models	ZY-3 satellite images	Pixel and feature level extraction of urban scenes	Overall accuracy kappa coefficient
Reversible jump Markov chain Monte Carlo sampler [12]	Synthetic image	Extraction of rivers, channels and roads	Completeness, correctness

Early feature extraction methods often used custom designed filters to extract local structures such as roads. These techniques were gradually replaced by boundary-based detection approaches that over the years were found to be more reliable in dealing with complicated spatial patterns. In recent years, the paradigm of machine learning has been dominating the feature extraction. A typical example is the framework which combines machine learning techniques with Euclidean distance measures in [9]. The main input of this method is satellite images. The system is trained to recognize certain blocks of features in the image using algorithms. It can have linear features like rivers. In this case, the REPTree classifier is particularly useful, since it builds the decision trees incrementally to find and improve the relevant features.

Saliency-based analysis has also been a powerful tool in separating prominent structures from the noise in the background. Saliency maps can be created by performing K-means clustering of multispectral images which emphasize the regions of interest. A similar approach is followed for panchromatic images where co-occurrence histograms are used to highlight the salient features of selected areas. After obtaining the individual saliency maps, we propose a fusion mechanism to combine them. This way irrelevant background is suppressed and meaningful features are highlighted. This integrated approach greatly enhances the overall interpretability of the image.

Another promising direction is the angular feature based classification methods for high resolution urban imagery [11]. The techniques are based on three different Angular Difference Features (ADFs) for capturing pixel and feature level variations. The idea is that man-made structures (rooftops, edges of buildings, etc.) have different geometric signatures than natural landscapes. The disadvantage of angle based methods is that the height or the prominence of structures protruding from the ground may be underestimated due to visual distortions in the high resolution imagery. This may cause some features to be ambiguous, especially in the shadowed areas, and sometimes cause mismatches or misinterpretations in the process of classification.

Image Segmentation

Image segmentation is a process of partitioning an image into meaningful regions to identify different objects and their boundaries accurately [13]. Various optimization based segmentation techniques have been successfully applied to colour image segmentation. The Cuckoo Search (CS) algorithm is one such technique. In Reference [14], an improvement for the traditional CS method was proposed by adding McCulloch's method based on the Lévy flight distributions to improve the exploration behaviour of the algorithm. Heuristic optimization algorithms like Particle Swarm Optimization (PSO), Artificial Bee Colony Optimization (ABC) and CS have been extensively applied. Yet, their

computational cost usually grows very fast with the number of segmentation thresholds, which limits their efficiency for complex tasks.

One important difference in the modified CS framework is that it uses McCulloch's approach to generate more stable random numbers, which results in better convergence properties. The method can be extended to include multilevel thresholding which improves the segmentation results for more heterogeneous imagery. In particular, pixel-based clustering is often used in remote sensing applications to address ambiguous or visually similar regions when uniformly illuminated scenes are concerned. One clustering method that can be used is the Hopfield Neural Network. This approach is better if you require a reasonable number of clusters. Table 3 summarizes the performance evaluation of different segmentation algorithms using standard metrics.

Table 3. Image segmentation methods and their performance metrics.

Method	Test Images	Performance Metrics
Cuckoo Search, McCulloch's method [14]	Pleiades satellite images	Recall, Precision, Root Mean Square Error
Markov Random Filed method [15]	QuickBird bands with four spectral bands	Recall, Precision
Deep convolutional Neural Network [16]	IR images	Recall, precision, F-measure.
Levy flight firefly algorithm [17]	Multiband satellite images from NASA	Overall accuracy kappa coefficient
Graph based segmentation, Gabor filter [18]	Quickbird satellite images	Completeness, correctness

Among the image segmentation techniques, histogram based thresholding is still one of the most popular methods. The simplest bi-level thresholding uses only one threshold to classify the pixels into two classes. While the multilevel thresholding uses several thresholds to divide the image into multiple regions [19]. The threshold is usually selected according to optimization criteria, for example the maximum variance between classes or the minimum classification error [20]. The comparative studies have always shown that the multilevel thresholding is more effective for the segmentation than the bi-level thresholding, especially for the images with complicated intensity distribution. Further improvements are possible by including Markov Random Field (MRF) models as in [15]. MRF segmentation improves the stability of classification and allows the use of minimum distance classification schemes with smaller pixel clusters to estimate likelihoods and neighbourhood interactions. After segmentation, a supervised Random Forest classifier can be applied to each extracted segment, for which it is shown that good classification is possible even with small feature sets.

In [16], a Deep Convolutional Neural Network (DCNN) architecture for the segmentation of high resolution satellite imagery is proposed. The method is a combination of edge detection and SEGNET architecture. The SEGNET architecture takes an encoder-decoder approach for generating pixel-level labels across different classes. The method can produce detailed and accurate segmentations, but is computationally expensive due to the rather large decoder component. This complexity creates practical challenges for

researchers and practitioners who work in resource-constrained settings or with large scale remote sensing datasets [21].

A major limitation of many segmentation methods is that the extracted boundaries are frequently fuzzy or ill-defined. For this problem, a firefly optimization based on the fuzzy entropy function is proposed [17]. In this method, the fuzziness entropy is used to measure the contrast of adjacent regions and the threshold values are obtained by minimizing a fitness function so that the resulting segments have the same entropy. The firefly algorithm is founded on the luminescent communication behaviour of fireflies where each individual is attracted to other individuals depending on the perceived brightness. This mechanism can search the whole space efficiently, and it is suitable for getting the optimal threshold sets for the difficult segmentation tasks. The complex building arrangement and textured environment make road network extraction in dense urban areas more challenging. Gabor filters are used to improve the separability of structural patterns by enhancing textural cues before segmentation. The enhanced scene is then provided to a graph based segmentation framework [18]. This approach consists of two major steps including graph representation of the segmented image and region division and merging based on colour and shape similarities. In [22] further improvements are shown by using a Cuckoo Search (CS) algorithm based on a Tsallis entropy for multilevel colour image thresholding. The Tsallis entropy is a flexible method of thresholding, based on global properties of the image histogram, that contains an adjustable parameter. But the approach is computationally expensive because of the complexity of the algorithm.

The goal of object based image analysis is to find segmentation methods that lead to the best possible overall classification performance. It usually has two steps: segmentation of the image into homogeneous regions and classification of the image regions. The quantitative evaluation is usually performed by supervised methods. Qualitative assessments are easier to perform, but subjective. The quality of segmentation is often measured using the Under-Segmentation Error (USE) and Over-Segmentation Error (OSE) where each reference segment is evaluated independently and errors are summed to obtain a global performance metric. Other metrics for evaluation are also used, such as pixel accuracy (ratio of correctly classified pixels), precision (fraction of predicted positives that are correct) and recall (fraction of ground-truth objects that are detected).

Change Detection

Detecting change is important for monitoring change in a geographic area, particularly since satellite observations exist for most areas [46]. Methods for change detection are generally classified into four groups: algebraic methods, transformative methods, Geographic Information System (GIS)-based methods and advanced model-based methods. Algebraic techniques are used to quantify visual differences with mathematical operations. Common techniques include image differencing, Change Vector Analysis (CVA) and Principal Component Analysis (PCA). Monitoring vegetation growth is very useful with image differencing. PCA is often used for analysis of variability in water quality and vegetation composition. Transform based methods improve temporal differences by applying mathematical transformations and Tasseled Cap Transformation (TCT) is a popular transformation for this purpose. TCT is particularly useful for deforestation impacts, image differencing is effective in the tropics and TCT provides reliable change monitoring in urban areas. A comprehensive review of such change detection strategies can be found in [47].

Techniques based on GIS offer an integrated environment for processing, comparing and interpreting multi-temporal images. These techniques align historical and modern maps to better detect and classify change. For example, Reference [48] describes a GIS-based

strategy to analyze the change of vegetation cover and Reference [49] put forward a GIS-based framework to identify land cover change. Apart from these standard approaches, more advanced techniques include spectral unmixing and reflectance models where physical parameters are obtained from the reflectance values extracted from satellite images in order to better understand the underlying surface properties. [50] proposes using spectral mixture models for temporal unmixing in the analysis of vegetation change. Digital Surface Models (DSMs), which reflect the characteristics of the surface elevation, are also useful in monitoring the urban development and vegetation mapping. [51] shows the identification of structural changes in urban areas by using DSMs.

The maximum likelihood classifier is widely used in the first step of classification of land cover categories in remote sensing imagery [52]. After verifying the classified groups, image differencing or statistical change detection with probability matrices is used to detect changes between different time periods. Maximum likelihood classification is highly dependent on the accuracy of training data and misclassification of pixels can reduce the overall efficiency. Multi-Step Image Matching (MSIM) has been proposed for the detection in rapidly evolving environments, which combines fast clustering, feature detection and boundary matching to detect changes in disaster-prone regions [55]. Object-level analysis, as shown in Ref. [56], also enables tracking spatial and temporal changes with graph-based methods. These methods complement classifier-based methods that often fail to capture complex temporal dynamics. Correlation coefficients are commonly calculated in graph-based analysis in quantitative evaluation. Usually some important metrics are used to evaluate the performance of the change detection algorithm. The overall accuracy (OA) is calculated as the ratio of the number of correctly detected changed pixels and unchanged pixels. False Alarm (FA) is the number of wrongly detected changed pixels over the total number of unchanged pixels. Missed Alarm (MA) is the ratio of missed changes. These metrics together can be used to evaluate the performance of a change detection method for subtle and large changes in satellite images comprehensively.

Image Classification

This is especially important given the growing availability and complexity of Earth observation data [66,67]. The correct classification of images is important for the structuring and understanding of the satellite image. Image classification is a pattern recognition process, where pixels or image segments are classified into pre-defined classes based on similarity measures. Classification methods are divided into three categories: supervised, unsupervised and post-classification methods. Supervised classification uses training data to help the algorithm find relevant patterns. Such techniques may be parametric, where classes are assigned based on algebraic assumptions, like Bayesian classifiers and decision trees, or non-parametric, where no explicit probability density functions are required. The latter category includes methods such as k-Nearest Neighbour (k-NN) and logistic regression [68]. In contrast, unsupervised classification groups pixels based on their similarities, without any knowledge of the class labels. Hierarchical clustering or partition-based methods can be used for unsupervised classification.

Two old problems plague the classification of satellite images. One is the mixed pixels, i.e., pixels with multiple land cover types. The other is the difficulty of processing large high dimensional data. Among the popular solutions for land use and land cover mapping, the Random Forest classifier has been receiving attention. It is an ensemble method that takes the predictions of multiple decision trees and combines them to give a final class assignment. For example, Reference [69] proposes a Random Forest-based classification of wetlands using optimized geometric and spectral features from multi-sensor imagery.

Reference [67] takes it one step further by using a genetic algorithm to optimize the feature selection and tree parameters for better land cover classification. Spectral unmixing, or soft classification, is a common method for solving the mixed pixel problem. This method decomposes the spectral signature of each pixel to estimate the fractional composition of each pixel in different land cover classes. In [70] we use spectral unmixing to get fractional abundances to better represent the classes. A mixed supervised-unsupervised unmixing approach for the classification of multispectral images is presented in Reference [71]. In Reference [72], a Multiple Endmember Spectral Mixture Analysis (MESMA) technique is proposed for urban land cover mapping, where a class can be modelled by multiple endmembers to account for intra class variability. Neural network classifiers also play an important role in remote sensing applications. Radial basis function neural networks (RBFNNs) have advantages like robustness to noise and flexibility as they have many tunable parameters [73]. Classification Errors Correction (CEC) is a post-classification refinement technique to correct the misclassifications remaining after the initial analysis by combining highly correlated spectral bands using Discrete Wavelet Transform (DWT) for better accuracy [74]. Deep learning approaches have further pushed the field forward, e.g. [75] introduces a deep feature learning pipeline where the images are decomposed into a number of scales, each scale used as a training input for a Deep Convolutional Neural Network (DCNN). We employ spatial pyramid pooling (SPP) to speed up the training, which allows various input sizes with common parameters for the network.

Another important task in satellite image analysis, cloud detection, has been solved using Bayesian network classifiers [76]. This approach considers attributes related to the structure and appearance of the cloud. The classification performance is evaluated by k-fold cross validation for robust generalization. In Reference [77], the authors proposed an integrated framework (FSE-DST) that combines fuzzy evaluation and Dempster-Shafer theory to improve the classification reliability. The approach begins with an object-based classification, and then uses evidence fusion to increase the confidence of decision, resulting in more accurate and interpretable classification results. Hybrid methods are still important for improving the accuracy and reliability of satellite image classification.

Image Feature Detection

There has been an increase in the satellite images which helps to understand and analyze various applications. Feature recognition is gaining importance with the advancement in deep learning. **Table 6** represents the different feature detection methods and performance metrics.

Table 6. Feature Detection methods and their performance metrics.

Method	Test Images	Application	Performance Metrics
Fuzzy logic-based detection	Hyperspectral image	Edge detection	PSNR, MSE
Deep Convolutional Neural Network	QuickBird VHR image	Informal settlement detection	Overall accuracy
Invariant pixel detection, PCA	Landsat 7 ETM+ images	Cloud detection	Accuracy and F-measure
Active contour model	Google Earth images	Shadow detection and height estimation	Precision and recall

The ability to distinguish between features such as buildings, roads, vegetation and other landscape components is important for applications including environmental monitoring, urban analysis and disaster assessment. To detect correspondences between images, we need to detect salient points that are distinctive and repeatable in each image. The classification process involves comparing these feature descriptors with the ones obtained from the training samples and the class that has the maximum correspondence is taken as the best match. While global and texture based descriptors often work well for retrieving visually similar images from large databases, structure oriented local features are more appropriate for detailed classification tasks. The detection of nonlinear features such as informal urban settlements, oil spills or irregular natural formations still remains a challenging task. Thus the identification of salient points in such complex structures is of great importance for reliable classification results. Therefore, accurate image classification is based on robust feature detection.

Edge detection is an important feature extraction process and extensively used for outline of an object in satellite images. Commonly used edge detectors are Prewitt, Laplacian and Laplacian of Gaussian filters. Such filters are limited in the sense that they have a fixed edge thickness and are sensitive to the threshold settings. Linear time-invariant filters model edges as a sharp change in gray-level intensity. These filters provide gains in computational efficiency to overcome these limitations. Also, a fuzzy logic-based edge detection technique has been proposed to improve the detection of boundaries in complex scenes [78]. Stereo image matching can be used to generate a DSM for cloud detection that can be compared to a DEM to find seed points of clouds where differences in elevation are detected [82]. Next, we classify these cloudy regions by combining height information and spectral cues.

The image quality is degraded by the jitter caused by the satellite motion or sensor instability. Thus, it is important to identify and correct these distortions for the analysis to be accurate. In [83] a displacement model was proposed for detection of high altitude jitter in satellite imagery. The system applies bidirectional Kalman filtering and smoothing to mitigate the jitter effects, and then a re-imaging model to reconstruct the corrected outputs. Similarly, the ability of Fully Convolutional Networks (FCNs) to learn pixel-level dependencies [79] is also strong and it enables the discrimination between formal and informal settlements on the basis of morphological features. To identify these patterns, the spatial information is often extracted by means of wavelet transforms. Another application of the feature-based image analysis is the detection of oil spills, which is particularly significant for the maritime navigation and the environmental protection. Wide coverage SAR imagery has been used to develop machine learning driven feature selection methods for better oil spill identification [84]. In [85], aerial vehicle detection has been addressed with a transfer learning approach with a super-resolution framework. Detection is guided by sparse coefficients learnt from high and low resolution image patches. The first implementations utilize a linear SVM search technique, but they are not very efficient since it is difficult to match similar fragments of images. Dictionary learning and sparse coding overcome this limitation and allow the building of low resolution representations from shared coefficients of high resolution representations.

Comparison among different image processing techniques shows that their performance is highly dependent on the application context, noise characteristics and type of imagery being analyzed. Ultimately, the quality of segmentation and thus the overall analytical result is controlled by the precision of the processes of feature extraction, detection, fusion and change detection. The type of image noise determines which enhancement method is to be used. The structural complexity in terms of edges, contours and textures shows the need of robust and adaptive feature extraction methods. The choice of

techniques is based on the problem to be solved so that the results of image processing are consistent with the set objectives.

Literature Overview

The reviewed literature includes a wide range of satellite image processing techniques such as enhancement, feature extraction, segmentation, fusion, classification, change detection and evaluation methodologies. Each reference provides a unique view on how remotely sensed data can be analyzed, interpreted, or optimized for various real-world applications. To give a broad and systematic picture of these works, the table below collates all the 92 cited studies and summarizes their main ideas, methods, and applications. It is an organized representation that will help to understand the evolution of methods in different domains, understand thematic connections, and highlight the foundational studies that guide the following sections of this research.

Table 1 – Consolidated Reference Summary

Ref. No.	Topic / Category	Core Technique / Concept	Summary of Contribution	Application Area
[1]	Image Enhancement	Digital Image Processing Fundamentals	Establishes foundational concepts for contrast improvement, filtering, and spatial/frequency transformations.	General remote sensing & image preprocessing
[2]	Image Enhancement	Histogram Planting (Brightness Preservation)	Proposes a modified histogram-based method that enhances contrast without distorting average brightness.	Low-light satellite imagery
[3]	Image Enhancement	Modified Differential Evolution	Optimizes enhancement functions using evolutionary computation for better contrast while preventing artifacts.	High-resolution multispectral images
[4]	Image Enhancement	Fractional Differential Unsharp Masking	Applies fractional calculus to improve edges while maintaining natural image appearance.	Infrared and thermal images
[5]	Image Fusion	Rational Polynomial Coefficient Fusion	Combines multi-resolution images without requiring dense ground control points.	Multi-sensor fusion
[6]	Feature Extraction	SIFT (Scale-Invariant Features)	Identifies stable feature points under scale, orientation, and illumination variations.	Object recognition & remote sensing

[7]	Feature Extraction	Mutual Information-Based Matching	Uses statistical dependence between images to enhance multimodal feature matching accuracy.	Multisensor image registration
[8]	Feature Extraction	Scale-Invariant Keypoints	Identifies rotation/scale-stable keypoints for robust image comparison.	Pattern matching
[9]	Feature Extraction / ML	Machine Learning with REPTree	Uses decision-tree learning to extract meaningful spatial features from satellite images.	Landform mapping
[10]	Feature Extraction	Spectral Residual Saliency	Extracts visually prominent regions by analyzing spectral frequency behavior.	Cloud/ROI extraction
[11]	Feature Extraction	Angular Difference Features (ADF)	Uses angular responses to distinguish man-made structures from natural terrain.	Urban scene interpretation
[12]	Feature Extraction	RJ-MCMC Sampling	Searches parameter space for line-like features such as roads or rivers.	Large-area feature extraction
[13]	Image Segmentation	Seeded Region Growing	Separates regions through similarity-driven growth from seed points.	Land cover segmentation
[14]	Image Segmentation	Cuckoo Search Optimization	Uses Levy flights to optimize multilevel thresholds for complex segmentation.	Color satellite image segmentation
[15]	Segmentation & Classification	Markov Random Fields	Models pixel interactions to obtain smoother and more consistent segmentation.	Land-use mapping
[16]	Deep Learning Segmentation	SegNet	Encoder-decoder network performing pixel-level classification in high-resolution imagery.	Multi-class object segmentation
[17]	Segmentation	Firefly Optimization	Balances threshold selection using light-attraction behavior in swarm algorithms.	Multi-band satellite data
[18]	Segmentation	Graph Model + Gabor Filters	Uses texture features and graph partitioning	Urban scene segmentation

			for more accurate object delineation.	
[19]	Segmentation	Otsu Thresholding	Automatic thresholding maximizing between-class variability.	Gray-scale satellite segmentation
[20]	Segmentation	Minimum Error Thresholding	Selects thresholds that minimize probability of misclassification.	Multilevel classification
[21]	Semantic Segmentation	Fully Convolutional Networks	Converts classification CNNs into fully convolutional models for dense prediction.	Remote sensing semantic labeling
[22]	Segmentation	Tsallis Entropy	Uses generalized entropy for multilevel thresholding in complex imagery.	Color image segmentation
[46]	Change Detection	General Remote-Sensing Techniques	Establishes early framework for detecting temporal variations in multispectral images.	Vegetation/land change mapping
[47]	Change Detection	Comparative CD Strategies	Provides systematic review of change detection categories and methods.	Forest ecology
[48]	Change Detection	GIS-Driven Mapping	Integrates remote sensing with GIS layers for multi-temporal vegetation analysis.	Ecosystem monitoring
[49]	Change Detection	GIS Land Cover Update	Uses spatial overlays to determine urban and agricultural land cover change.	Urban planning
[50]	Change Detection	Spectral Mixture Modeling	Uses temporal unmixing to analyze fractional vegetation changes.	Forest and rangeland assessment
[51]	Change Detection	Digital Surface Models	Derives elevation-based indicators to detect newly built structures.	Urban expansion
[52]	Classification	Maximum Likelihood Method	Probabilistic classifier assigning pixels based on statistical class distributions.	Land cover classification
[53]	Time Series CD	Seasonal Trend Detection	Introduces multi-temporal analysis to detect gradual and abrupt forest disturbances.	Environmental monitoring

[54]	Time Series CD	BFAST Algorithm	Identifies multiple structural breaks in satellite-image time series.	Deforestation analysis
[55]	Change Detection	Multi-Step Image Matching	Combines clustering, boundary detection, and matching for rapid disaster change mapping.	Disaster response
[56]	Change Detection	Graph-Based Temporal Modeling	Tracks changes by comparing object graphs across time steps.	Object-based change detection
[66]	Classification	Remote Sensing Fundamentals	Describes foundations for interpretation, classification, and sampling of RS images.	Land cover mapping
[67]	Classification	Random Forest Optimization	Enhances land cover models by fine-tuning tree and feature selection parameters.	Multisensor classification
[68]	Classification	k-NN Classifier	Introduces instance-based learning for non-parametric classification.	Class label prediction
[69]	Classification	Random Forest for Wetlands	Uses geometric/spectral features for wetland differentiation.	Hydrological mapping
[70]	Classification	Spectral Unmixing	Breaks down mixed pixels into constituent component fractions.	Mixed land cover analysis
[71]	Classification	Hybrid Unmixing	Combines supervised & unsupervised unmixing to handle complex pixel mixtures.	Multispectral classification
[72]	Classification	MESMA	Adapts endmembers per pixel to improve urban material identification.	Urban land cover
[73]	Classification	RBF Neural Networks	Applies radial basis functions for robust classification even in noisy data.	Nonlinear classification
[74]	Classification	Error Correction via DWT	Removes misclassified pixels by analyzing wavelet-merged spectral bands.	Post-classification refinement
[75]	Classification	Deep Feature Learning	Utilizes pyramid pooling and deep CNNs	High-resolution classification

			for multiscale image recognition.	
[76]	Classification	Bayesian Cloud Detection	Classifies cloud types using probabilistic graphical models.	Cloud masking
[77]	Classification	Fuzzy + Dempster-Shafer	Combines fuzzy evaluation with evidence theory for reliable classification validation.	Decision support
[78]	Feature Detection	Fuzzy Edge Detection	Uses fuzzy rules to detect uncertain or weak edges in grayscale imagery.	Edge detection
[79]	Feature Detection	Fully Convolutional Networks	Learns pixel dependencies for outlining complex structures like settlements.	Informal settlement mapping
[82]	Feature Detection	DSM via Image Matching	Generates dense surface models by matching stereo satellite imagery.	Elevation modeling
[83]	Feature Detection	Jitter Detection using Kalman Filtering	Compensates for platform-induced jitter to stabilize satellite images.	Satellite motion correction
[84]	Feature Detection	ML-Based Oil Spill Detection	Identifies slick regions using machine learning on SAR imagery.	Maritime surveillance
[85]	Feature Detection	Transfer Learning for Vehicles	Detects small vehicles using sparse coding and super-resolution learning.	UAV / aerial detection
[86]	Metrics	Confusion Matrix Framework	Establishes rules for evaluating classification accuracy using confusion tables.	Accuracy assessment
[87]	Metrics	Classification Accuracy Standards	Provides accepted guidelines for validating thematic maps.	Land cover validation
[88]	Metrics	Precision–Recall Framework	Explains statistical interpretation of precision, recall, and F-measure.	Classifier evaluation
[89]	Metrics	Land Cover Change Case Study	Demonstrates multi-year change assessment using confusion matrices.	Watershed monitoring

[90]	Metrics	Multinomial Distribution	Describes statistical testing for multi-class classification outcomes.	Accuracy significance testing
[91]	Metrics	Quality Control Column Set	Introduces per-class quality thresholds during map validation.	Reference-based evaluation
[92]	Metrics	Information Retrieval Metrics	Defines F-measure and related evaluation metrics for classification.	Pattern recognition

The summarized reference underlines the richness and variety of research which has resulted in modern satellite image processing. The table describes the contribution of each method to the solution of specific problems caused by sensor differences, complex environments and computational limits regarding enhancement, segmentation, feature extraction, classification and change detection. The collective insights from these studies not only clarify the progression of techniques but also identify consistent gaps where current approaches are struggling such as mixed pixel handling, multi-sensor fusion reliability and fine-grained feature discrimination. Such observations reinforce the importance of a holistic, methodologically integrated framework in the design of new solutions. We extend this discussion in the next section by describing the steps and architectural decisions and experimental strategies taken in the present research.

Performance Evaluation

Quantitative and qualitative evaluation are available for analyzing the performance of different methods. Various performance metrics are calculated and compared with existing methods to evaluate the improvement in performance. A few of the computed parameters are explained in detail which include Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE), Feature Similarity Index (FSIM) and Structural Similarity Index (SSIM).

PSNR finds the quality of the final image. Equation (1) computes PSNR as:

$$\text{PSNR} = 10 * \log_{10} ((\text{MAX}_I^2) / \text{MSE})$$

where MSE is the Mean Square Error. Equation (2) calculates the MSE as:

$$\text{MSE} = (1 / (M * N)) * \sum_{i=1}^M \sum_{j=1}^N (I(i,j) - I'(i,j))^2$$

where $I(i,j)$ is the original image and $I'(i,j)$ represents the thresholded image. PSNR and MSE are mainly used to analyze the quality of compressed and reconstructed image. SSIM measures the structural similarity between the source and final image while FSIM measures the feature similarity between the source and final image. There are two features that can be viewed in FSIM. They are Phase congruency and Gradient magnitude. Phase congruency (PC) is a dimensionless quantity which is significant in local structure map and is a primary feature in FSIM. PC is contrast invariant with no effect of contrast information on human visual system. Gradient magnitude (GM) is another important feature in FSIM. PC and GM are complementary to one another in describing the image local quality. It is computed using Equation (3) as

$$\text{FSIM}(x) = (\sum_x [S_L(x) * PC_m(x)]) / (\sum_x PC_m(x)) \quad (3)$$

where x represents the whole image, $S_L(x)$ denotes the similarity in the two images and PC_m is the phase congruency map.

SSIM is calculated using Equation (4):

$$\text{SSIM}(x,y) = ((2\mu_x\mu_y + C1)(2\sigma_{xy} + C2)) / ((\mu_x^2 + \mu_y^2 + C1)(\sigma_x^2 + \sigma_y^2 + C2)) \quad (4)$$

where μ_x and μ_y represents the sample means of x and y , respectively; σ_x and σ_y give the sample variances of x and y , respectively; and σ_{xy} describes the sample correlation coefficient between x and y and x and y are local windows in the input images. SSIM and FSIM values vary between 0 and 1 with a value 1 representing best similarity. These values can analyze the extend by which the image has recovered after the image is distorted.

Performance evaluation in classification tasks relies on quantifying the number of correctly and incorrectly predicted instances across all classes. These counts are typically organized into a confusion matrix, a matrix-based representation that summarizes classifier performance with respect to a designated test dataset. The confusion matrix is indexed by the true class labels on one axis and the predicted class labels on the other, allowing a clear visualization of how well the classifier distinguishes among categories [86]. It provides a direct comparison between observed and predicted outcomes, making it possible to identify whether mislabeling has occurred and to measure the correctness of predictions [87]. Correct classifications appear along the main diagonal, whereas misclassifications are positioned in the off-diagonal elements, simplifying the task of locating errors within the dataset. From this matrix, four fundamental metrics True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are derived. TP denotes instances correctly labeled as belonging to the positive class; TN denotes correct identification of negative cases; FP refers to negative samples incorrectly labeled as positive; and FN corresponds to positive samples mislabeled as negative [88].

For example, Reference [89] analyzes the land cover change detection for a watershed using Landsat 7 ETM+ imagery of 2007 and 2018. The confusion matrix has six land cover classes, i.e. forest, water bodies, agricultural land, grassland, bare land and urban areas. Out of 1470 evaluated pixels, 60 pixels were correctly classified as water bodies, 20 pixels each as forests and agricultural land, 290 pixels as grassland, 90 pixels as urban areas and 750 pixels as barren land. The rest of the pixels are misclassified. This causes the classification errors. For example, for the water bodies, TP is the pixels that are correctly classified as water, TN is all the non-water classes that are correctly classified, FP is the non-water pixels that are incorrectly classified as water and FN is the water pixels that are assigned to other classes.

The Overall Accuracy (OA) is the sum of diagonal entries (correctly classified pixels) divided by the total number of pixels in the confusion matrix. In the comparison of two classified maps, the values of the confusion matrix are often assumed to be multinomially distributed, assuming random allocation of classes [90]. However, this assumption breaks down when reference truth data are available, since true class labels are fixed within columns and cannot exist outside of their assigned categories. To overcome this limitation, the Quality Control Column Set (QCCS) framework [91] has been proposed that allows for quality assessment per category and checks whether predefined accuracy requirements are met.

Other evaluation metrics like Accuracy, Recall, Precision and F-Measure are also computed from the confusion matrix [92]. Accuracy: It is the ratio of correctly predicted observation to the total observation. It is useful when the data set is balanced. Recall is a measure of the completeness of positive predictions. In a binary classification setting, recall is the ratio of true positives to the sum of true positives and false negatives. For multiclass imbalanced datasets, recall is defined as the sum of true positives across all classes divided by the sum of true positives and false negatives for the classes. Its value ranges from 0 to 1, with 1 being a perfectly complete. Precision is the ratio of correctly predicted positive observations to all predicted positive observations. For binary problems it is true positives over the sum of true positives and false positives. In the multiclass imbalanced settings, the precision is calculated as follows: , where TP is the true positive and FP is the false positive. Imbalanced datasets, characterized by an unequal distribution of classes,

present challenges for classification algorithms, as most machine learning algorithms assume equal representation of classes. Such imbalance is common in real-world applications like fraud detection and spam filtering. Recall and precision are defined in (5) and (6), respectively.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (5)$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (6)$$

where TP or True Positive is when the sample is correctly classified as positive, FP or False Positive is when the sample is incorrectly classified as positive and FN or False Negative is when the sample is incorrectly classified as negative.

F-measure is a parameter which decides the image accuracy. Equation (7) computes F-measure as:

$$\text{F_measure} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (7)$$

A high value of recall and precision guarantees a high F-measure. Precision, recall and F-measure are applicable mainly in pattern recognition and classification.

The significance of the confusion matrix lies in the fact that it can point out the errors made by the classifier.

ROC or Receiver Operating Characteristic curve is a graph representing the classification performance. Area Under Curve (AUC) is a measure of the two-dimensional area under ROC curve. ROC is plotted between True Positive Rate (TPR) and False Positive Rate (FPR). TPR is the ratio of the positive samples that are correctly considered as positive w.r.t all positive samples. FPR is the ratio of negative samples which are misclassified as positive w.r.t all negative samples. AUC ranges between 0 and 1 and closer the value to 1, better is the classification performance. Similar to ROC curve is precision-recall curve which is focused towards the performance of the minority class. It can analyze a single classifier but challenging for classifier comparisons. Precision-Recall AUC can also be computed and is useful for imbalanced classification.

Conclusion

The paper reviews in detail methods for processing satellite images for feature extraction, segmentation, classification, enhancement and change detection, and their application to a wide range of remote sensing problems. The review shows that the performance of these techniques is very sensitive to image characteristics, computational complexity and the specific requirements of the application domain. Traditional methods like histogram based enhancement, clustering and algebraic change detection are relevant but may have limitations in mixed pixels, noisy inputs or large scale datasets. Advanced approaches incorporating swarm intelligence, fuzzy logic, spectral unmixing and deep learning architectures have been developed to improve the analysis of complex spatial patterns, the recognition of nonlinear features, and to achieve higher levels of automation and accuracy.

The study also found problems with segmentation quality, boundary ambiguity, and classifier performance that affected the reliability of the satellite-derived information. While evaluation techniques like confusion matrices, accuracy metrics, and probabilistic quality control frameworks offer valuable insights into the effectiveness of classification, they also emphasize the need for robust and balanced data sets. Moreover, the combination of machine learning and optimization algorithms has emerged as a promising direction to handle problems of feature variability, data imbalance, and multi-sensor information fusion.

The results indicate that there is no best technique. In general the selection of an appropriate processing technique should be dictated by the nature of the imagery, the presence of noise and the objectives of the analysis. Improved sensor technology, as well as the development of artificial intelligence and image processing techniques will enhance

the ability to extract meaningful information from satellite data. Future work should be on hybrid and adaptive frameworks that combine the strengths of the conventional and modern methods for more accurate, scalable and application-specific solutions for remote sensing problems.

References

- [1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Pearson, 2008.
- [2] N. Sengee and H. Choi, "Brightness preserving histogram equalization," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 4, pp. 1803–1810, 2008.
- [3] X. Li and Y. Wang, "Image enhancement using modified differential evolution algorithm," *Expert Systems with Applications*, vol. 37, no. 7, pp. 6025–6032, 2010.
- [4] C. Fu, Y. Zhao, and M. Zhou, "Fractional differential and its application in image enhancement," *Signal Processing*, vol. 92, no. 9, pp. 2143–2151, 2012.
- [5] P. Lewis et al., "Image fusion using rational polynomial coefficients," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 78, pp. 109–118, 2013.
- [6] D. Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, vol. 60, no. 2, pp. 91–110, 2004.
- [7] F. Maes et al., "Multimodality image registration by maximization of mutual information," *IEEE Trans. Medical Imaging*, vol. 16, no. 2, pp. 187–198, 1997.
- [8] D. G. Lowe, "Object recognition using scale invariant features," *Proc. ICCV*, 1999.
- [9] J. Quinlan, "C4.5 decision trees and improvements," *Machine Learning*, Springer, 1993.
- [10] X. Hou and L. Zhang, "Saliency detection using spectral residual," *CVPR*, 2007.
- [11] J. A. Richards, "Angular feature-based classification of urban structures," *IEEE Trans. Geoscience and Remote Sensing*, vol. 38, no. 6, pp. 2643–2654, 2000.
- [12] J. Green and P. Revesz, "Reversible jump Markov chain Monte Carlo for image feature extraction," *Pattern Recognition*, vol. 35, pp. 999–1012, 2002.
- [13] R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.
- [14] X.-S. Yang and S. Deb, "Cuckoo search: A metaheuristic optimization algorithm," *Nature Inspired Cooperative Strategies*, 2009.
- [15] J. Sethian, "Markov random fields for image segmentation," *SIAM Review*, vol. 39, no. 1, pp. 1–17, 1997.
- [16] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. PAMI*, 2017.
- [17] X. Yang, "Firefly algorithm for multi-thresholding," *Applied Soft Computing*, vol. 12, no. 11, pp. 3632–3643, 2012.
- [18] Z. Wang and B. Zheng, "Graph-based segmentation using Gabor filters," *Pattern Recognition Letters*, vol. 32, no. 15, pp. 2010–2023, 2011.
- [19] Otsu N., "A threshold selection method from gray-level histograms," *IEEE Trans. SMC*, 1979.
- [20] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognition*, 1986.
- [21] J. Long et al., "Fully convolutional networks for semantic segmentation," *CVPR*, 2015.
- [22] S. Sahoo et al., "A survey on multilevel thresholding," *Computer Vision and Image Understanding*, 2015.
- [46] A. Singh, "Digital change detection techniques using remotely-sensed data," *International Journal of Remote Sensing*, 1989.
- [47] P. Coppin and M. Bauer, "Digital change detection in forest ecosystems," *Remote Sensing Reviews*, 1996.
- [48] J. Franklin, "Mapping vegetation using GIS and remote sensing," Wiley, 2001.
- [49] Y. Chen et al., "Land cover change detection using GIS and remote sensing," *Remote Sensing of Environment*, 2003.

- [50] Roberts et al., "Temporal unmixing for vegetation analysis using spectral mixture models," *Remote Sensing of Environment*, 1998.
- [51] P. Gamba et al., "Urban DSM for change detection," *IEEE Trans. Geoscience and Remote Sensing*, 2000.
- [52] J. Richards, "Maximum likelihood classification," *Remote Sensing Digital Image Analysis*, Springer, 2013.
- [53] J. Verbesselt et al., "Forest change detection using time-series," *Remote Sensing of Environment*, 2010.
- [54] J. Verbesselt et al., "BFAST: Monitoring change in time-series of satellite imagery," *Remote Sensing of Environment*, 2012.
- [55] Y. Zhang, "Multi-step image matching for change detection," *Photogrammetric Engineering*, 2001.
- [56] X. Huang, "Object-based temporal analysis for change detection," *IEEE GRSL*, 2011.
- [66] T. M. Lillesand, R. W. Kiefer, and J. Chipman, *Remote Sensing and Image Interpretation*, Wiley, 2008.
- [67] Gislason et al., "Random forests for land cover classification," *Remote Sensing of Environment*, 2006.
- [68] T. Cover and P. Hart, "Nearest neighbor classification," *IEEE Trans. Information Theory*, 1967.
- [69] J. Corcoran et al., "Wetland classification using random forests," *Remote Sensing*, 2013.
- [70] N. Keshava, "Spectral unmixing of multispectral images," *IEEE Signal Processing Magazine*, 2003.
- [71] Plaza et al., "Hybrid spectral unmixing for land cover classification," *IEEE Trans. Geoscience and Remote Sensing*, 2009.
- [72] Roberts et al., "MESMA for urban land cover," *Remote Sensing of Environment*, 1998.
- [73] Moody and Darken, "Radial basis function networks," *Neural Computing Surveys*, 1989.
- [74] F. Melgani, "Classification error correction using wavelets," *IEEE GRSL*, 2004.
- [75] He et al., "Deep feature learning for satellite image classification," *CVPR*, 2014.
- [76] S. Kokhanovsky, "Cloud classification using Bayesian networks," *Atmospheric Research*, 2004.
- [77] Shafer and Dempster, "A mathematical theory of evidence," *Princeton Univ. Press*, 1976.
- [78] D. Dubois and H. Prade, "Fuzzy logic approaches in image edge detection," *IEEE Trans. SMC*, 1992.
- [79] J. Long et al., "Fully Convolutional Networks for dense prediction," *CVPR*, 2015.
- [82] M. Hollaus et al., "DSM generation from image matching," *ISPRS Journal of Photogrammetry*, 2010.
- [83] J. Shi et al., "Jitter detection in satellite imagery using Kalman filters," *IEEE Trans. Aerospace and Electronic Systems*, 2012.
- [84] Solberg et al., "Machine-learning-based oil spill detection," *IEEE Trans. Geoscience and Remote Sensing*, 2006.
- [85] Zou et al., "Vehicle detection in aerial images using transfer learning," *IEEE Trans. Image Processing*, 2015.
- [86] J. Stehman, "Confusion matrices in classification accuracy," *Remote Sensing of Environment*, 1997.
- [87] Congalton, "Accuracy assessment of remote sensing-derived classifications," *Photogrammetric Engineering*, 1991.
- [88] Powers, "Evaluation: Precision, recall, F-measure," *Journal of Machine Learning Technologies*, 2011.

- [89] J. Lu et al., "Land cover change mapping using Landsat ETM+," Remote Sensing, 2016.
- [90] Agresti and Coull, "Statistics for multinomial distributions," The American Statistician, 1998.
- [91] Pontius and Millones, "Quality-Control Column Sets (QCCS)," Landscape Ecology, 2011.
- [92] van Rijsbergen, Information Retrieval, 2nd ed., Butterworths, 1979.