

Article

Evaluation of Multiple Classifier Systems for Landslide Identification in LANDSAT Thematic Mapper (TM) Images

Luiz Augusto Manfré ^{1,*}, Rodrigo Affonso de Albuquerque Nóbrega ²
and José Alberto Quintanilha ³

¹ Transportation Engineering Department of Polytechnic College, University of Sao Paulo, São Paulo 05507-070, Brazil

² Cartography Department of Geosciences, Institute of Federal, University of Minas Gerais, Belo Horizonte 31270-901, Brazil; raanobrega@ufmg.br

³ Department of Transportation Engineering, Polytechnic School of the University of Sao Paulo, São Paulo 05507-070, Brazil; jaquinta@usp.br

* Correspondence: luizmanfre@gmail.com; Tel.: +55-11-3091-5504

Academic Editors: Jason K. Levy and Wolfgang Kainz

Received: 24 May 2016; Accepted: 1 September 2016; Published: 13 September 2016

Abstract: Landslide scar location is fundamental for the risk management process, e.g., it allows mitigation of these areas, decreasing the associated hazards for the population. Remote sensing data usage is an essential tool for landslide identification, mapping, and monitoring. Despite its potential use for landslide risk management, remote sensing usage does have a few drawbacks. The aforementioned events commonly occur at high steep slope regions, frequently associated with shadow occurrence in satellite images, which impairs the identification process and results in low accuracy classifications. In this sense, this paper aims to evaluate the accuracy of different ensembles of multiple classifier systems (MCSs) for landslide scar identification. A severe landslide event on a steep slope with a high rainfall rate area in the southeast region of Brazil was chosen. Ten supervised classifiers were used to identify this severe event and other possible features for the LANDSAT thematic mapper (TM) from June of 2000. The results were evaluated, and nine MCSs were constructed based on the accuracy of the classifiers. Voting was applied through the ensemble method, coupled with contextual analysis and random selection tie-breaker methods. Accuracy was evaluated for each classification ensemble, and a progressive enhancement in the ensemble accuracy was noted as the least accurate classifiers were removed. The best accuracy for landslide identification emerged from the ensemble of the three most accurate classification results. In summary, MCS application generally improved the classification quality and led to fewer omission errors, coupled with a better classification percentage for the ‘landslide’ class. However, the MCS ensemble algorithm selection must be customized to the purpose of the classification. It is crucial to assess single accuracy indicators of each algorithm to ascertain those with the most consistent performance regarding the final results.

Keywords: multiple classifier system; landslide scars; accuracy evaluation

1. Introduction

Several regions in the world are affected by high rainfall rates over short periods of time, which are conducive to natural disasters [1]. In mountainous regions, these events favor the occurrence of landslides [2,3] and present risks to the population that lives in or travels through these regions. As a result of construction and changes to the soil structure, road edges favor this type of event [4,5]. A previous study [6] analyzed the frequency and distribution of landslides in three hydrographic

basins and found that road construction is the most common cause of landslides. Monitoring and identifying locations that are prone to landslides is extremely complicated because the teams that are responsible for monitoring natural disasters are usually small, whereas the areas that need to be monitored are large [7]. Despite these difficulties, mapping and mitigating the damage caused by landslides is important for ensuring the population's safety [8].

The use of remote sensing data offers great potential for monitoring and managing landslides. Remote sensing is a good alternative to the mapping of landslide scars because it can cover large regions and it allows for rapid analyses [9]. However, landslides usually occur in mountainous regions, where the effects of topography, combined with the shadows of the relief and the vegetation, can hide landslide scars [10]. Therefore, remote sensing techniques should be developed to address these problems.

Several resources can be used to obtain satellite image information. Enhancement techniques can facilitate object identification [11], whereas classification algorithms use statistical information of the images to separate classes of interest [12]. However, the criteria must be well defined and appropriate for the class/object of interest, to minimize the possibility of inaccurate results [13]. In addition, all algorithms require parameter adjustments based on the main purpose of the classification to achieve the best performance [14]. However, despite all preventative measures, the divergence between the results of different algorithms may be extremely high.

For this reason, there has been a search for alternatives that combine classifications to achieve a realistic final result, that has been discussed in the literature [15–21]. Known in the literature as multiple classifier systems (MCSs) [22] or classifier ensembles [23], combined classifications can be employed using several approaches. There are at least three categories of classification ensembles: algorithms that are based on the manipulation of training samples [24], concatenation combinations [25], and parallel combinations [26].

In a literature review of the use of MCSs, it was shown [27] that the efficiency of classification ensembles is based on continually improving the accuracy of the results. The authors emphasize that selecting the most appropriate ensemble strategy for the classification purpose is fundamental for the use of MCSs. According to [28], the ensemble of algorithms must be developed cautiously. The most important step for obtaining good results is the selection of the most appropriate algorithms to solve the problem.

Previous evaluations of the classifiers to be used are crucial for understanding their performance while obtaining the desired result. According to [29,30], an integrated analysis of various accuracy indicators is important in order to better understand the results of a classification.

This study is aimed at identifying the best classifier ensembles for mapping landslide scars in mountainous regions in medium spatial resolution satellite images (30 m). In addition, this paper evaluates the contribution of diversity measures of classifications to the final result of the combination accuracy of the algorithms.

To meet the proposed objective, this paper presents an introduction into MCSs and landslide identification; in the material and methods section, the study area characteristics are presented, followed by the procedures that were used. The results section presents the classification results for each algorithm and their accuracy assessment, as well as the best MCS classification and the accuracy analysis for each MCS. The results are followed by a discussion and then conclusions.

2. Material and Methods

2.1. Study Area

The study area is located in the Serra do Mar mountain range in the state of São Paulo, Brazil. A square of 144 km² was defined to include a severe landslide that occurred in December 1999 at the hydrographic basin of the Pilões River. The landslide occurred after four days of heavy rain with 230 mm of total precipitation. The landslide displacement affected 700 m of the Anchieta Highway

at approximately kilometer 42 of the highway [31]. Figure 1 presents the study area location (Black Square) and the Pilões River landslide location. The selected square encompasses two important roads that connect the city of Sao Paulo to the largest port in Latin America. The region was also chosen because of the difficulty for landslide scar mapping, as it presents a complex landscape with many shadows.



Figure 1. Location of the study area (Black Square) and a picture of the landslide scar at the Pilões watershed. Source: Landslide picture from [32].

The region has an annual precipitation of more than 3000 mm. The most intense rainfall occurs between November and March. The rainiest months have experienced total rainfalls of more than 1000 mm [33]. Therefore, this event was used as a sample to assess the capability of the methodology and identify other landslide scars in the study area.

2.2. Procedures

Figure 2 summarizes the methodology that was used in this study in a flowchart that lists the steps for the digital image processing.

First, the study area square was extracted from the LANDSAT TM 5 scene taken on 25 June 2000, and was pre-processed to improve its characteristics. The image was corrected to reflectance, registered to the Digital Elevation Model, and was orthorectified using the Rational Polynomial Coefficients algorithm. To facilitate the visual identification of landslide scars, which increased the quality of the training areas, enhancement techniques were used to highlight the landslide area. Color conversion from red, green, and blue (RGB) to hue, saturation, and value (HSV) was adopted, and bands 3 (red), 4 (near infrared), and 5 (mid-infrared) were used. According to [34], the ensemble selection increases the differences between the soil and vegetation. The conversion to the HSV color space was used to soften the shadow effects of the topography [35]. Figure 3 shows the sample scene in the HSV color space and highlights the landslide area.

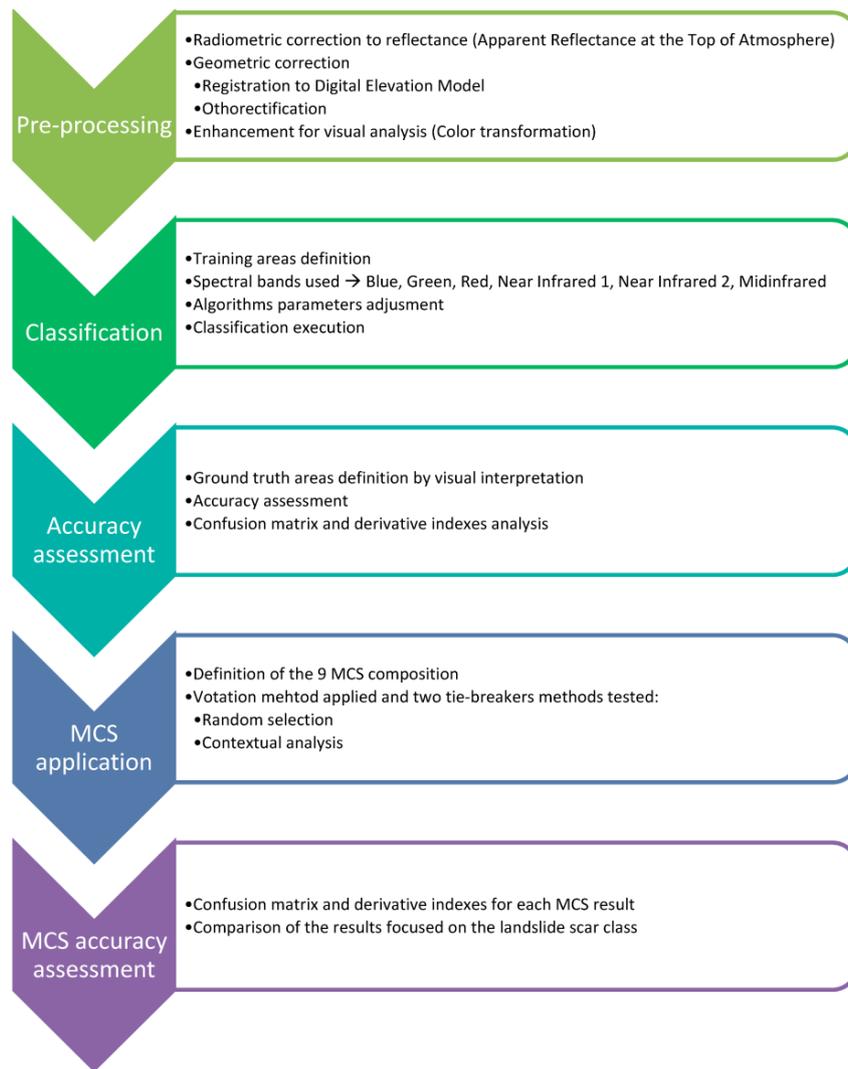


Figure 2. Flowchart of the methodology used in this study.

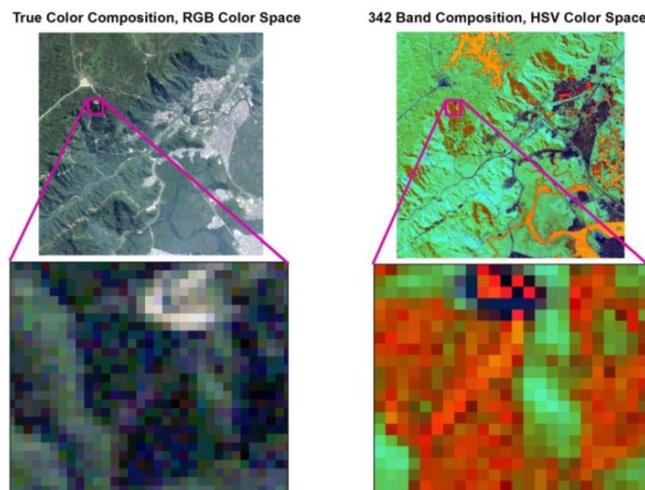


Figure 3. Sample of the LANDSAT 219-077 scene (25 June 2000) displayed in the RGB and HSV color space.

To proceed with the classification, training areas were defined based on the user's previous knowledge about the area, especially of the area of the known landslide at the Pilões watershed, using the HSV color image. The study area was classified by 10 different classification algorithms. The six bands of the LANDSAT TM 5 satellite were used to prevent excluding any available information.

This study used 10 commonly used supervised classification algorithms in the literature [21,36–40] to classify the LANDSAT scenes. In addition, a large number of algorithms were applied to achieve diverse classification results. According to the theoretical background of classifier systems, diversity measures play an important role in the final result [27,41].

All algorithms required parameter adjustments to achieve maximum performance [42]; therefore, the best adjustments were considered to separate the classes of landslide scars. These parameters were defined based on the characteristics of each algorithm as well as the characteristics of the class to be identified. The 'landslide' class has extremely specific characteristics; however, it is mistaken for other classes in some of the LANDSAT 5 bands. Thus, the parameters of this class were highly restrictive; i.e., they allowed for the minimum variation within the class.

The parameters of the supervised classification algorithms were adjusted based on their responses to the classification of the landslide area along the Pilões River. The decision tree algorithm was applied based on the rules that were established by the J48 classifier in the WEKA 3.6 data mining software. J48 is a powerful classifier that is used in remote sensing [43–45]. The algorithms that were used are described below:

- Supervised
 - Support Vector Machine (SVM)—Kernel Type = Radius Bias Function, Gamma in Kernel Function = 0.167, Penalty Parameter = 500, Classification Probability Threshold = 0
 - Neural Network (NN)—Activation = Logistic, Training Threshold Contribution = 0.9, Training Rate = 0.5, Training Momentum = 0, Training RMS Exit Criteria = 0.1
 - Binary Encoding (BE)—Minimum Encoding Threshold for Landslide class = 3
 - Spectral Information Divergence (SID)—Maximum Divergence Threshold for Landslide class = 0.005
 - Spectral Angle Mapper (SAM)—Maximum Angle for Landslide class = 0.047
 - Maximum Likelihood Classification (MLC)—Probability Threshold for Landslide class = 0.05
 - Mahalanobis Distance (MHD)—Maximum Distance Error = 1.5
 - Minimum Distance (MD)—Maximum Standard Deviation from Mean for Landslide class = 1.1, Maximum Distance Error = 10
 - Parallelepiped (PAR)—Maximum Standard Deviation from Mean for Landslide Class = 0.72
- Decision Tree (DT)
 - J-48 Data Mining

The classification results for each algorithm were evaluated based on their accuracy, and 9 ensembles from the 10 classifiers were defined using fewer algorithms with greater accuracy. The classification ensembles were created using the majority voting principle. In the case of a tie, the pixel-based context analysis or the random selection method was used. The accuracy of each of the ensembles was evaluated, and an analysis of the accuracy evolution was performed.

The performances of all of the classification results were evaluated based on ground truth areas defined by visual interpretation. The following accuracy indicators were used: the kappa index, overall accuracy, probability of correct classification for the 'landslide' class, and commission errors and omission errors for the "landslide" class [46,47].

Based on the evaluation of the accuracy of the classification results, it was possible to define the classifiers with the best performance as well as to progressively define the algorithm ensembles, which

eliminated the worst results from the ensemble. Nine classifier ensembles were defined: one consisted of 10 algorithms, and the others progressively excluded the algorithms with the worst performances based on the accuracy indicators. Table 1 shows the algorithms that were used in each of the ensembles.

Table 1. Classifiers used in each ensemble.

| Number of Classifiers | Classifiers |
|-----------------------|---|
| 10 | BE, DT, MHD, MD, MLC, NN, PAR, SAM, SID and SVM |
| 9 | DT, MHD, MD, MLC, NN, PAR, SAM, SID and SVM |
| 8 | MHD, MD, MLC, NN, PAR, SAM, SID and SVM |
| 7 | MHD, MLC, NN, PAR, SAM, SID and SVM |
| 6 | MHD, MLC, NN, PAR, SAM and SVM |
| 5 | MHD, MLC, NN, SAM and SVM |
| 4 | MD, MLC, NN, and SVM |
| 3 | MLC, SVM and NN |
| 2 | SVM and NN |

The ensembles were selected based on the best accuracy levels of the classifiers. According to [28], the algorithm selection is crucial for improving the accuracy in a multi-classifier system. Therefore, the ensemble does not use the greatest error sources and has greater potential for improving the final results.

To design these ensembles, all of the classifications were exported to a table format, and the matrices were converted into a single column of 159,201 rows with one row for each pixel. A voting analysis was performed to compute the classification of each algorithm i , and the most recurring class k was attributed to pixel j [19,25]. If a tie occurred between two or more classes k , two strategies were adopted: the random selection method [48] and the nearest neighbor analysis [49]. Therefore, the value F_{ij} was calculated as the final result for each of the tie breaking strategies.

The accuracy of the classifier ensembles and of the classification algorithms was evaluated. Finally, the results of the accuracy indicators for the isolated classifiers and for the 18 ensembles that were developed (9 using the random selection method and 9 using the context analysis method for tie breaking) were compared.

3. Results

Figure 4 shows the results of the classifications for each of the 10 algorithms. The classification results vary widely, especially those for the ‘landslide’ class. The results of some of the classifiers were mixed for the other classes because of the priority that was given to the ‘landslide’ class in the adjustment of the algorithm parameters.

Table 2 briefly shows the evaluation of the classifier accuracy. The commission and omission errors, as well as the probability of correct classification, are presented only for the ‘landslide’ class. The SVM, NN, and MLC classifiers outperformed the others.

The results obtained by the MLC, NN, and SVM algorithms are important for evaluating the kappa coefficient and the overall accuracy. An evaluation of the commission and omission errors shows that the commission errors were always high (greater than 0.6) and that the lowest omission errors were achieved by the MLC and NN algorithms. These results are due to the high spectral similarity between the “landslide” class and the “vegetation” and “water” classes (especially due to the shadows of the topography). The best probability of correct classification combined with the lowest commission errors in the ‘landslide’ class were obtained by the NN, SVM, and MLC algorithms. Other algorithms, such as Binary Encoding, achieved good probabilities of correct classification for the “landslide” class; however, large commission errors were observed. Algorithms such as SAM had low kappa coefficients, but the probability of correct classification and the commission errors for the “landslide” class were better.

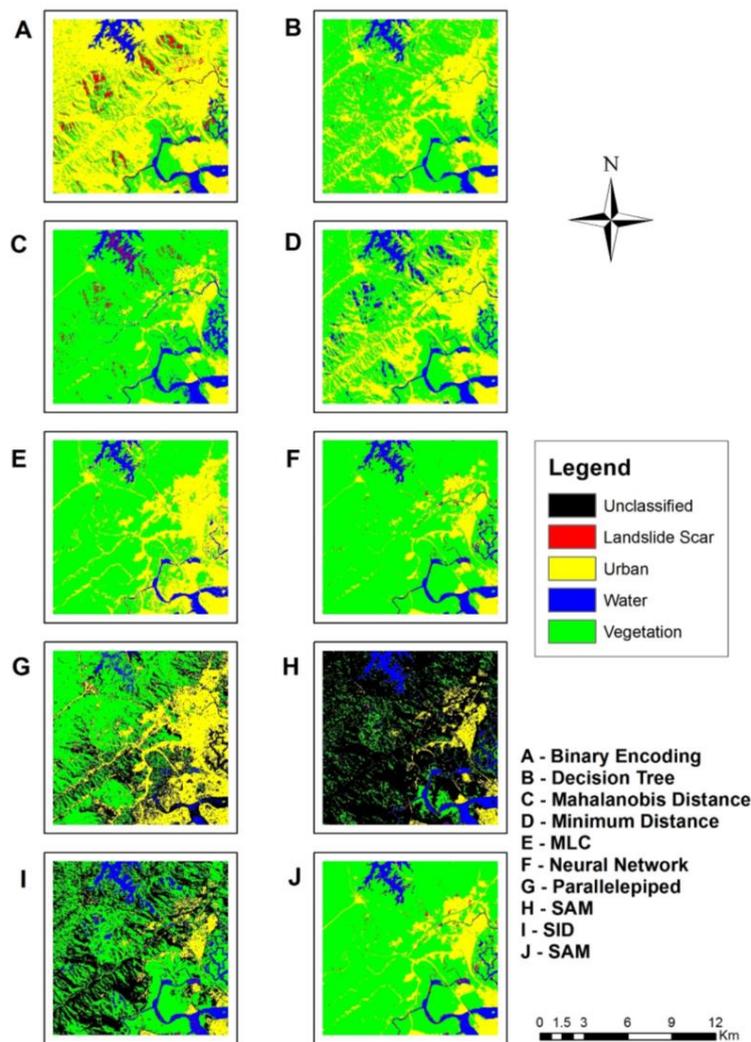


Figure 4. Classification maps for each algorithm. (A) Binary Encoding; (B) Decision Tree; (C) Mahalanobis Distance; (D) Minimum Distance; (E) Maximum Likelihood Classification; (F) Neural Networks; (G) Parallelepiped; (H) SAM; (I) SID; (J) SVM.

Table 2. Accuracy evaluation for each of the 10 classifiers.

| Classifiers | Kappa | Overall Accuracy % | Landslide Class | | |
|-----------------------------------|--------|--------------------|---------------------|-------------------|-----------|
| | | | Commission Errors % | Omission Errors % | % Correct |
| Binary Encoding | 0.3424 | 48.5433 | 93.81 | 48.77 | 61.23 |
| Decision Tree (J48) | 0.8081 | 89.4852 | 82.22 | 77.78 | 22.22 |
| Mahalanobis Distance | 0.7527 | 85.6517 | 94.14 | 54.94 | 43.06 |
| Minimum Distance | 0.5561 | 71.7515 | 80.85 | 62.5 | 37.5 |
| Maximum Likelihood Classification | 0.9359 | 96.6484 | 59.3 | 48.61 | 51.39 |
| Neural Network | 0.9441 | 97.0865 | 59.09 | 25.00 | 75.00 |
| Parallelepiped | 0.638 | 79.759 | 82.81 | 69.44 | 30.56 |
| Spectral Angle Mapper | 0.3002 | 37.678 | 60.00 | 55.56 | 44.44 |
| Spectral Information Divergence | 0.4673 | 62.322 | 70.11 | 63.89 | 36.11 |
| Support Vector Machine | 0.9325 | 96.4513 | 66.49 | 13.89 | 86.11 |

Table 3 shows the evaluation of the results of different classifier combinations. Figures 5–9 show the evolution of the accuracy indicators for the classifier ensembles and the comparison with the results of the best classifier (kappa coefficient, overall accuracy, commission errors, omission errors, and probability of correct classification for the ‘landslide’ class, respectively).

Table 3. Classifier ensembles accuracy evaluation.

| Tie Breaking | Classifiers Number | Kappa | Overall Accuracy % | Landslide Class | | |
|-------------------------|---------------------|--------|--------------------|---------------------|-------------------|-----------|
| | | | | Commission Errors % | Omission Errors % | % Correct |
| Random Selection Method | 10 | 0.8692 | 93.404 | 61.18 | 35.29 | 64.71 |
| | 9 | 0.93 | 96.6313 | 59.21 | 39.22 | 60.78 |
| | 8 | 0.9523 | 97.7385 | 59.77 | 31.37 | 68.63 |
| | 7 | 0.9441 | 97.338 | 52.63 | 29.41 | 70.59 |
| | 6 | 0.963 | 98.2568 | 52.56 | 27.45 | 72.55 |
| | 5 | 0.9689 | 98.5395 | 50.6 | 19.61 | 80.39 |
| | 4 | 0.9704 | 98.6101 | 49.35 | 23.53 | 76.47 |
| | 3 | 0.9694 | 98.563 | 50.6 | 11.11 | 88.89 |
| | 2 | 0.9584 | 98.0448 | 60.4 | 21.57 | 78.43 |
| | Contextual Analysis | 10 | 0.9147 | 95.8539 | 57.14 | 29.41 |
| 9 | | 0.9558 | 97.8998 | 59.26 | 26.67 | 73.33 |
| 8 | | 0.9659 | 98.3966 | 50 | 24 | 76 |
| 7 | | 0.9622 | 98.2097 | 53.49 | 21.47 | 78.43 |
| 6 | | 0.9622 | 98.2097 | 53.49 | 21.57 | 78.43 |
| 5 | | 0.9719 | 98.6805 | 48.15 | 17.65 | 82.35 |
| 4 | | 0.9709 | 98.6337 | 48.75 | 19.61 | 80.39 |
| 3 | | 0.9704 | 98.6101 | 50.59 | 9.52 | 90.48 |
| 2 | | 0.9664 | 98.4217 | 53.68 | 13.73 | 86.27 |

The best results for the “landslide” class were achieved by the ensemble of three classifiers (SVM, NN, and MLC). However, for the overall classification, the best results were achieved by the ensemble of five classifiers. In addition, for all of the ensembles, the contextual analysis method provided better results than the random selection method for tie breaking.

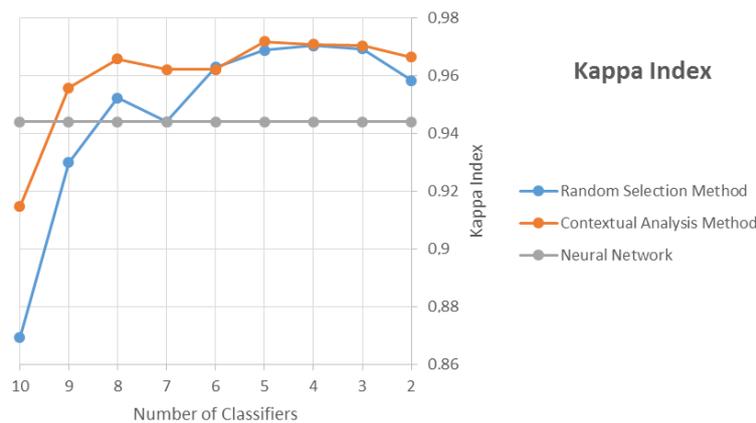


Figure 5. Evolution of the kappa coefficient for the results of the classifier ensembles.

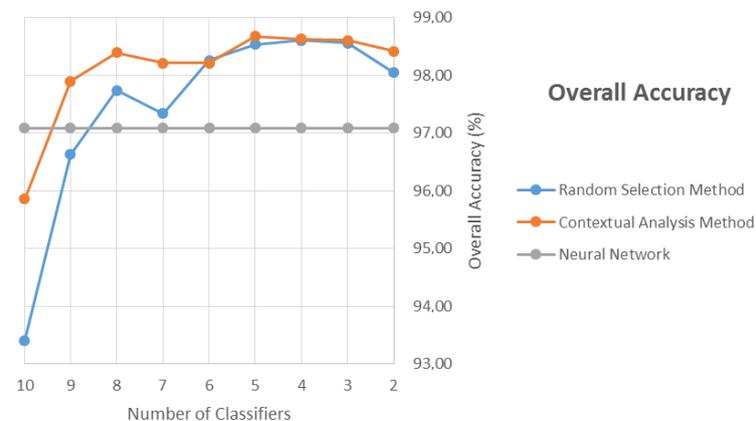


Figure 6. Evolution of the overall accuracy for the classifier ensembles.

Figure 5, which shows the evolution of the kappa coefficient in relation to the classifier ensembles, indicates that the performance of the ensemble improves significantly as the first three algorithms were eliminated, reaches its maximum performance with the ensemble of five algorithms, and decreases slightly for the ensemble of two algorithms. Furthermore, the results of the ensembles are only higher than the highest kappa coefficient of the isolated classifiers (in this case, NN) after the ensemble of nine classifiers. The same evolution is observed for the overall accuracy in Figure 6.

An analysis of the commission errors for the “landslide” class (Figure 7) indicates that the best result is achieved through the contextual analysis approach that consists of five algorithms. A reduction in the commission errors is observed until the five best algorithms are used, and the values then increase (four, three, and two algorithms). A comparison of the best results of the algorithms in isolation indicated that SVM outperformed the others; however, only the classifications that consisted of ten and two algorithms in the random selection approach had worse results.

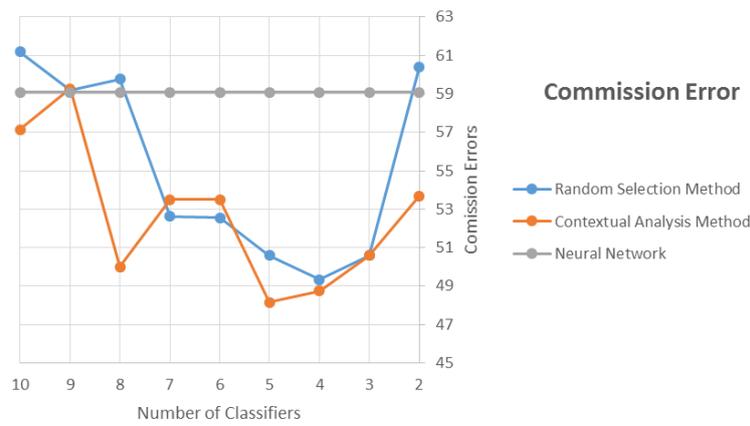


Figure 7. Evolution of the commission errors for the classifier ensembles.

In the analysis of the omission errors for the “landslide” class (Figure 8), the SVM algorithm outperformed the others and was only surpassed by the ensembles with three algorithms. However, the ensemble of 10 algorithms did not provide the worst results; when the random selection approach was used for tie breaking, the result of this ensemble was only surpassed after seven algorithms were used. This may be due to the use of the Binary Encoding algorithm, which resulted in high commission errors but low omission errors. This result is only compensated for when some of the classifiers are eliminated. Although these classifiers give lower commission errors, the omission errors are higher; therefore, other classes are associated with those that are related to landslide areas.

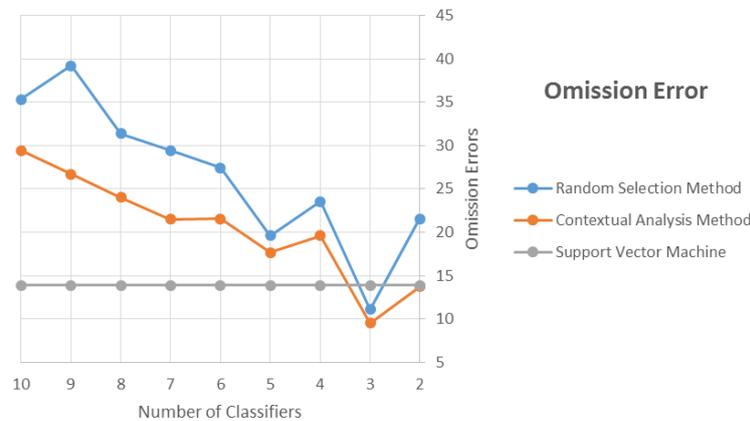


Figure 8. Evolution of the omission errors for the classifier ensembles.

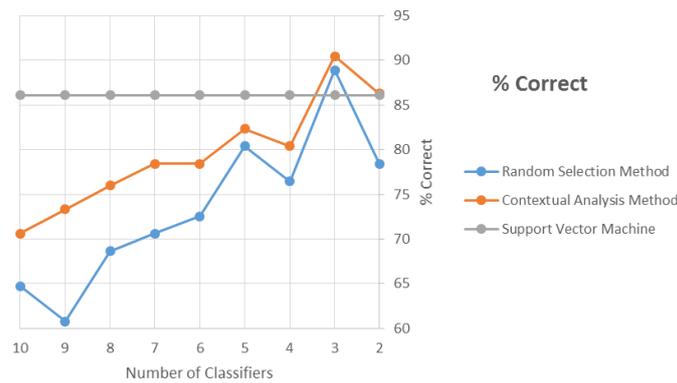


Figure 9. Evolution of the probability of correct classification for the classifier ensembles.

The analysis of the probability of correct classification for the “landslide” class reveals a clear tendency towards improved accuracy as the poorly performing algorithms are eliminated. The ensemble of three algorithms yields the best result when the contextual analysis approach is used for tie breaking. Moreover, the ensembles of three algorithms (SVM, NN, and MLC) are the only ones that produce better results than those obtained by the SVM algorithm in isolation.

Table 4 shows the confusion matrix that was designed with the 10 classifiers, and the final voting result is considered to be ‘true’. Considerable confusion between the “landslide”, “water”, and “vegetation” classes is observed.

Table 4. Ensemble of 10 classifiers confusion matrix. Evaluation of the misclassification among the different classifiers used in this analysis.

| From/To | Unclassified | Landslides | Urban | Water | Vegetation |
|--------------|--------------|------------|--------|--------|------------|
| Unclassified | 0.4123 | 0.1727 | 0.2705 | 0.0663 | 0.2779 |
| Landslide | 0.0141 | 0.4408 | 0.0033 | 0.014 | 0.0082 |
| Urban | 0.2466 | 0.029 | 0.6 | 0.0037 | 0.0948 |
| Water | 0.015 | 0.1811 | 0.0033 | 0.9076 | 0.0094 |
| Vegetation | 0.3121 | 0.1764 | 0.1229 | 0.0083 | 0.6097 |

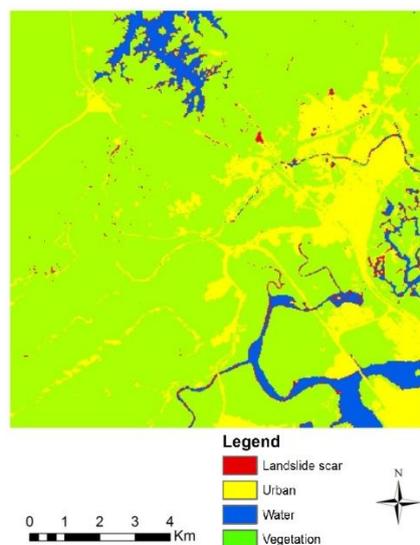


Figure 10. Final classification consisting of three algorithms (SVM, NN, and MLC) and the contextual analysis approach for tie breaking.

Therefore, the ensemble with three classifiers (SVM, NN, and MLC) that used the contextual analysis approach for tie breaking provided the best accuracy indicators. Figure 10 shows the map with the final classification and clearly shows the areas that are classified as landslides.

In addition to the known landslide scar, other relevant features and scars were identified by the methodology: a landslide scar associated with a road (A), another big landslide scar occurred close to an urban area (B) and a steep slope of exposed soil near an oil refinery (C) are presented in Figure 11. In addition to the events above, other landslide features were also identified, such as isolated scars in the mountain and severe erosion areas associated with rivers and reservoir borders and with an oil pipeline along the mountain.

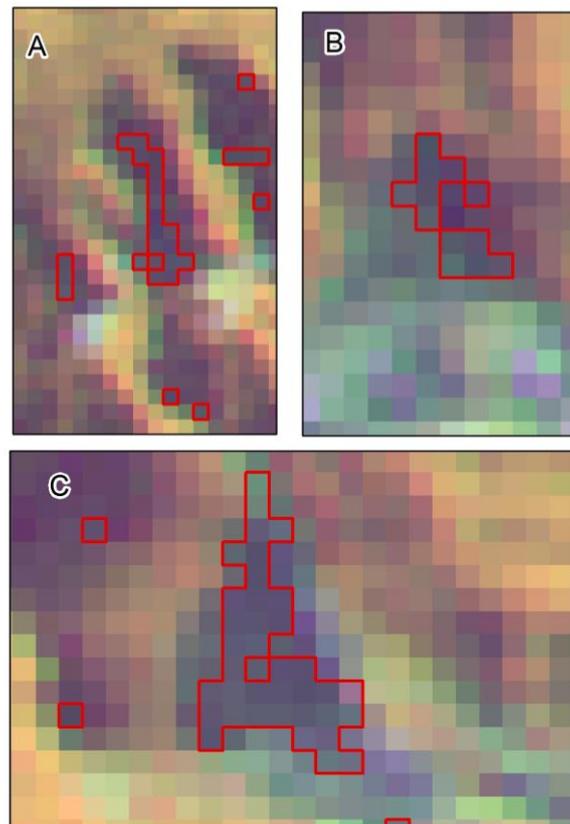


Figure 11. Additional landslide features identified in the study area. (A) landslide events occurred close to a road segment; (B) a landslide detected nearby an urban area; (C) a steep slope of exposed soil near an oil refinery.

4. Discussion

The use of classifier combination techniques efficiently yielded higher quality results by jointly evaluating the overall accuracy, kappa coefficient, omission and commission errors, and the probability of correct classification for the ‘landslide’ class. High accuracy classification in areas of uneven topography is an important issue in optical remote sensing [50], especially when the target of interest may be located in the shadows from the relief. Alternatives such as the use of enhancement techniques with color conversion or the use of vegetation indices [51] facilitate the visual identification of these areas; however, they may not yield satisfactory results for the entire scene. The use of classifier combinations provided better accuracy indicators, introduced fewer omission and commission errors, and showed greater potential for identifying other scars in the studied scene when compared to the results of [52].

The evaluation of the isolated classification results of each of the algorithms demonstrated that some performed well in addressing the problem of landslide identification. Other algorithms had very low accuracies and proved to be inefficient for landslide identification. The DT, SVM, NN, and MLC algorithms produced kappa coefficients greater than 0.8 for the overall results of the classifications. The other classifiers had intermediate or low kappa coefficients and overall accuracies, although these results are reasonable for the “landslide” class, such as in the cases of the SAM and MDC algorithms.

The confusion matrix for the 10 classifications (Table 4) shows the confusion between the classifiers for each of the classes. The “landslide” class is especially confused with the “unclassified”, “water”, and “vegetation” classes. The confusion between the classifiers indicates that the “landslide” class has characteristics that overlap with these other classes, which requires more elaborate classification strategies. According to the literature [27,53], the use of a greater diversity measure between the classifiers improves the ensemble performance; however, some algorithms contribute negatively by increasing the error sources for the final classification.

The evaluation of the results from the combinations indicates that they become more efficient when the algorithm with the worst performance is eliminated. These results are similar to those of previous reports [54,55], which stated that the most important step in the use of combined classifications is the algorithm selection. The omission error for the “landslide” class was the only accuracy indicator in which all of the combinations outperformed the isolated classifier with the best performance, which was the SVM algorithm. This result is related to the high inclusion of pixels in the “landslide” class by various algorithms and leads to high commission errors. In addition, the probability of correct classification in the “landslide” class for the combination of 10 classifiers is lower than the isolated performance of the SVM algorithm.

In general, the use of some classification algorithms negatively affected the results, especially for the “landslide” class. For the overall classification of the LANDSAT sample scene, the best results were achieved by the combinations of five, four, and three algorithms. Although the selection of the algorithm focused on improving the performance of the “landslide” class, the accuracy of the entire classification increases with the elimination of some algorithms.

According to previous studies [27,41], the diversity of the classifiers is important to the classification result, and the combination of similar classifications does not improve the results. However, this result was not observed in this study for the accuracy indicators, especially for the “landslide” class. The use of some algorithms introduces errors into the classification combinations and reduces the accuracy for the class in question [56]. In this case, it is important to evaluate the individual performance of each classifier for the class of interest because the purpose is to increase the accuracy of the specific class. For the purposes of this study, the SVM, NN, and MLC classifiers performed adequately and had lower commission and omission errors and high classification percentages for the “landslide” class. Based on the results of these three algorithms, it was possible to develop a classifier combination to achieve accurate results for landslide mapping.

However, for the combination of two algorithms, all of the accuracy indicators produced results that were inferior to the combination of three algorithms. This result is crucial because it clearly shows that selecting classifiers and adding diversity measures is important for improving the final classification results. However, the information that is used for the classification must have an appropriate accuracy in order to not negatively influence the final result of the combination.

Therefore, the classification must be evaluated based on several accuracy indicators and not only on the kappa coefficient or the overall accuracy [57]. This is especially true for this study, which focused on achieving the best performance for a specific class. For this study, the analysis of the commission and omission errors is extremely important because the purpose is to accurately identify the landslide areas and to not overestimate the mapping of these areas. The results of these accuracy indicators for the classification ensembles were the best of all of the classifiers, especially for the MCS that consists of the three best algorithms (SVM, NN, and MLC).

The use of the random selection method as a decision mechanism in the case of a tie is appropriate; however, statistical tools are not more efficient than complementary spatial information for tie breaking. The use of the geographic-contextual analysis method resulted in more accurate classifications, including appropriate classifications to solve the problem of interest. This result was expected because the use of spatial information influences the decision and avoids randomness.

A comparison of the results that were obtained by combining three algorithms with those reported previously [52] reveals large differences, especially in the omission error and the probability of correct classification for the “landslide” class. The combined use of three algorithms with good accuracy indicators resulted in good performance for the classification of the “landslide” class. Unlike results that have been reported in the literature [58–62], the inclusion of diversity measures negatively affected the classification results, including both the whole scene classification and the classification for the “landslide” class. According to [63], each classifier has better performance for specific cases; therefore, the combination of the three best classifiers for the “landslide” class yielded a better result for this class.

As has been mentioned by several authors [64–66], the use of an MCS generally improved the quality of the classification and resulted in fewer omission errors as well as in a better classification percentage for the ‘landslide’ class. However, the strategy of using algorithms for the MCS ensemble must be analyzed and adapted to the purpose of the classification. It is important to individually evaluate the accuracy indicators of each algorithm to identify those with the performance that is consistent with the final results.

5. Conclusions

In summary, this article presents an evaluation of MCSs focused on the classification accuracy of one single class. In this sense, the applied methodology demonstrates that the inclusion of diversity measures in the classifier ensemble is important for improving the classification. However, the classifiers that are used must be evaluated in order to avoid the introduction of sources of error into the combination.

The study area presents a known severe landslide scar, and the application of the methodology identified other features of interest, considering erosive and landslide perspectives, which were not all identified by the isolated results of the algorithms. This result indicates that the MCS enhanced the result of the classification and improved the identification of landslides through LANDSAT data.

Furthermore, the usage of contextual analysis as a tie-breaker resulted in better accuracy classification, demonstrating that spatial analysis adds more knowledge than statistical approaches.

The presented methodology was applied to a specific region and some characteristics must be considered before its direct replication. The study area presents a dynamic that facilitate the occurrence of severe landslide events, which allows for the usage of Landsat imagery. Other satellites can provide better results, depending on the study area specification. However, the correct usage of MCS enhances the classification result and facilitates the identification of landslide scars.

Acknowledgments: To FAPESP and CNPq for the scholarships provided for the development of this study. To the Geoprocessing Lab for the availability of the software and hardware used in this article. To Mariana Abrantes Giannotti, Ailton Luchiari, Eduardo Macedo Soares, and Amarillis Lucia Casteli Figueiredo Gallardo for reviewing this paper.

Author Contributions: Luiz Augusto Manfré. and José Alberto Quintanilha conceived and designed the experiments; Luiz Augusto Manfré performed the experiments; Luiz Augusto Manfré, José Alberto Quintanilha and Rodrigo Affonso de Albuquerque Nóbrega analyzed the data; Luiz Augusto Manfré, Rodrigo Affonso de Albuquerque Nóbrega and José Alberto Quintanilha wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|-----|-----------------------------------|
| TM | Thematic Mapper |
| MCS | Multiple Classifier System |
| SVM | Support Vector Machine |
| SAM | Spectral Angle Mapper |
| NN | Neural Net |
| BE | Binary Encoding |
| SID | Spectral Information Divergence |
| MLC | Maximum Likelihood Classification |
| MD | Minimum Distance |
| MHD | Mahalanobis Distance |
| PAR | Parallelepiped |
| DT | Decision Tree |

References

- Borga, M.; Stoffel, M.; Marra, F.; Jakob, M. Hydrogeomorphic response to extreme rainfall in headwater systems: Flash floods and debris flow. *J. Hydrol.* **2014**, *518*, 194–206. [[CrossRef](#)]
- Petley, D. Global pattern of loss of life from landslides. *Geology* **2012**, *40*, 927–930. [[CrossRef](#)]
- Wang, H.; Zhang, Y.; Hu, H. A Study on the relationship between the occurrence of Landslides and Rainfall. In Proceedings of the 2nd International Conference on Electric Technology and Civil Engineering, Yichang, China, 8–20 May 2012; pp. 200–203.
- Di Martino, L.; Masciocco, L.; Ricca, G.; Toja, M. Relationships between landslides phenomena and road network: An example from hilly region of Asti Province (North-Western Italy). *Eng. Geol. Soc. Territ.* **2014**, *2*, 1049–1053.
- Kumar, K.; Jangpangi, L.; Gangopadhyay, S. Highway vs. landslides and their consequences in Himalaya. *Landslide Sci. Safer Geoenvironment* **2014**, *1*, 389–395.
- Guthrie, R.H. The effects of logging on frequency and distribution of landslides in three watersheds on Vancouver Island, British Columbia. *Geomorphology* **2002**, *43*, 273–292. [[CrossRef](#)]
- Guzzetti, F.; Mondini, A.C.; Cardinali, M.; Fiorucci, F.; Santangelo, M.; Chang, K.T. Landslide inventory maps: New tools for an old problem. *Earth Sci. Rev.* **2012**, *112*, 42–66. [[CrossRef](#)]
- Manconi, A.; Casu, F.; Ardizzone, F.; Bonano, M.; Cardinali, M.; De Luca, C.; Gueguen, E.; Marchesini, I.; Parise, M.; Vennari, C.; et al. Brief Communication: Rapid mapping of landslide events: The 3 December 2013 Montescaglioso landslide, Italy. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 1835–1841. [[CrossRef](#)]
- Tofani, V.; Hong, Y.; Singhroy, V. Introduction: Remote sensing techniques for landslide mapping and monitoring. *Landslide Sci. Safer Geoenvironment* **2014**, *1*, 301–303.
- Nichol, J.; Wong, M.S. Satellite remote sensing for detailed landslide inventories using change detection and image fusion. *Int. J. Remote Sens.* **2005**, *26*, 1913–1926. [[CrossRef](#)]
- Liu, J.G.; Mason, P. *Essential image processing and GIS for Remote Sensing*; Wiley-Blackwell: New York, NY, USA, 2009; p. 450.
- Lu, D.; Weng, Q.; Moran, E.; Li, G.; Hetrick, S. Remote sensing image classification. In *Advances in Environmental Remote Sensing: Sensors, Algorithms, and Applications, Part II*; CRC Press/Taylor & Francis Group Publishers: Boca Raton, FL, USA, 2011; pp. 219–240.
- Richards, J.A. *Remote Sensing Digital Image Analysis*; Springer: Berlin, Germany, 1999.
- Schott, J.R. *Remote Sensing*; Oxford University Press: Oxford, UK, 2007; p. 513.
- Steele, B.M. Combining multiple classifiers: An application using spatial and remotely sensed information for land cover type mapping. *Remote Sens. Environ.* **2000**, *74*, 545–556. [[CrossRef](#)]
- Briem, G.; Benediktsson, J.; Sveinsson, J. Multiple classifiers applied to multisource remote sensing data. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2291–2299. [[CrossRef](#)]
- Benediktsson, J.A.; Chanussot, J.; Fauvel, M. Multiple classifier systems in remote sensing: From basics to recent developments. In *Multiple Classifier Systems*; Springer: Berlin Germany, 2007; pp. 501–512.
- Doan, H.T.X.; Foody, G.M. Increasing soft classification accuracy through the use of an ensemble of classifiers. *Int. J. Remote Sens.* **2007**, *28*, 4609–4623. [[CrossRef](#)]

19. Foody, G.M.; Boyd, D.S.; Sanchez-Hernandez, C. Mapping a specific class with an ensemble of classifiers. *Int. J. Remote Sens.* **2007**, *28*, 1733–1746. [[CrossRef](#)]
20. Waske, B.; Van der Linden, S.; Benediktsson, J.A.; Rabe, A.; Hostert, P. Sensitivity of support vector machines to random feature selection in classification of hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2880–2889. [[CrossRef](#)]
21. Smits, P.C.; Dellepiane, S.G.; Schowengerdt, R.A. Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach. *Int. J. Remote Sens.* **1999**, *20*, 1461–1486. [[CrossRef](#)]
22. Giacinto, G. An approach to the automatic design of multiple classifier systems. *Pattern Recognit. Lett.* **2001**, *22*, 25–33. [[CrossRef](#)]
23. Kuncheva, L.I.; Whitaker, C.J. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Mach. Learn.* **2003**, *51*, 181–207. [[CrossRef](#)]
24. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [[CrossRef](#)]
25. Xu, L.; Krzyzak, A.; Suen, C.Y. Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Trans. Syst. Man Cybern.* **1992**, *22*, 418–435. [[CrossRef](#)]
26. Rahman, A.F.R.; Fairhurst, M.C. Serial combination of multiple experts: A unified evaluation. *Pattern Anal. Appl.* **1999**, *2*, 292–311. [[CrossRef](#)]
27. Du, P.; Xia, J.; Zhang, W.; Tan, K.; Liu, Y.; Liu, S. Multiple classifier system for remote sensing image classification: A review. *Sensors* **2012**, *12*, 4764–4792. [[CrossRef](#)] [[PubMed](#)]
28. Abe, B.T.; Olugbara, O.O.; Marwala, T. IAENG transactions on engineering technologies. In *Classification of Hyperspectral Images Using Machine Learning Methods*; Springer Netherlands: Dordrecht, The Netherlands, 2014; pp. 555–569.
29. Pontius Junior, R.G.; Millones, M. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429. [[CrossRef](#)]
30. Iiames, J.S.; Congalton, R.G.; Lunetta, R.S. Analyst variation associated with land cover image classification of Landsat ETM+ data for the assessment of coarse spatial resolution regional/global land cover products. *GIScience Remote Sens.* **2013**, *50*, 604–622.
31. Ogura, A.T. *Relação e Previsibilidade de Eventos Meteorológicos Extremos Deflagadores de Acidentes de Movimentos de Massa na Serra do Mar*; IPT (Instituto de Pesquisa e Tecnologia): São Paulo, Brazil, 2006.
32. Manfre, L.A. Identificação e Mapeamento de Áreas de Deslizamento Associadas a Rodovias Utilizando Imagens de Sensoriamento Remoto. Ph.D. Thesis, Escola Politécnica da Universidade de São Paulo, São Paulo, Brazil, 2015.
33. Wolle, C.M. Análise dos Escorregamentos Translacionais Numa Região da Serra do Mar no Contexto de uma Classificação de Mecanismos de Instabilização de Encostas. Ph.D. Thesis, Escola Politécnica da Universidade de São Paulo, São Paulo, Brazil, 1988.
34. Toledo, J.M.A. Crop Discrimination Using Harmonic Analysis of EVI Modis Time-Series Data. Master's Thesis, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brazil, 2008.
35. Ma, H.; Qin, Q.; Shen, X. Shadow segmentation and compensation in high resolution satellite images. In Proceedings of the International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 7–11 July 2008; Volume 2, pp. 1036–1039.
36. Kanungo, T.; Mount, D.M.; Netanyahu, N.S.; Piatko, C.D.; Silverman, R.; Wu, A.Y. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 881–892. [[CrossRef](#)]
37. Lo, C.P.; Choi, J. A hybrid approach to urban land use/cover mapping using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images. *Int. J. Remote Sens.* **2004**, *25*, 2687–2700. [[CrossRef](#)]
38. Nemmour, H.; Chibani, Y. Multiple support vector machines for land cover change detection: An application for mapping urban extensions. *ISPRS J. Photogramm. Remote Sens.* **2006**, *61*, 125–133. [[CrossRef](#)]
39. Otukey, J.R.; Blaschke, T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, S27–S31. [[CrossRef](#)]
40. Petropoulos, G.P.; Vadrevu, K.P.; Xanthopoulos, G.; Karantounias, G.; Scholze, M. A comparison of spectral angle mapper and artificial neural network classifiers combined with Landsat TM imagery analysis for obtaining burnt area mapping. *Sensors* **2010**, *10*, 1967–1985. [[CrossRef](#)] [[PubMed](#)]

41. Ranawana, R.; Palade, V. Multi-classifier systems: Review and a roadmap for developers. *Int. J. Hybrid Intell. Syst.* **2006**, *3*, 1–41. [[CrossRef](#)]
42. Testud, J.; Oury, S.; Black, R.A.; Amayenc, P.; Dou, X. The concept of “normalized” distribution to describe raindrop spectra: A tool for cloud physics and cloud remote sensing. *J. Appl. Meteorol.* **2001**, *40*, 1118–1140. [[CrossRef](#)]
43. Tso, G.K.F.; Yau, K.K.W. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy* **2007**, *32*, 1761–1768. [[CrossRef](#)]
44. Ghose, M.K.; Pradhan, R.; Ghose, S.S. Decision tree classification of remotely sensed satellite data using spectral separability matrix. *Int. J. Adv. Comput. Sci. Appl.* **2010**, *1*, 93–101.
45. Wu, F.; Zhan, J.; Yan, H.; Shi, C.; Huang, J. Land cover mapping based on multisource spatial data mining approach for climate simulation: A case study in the farming-pastoral ecotone of North China. *Adv. Meteorol.* **2013**, *2013*. [[CrossRef](#)]
46. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16. [[CrossRef](#)]
47. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1999**, *37*, 35–46. [[CrossRef](#)]
48. Ho, T.K.; Hull, J.J.; Srihari, S.N. On multiple classifier systems for pattern recognition. In Proceedings of the 11th IEEE International Conference on Pattern Recognition, Hague, The Netherlands, 30 August–3 September 1992; p. 84.
49. Ben Abdallah, A.C.; Frigui, H.; Gader, P. Adaptive local fusion with fuzzy integrals. *IEEE Trans. Fuzzy Syst.* **2012**, *20*, 849–864. [[CrossRef](#)]
50. Tian, G.; Liu, J.; Xie, Y.; Yang, Z.; Zhuang, D.; Niu, Z. Analysis of spatio-temporal dynamic pattern and driving forces of urban land in China in 1990s using TM images and GIS. *Cities* **2005**, *22*, 400–410. [[CrossRef](#)]
51. Lorenzi, L.; Melgani, F.; Mercier, G. A complete processing chain for shadow detection and reconstruction in VHR images. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3440–3452. [[CrossRef](#)]
52. Manfré, L.A.; Shinohara, E.J.; Silva, J.B.; Siqueira, R.N.P.; Giannotti, M.A.; Quintanilha, J.A. Method for landslides identification at the Sao Paulo state coast, Brazil. *Geociências* **2014**, *33*, 172–180.
53. Woźniak, M.; Graña, M.; Corchado, E. A survey of multiple classifier systems as hybrid systems. *Inf. Fusion* **2014**, *16*, 3–17. [[CrossRef](#)]
54. Kittler, J.; Hatef, M.; Duin, R.; Matas, J. On combining classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *20*, 226–239. [[CrossRef](#)]
55. Lysiak, R.; Kurzynski, M.; Woloszynski, T. Optimal selection of ensemble classifiers using measures of competence and diversity of base classifiers. *Neurocomputing* **2014**, *126*, 29–35. [[CrossRef](#)]
56. Moreno-Seco, F.; Inesta, J.; De Leon, P.; Mico, L. Comparison of classifier fusion methods for classification in pattern recognition tasks. In *Structural, Syntactic, and Statistical Pattern Recognition*; Springer: Berlin, Germany, 2006; pp. 705–713.
57. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* **2002**, *80*, 185–201. [[CrossRef](#)]
58. Krogh, A.; Vedelsby, J. Neural network ensembles, cross validation, and active learning. *Adv. Neural Inf. Process. Syst.* **1995**, *7*, 231–238.
59. Brown, G.; Wyatt, J.; Harris, R.; Yao, X. Diversity creation methods: A survey and categorisation. *Inf. Fusion* **2005**, *6*, 5–20. [[CrossRef](#)]
60. Michail, P.; Benediktsson, J.A.; Ioannis, K. The effect of classifier agreement on the accuracy of the combined classifier in decision level fusion. *IEEE Trans. Geosci. Remote Sens.* **2005**, *39*, 2539–2546.
61. Chandra, A.; Yao, X. Evolving hybrid ensembles of learning machines for better generalisation. *Neurocomputing* **2006**, *69*, 686–700. [[CrossRef](#)]
62. Foody, G.M. Classification accuracy comparison: Hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. *Remote Sens. Environ.* **2009**, *113*, 1658–1663. [[CrossRef](#)]
63. Giacinto, G.; Roli, F.; Fumera, G. Selection of image classifiers. *Electron. Lett.* **2000**, *36*, 420–422. [[CrossRef](#)]
64. Koc-San, D. Evaluation of different classification techniques for the detection of glass and plastic greenhouses from WorldView-2 satellite imagery. *J. Appl. Remote Sens.* **2013**, *7*. [[CrossRef](#)]

65. Waske, B.; Van Der Linden, S.; Oldenburg, C.; Jakimow, B.; Rabe, A.; Hostert, P. imager—A user-oriented implementation for remote sensing image analysis with Random Forests. *Environ. Model. Softw.* **2012**, *35*, 192–193. [[CrossRef](#)]
66. Liu, J.; Pan, Y.; Zhu, X.; Zhu, W. Using phenological metrics and the multiple classifier fusion method to map land cover types. *J. Appl. Remote Sens.* **2014**, *8*, 083691. [[CrossRef](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).