

Article

Crop Classification in Satellite Images through Probabilistic Segmentation Based on Multiple Sources [†]

Oscar S. Dalmau ^{1,*}, Teresa E. Alarcón ^{2,*} and Francisco E. Oliva ^{2,†}

¹ Centro de Investigación en Matemáticas, CIMAT, 36023 Guanajuato, Mexico

² Centro Universitario de los Valles, Universidad de Guadalajara, 46600 Ameca, Jalisco, Mexico; francisco.oliva@profesores.valles.udg.mx

* Correspondence: dalmau@cimat.mx (O.S.D.); teresa.alarcon@profesores.valles.udg.mx (T.E.A.); Tel.: + 52-473-732-7155 (O.S.D.)

[†] This paper is an extended version of our paper published in Oliva, F.E.; Dalmau, O.S.; Alarcón, T.E. Classification of Different Vegetation Types Combining Two Information Sources Through a Probabilistic Segmentation Approach. In Proceedings of the 14th Mexican International Conference on Artificial Intelligence (MICAI), Morelos, Mexico, 25–31 October 2015.

[‡] These authors contributed equally to this work.

Received: 11 May 2017; Accepted: 8 June 2017; Published: 13 June 2017

Abstract: Classification methods based on Gaussian Markov Measure Field Models and other probabilistic approaches have to face the problem of construction of the likelihood. Typically, in these methods, the likelihood is computed from 1D or 3D histograms. However, when the number of information sources grows, as in the case of satellite images, the histogram construction becomes more difficult due to the high dimensionality of the feature space. In this work, we propose a generalization of Gaussian Markov Measure Field Models and provide a probabilistic segmentation scheme, which fuses multiple information sources for image segmentation. In particular, we apply the general model to classify types of crops in satellite images. The proposed method allows us to combine several feature spaces. For this purpose, the method requires prior information for building a 3D histogram for each considered feature space. Based on previous histograms, we can compute the likelihood of each site of the image to belong to a class. The computed likelihoods are the main input of the proposed algorithm and are combined in the proposed model using a contrast criteria. Different feature spaces are analyzed, among them are 6 spectral bands from LANDSAT 5 TM, 3 principal components from PCA on 6 spectral bands and 3 principal components from PCA applied on 10 vegetation indices. The proposed algorithm was applied to a real image and obtained excellent results in comparison to different classification algorithms used in crop classification.

Keywords: probabilistic segmentation; remote sensing; likelihood; vegetation indices; histogram

1. Introduction

The segmentation process allows us to divide the image into significant parts according a certain criterion. Clustering algorithms such as k-means, fuzzy c-means and the Iterative Self-Organizing Data Analysis Technique (ISODATA) algorithm [1] have been used successfully for segmentation problems, however these methods per se, do not consider the contextual information for a pixel, what is necessary to obtain a good segmentation. A very effective approach for including the features of the pixel neighborhood is the Bayesian estimation together with the Markov Random Field (MRF), [2–9]. With this approach, a label field is computed assuming that a dependence exists between all probability distributions of the pixels belonging to the same neighborhood. This assumption is specified by considering a Markovian priori distribution. Gauss Markov Measure Field (GMMF) [4] is one of the models that combines Bayesian estimation with Markov Random Field and it is used in many

classification tasks, [5–8,10–13]. One of the main difficulties for GMMF, as for all methods based on the combination of Bayesian estimation and MRF for image segmentation, is the likelihood computation. In the case of 1D or 3D feature spaces the likelihood can be computed based on the corresponding normalized histograms. However, the computation of the likelihood becomes a very hard problem when the number of features increases. In the case of crops classification for satellite images the number of features is high, which is a limitation for the direct use of the GMMF [4]. On the other hand, crops classification is a complex task due to the similarity of the spectral signatures among different crops. Hence, the selection of the feature space, i.e., information sources, is a key step in this research, so that we can use the GMMF as a classifier that allows us to incorporate contextual information and to reach good classification results.

In [14] authors considered a pixel based image approach in order to segment 5 different land cover types in Russia. The experimental work included the minimum euclidean distance, the box classifier, Mahalanobis distance, the maximum likelihood classifier and clustering techniques. The feature space was composed of blue, green and red bands. The best performance was achieved by the maximum likelihood. Authors in [15] used three different vegetation indices: the Normalized Difference Vegetation Index (NDVI) [16], the Green Normalized Difference Vegetation Index (GNDVI) [17], and the Normalized Difference Red Edge Index (NDRE) [18] for crop classification in the region located in Turkey. All indices were computed taking into consideration the spectral bands obtained from the RapidEye satellite, which is the first high-resolution multispectral satellite system incorporating the red-edge band which is sensitive to vegetation chlorophyll [15]. Four different feature spaces were studied in the research in [15]: the first spectral space contains NDVI, GNDVI and NDRE indices, the others three feature spectral spaces are composed of only two of the three indices included in the first space. For crop classification, the support vector machine method [19] was used. The experimental work in [15] demonstrated that the spectral index NDRE provided the highest contribution to the support vector machine classifier and that the space composed of three spectral indices outperformed the rest of studied feature spaces. The proposal in [20] used ASTER data and studied different feature spaces in order to classify the sugarcane vegetation in Uttarkhand, India. The analyzed spaces were: feature space determined by the three bands located in the VNIR (visible and near infrared) region of the spectrum; and a combination between VNIR bands and NDVI transformation. The classifier used in this work was the maximum likelihood method and achieved the highest classifier precision using the combination of VNIR bands and NDVI index. The main drawback in [20] was the similarity of spectral signatures among sugarcane vegetation classes. The methodology described in [21] utilizes the green, red, near and short wave infrared bands from SPOT-5 satellite in order to identify crops in Texas. With the SPOT-5 data, the authors studied the performance of 5 classifiers: maximum euclidean distance, Mahalanobis distance, maximum likelihood, spectral angle mapper [22] and SVM [19]. The experimental work proved that the SPOT-5 data in conjunction with maximum likelihood and SVM allows to estimate crop areas. Researchers in [23] carried out a performance analysis of the following supervised classifiers: maximum likelihood, parallelepiped and Mahalanobis distance. As in [21] the feature space was built from SPOT-5 data. The goal of the research was the identification of the different land classes in Barcelona, Spain. Maximum likelihood outperformed other classifiers for almost all the objects present in the study image, however for water bodies the Mahalanobis distance led to best accuracy results. The selection of the spectral bands and the right classifier was a critical step [24–26]. Unlike the previous works, which are pixel-based, in [27] the researchers elaborated an object based image algorithm. The purpose of the algorithm was to identify 13 crop types in California, USA. The spectral information was taken from ASTER satellite, during three different growing-seasons periods. The proposal in [27] combines an object based approach with the decision tree rules. The feature space is composed of several vegetation indices derived from the visible, near and short waves bands, together with textural features extracted from the spectral bands. The experimental work evidenced that the textural features improved the discrimination among heterogeneous permanent crops. In this research also is emphasized the importance of NIR and

SWIR bands for crop identification and it is studied the contribution of each feature to the classification accuracy. All the reported works, either pixel based or object based approaches face the same challenges in best discriminating a class: how do we select the appropriate feature space and how do we select the right classifier?

This work is an extension of our conference paper [28] and we propose a supervised approach for crop classification, in which multiple sources of information or features spaces can be combined, including contextual information. To this end, we propose a generalization of the GMMF model that allows us the combination of multiple sources of information and also considers the spatial or contextual information.

The structure of this paper is as follows: Section 2 details about the previous works, Section 3 explains the proposal, Section 4 is dedicated to the description and discussion of the experimental work and finally Section 5 contains our conclusions.

2. Previous Works

In Ref. [29], a supervised algorithm for crop classification in satellite images was proposed. The algorithm includes a segmentation step achieved through the GMMF model [4], in which the computation of the likelihood is an important ingredient. Equation (1) represents the GMMF approach proposed in [4].

$$p^* = \arg \min_p \sum_{r \in \mathcal{L}} \sum_{k \in \mathcal{K}} (p_k(r) - v_k(r))^2 + \lambda \sum_{s \in \mathcal{N}_r} (p_k(r) - p_k(s))^2, \quad (1)$$

where $v_k(r)$ is the likelihood of pixel r to belong to the class k , $\lambda > 0$ is a regularization parameter, \mathcal{N}_r represents a set of neighboring pixels to the pixel r and p^* is the estimated probability distribution field, in this way $p^*(r)$ allows us to classify the pixel r by maximizing $p_k^*(r)$ over k . The solution of the optimization problem (1) can be obtained by the following Gauss-Seidel scheme:

$$p_k(r) = \frac{v_k(r) + \lambda \sum_{s \in \mathcal{N}_r} p_k(s)}{1 + \lambda |\mathcal{N}_r|}. \quad (2)$$

The final segmentation is obtained by using ‘the winner takes it all’ strategy, i.e., given the vector field p^* , Equation (1), the segmentation is computed with the following equation:

$$s(r) = \arg \max_{k \in \mathcal{K}} p_k(r), \quad \forall r \in \mathcal{L}. \quad (3)$$

Authors in [29] estimated the likelihood by means of 3D histograms in the feature space composed of green, red and near infrared bands. It is well known that the information contained in these three bands allows to recognize crop patterns, [30]. The histograms were computed based on expert information.

Figure 1 represents the spectral signature of five crops studied in [28,29]. Observe the similarity of the response among all crops. This fact explains the precision errors of the proposed algorithm in [29]. For addressing this problem the authors in [28] considered using two different feature spaces (sources of information): the feature space composed of green, red and near infrared bands and the space based on the three principal components of 10 vegetation indices [15,16,31–37]. From these two feature spaces they constructed histograms, and from these histograms they estimated the corresponding likelihood. Different from [29], authors in [28] introduced a modification in data term of the GMMF approach in order to select at each pixel r the likelihood with the least entropy:

$$p^* = \arg \min_p \sum_{r \in \mathcal{L}} \sum_{k \in \mathcal{K}} \sum_{i=1}^2 \omega_i(r) (p_k(r) - v_k^i(r))^2 + \lambda \sum_{s \in \mathcal{N}_r} (p_k(r) - p_k(s))^2, \quad (4)$$

where $v^i(r)$ is the likelihood that comes from the i -th source and the weight function $\omega_i(r) \in \{0, 1\}$ is given by Equation (5)

$$\omega_i(r) = \begin{cases} 1 & \text{if } \mathcal{E}(v^i(r)) < \mathcal{E}(v^{3-i}(r)) \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

$i \in \{1, 2\}$ and $\mathcal{E}(\cdot)$ is an entropy measure [38]. The authors in [28] used Gini impurity index, i.e.,

$$\mathcal{E}(f) = 1 - f^T f, \quad (6)$$

such that $\mathbf{1}^T f = 1, f \succeq 0$.

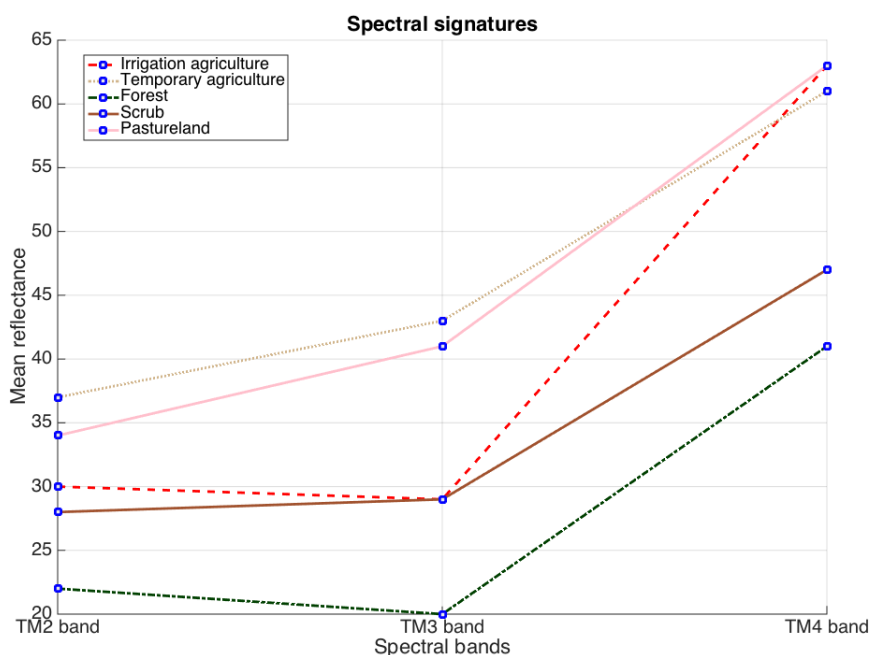


Figure 1. Mean reflectance values for the TM432 bands for each vegetation type under study.

From Equation (5) it is understandable that only the likelihood that leads to a probability distribution with a lowest entropy is included in the data term for the corresponding pixel, i.e., once the likelihoods for the two feature spaces are computed, the weight function (5) selects only one likelihood for each pixel. The previous idea [28] contributed to diminish the misclassification in the recognition of different crops. The inclusion of a new feature space improved the classification process with respect to the results obtained in [29]. Note that in practice, the functional in (4) selects only one feature space at each pixel. In this work we present an extension of our researches in [28,29] and we propose a generalization of the GMMF model that allows us the fusion of multiple sources of information for classification tasks. We evaluate the behavior of the proposal for crop recognition task using satellite images, however the proposal can be applied to other classification problems.

3. The Proposal

The segmentation approach is composed of three stages: training, segmentation and validation. In Figure 2 are illustrated the stages of the proposal. In the next sections we explain each stage.

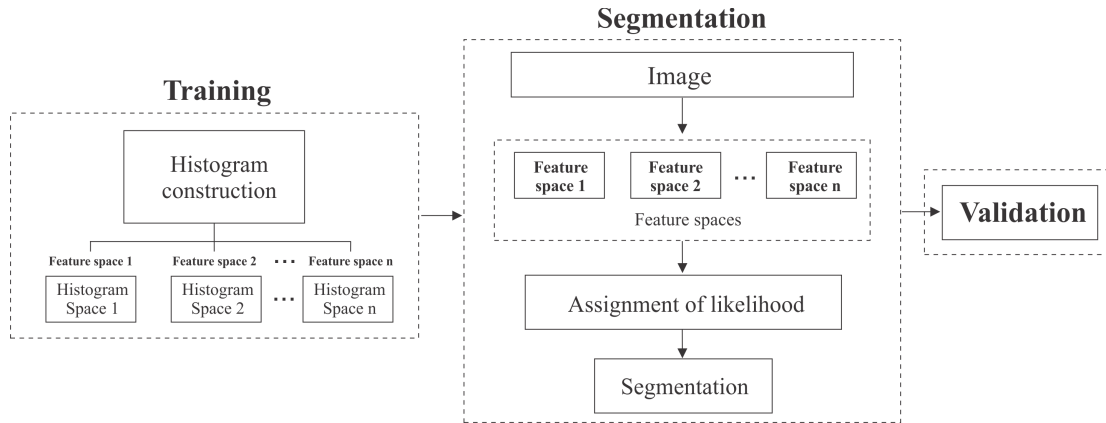


Figure 2. Stages of the proposed algorithm.

3.1. Training Stage

The training stage is based on histograms construction [28,29]. We assume that the information about the classes has been provided by an expert, for example, by making scribbles on training images for each class. Let us denote the normalized histogram as:

$$h(x;k) \propto N(x;k), \quad (7)$$

$$\sum_x h(x;k) = 1, \forall k \in \mathcal{K}, \quad (8)$$

where $\mathcal{K} = \{1, 2, \dots, K\}$ and K is the number of classes, x corresponds to a feature vector. For example, if one considers three spectral bands for satellite images; $h(x;k)$ is a 3D histogram corresponding to the class k . The computed histograms allows us to estimate the likelihood of pixels of an image to belong to a class with respect to a feature space. This is an important step, because the computation of the likelihood is a required element for Markov Measure Field segmentation models [5–7,28,29,39,40].

Authors in Ref. [28] considered the use of only two feature spaces, now we propose to use multiple feature spaces or sources of information. Let us denote the number of feature spaces as N_f . Then, we build the normalized histograms $h^f(\cdot, k)$ for each feature space $f \in \mathcal{F} = \{1, 2, \dots, N_f\}$ and class $k \in \mathcal{K}$. Let us define the mapping function $x^f : \mathcal{L} \subset \mathbb{Z}^2 \rightarrow \mathbb{R}^{|f|}$ that allows us to extract for each site $r \in \mathcal{L}$ of an image the feature vector $x^f(r) \in \mathbb{R}^{|f|}$ where $|f|$ represents the size of the feature space f . Therefore, $h^f(x^f(r), k)$ can be interpreted as the probability of the feature vector $x^f(r)$ associated to pixel r given the class k and the feature space f .

3.2. Segmentation Stage through Multiples Sources and Probabilistic Approach

Here we assume that we have already trained our algorithm, i.e., we know the histograms $h^f(\cdot, k)$ for each feature space f and for class k .

The segmentation procedure is as follows:

1. For a given image, we compute the feature spaces provided by the mapping $x^f(r)$, $\forall r \in \mathcal{L}$, $f \in \mathcal{F}$.
2. Then, the likelihood is assigned according to the following equation

$$v_k^f(r) \propto h^f(x^f(r); k) \quad (9)$$

such that $\sum_{k \in \mathcal{K}} v_k^f(r) = 1, \forall r \in \mathcal{L}, f \in \mathcal{F}$. The likelihood is obtained by normalizing $h^f(x^f(r); k)$ with respect to the classes.

3. Here we propose a robust GMMF approach that generalizes the proposal in [28] including more feature spaces and with weight functions in both the data and regularization terms. Equations (10)–(12) indicate the modifications.

$$p^* = \arg \min_p \sum_{r \in \mathcal{L}} \sum_{k \in \mathcal{K}} \sum_{f \in \mathcal{F}} \omega_f(r) (p_k(r) - v_k^f(r))^2 + \lambda \sum_{s \in \mathcal{N}_r} \omega_{rs} (p_k(r) - p_k(s))^2, \quad (10)$$

$$\omega_f(r) = \frac{\omega_f^s(r)}{\sum_{f \in \mathcal{F}} \omega_f^s(r)} \quad (11)$$

$$\omega_f^s(r) = \frac{\mu}{\mu + \mathcal{U}_f(r)} \quad (12)$$

where $\mathcal{U}_f(r) \in [0, 1]$ is an uncertainty measure of the information source f at pixel r , for example, the measure of entropy (5) used in [28]; $\lambda > 0$ is the regularization parameter, $\mu > 0$ controls the relative importance of the likelihood for different sources. When μ is very large, the contribution of the likelihood for all sources tends to be the same and when it is close to zero, the functional in (10) tends to select the likelihood corresponding to the lowest uncertainty, i.e., this tends to the solution proposed in [28] when using the entropy as uncertainty measure. The weight $\omega_{rs} \in [0, 1]$ function allows to control the edges between classes, here we use

$$\omega_{rs} = \frac{\mu}{\mu + \|u(r) - u(s)\|_2^2} \quad (13)$$

$$u_k(r) \stackrel{\text{def}}{=} \sum_{f \in \mathcal{F}} \omega_f(r) v_k^f(r) \quad (14)$$

such that $\omega_{rs} \approx 1$ if the sites r, s very probably belong to the same class and $\omega_{rs} \approx 0$ otherwise. The solution of the optimization problem (10) yields the following Gauss-Seidel scheme:

$$p_k(r) = \frac{u_k(r) + \lambda \sum_{s \in \mathcal{N}_r} \omega_{rs} p_k(s)}{1 + \lambda \sum_{s \in \mathcal{N}_r} \omega_{rs}}. \quad (15)$$

which is similar to the Gauss-Seidel scheme (2), but now the term $u_k(r)$, Equation (14), is a mixture term that allows us to combine or fuse different likelihoods. In addition, the above formula also includes function weights to control the edges between classes. Note that, the first term in the numerator of the Equation (15) is a convex linear combination of likelihoods derived from different sources. Feature spaces with lower uncertainty have a greater $\omega_f(r)$ and therefore they have a greater contribution to the data term in Equation (10).

3.3. Validation

For validation purposes we use the measure of agreement Cohen's kappa [41], the precision by class and the overall accuracy, see Section 4.4 for details. These comparison measures use the information derived from the confusion matrix [42,43]. This matrix is built through the manual segmentation given by an expert (ground truth) and the results obtained by the algorithms.

4. Experiments and Discussion

In this section we describe the experimental work and analyze the results obtained by our proposal when applied to crop classification task in satellite images.

4.1. Study Area

The study area is located in western Mexico at coordinates Lat. 20°39'58" N, Long. 103°21'7" W, with an altitude of 1550 m above sea level, in the Atemajac Valley [44]. We study five types of vegetation located in this region, see Table 1.

Table 1. Vegetation types used for the study.

Class	Vegetation Name
C1	Irrigation agriculture
C2	Temporary agriculture
C3	Forest
C4	Scrub
C5	Pastureland

4.2. Data Sources

Data used in the experimental work are from Landsat-5 Thematic Mapper (TM) satellite imagery. The spectral band information for this satellite appears in Table 2.

Table 2. Spectral bands of the Landsat-5 TM Sensor.

TM Bands	Wavelength (μ)	Features
TM1	0.45–0.52	B (Blue)
TM2	0.52–0.60	G (Green)
TM3	0.63–0.69	R (Red)
TM4	0.76–0.90	near infrared
TM5	1.55–1.75	mid-infrared
TM6	10.4–12.50	thermal infrared
TM7	2.08–2.35	mid-infrared

The studied image has resolution of 30 m and 2^8 radiometric resolution and corresponds to 1 March 2011. Data was delivered in level 1T, in which geometric correction was applied [45].

The image was obtained through the USGS Global Visualization Viewer site [46]. For error analysis, the image was manually segmented by experts Oliva et al. [28]. The studied image and the ground truth appear in Figure 3.

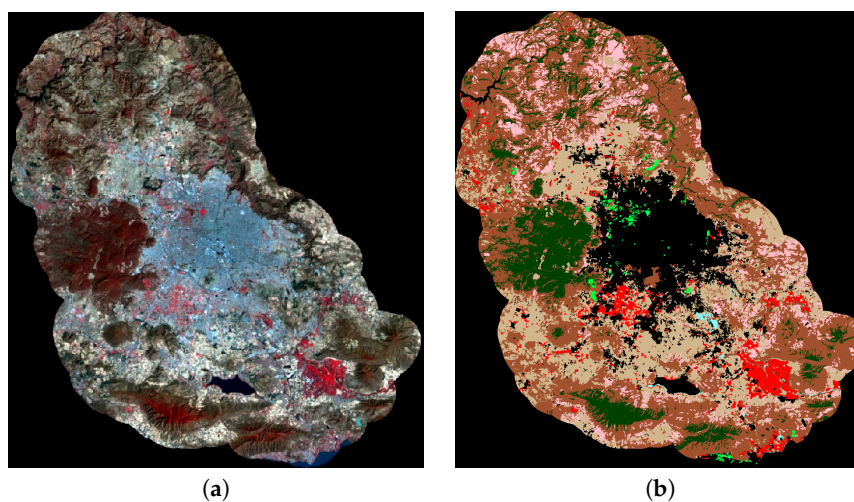


Figure 3. (a) Studied image; (b) ground truth given by an expert.

4.3. Studied Feature Spaces

Three feature spaces were analyzed in our work:

1. **Space 1:** it is composed of information from TM2, TM3 and TM4 bands due to it is well known [30] that these three bands contain relevant data for crop detection. The selected bands are preprocessed through a bilateral filter [47].
2. **Space 2:** it contains the first three principal components from the PCA applied on 10 vegetation indices, see Table 3. Such indices are based on mathematical operations on spectral bands and they allow to enhance the information related to vegetation. In order to compute the indices we calculate the reflectance values, ρ , corresponding to the acquired images, using the algorithms in [48], see also the procedures given in [49–52]. The included indices appear in Table 3. Symbols ρ_r, ρ_g, ρ_b and ρ_{NIR} denote the reflectance values for the red, blue, green and infrared bands respectively. We followed the recommendation given in [35], and set $L = 0.5$ to compute the SAVI and SARVI expressions. For SARVI we considered $\gamma = 1$ as authors in [34].

Table 3. Explored Vegetation indices. In the equations below ρ_r, ρ_g, ρ_b and ρ_{NIR} denote the reflectance values for the red, blue, green and infrared bands respectively and $\rho_{rb} = \rho_r - \gamma(\rho_b - \rho_r)$.

Name VI	Formula	References
MSR	$\frac{\rho_{NIR} - 1}{\frac{\rho_r}{\sqrt{\frac{\rho_{NIR}}{\rho_r} + 1}}}$	[31]
CI	$\frac{\rho_g}{\rho_{NIR} - 1}$	[32]
midrule NDVI	$\frac{\rho_{NIR} - \rho_r}{\rho_{NIR} + \rho_r}$	[15]
GNDVI	$\frac{\rho_{NIR} - \rho_g}{\rho_{NIR} + \rho_g}$	[16]
EVI	$2.5 \left[\frac{\rho_{NIR} - \rho_r}{1 + \rho_{NIR} + 6\rho_r - 7.5(\rho_b)} \right]$	[33]
SARVI	$\frac{(1+L)(\rho_{NIR} - \rho_{rb})}{(\rho_{NIR} + R_{rb} + L)}$	[34]
RDVI	$\frac{\rho_{NIR} - \rho_r}{\sqrt{\rho_{NIR} + \rho_r}}$	[31]
SAVI	$\frac{(1+L)(\rho_{NIR} - \rho_r)}{\rho_{NIR} + \rho_r + L}$	[35]
MSAVI	$\frac{1}{2} \left[(2\rho_{NIR} + 1) - \sqrt{(2\rho_{NIR} + 1)^2 - 8(\rho_{NIR} - \rho_r)} \right]$	[36]
WDRVI	$\frac{\alpha \times \rho_{NIR} - \rho_r}{\alpha \times \rho_{NIR} + \rho_r}$	[37]

3. **Space 3:** this space contains three principal components from PCA applied on the spectral bands TM1, TM2, TM3, TM4, TM5 and TM7, see Table 2. Although TM2, TM3 and TM4 bands more accurately describes information related to vegetation [30], in this investigation PCA on all six mentioned bands is applied in order to include information from other spectral regions.

4.4. Comparison Measures

The comparison measures, for validation purposes, are based on the confusion matrix C which is used in multi-class classification problems [42,43] to assess the performance of algorithms. Here, we assume that the rows of the confusion matrix correspond to the actual classes (the “ground truth”) and the columns correspond to the predicted classes. Then, the entry C_{ij} of the confusion matrix C represents the number of data of the class i that are predicted in class j . The performance measure

we used for comparing the algorithms are: the overall accuracy, the precision for each class j and the Cohen's kappa, i.e.,

$$\text{Overall accuracy} = \frac{\sum_{j=1}^K C_{jj}}{\sum_{i=1}^K \sum_{j=1}^K C_{ij}}, \quad (16)$$

$$\text{Precision}_j = \frac{C_{jj}}{\sum_{i=1}^K C_{ij}}, \quad (17)$$

$$\text{Kappa} = \frac{p_o - p_e}{1 - p_e}, \quad (18)$$

where

$$p_o = \text{Overall accuracy,}$$

$$p_e = a^T b,$$

$$a_j = \frac{\sum_{i=1}^K C_{ij}}{\sum_{i=1}^K \sum_{j=1}^K C_{ij}},$$

$$b_i = \frac{\sum_{j=1}^K C_{ij}}{\sum_{i=1}^K \sum_{j=1}^K C_{ij}}.$$

The overall accuracy (16) is the number of correct classifications divided by the total number of classified data. The precision (17) is a measure of the accuracy of a specific class j . Thus, the precision is the ratio between the number of correctly predicted data of the class j divided by the total number of classified data in the class j . Cohen's kappa (18) is a statistical measure of inter-rater agreement or inter-annotator agreement for 2 raters [41]. Cohen's kappa measures the agreement between two raters who each classify N items into K mutually exclusive categories, see Table 4 for the interpretation of the values of Kappa index [53]. In Ref. [53] the reader can find more details about the measure of agreement Kappa.

Table 4. Interpretation of Cohen's kappa measure.

Cohen's Kappa	Interpretation
<0	Poor agreement
0.00–0.20	Slight agreement
0.21–0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–1.00	Almost perfect agreement

4.5. Results and Discussion

We carried out several experiments where we combined the feature spaces described in the Section 4.3. Table 5 gives a brief description about all possible feature space combinations. Column 1 indicates the number of combination and column 2 specifies the feature spaces included in the combination.

We conducted 11 experiments with real data described in Section 4.2. The first three experiments considered only one feature space with the standard GMMF algorithm [29]. The next four experiments were performed using the lowest entropy model [28], see Equations (4) and (5) and the space combinations 4–7 given in Table 5. The last four experiments also used the combinations 4–7, and the segmentation process was conducted using the fusion sources model proposed in this work, see Equations (10)–(12).

Table 5. All possible combinations of feature spaces.

N. Combination	Feature Space Combination
1	Space 1
2	Space 2
3	Space 3
4	Space 1 + Space 2
5	Space 1 + Space 3
6	Space 2 + Space 3
7	Space 1 + Space 2 + Space 3

Tables 6–8 summarize the numerical results of all 11 experiments. These tables present the numerical information about the classification of 5 different crops, see Table 1, the precision of classification for each class, the overall accuracy and Kappa coefficient. We note that the results of the first row of Tables 6 and 7 differ from those presented in Ref. [28,29]. This is due to, in this results, we only consider sites in the image that correspond to crops in the five categories of interest, i.e., only pixels in the region of interest.

Table 6 shows the results obtained using the standard GMMF algorithm with different features spaces [29,54]. Experiments 1–3 consider the information from Space 1, Space 2 and Space 3 respectively, see Table 5. Note that best result is obtained when using the first three principal components based on all spectral bands, except the TM6 band (Experiment 3). Note also that, the first three principal components using the 10 spectral indexes (Experiment 2) achieved a similar performance to the one obtained with the typical combination of TM2, TM3 and TM4 bands (Experiment 1). This means that for the crop classification using GMMF, the amount of information is relevant.

Table 6. Numerical results of experiments using the GMMF model with only one feature space.

Experiment	Precision					Overall Accuracy	Cohen's Kappa
	C1	C2	C3	C4	C5		
1	0.86	0.83	0.89	0.85	0.63	0.8331	0.7520
2	0.84	0.81	0.89	0.87	0.62	0.8324	0.7528
3	0.83	0.82	0.91	0.89	0.65	0.8506	0.7801

Table 7 summarizes numerical results for the experiments in which we considered the lowest entropy model [28], see Equations (4) and (5). Experiment 4 combines the Space 1 and the Space 2, Experiment 5 takes into account the Space 1 and the Space 3, Experiment 6 includes Space 2 and Space 3 and Experiment 7 uses all studied spaces. Observe that the combination of different feature spaces together with the lowest entropy model allowed to improve the precision of crop recognition in general, and hence the overall accuracy and Kappa index increased. Table 7 also confirms that not only the information from TM2, TM3 and TM4 (Space 1) is relevant from crop recognition, but also the information derived from other bands.

Table 8 shows the numerical results of the experiments 8–11. These four experiments considered the same feature space combination as in the experiments 4–7 in Table 7, but using the fusion sources model (10)–(12). Note that the Experiment 9 achieved the best performance. This experiment considered the combination of Space 1 and Space 3.

Although the results of the experiments in Table 8 are similar in comparison with the results obtained by means of lowest entropy model, see Table 7, they have a better performance than those obtained by lowest entropy model. In general, the new proposal obtained the best performance compared with our previous conference papers. Furthermore, the study of feature spaces performed in this work also allowed us to improve the performance of our previous work.

Table 7. Numerical results of segmentation experiments using the modified GMMF with minimum entropy criteria.

Experiment	Precision					Overall Accuracy	Cohen's Kappa
	C1	C2	C3	C4	C5		
4	0.86	0.83	0.92	0.87	0.67	0.8472	0.7726
5	0.89	0.84	0.94	0.88	0.70	0.8624	0.7955
6	0.85	0.83	0.91	0.89	0.64	0.8485	0.7770
7	0.87	0.84	0.93	0.88	0.67	0.8588	0.7904

Table 8. Numerical results of experiments with the GMMF model of fusion of different sources with $\mu = 1$.

Experiment	Precision					Overall Accuracy	Cohen's Kappa
	C1	C2	C3	C4	C5		
8	0.88	0.83	0.92	0.87	0.67	0.8497	0.7767
9	0.90	0.84	0.94	0.89	0.71	0.8649	0.7993
10	0.88	0.82	0.91	0.89	0.66	0.8517	0.7812
11	0.90	0.83	0.93	0.89	0.69	0.8600	0.7923

Our experimental work corroborated that the vegetation indices enhance the vegetation information, and they are able to distinguish between vegetation and non-vegetation. However, for the classification task, they are not sufficient as an information source to discriminate different types of crops.

Though more information sources we have a greater computation time, for this task we suggest using more feature spaces in order to improve the accuracy of the classification process. For that reason the proposal described in Section 3.2 considers a fusion of multiples feature spaces.

In this work we used feature spaces generated from the pixel information and different spectral bands, however the proposal is more general and it allows to include other feature spaces and not only punctual information, but also local information derived from different image modalities. The study of local information and the fusion of different image modalities is a part of our future work.

On the other hand, Table 9 shows a comparative analysis of different classification methods that have been widely used in the context of crop classification problems. In the experiment we use the real image described in Section 4.2. Here we compare different algorithms implemented in the software MultiSpec, funded by NASA, that is available online [55]. MultiSpec is a freeware multispectral image analysis software developed at Purdue University and the latest release date is on March 2017. In this experiment, we include the following methods: Minimum Euclidean Distance (MED) [56], Maximum Likelihood Classifier (ML) [57], Fisher Linear Likelihood (FLL) [58] and (ESS) [59,60] because they reached the best performance results. We also include two versions of the one vs all multiclass SVM method, the linear and the non-linear Radial Basis Function alternatives, using the winner-takes-all strategy. In this case, we use the matlab built-in function for SVM. In the comparison study we also consider the original version of the probabilistic segmentation approach described in [28,29]. Additionally, for a fair comparison, we include the performance analysis of these algorithms using the best results reached in the feature space study in this work, see Tables 6 and 7, denoted as MICA1 2014* and MICA1 2015* in Table 9. Finally, we include the results achieved by our proposal based on the fusion of information sources. Note that for our proposal, and in general for probabilistic segmentation approaches, the natural feature space is the space of probabilistic distributions or the likelihood space, unlike others methods under study where the natural space is the original image. A detailed comparison study of feature spaces and classification methods for crop classification is a very interesting task, however, this is out of the scope of the present research and we leave it for future work.

Table 9. Numerical results of different classification methods. MICAI 2014* and MICAI 2015* denote the best results of the proposals in [28,29] under the feature space study conducted in this work.

Method	Feature Space	C1	C2	C3	C4	C5	Overall Accuracy	Kappa
MED [61]	Landsat-5 TM	0.82	0.73	0.54	0.80	0.29	0.6244	0.4781
ML [62]	Landsat-5 TM	0.65	0.73	0.72	0.76	0.40	0.6928	0.5558
FLL [63]	Landsat-5 TM	0.88	0.74	0.71	0.75	0.46	0.7105	0.5743
ESS [60]	Landsat-5 TM	0.76	0.70	0.80	0.74	0.56	0.7257	0.5811
SVM linear	Landsat-5 TM	0.77	0.74	0.96	0.73	0.19	0.7498	0.6071
SVM rbf	Landsat-5 TM	0.79	0.83	0.90	0.90	0.66	0.8536	0.7859
MICAI 2014 [29]	Space 1	0.86	0.83	0.89	0.85	0.63	0.8331	0.7520
MICAI 2014*	Space 3	0.83	0.82	0.91	0.89	0.65	0.8506	0.7801
MICAI 2015 [28]	Spaces 1 & 2	0.86	0.83	0.92	0.87	0.67	0.8472	0.7726
MICAI 2015*	Spaces 1 & 3	0.89	0.84	0.94	0.88	0.70	0.8624	0.7955
Proposal	Spaces 1 & 3	0.90	0.84	0.94	0.89	0.71	0.8649	0.7993

Based on the comparison results in Table 9, we note that our proposal, in general, outperforms the remainder analyzed methods. The precision, overall accuracy and the Kappa index attained a better performance with respect to our previous conference papers [28,29]. Observe, that the non-linear SVM version with Radial Basis Function led to the best classification for class C4, however the best classification for the rest classes was reached by our proposal. On the other hand, the analysis of the feature spaces developed in this work, also allowed to improve the results of the probabilistic methods in [28,29]. The method MICAI 2014* corresponds to the GMMF algorithm [29] using the Space 3 studied here. Note that the overall accuracy and Kappa coefficient increased in comparison to the Space 1 used in the original paper in Ref. [29]. Similarly, MICAI 2015* refers to the proposal in [28] using the feature space combination Space 1 and Space 3. Note that performance measures of classification also increased.

For illustrative purposes, the classification maps of the three most accurate methods given in Table 9 are shown in Figure 4.

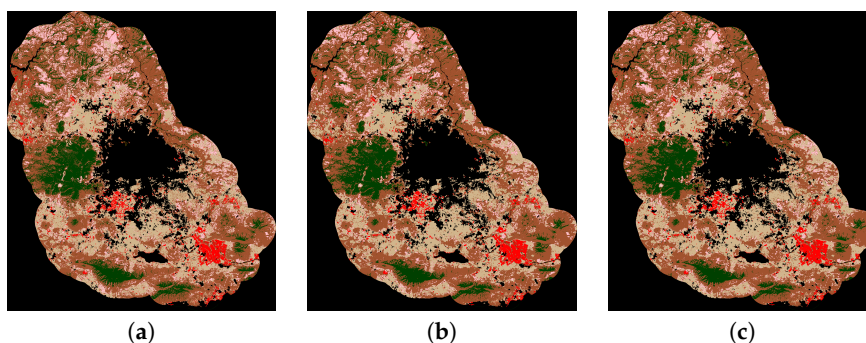


Figure 4. Classification maps of the three most accurate methods given in Table 9: (a) non-linear SVM version with Radial Basis Function; (b) MICAI 2015* and (c) Our proposal.

5. Conclusions

The selection of the feature space is a challenge for any image classification task. Our proposal provides a framework in which different information sources are combined taking into account the entropy of their probability distribution. To this end, we propose a generalization of the Gauss Markov Measure Field model, in which both, data and regularization terms contain a weight function, that allows to combine different feature spaces, in a such way that information sources with lowest entropy probability distribution have a greater contribution in the classification process. The proposal considers the punctual and contextual spatial information. The weight function of the second term allows us to control the edges between classes obtaining a robust potential based on the likelihoods.

The performance of the proposed approach is evaluated in satellite images for classifying different crops. Although we focused on classification problems in satellite images, the proposal can be used for any image classification task, in which it is necessary to combine several feature spaces, this is a potential research in the future. We also studied different spaces and all possible combinations. The achieved precision indices: overall accuracy and Kappa coefficient demonstrated the good behavior of our approach. Experiments with real images show that the proposed algorithm obtains excellent results compared with algorithms used in the context of satellite image for crop classification. Finally, an interesting problem that emerges from this work is to carry out a detailed comparison study of different features spaces and classification methods for crop classification. We leave this study for future work.

Acknowledgments: We wish to acknowledge Maximiliano Bautista Andalón and Ana Teresa Ortega Minakata members of Land Information Institute of Jalisco (IITEJ) for providing the required information and ground truth images for this work. This work was supported in part by CONACYT (Mexico), Grant 258033.

Author Contributions: Oscar S. Dalmau and Teresa E. Alarcón conceived the theoretical bases of the paper, designed, implemented algorithms, performed the experiments, and wrote the paper; Francisco E. Oliva provided the images and figures in the paper, implemented algorithms and wrote part of the codes, performed the experiments and studied the state of the art. All authors reviewed the paper and gave constructive advices.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jain, A.; Dubes, R.C. *Algorithm for Clustering Data*; Prentice Hall: Upper Saddle River, NJ, USA, 1988.
2. Li, S.Z. *Markov Random Field Modeling in Image Analysis*; Springer: Tokyo, Japan, 2001.
3. Geman, S.; Geman, D. Stochastic relaxation, Gibbs distributions and bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *6*, 721–741.
4. Marroquin, J.L.; Rivera, M.; Nakamura, M. Gauss-Markov measure field models for low-level vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 337–348.
5. Cedeño, O.D.; Rivera, M.; Mayorga, P.P. Computing the Alpha-Channel with Probabilistic Segmentation for Image Colorization. In Proceedings of the IEEE 11th International Conference on Computer Vision (ICCV 2007), Rio de Janeiro, Brazil, 14–21 October 2007; pp. 1–7.
6. Dalmau, O.; Rivera, M. A General Bayesian Markov Random Field Model for Probabilistic Image Segmentation. In *Combinatorial Image Analysis: Lecture Notes in Computer Science*; Wiederhold, P., Barneva, R., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5852, pp. 149–161.
7. Dalmau, O.; Rivera, M. Beta-Measure for Probabilistic Segmentation. In *Advances in Artificial Intelligence Part I, Proceedings of the 9th Mexican International Conference on Artificial Intelligence (MICAI 2010)*, Pachuca, Mexico, 8–13 November 2010; Springer: Berlin/Heidelberg, Germany, 2010; pp. 312–324.
8. Marroquin, J.L.; Arce, E.; Botello, S. Hidden Markov measure field models for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 1380–1387.
9. Rivera, M.; Ocegueda, O.; Marroquin, J.L. Entropy controlled gauss-markov random measure field models for early vision. In Proceedings of the Variational, Geometric, and Level Set Methods in Computer Vision (VLSM), Beijing, China, 16 October 2005; Volume 3572, pp. 137–148.
10. Marroquin, J.L.; Vemuri, B.C.; Botello, S.; Calderon, E.; Fernandez-Bouzas, A. An accurate and efficient Bayesian method for automatic segmentation of brain MRI. *IEEE Trans. Med. Imaging* **2002**, *5*, 934–945.
11. Gijbels, I.; Lambert, A.; Qiu, P. Edge-preserving image denoising and estimation of discontinuous surfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1075–1087.
12. Ocegueda, O.; Shah, S.K.; Kakadiaris, I.A. Which parts of the face give out your identity? In Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 20–25 June 2011; pp. 641–648.
13. Ocegueda, O.; Fang, T.; Shah, S.K.; Kakadiaris, I.A. 3D face discriminant analysis using Gauss-Markov posterior marginals. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 728–739.
14. Bogoliubova, A.; Tymkow, P. Land cover changes and dynamics of yuntolovsky reserve. *Electron. J. Pol. Agric. Univ.* **2014**, *17*. Available online: <http://www.ejpau.media.pl/volume17/issue3/art-03.html> (accessed on 13 June 2017).

15. Ustuner, M.; Sanli, F.; Abdikan, S.; Esetlili, M.; Kurucu, Y. Crop type classification using vegetation indices of RapidEye imagery. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey, 29 September–2 October 2014; pp. 195–198.
16. Tucker, C.J. Detection of Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150.
17. Gitelson, A.A.; Merzlyak, M.N. Detection of Red Edge Position and Chlorophyll Content by Reflectance Measurements Near 700 nm. *J. Plant Physiol.* **1996**, *148*, 501–508.
18. Barnes, E.M.; Clarke, T.R. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In Proceedings of the Fifth International Conference on Precision Agriculture, Bloomington, MN, USA, 16–19 July 2000.
19. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: Berlin, Germany, 2013.
20. Vikesh, K.; Vinod, K.; Kamal, J. Development of Spectral Signature and Classification of Sugarcane using ASTER DATA. *Int. J. Comput. Sci. Commun.* **2010**, *1*, 245–251.
21. Yang, C.; Everit, J.H.; Murden, D. Evaluating high resolution {SPOT} 5 satellite imagery for crop identification. *Comput. Electron. Agric.* **2011**, *75*, 347–354.
22. Kruse, F.A.; Lefkoff, A.B.; Boardman, J.B.; Heidebrecht, K.B.; Shapiro, A.T.; Barloon, P.J.; Goetz, A.F.H. The Spectral Image Processing System (SIPS)—Interactive Visualization and Analysis of Imaging Spectrometer Data. *Remote Sens. Environ.* **1993**, *44*, 145–163.
23. Nawaz, A.; Iqbal, Z.; Ullah, S. Performance analysis of supervised image classification techniques for the classification of multispectral satellite imagery. In Proceedings of the ICASE 2015—4th IEEE International Conference on Aerospace Science and Engineering, Islamabad, Pakistan, 2–4 September 2015.
24. Dash, M.; Liu, H. Feature selection for classification. *Intell. Data Anal.* **1997**, *1*, 131–156.
25. Luis Carlos Molina, A.L.B.; Nebot, A. Feature selection algorithms: A survey and experimental evaluation. In Proceedings of the IEEE International Conference on Data Mining (ICDM), Maebashi City, Japan, 9–12 December 2002.
26. Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238.
27. Pena-Barragan, J.M.; Ngugi, M.K.; Plant, R.E.; Six, J. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens. Environ.* **2011**, *115*, 1301–1316.
28. Oliva, F.E.; Dalmau, O.S.; Alarcón, T.E. Classification of Different Vegetation Types Combining Two Information Sources Through a Probabilistic Segmentation Approach. In Proceedings of the Mexican International Conference on Artificial Intelligence, Cuernavaca, Mexico, 25–31 October 2015; Volume 9414, pp. 382–392.
29. Oliva, F.E.; Dalmau, O.S.; Alarcón, T.E. A Supervised Segmentation Algorithm for Crop Classification Based on Histograms Using Satellite Images. In Proceedings of the Mexican International Conference on Artificial Intelligence, Tuxtla Gutiérrez, Mexico, 16–22 November 2014; Volume 8856, pp. 327–335.
30. Sabins, F.F. *Remote Sensing: Principles and Applications*; Waveland Press Inc.: Long Grove, IL, USA, 2007.
31. Jordan, C. Derivation of leaf area index from quality of light on the forest floor. *Ecology* **1969**, *50*, 663–666.
32. Gitelson, A.A.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282.
33. Huete, A.R.; Liu, H.; Batchily, K.; Van Leeuwen, W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451.
34. Kaufman, Y.; Tanre, D. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *Geosci. Remote Sens.* **1992**, *30*, 261–270.
35. Jordan, C. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309.
36. Broge, N.H.; Leblanc, E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* **2001**, *76*, 156–172.
37. Gitelson, A.A. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *J. Plant Physiol.* **2004**, *161*, 165–173.
38. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.

39. Marroquin, J.L.; Velasco, F.A.; Rivera, M.; Nakamura, M. Gauss-Markov Measure Field Models for Low-Level Vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 337–348.
40. Rivera, M.; Dalmau, O.S. Variational Viewpoint of the Quadratic Markov Measure Field Models: Theory and Algorithms. *IEEE Trans. Image Process.* **2011**, *21*, 1246–1257.
41. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46.
42. Kohavi, R.; Provost, F. Guest Editors' Introduction: On Applied Research in Machine Learning. *Mach. Learn.* **1998**, *30*, 127–132.
43. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437.
44. Pulido, H.G.; Bautista, A.M.; Guevara, R.M. *Jalisco Territorio y Problemas de Desarrollo*; Instituto de Información Estadística y Geográfica: Zapopan, Mexico, 2013.
45. Northrop, A.; Team, L.S. *Ideas-Lansat Products Description Document, Technical Report*; Telespazio VEGA UK Ltd.: Luton, UK, 2015.
46. Visualization Viewer Site. Available online: <http://glovis.usgs.gov/> (accessed on 13 June 2017).
47. Tomasi, C.; Manduchi, R. Bilateral Filtering for Gray and Color Images. In Proceedings of the Sixth International Conference on Computer Vision, Bombay, India, 7 January 1998; pp. 839–846.
48. Chen, D.; Huan, J.; Jackson, T.J. Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. *Remote Sens. Environ.* **2005**, *98*, 225–236.
49. Jiang, H.; Feng, M.; Zhu, Y.; Lu, N.; Huang, J.; Xiao, T. An Automated Method for Extracting Rivers and Lakes from Landsat Imagery. *Remote Sens.* **2014**, *6*, 5067–5089.
50. Rokni, K.; Ahmad, A.; Selamat, A.; Hazini, S. Water Feature Extraction and Change Detection Using Multitemporal Landsat Imagery. *Remote Sens.* **2014**, *6*, 4173–4189.
51. Yashon, O.; Tateishi, R. A water index for rapid mapping of shoreline changes of five East African Rift Valley lakes: an empirical analysis using Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2006**, *27*, 3153–3181.
52. Ji, L.; Zhang, L.; Wylie, B. Analysis of Dynamic Thresholds for the Normalized Difference Water Index. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1307–1317.
53. Viera, A.; Garrett, J. Understanding interobserver agreement: The kappa statistic. *Fam. Med.* **2005**, *37*, 360–363.
54. Marroquin, J.L.; Botello, S.; Calderon, F.; Vemuri, B.C. The MPM-MAP Algorithm for image segmentation. *Pattern Recognit.* **2000**, *1*, 303–308.
55. MultiSpec Software. Available online: <https://engineering.purdue.edu/~biehl/MultiSpec/> (accessed on 13 June 2017).
56. Su, B.; Noguchi, N. Agricultural land use information extraction in Miyajimanuma wetland area based on remote sensing imagery. *Environ. Control Biol.* **2012**, *50*, 277–287.
57. Omkar, S.N.; Senthilnath, J.; Mudigere, D.; Kumar, M.M. Crop classification using biologically-inspired techniques with high resolution satellite image. *J. Indian Soc. Remote Sens.* **2008**, *36*, 175–182.
58. Karakahya, H.; Yazgan, B.; Ersoy, O.K. A Spectral-Spatial Classification Algorithm for Multispectral Remote Sensing Data. In Proceedings of the Artificial Neural Networks and Neural Information Processing, Istanbul, Turkey, 26–29 June 2003; pp. 1011–1017.
59. Kettig, R.L.; Landgrebe, D.A. Computer Classification of Remotely Sensed Multispectral Image Data by Extraction and Classification of Homogeneous Objects. *IEEE Trans. Geosci. Electron.* **1976**, *14*, 19–26.
60. Landgrebe, D. The Development of a Spectral-Spatial Classifier for Earth Observational Data. *Pattern Recognit.* **1980**, *12*, 165–175.
61. Wolfowitz, J. The Minimum Distance Method. *Ann. Math. Statist.* **1957**, *28*, 75–88.
62. Hazewinkel, M. Maximum likelihood method. In *Encyclopedia of Mathematics*; Springer: Berlin, Germany, 2001.
63. Mika, S.; Ratsch, G.; Weston, J.; Scholkopf, B.; Mullers, K.R. Fisher discriminant analysis with kernels. In Proceedings of the Neural Networks for Signal Processing IX, Madison, WI, USA, 25 August 1999; pp. 41–48.

