Evolutionary Approach for Detection of Buried Remains Using Hyperspectral Images

León Dozal, José L. Silván-Cárdenas, Daniela Moctezuma, Oscar S. Siordia, and Enrique Naredo

Abstract
Hyperspectral imaging has been successfully utilized to locate clandestine graves. This study applied a Genetic Programming technique called Brain Programming (BP) for automating the design of Hyperspectral Visual Attention Models (H-VAM), which is proposed as a new method for the detection of buried remains. Four graves were simulated and monitored during six months by taking in situ spectral measurements of the ground. Two experiments were implemented using Kappa and weighted Kappa coefficients as classification accuracy measures for guiding the BP search of the best H-VAM. Experimental results demonstrate that the proposed BP method improves classification accuracy compared to a previous approach. A better detection performance was observed for the image acquired after three months from burial. Moreover, results suggest that the use of spectral bands that respond to vegetation and water content of the plants and provide evidence that the number of buried bodies plays a crucial role on a successful detection.

Introduction
Locating unmarked graves represents a complicated and time-consuming forensic problem because their locations are often remote and the burial time is generally unknown (Siegel and Saukko, 2013). The research on the detection of clandestine graves through multi- and hyperspectral images is incipient, yet has proven to be one of the most challenging forensic problems. This is an important area of work, since airborne hyperspectral data can enable searching over a large area that is otherwise inaccessible by foot; especially because, in principle, any area of the Earth can be mapped by hyperspectral imaging, be it with aircraft or satellites (Ross et al., 2005).

Several studies have tested the potential of multispectral and hyperspectral images with varying results. Kalacskai and Bell (2006) were among the first that demonstrated the potential of remote sensing as a tool for locating heretofore unknown mass graves. Afterwards, Kalacskai et al. (2009) analyzed the in situ and airborne spectral reflectance of a set of animal mass graves and identically constructed false graves. Their results indicated that the reflectance spectra of grave are readily distinguishable from false grave at both scales. In addition, they observed that vegetation regeneration was severely inhibited by cattle carcasses for up to a period of 16 months. Caccianiga et al. (2012) studied the effects of decomposition of buried swine carcasses on soil and vegetation structure and composition as a tool for detecting clandestine graves. They found that soil disturbance was the main factor affecting plant cover, while the role of decomposition seemed to be much less critical. Leblanc et al. (2014) performed a blind-test of the potential for airborne hyperspectral imaging technology to locate buried remains of pig carcasses. They were able to predict two single graves, within GPS error (10 m), whose location they did not know. Recently, Silván-Cárdenas et al. (2017) studied some methods for detecting clandestine graves using hyperspectral data collected on ground. Through a controlled experiment using buried carcasses of pigs, demonstrated that hyperspectral data have potential for detecting buried remains only after three months from burial. Furthermore, that the critical spectral regions for graves detection are the NIR and SWIR1 spectral regions, some of which were so narrow (10 nm) that stressed the need for hyperspectral sensing.

The method of acquisition of hyperspectral images is equally important than the process of pattern recognition for detection of graves based on such information. In this sense, some techniques of evolutionary computation have been successfully applied for selection and combination of spectral bands aiming at different applications such as classification of vegetation species, soil mineral identification, synthesizing spectral indices, estimate pasture mass and quality, and precision farming, to mention just a few (Ross et al., 2005; Chion et al., 2006; Albaracín et al., 2016; Zhuo et al., 2006; Li et al., 2011, Kawamura et al., 2010, Puente et al., 2011, Ullah et al., 2012, Davis et al., 2006, Landry et al., 2006; Kawamura et al., 2010; Awuley and Ross, 2016). On the other hand, currently, visual attention models have been designed for the spatial and spectral analysis of hyperspectral images with applications such as detection of prominence, visualization and interpretation, and detection of objects (Le Moan et al., 2011and 2013; Wang, 2013; Liang et al., 2013; Cao et al., 2015; Zhang et al., 2017).

In this study an evolutionary technique is proposed based on genetic programming, known as Brain Programming (BP), for optimizing a so-called Hyperspectral Visual Attention Models (H-VAM) for graves detection.

Problem Statement
The present work addresses the problem of detection of clandestine graves as a problem of classification of hyperspectral images. The image classification problem can be stated in formal terms as follows. Suppose we want to classify each pixel in an image into one of N classes, let say C1, C2, ..., CN. Then, decision rules must be established to enable assignment of any given pixel to these classes (Varshney and Arora, 2004).

When working with hyperspectral images, some issues arise due to the high dimensionality of this type of images, e.g., Hughes phenomenon, high information redundancy in spectral and spatial domains, need for finding features that increase discrimination between classes and high computational resources required in the classification process.

For this reason, a compelling need to reduce the dimension of data exists. The methods for reduction of dimensionality can be roughly divided into two categories: feature extraction...
and feature selection (Li et al., 2014). The former refers to a process to transform the high-dimensional feature space to a low-dimensional space through linear or no linear combinations of original bands. Whereas, the latter refers to a selection of an optimal subset of features, through a combinatorial optimization (Yoon and Park, 2015). At this point, it should be noted that classification accuracy is highly impacted by feature selection or extraction methods.

In this study, we tested the BP algorithm as a feature selection method. Much research has been done on feature selection from an evolutionary computation approach (see the Brain Programming Implementation Section), but little work has been done around feature extraction methods incorporating both spatial and spectral information through such approach. In order to integrate spectral and spatial information in our solution to the grave detection problem, we use an H-VAM optimized with a BP algorithm. The H-VAM allows building a saliency map that corresponds to a classified image.

### Visual Attention from a Computational Standpoint

The concept of visual attention or visual saliency comes from the study of the human visual system and this consists in the ability to efficiently find objects or regions that stand out relative to their neighboring parts in a scene (Borji and Itti, 2013). Therefore, visual attention is an inherently active and selective process by which people attend to a subset of the available information for further processing along the visual pathway (Zhang et al., 2008). For this reason, when it comes to computer science, the notion of saliency is closely related to visual attention in color or gray scale images (Le Moan et al., 2013). In a broad sense, the concept of saliency is not exclusive of the vision process. Recently, Le Moan et al. (2013) coined the term “spectral saliency” to name the extent to which a certain group of pixels stands out in an image, in terms of reflectance rather than in terms of colorimetric attributes. In principle, any kind of data can be considered, although, until now, only a few studies have used other than visible spectrum images.

In 1998, the first computer models of visual attention appeared. The most influential model has been that of Itti et al. (1998); they introduced the concept of “saliency map” that is an image in gray scales where the larger the value of a pixel, the more prominent it is. Since then, the saliency map has been utilized in various applications, such as object recognition, object detection, segmentation, and tracking. The reader is referred to Borji and Itti (2013) for a comprehensive review of the state of the art in visual attention modeling. There is a growing interest in using saliency attention models for multi and hyperspectral images. Considering the high dimensionality of hyperspectral images, traditional Visual Attention Model (VAM), such as that of Itti et al. (1998), cannot be directly applied to hyperspectral imagery (Cao et al., 2015). Nowadays, there are few VAMs dealing with the spectral saliency (Le Moan et al., 2011and 2013; Wang, 2013; Liang et al., 2013; Cao et al., 2015; Zhang et al., 2017).

Most of the methods adopted saliency detection based on spectral signatures. In this way, feature extraction is only operated in the spectral domain, but the spatial distribution of targets has not been fully explored. However, the spatial distribution has proved to be very valuable for image analysis in remote sensing and computer vision communities (Cao et al., 2015).

### Introduction to Brain Programming

Brain Programming (BP) is an evolutionary computing technique based on Genetic Programming (GP), which is a technique to solve optimization problems based on Darwin’s theory of evolution; and the well-known principle of natural selection, proposed by Koza (1992).

BP evolves parts of computer programs, in this case, parts of the H-VAM. The evolvable parts of the H-VAM are represented as syntax tree structures that can solve a particular sub-problem, these structures are known as individuals. BP starts with an initial population of individuals that are randomly generated. Each tree that is part of the individual is produced from a set of functions, \( F = \{ f_1, f_2, \ldots, f_n \} \) and a set of terminals \( T = \{ t_1, t_2, \ldots, t_n \} \). It is important to note that each tree can have different sets \( F \) and \( T \), unlike \( GP \); this allows the function specialization of each tree. Functions correspond to nodes with children in the tree whereas terminals correspond to leaves in the tree. The randomly generated initial population is evolved to find new individuals that can give a better solution to a particular problem. First, individuals are evaluated using a fitness function. Besides, the fitness function is defined as a criterion for reproduction probability, each individual is given a probability for reproduction directly proportional to its fitness. Next, individuals are selectively recombined according to their probability of reproduction to produce a new population of potentially better individuals. These new individuals are created by means of operations called genetic operators, which are applied with an a priori probability. The principal genetic operators in BP are chromosomal crossover (two individuals exchange complete trees), gene crossover (two individuals exchange a sub-tree of each of their trees), chromosomal mutation (one individual arbitrarily changes each one of its trees by a new one randomly generated) and gene mutation (one individual arbitrarily changes one sub-tree of each one of its trees by a new one randomly generated). For a more detailed explanation of genetics operations of BP see (Dozal et al., 2014). During evolution, each new iteration, known as generation, will contain improved solutions and the
process is repeated until some stop criterion is met. This may be a pre-specified level of fitness, convergence, or a maximum number of generations (Davis et al., 2006).

The primary objective of BP is to be able to simply define a task and have the machine independently learn to perform it. BP provides the framework for the machine to evolve its own algorithms, delivering a solution as a computer program or as data representing a computer program (Banzhaf et al., 1998).

In this way, GP has been very successful at evolving novel and unexpected ways of solving problems, generally producing a number of instances that are competitive with human-produced results (Koza, 2010). The GP paradigm is worth exploring because it lends itself well to feature extraction in combination with feature selection (Iglesia, 2013). The executable nature of the evolved individuals makes GP ideally suited for implementing multispectral analysis and other remote sensing applications, such as classifiers and mathematical indices that enhance the signal of several features over the Earth’s surface by means of supervised learning approaches (Agnelli et al., 2002). For a more detailed description of the BP approach see (Dozal et al., 2014).

Materials and Methods

Field Data Collection

Four graves were simulated in a terrain near to Yautepec, Morelos, México. The terrain was divided into a grid of 19-by-19 squares. Ten carcasses of pigs (pístrain breed) were distributed into four graves. The spatial distribution of these graves is shown in Figure 1 with labels F2, F4, F6, and F7. Each grave size is 2 m-by-2 m wide by 1.1 to 1.5 m depth. Carcasses were distributed as shown in Table 2. In addition to these four graves, three empty pits were created with the same dimensions, in order to simulate soil disturbances caused by other non-burial process such as plow or constructions. The graves were kept undisturbed during the whole six months measurement period.

After burial, from 12 February to 29 July 29, the soil reflectance was measured every fourteen days approximately, by using a full-range field spectroradiometer (Field Spec 4 Std Res by ASD Inc. @ 350-2500nm) in each one of the 19-by-19 squares. Measurements were taken using the natural illumination, the time of data acquisition was approximately within 11:00 and 13:00 hours local time so that the sun position did not vary a lot, and whenever the sky conditions was clear or mostly clear, the winds blew slowly (0 to 5 kmh), and there was no rainfall. As a result, twelve hyperspectral images (H) were formed with the ground sampling distance of 1 m. By July, the vegetation had already over passed the mesh height so that some measurements were carried out under the canopy of vegetation. This was actually the main reason to stop monitoring, as measurements under the canopy are not compatible with airborne or satellite measurements (Silván-Cárdenas et al., 2017).

After the acquisition of the ground measurements, the data were preprocessed prior to classification; this process is explained in the next Section.

Hyperspectral Image Preprocessing

Before analyzing the data for classification and graves detection, it is necessary to apply a preprocessing due to technical conditions we describe next. A spectral misalignments occurs because the spectroradiometer combines three sensors for the measurement of the spectral signature, that are, VNIR: 350-1000 nm, SWIR1: 1001-1800 nm and SWIR2: 1801-2500 nm, this is 2,151 bands. If the spectroradiometer is not warmed up sufficiently, the measurements around the edge of the spectral regions of the three sensors do not match. This was observed only in a couple of images and the problem was corrected by fixing the spectral region of VNIR sensor and applying an offset to the reflectance captured by the SWIR1 and SWIR2 sensors. Another problem derived from bad calibration and changing light conditions is banding; some images presented this issue in which the pixels of an entire line appear lighter or darker than those of neighbor lines. This was corrected by multiplying the reflectance values of banding line by a coefficient that was estimated using the average reflectance of neighbor lines. It is a common practice to eliminate spectral bands of low atmospheric transmissivity in which reflectance values fluctuate erratically. Such bands are: 1350-1480 nm, 1780-2032 nm and 2450-2500 nm, the incident radiation on the soil, in these wavelengths, is almost null because water vapor absorbs almost completely the electromagnetic radiation. Finally, due to the changing illumination conditions among the measurements it becomes necessary to normalize them. We identified a location in the grid that falls on a piece of concrete, such location was considered as reflectance invariant through the time. Normalization consisted on adjusting the spectral range from 400 to 500 nm (blue band) of all images at that location. Taking as reference the first image acquired on 12 February, images were re-calibrated through the empirical line method, so that following calibration all images presented similar reflectance in the blue band for the spectrally-invariant point (Silván-Cárdenas et al., 2017). Images are hereafter identified by acquisition date in the format MMDDYY.

The H-VAM Algorithm

This study proposed a modified VAM (H-VAM) based on the Iitti et al. (1998) model. We adapted the model to the intrinsic characteristics of H 5. Therefore, the H-VAM algorithm is an abstraction of the natural process of visual attention, which

![Figure 1. Spatial distribution of graves.](Image 327x493 to 561x731)

<table>
<thead>
<tr>
<th>Grave ID</th>
<th>Depth [m]</th>
<th>No. Carcasses</th>
<th>Weight [Kg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>1.2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F2</td>
<td>1.2</td>
<td>3</td>
<td>255</td>
</tr>
<tr>
<td>F3</td>
<td>1.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F4</td>
<td>1.0</td>
<td>2</td>
<td>170</td>
</tr>
<tr>
<td>F5</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F6</td>
<td>1.2</td>
<td>1</td>
<td>85</td>
</tr>
<tr>
<td>F7</td>
<td>1.1</td>
<td>4</td>
<td>340</td>
</tr>
</tbody>
</table>
can be applied to find the location of buried remains in a hyperspectral image.

The H-VAM algorithm takes advantage of the knowledge about the task to be performed and the intrinsic properties of the objects in the scene, such as spectral reflectance. The H-VAM program is based on specialized functions, known as visual operators, which are in charge of extracting characteristics of the observed scene and also of its subsequent combination. All this, for highlighting the necessary information for the task at hand, in this case the detection of buried remains. Figure 2 shows the diagram of the H-VAM algorithm, which is explained in detail in the following subsections.

Feature Extraction and Visual Maps
As stated in Figure 2 the scene recorded through a full-range spectroradiometer corresponds to a \( H \) with 1,719 wavelengths, which is input to the H-VAM algorithm. The \( H \) is used to extract different kinds of characteristics by using specialized operators according to the type of information to be extracted.

The matter of how spectral variations are represented mathematically and conceptually is an important first step in defining how the extraction of the desired information should proceed (Landgrebe, 1999). In this trace, the present study tackles feature extraction from both the spectral and spatial domains. The feature extraction of H-VAM algorithm is implemented through functions called Visual Operators (VO). These are mathematical functions, which are in charge to process the \( H \) with the aim of extract specific characteristics in the form of spectral or spatial features. Each VO is defined as a transformation \( I_{H} \rightarrow VM_{d} \) where \( H \) represents the input and \( VM_{d} \) is the resulting image, known as visual map, along a spectral or spatial domain \( d \).

In this study, \( H \)s are defined as: \( H = \{ \rho_{\lambda}, \lambda \in W \} \) where \( \rho_{\lambda} \) is the reflectance value for a wavelength \( \lambda \) and \( W \) is the wavelength range from 350 nm to 2,500 nm with a resolution of 1 nm; thus \( H \) is conformed by a grid of 19-by-19 pixels and 1,719 spectral bands. In the first place, the input image \( H \) is transformed by a VO resulting in a Visual Map (VM). It is important to realize that the transformed images \( VM_{d} \) may not retain the physical interpretation as the input reflectance image (Li et al., 2014). The structure of \( VO_{\text{spectral}} \) and \( VO_{\text{spatial}} \) is described in the followings sections.

Spectral Features Extraction
The spectral feature extraction is in charge to obtain information from the relationships among spectral bands. It is possible to obtain a lot of information by combining spectral bands; the wavelength relationship between bands can be as important as the reflectance value of the bands (Khanna et al., 2007). The relations among spectral responses to different wavelengths are defined by mathematical transformations. The spectral indices are a good example; they are based on combinations of two or more bands that enhance specific spectral properties.

Remote sensing experts analyze the reflectance properties of Earth’s surface and the relationships between bands instead of the absolute value of reflectance at any one of those bands, then they design indices to extract information to characterize it (Khanna et al., 2007). From a computational standpoint, the operator \( VO_{\text{spectral}} \) states a transformation from the \( H \) to a visual map \( VM_{\text{spectral}} \) and it is defined as follows:

\[
VO_{\text{spectral}} : H \rightarrow VM_{\text{spectral}}.
\]

The pixel values of the resulting \( VM_{\text{spectral}} \) represent the prominence of the extracted spectral features.

Spatial Features Extraction
In this study, the spatial feature extraction refers to the process specialized on obtaining morphological intrinsic features from the objects in the image, in other words, the structural information. Spatial information comprises: edge detection, suppression of selected image structures based on shape, orientation, and size criteria, among others. Accordingly, this method searches for objects with a specific spatial arrangement of pixels rather than single pixels with a specific spectral signature (Soille and Pesaresi, 2002). Mathematical morphological operators can tackle the analysis of spatial data in an image. According to Soille (2003), Mathematical Morphology (MM) can be defined as a theory for the analysis of spatial structures. In this sense, the extensive use of MM in geosciences and remote sensing research areas has proved as a powerful tool for the spatial analysis of \( H \), s and for a wide range of applications such as segmentation of high resolution satellite imagery (Pesaresi and Benediktsson, 2001); extraction of information about size, shape and the orientation of image structures in single-band remote sensing of urban areas (Benediktsson et al., 2003); segmentation and classification of hyperspectral images (Tarabalka et al., 2010; Noyel et al., 2007; Plaza et al., 2009) to mention a few examples. For a comprehensive survey of literature regarding the use of MM in the field of geoscience and remote sensing, see Soille and Pesaresi (2002).

Former methods to analyze hyperspectral data were focused on a per-pixel processing, by merely utilizing the spectral information from thousands of spectral channels. A major weakness of such methods is the lack of information about the spatial organization of the pixels as hyperspectral data are not treated as images (Plaza et al., 2009). Later, with the aim to improve the results obtained by processing techniques of hyperspectral data, many studies remarked on the importance of integrating the spatial information (Plaza et al., 2009; Tarabalka et al., 2010).

In this work, the \( VO_{\text{spatial}} \) is defined as follows:

\[
VO_{\text{spatial}} : H \rightarrow VM_{\text{spatial}}.
\]

The pixel values of the resulting \( VM_{\text{spatial}} \) represent the prominence of the extracted spatial features.

Once the VMs are obtained, the H-VAM must compute an additional map, known as Conspicuity Map (CM), for each feature domain. The computation of CMs is as in Walter and Koch (2006) model and they are obtained by means of a center-surround function. Its main task is to measure the differences between firing rates at the center (c) and surroundings (s) of the VMs with the aim of making the feature extraction scale-invariant and normalizing CMs. The H-VAM utilizes
values of \( c = \{1, 2, 3\} \) and \( s = \{c + 1, c + 2\} \). At this point, we have \( \text{CM}_{\text{spatial}} \) and \( \text{CM}_{\text{spectral}} \) normalized. After that, the H-VAM model combines the CMs to obtain a single saliency map as explained next.

**Feature Integration and Saliency Map**

The final stage of the H-VAM process is the fusion of features into one single Saliency Map (SM). The combination of CMs is a hard problem because CMs belong to different and unrelated features. We consider that the problem must be addressed regarding the purpose of the task at hand. Indeed, in this work we adopted a goal-driven paradigm to guide the search of a suitable combination of features, where the main criterion to guide the suitable combination of characteristics is the classification accuracy. The integration of CMs is accomplished by means of a Feature Integration operator, \( F_I \), as follows:

\[
F_I : \text{CM}_{\text{spatial}} \times \text{CM}_{\text{spectral}} \rightarrow \text{SM}.
\]  

As a result, \( SM \) represents the prominence of the combined features, understood as a function of the extracted features. Finally, \( SM \) is binarized (BSM) using a prominence threshold equal to zero, \( th = 0 \). All values of pixels that are less or equal to zero are assigned the value of 0, while those greater than zero are assigned the value of 1. This binarization allows to compare, in a direct way, the BSM with the binary image of ground truth.

**Brain Programming Implementation**

This section describes the details of the implementation of BP for this particular problem. There are some works related with classification of multi- and hyperspectral images, which utilize evolutionary computation techniques as genetic algorithms [Zhuo et al., 2008; Li et al., 2011; Kawamura et al., 2010; Ullah et al., 2012] and genetic programming [Davis et al., 2006; Landry et al., 2006; Awuley and Ross, 2016].

**Genetic Representation of the H-VAM**

In this paper, the genotype, namely computational representation of an H-VAM, is represented with a triplet of trees (Figure 3). In this way, all possible combinations of terminals and functions represent the search space of the problem. As we mentioned before, BP algorithm is designed to deal with arrays of specialized trees, unlike GP. The term “specialized” refers to the fact that each tree in the genotype has its own sets of functions and terminals that were carefully chosen according to the desired kind of features we attempt to extract. This allows each tree to concentrate in the extraction of specific characteristics, such as spectral or spatial features, or in the fusion of both. The main idea is to evolve \( \text{VO}_{\text{spectral}}, \text{VO}_{\text{spatial}}, \) and \( F_I \) operators, thereby, each one of the trees encodes one operator. In the rest of the paper we will refer to such encoded operations as Evolved Visual Operators (EVO) and Evolved Feature Integration (EFI). Below, we detail the genotype structure, including both EVOs and EFI, according to the features they extract.

**Extraction of H-VAM Spatial Features**

The operator \( \text{EVO}_{\text{spatial}} \) is in charge of extracting information derived from the spectral signature. For this reason we included to the set of terminals, which is expressed by \( T_{\text{spatial}} \), the spectral responses in the wavelength range from 350 nm to 2500 nm with a resolution of 1 nm, that is to say, 1,719 spectral images, represented by \( H_x \), after removing the bands of low transmissivity. These terminals could be sufficient but, with the purpose of exploit a priori knowledge about wavelengths relations, we decided to add some Spectral Indexes (SI) to the list of terminals. Table 3 lists the SIs that serve as terminals for \( \text{EVO}_{\text{spectral}} \), their formula and a brief description of what they were made for.

Spectral indices can be formally defined as follows:

\[
\text{SI} = H_x(x,y,\lambda) \rightarrow \mathbb{R}
\]  

where \( H_x(x,y,\lambda) \) is the spectral value of a pixel located at \( x, y \) coordinates in the spectral image corresponding to \( \lambda \) wavelength. Hence, SIs can be seen as pixel-wise functions that associate spectral bands with a scalar from which the pixel class is determined [Albarracin et al., 2016]. In general, any linear or nonlinear combination of spectral responses at relevant wavelengths can be considered as a spectral index.

In addition, it is necessary to fix the set of functions that determine the relationship among terminals. The function set of \( \text{EVO}_{\text{spectral}} \) \( F_{\text{spectral}} \) is shown in Table 4, most of them appear in the formulas of the spectral indices in Table 3. The functions: \( +k, -k, \times k, \div k, \sqrt[k]{\cdot}, [\cdot]^k \), represent the signalized arithmetic operations where one of the terms is a constant \( k \) with random values ranging from 0 to 100. Finally, we added the Normalized Difference Spectral Index (NDSI) defined as NDSI = \( (\rho_1 - \rho_2)/(\rho_1 + \rho_2) \) where \( \rho_1 \) and \( \rho_2 \) are spectral responses at different wavelengths. This normalization is practical for reducing atmospheric turbidity or other error sources as well as for enhancing and standardizing the spectral response to observed targets. In principle, NDSIIs with consecutive wavelengths is similar to the first derivative up to a normalization factor [Ionue et al., 2008].

Similarly to SIs, all EVOs are operations defined at pixel level, this means, we apply the function to the corresponding pixels of two spectral images. Thus, all these functions are defined pixel to pixel. Most of the time SIs are designed by specialists, with BP approach the SI creation is automated. As we mention in the first Section, similar approaches have been made in the past [Smith and Bull, 2005; Chion et al., 2008; Albarracin et al., 2016].

Next, we explain the extraction of spatial features and its relationship with mathematical morphology.

**Extraction of H-VAM Spatial Features**

Contrary to \( \text{EVO}_{\text{spectral}} \), \( \text{EVO}_{\text{spatial}} \) are neighborhood-based operators and they are designed to extract spatial information from \( H_x \). Unsurprisingly, the set of terminals of \( \text{EVO}_{\text{spatial}} \) operators, known as \( T_{\text{spatial}} \), is compounded by \( H_x \) wavelengths. On the other hand, the set of functions, \( F_{\text{spatial}} \) is formed by the morphological operations showed in Table 5, as we
Table 3. Spectral indices that are part of terminals set, $T_{\text{spectral}}$, of EVO${}_\text{spectral}$.

<table>
<thead>
<tr>
<th>Spectral Indices</th>
<th>Formulation</th>
<th>Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmospherically Resistant Vegetation Index</td>
<td>$\text{ARVI} = -0.18 + 1.17 \times \text{NDVI}(\rho_{830} - \rho_{680})$</td>
<td>Vegetation. It has self-correction properties for the atmospheric effects</td>
</tr>
<tr>
<td>Enhanced Vegetation Index</td>
<td>$\text{EVI} = 2.5 \times (\rho_{\text{near IR}} - \rho_{645}) / (\rho_{645} + 6 \times \rho_{490} + 7.5 \times \rho_{680} + 1)$</td>
<td>Vegetation. It is sensible in high biomass regions and reduces atmosphere</td>
</tr>
<tr>
<td>Leaf Water Content Index</td>
<td>$\text{LWCI} = \log[1 - (\rho_{860} - \rho_{1640})] / \log[1 - (\rho_{860}^\text{mol} - \rho_{1640}^\text{mol})]$</td>
<td>Detects the water stress and determines the Relative Water Content in plants</td>
</tr>
<tr>
<td>Moisture Stress Index</td>
<td>$\text{MSI} = \rho_{1600} - \rho_{1650}$</td>
<td>Provide an indication of relative leaf/canopy moisture, see (Rock et al., 1983, Hunt and Rock, 1989).</td>
</tr>
<tr>
<td>Normalized Difference Infrared Index</td>
<td>$\text{NDII} = \rho_{860} - \rho_{1600}$</td>
<td>Estimate the moisture content in the root zone of vegetation (Srivongsitanon et al., 2016).</td>
</tr>
<tr>
<td>Normalized Difference Moisture Index</td>
<td>$\text{NDMI} = \rho_{1600} - \rho_{1650}$</td>
<td>Canopy water content. It was utilized in detection of forest disturbances (Jin and Sader, 2005).</td>
</tr>
<tr>
<td>Normalized Difference Mud Index</td>
<td>$\text{NMDI} = \rho_{1600} - \rho_{1650}$</td>
<td>Muddy or very shallow water (Bernstein et al., 2012).</td>
</tr>
<tr>
<td>Normalized Pigment Chlorophyll Ratio Index</td>
<td>$\text{NPCI} = \rho_{440} - \rho_{430} / \rho_{660} + \rho_{430}$</td>
<td>Ratio between total pigments and chlorophyll a. This ratio decrease in healthy and rise in stressed or senescing aquatic vegetation (Peñuelas et al., 1993, 1994).</td>
</tr>
<tr>
<td>Normalized Difference Vegetation Index</td>
<td>$\text{NDVI} = (\rho_{860} - \rho_{680}) / (\rho_{860} + \rho_{680})$</td>
<td>Green biomass, Leaf Area Index, Vegetation (Ionue et al., 2008, Peñuelas et al., 1995).</td>
</tr>
<tr>
<td>Normalized Difference Water Index</td>
<td>$\text{NDWI} = (\rho_{645} - \rho_{1240}) / (\rho_{860} + \rho_{1240})$</td>
<td>Soil/vegetation water content from space (Gao, 1996).</td>
</tr>
<tr>
<td>Normalized Multi-band Drought Index</td>
<td>$\text{NDMI} = \rho_{680} - (\rho_{1640} - \rho_{2130}) / (\rho_{860} + (\rho_{1640} - \rho_{2130})$</td>
<td>Soil/vegetation water content from space (Wang and Qu, 2007).</td>
</tr>
<tr>
<td>Normalized Phaeophytinization Index</td>
<td>$\text{NPQI} = \rho_{440} - \rho_{430} / \rho_{440} + \rho_{430}$</td>
<td>Chlorophyll degradation. It is sensitive to mite attack effects on apple trees (Peñuelas et al., 1995).</td>
</tr>
<tr>
<td>Photochemical Reflectance Index</td>
<td>$\text{PRI} = \rho_{531} - \rho_{570} / \rho_{531} + \rho_{570}$</td>
<td>Photosynthetic radiation use efficiency of plant leaves (Ionue et al., 2008).</td>
</tr>
<tr>
<td>Shortwave Infrared Water Stress Index</td>
<td>$\text{SIWSI} = (\rho_{1640} - \rho_{680}) / (\rho_{1640} + \rho_{680})$</td>
<td>Soil/Canopy water content (Fensholt and Sandholt, 2003).</td>
</tr>
<tr>
<td>Simple Ratio Pigment Index</td>
<td>$\text{SRPI} = \rho_{440} / \rho_{680}$</td>
<td>Ratio between carotenoid and chlorophyll a concentrations (Car/Chl).</td>
</tr>
<tr>
<td>Soil-Adjusted Vegetation Index</td>
<td>$\text{SAVI} = 1.5 \times (\rho_{680} - \rho_{690}) / (\rho_{680} + \rho_{690} + 0.5)$</td>
<td>Vegetation, minimizing soil-induced variations (Huete, 1988, Ionue et al., 2008).</td>
</tr>
<tr>
<td>Structure Insensitive Pigment Index</td>
<td>$\text{SPI} = \rho_{680} - \rho_{440} / \rho_{680} + \rho_{440}$</td>
<td>Ratio between carotenoid and chlorophyll a concentrations (Car/Chl). This ratio characterizes the physiology and phenology of vegetation (Peñuelas et al., 1995b).</td>
</tr>
<tr>
<td>Water Band Index</td>
<td>$\text{WBI} = \rho_{700} / \rho_{860}$</td>
<td>Plant Water Concentration (Peñuelas et al., 1994). It is the reciprocal of WI.</td>
</tr>
<tr>
<td>Water Index</td>
<td>$\text{WI} = \rho_{580} / \rho_{700}$</td>
<td>Plant Water Concentration (Peñuelas et al., 1997).</td>
</tr>
</tbody>
</table>

Table 4. Sets of functions and terminals utilized by BP to create $EVO_{\text{spectral}}$.

$F_{\text{spectral}} = \{+, -, \times, \div, \bot, \top, |\cdot|, \inf\}, \sup\}, \text{NDVI}\}$

$T_{\text{spectral}} = \{H, \cup, \text{ARVI}, \text{EV I}, \text{LWCI}, \text{MSI}, \text{NDII}, \text{NDMI}, \text{NPCI}, \text{NDVI}, \text{NDWI}, \text{NMDI}, \text{NPQI}, \text{PRI}, \text{SIWSI}, \text{SRPI}, \text{SAVI}, \text{SPI}, \text{WBI}, \text{WI}\}$

mentioned in Section 2.3.1, MM is an important ally to extract structural information from images. Fundamentally, morphological operators utilize a Structuring Element (SE) for scanning the image and investigate the morphological properties of the structures contained in the image, e.g., highlighting the structures with similar shape and size to the SE. In general, a SE is the elementary $n$-dimension neighborhood zone centered at pixel $x$, known as origin, and $\Psi(f)$ is the morphological transformation $\Psi$ of image $f$ using $B$ as an SE. The origin allows the positioning of the SE at every pixel in the image. Usually, a SE is a two-dimensional neighbor-
Table 5. Morphological operators with their corresponding brief description.

<table>
<thead>
<tr>
<th>Morphological Functions</th>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erosion</td>
<td>( b_b(f) = \land_{b \in B} f_{-b} )</td>
<td>Erosion of a grayscale image ( f ) using a structuring element ( B ) is defined as the minimum of the translations of ( f ) by the vectors ( -b ) in ( B ). To put it another way, the eroded value of a pixel ( x ) is the minimum value of the image in the neighborhood defined by the SE when its origin is at ( x ).</td>
</tr>
<tr>
<td>Dilation</td>
<td>( \delta_b(f) = \lor_{b \in B} f_{-b} )</td>
<td>Dilation of a grayscale image ( f ) using a structuring element ( B ) is defined as the maximum of the translations of ( f ) by the vectors ( -b ) in ( B ). In other words, the dilated value at a given pixel ( x ) is the maximum value of the image in the window defined by the SE when its origin is at ( x ).</td>
</tr>
<tr>
<td>Opening</td>
<td>( \gamma_b(f) = \delta_b(\varepsilon_b(f)) )</td>
<td>Opening (( \gamma )) of the grayscale image ( f ) is defined as the result of erosion by a structuring element ( B ) followed by the dilation with the reflected ( B' ). Thus, ( \gamma ) eliminates peaks that are smaller than the SE.</td>
</tr>
<tr>
<td>Closing</td>
<td>( \phi_b(f) = \varepsilon_b(\delta_b(f)) )</td>
<td>Closing (( \phi )) of the grayscale image ( f ) is defined as the result of dilation by a structuring element ( B ) followed by the erosion with the reflected ( B' ). Hence, ( \phi ) fills valleys that are smaller than the SE.</td>
</tr>
<tr>
<td>White top-hat</td>
<td>( WTH_b(f) = f - \gamma_b(f) )</td>
<td>White top-hat (WTH) of a grayscale image ( f ) is defined as the difference between the original grayscale image ( f ) and its opening ( \gamma ).</td>
</tr>
<tr>
<td>Black top-hat</td>
<td>( BTH_b(f) = \phi_b(f) - f )</td>
<td>Black top-hat (BTH) of a grayscale image ( f ) is defined as the difference between the closing ( \phi ) of the original image and the original image.</td>
</tr>
</tbody>
</table>

Watershed transformation 

The basic notions of watershed transform was first introduced by Beucher and Lantuejoul, see (Beucher and Lantuejoul, 1979), and it’s an algorithm for segmenting images, for more details about watershed transform, see (Meyer and Beucher, 1990). In this work, we utilize the algorithm proposed by Meyer (1993).

![Structuring elements](image)

Figure 4. Structuring elements that will be used during evolution. Only SEs with 7x7 neighborhood are shown; the darker pixel is the origin pixel \( x \), and the gray area indicates the shape of the SE.

sizes: 3x3, 5x5, and 7x7 pixels. The 3x3 and 5x5 diamond and octagon SEs are equal to 3x3 and 5x5 disk SEs, thus we utilized only the 7x7 size for diamond and octagon SEs. Also, the neighborhood size of rectangle SE is restricted to 3x5 and 5x7 pixels, because 1x3 neighborhood is the same as a line of size 3 and 0° of orientation. In addition to size, line shape SEs have three orientations, 0°, 45°, and 90°.

The fundamental operators in MM are erosion (\( \varepsilon \)) and dilation (\( \delta \)) because all other morphological operators can be expressed in terms of them. A SE is used for inspecting the image, as a result, the application of erosion/dilation operator gives an eroded/dilated image which shows the locus of points where the SE fits/hits the objects in the image. For this reason, the shape and size of the SE is usually chosen according to the geometric properties of the relevant and irrelevant image structures (Soille, 2003). Rotationally symmetric SEs are often employed for the opening transformations because the results do not depend on the orientation of the image.

Soille and Pesaresi (2002) showed an analysis of the suitability of MM processing for earth observation imagery. Also, morphological processing provides a good, simplified approach to perform joint spatial/spectral classification of urban hyperspectral data (Plaza et al., 2009). Regarding the desirability of using MM for the extraction of spatial features from HSs, see (Sternberg, 1986). Table 5 shows some morphological operators used in geoscience and remote sensing applications. For a formal and more detailed explanation of each MM operators, see Soille (2003).

In summary, 1,719 bands that constitute an H compound the \( T_{\text{spatial}} \) set. Whereas that, the set \( F_{\text{spatial}} \) of functions is compounded by the morphological operators in Table 6.

Note that each operator can be used together with any of the SEs described above, this is nineteen different SEs, resulting in a total of 116 morphological operators. Additionally, we added arithmetic functions to allow the \( B^p \) algorithm to establish relationships among morphological operators. Table 6 shows the set of functions and terminals of \( EVO_{\text{spatial}} \).

Fitness Function and Accuracy Evaluation

The features selection problem involves the selection of a subset of \( d \) features from a total of \( D \) features, based on a given optimization criterion \( F(X) \), where \( X \) denotes the subset of selected features (Oh et al., 2004).

In hyperspectral imaging literature, the feature selection problem is commonly known as band selection. Band selection consists in reducing the dimensionality of hyperspectral images and finding the useful wavelengths and spectral indices for analysis, classification and regression (Yoon and Park, 2015). In this study, \( F \) is represented by the fitness function used to assess \( X \). The objective is to find a subset, \( X \), that yields the highest accuracy rate of graves classification. Graves detection problem leads to unconstrained combinatorial optimization in which the classification accuracy is the search criterion. Owing to the capability of evolutionary computation to efficiently search in large spaces, such important set of heuristic algorithms are well suited to approaching feature selection (Iglesia, 2013). In this aspect, evolutionary computation offers an alternative to the well-known Principal Components Analysis or Random Forest to deal with high dimension of hyperspectral data (Plaza et al., 2009).

We can say that, each individual genotype represents an H-VAM, although we know that it only encodes the operators of extraction and integration of features. Each H-VAM is evaluated with a set of hyperspectral images, called training set, the resulting BSMs are compared against a ground truth map and the classification accuracy is calculated from these values. In this
The results obtained during this study are discussed below.

Experimental Results

For the experiments, we apply a four-fold cross-validation method; the set of the twelve hyperspectral images (see the previous Section), was randomly divided into four subsets. Furthermore, two experiments were carried out, one of them with 2, 3, and the other with 2, 3.

Experiments with Kappa

Structural Analysis of Solutions

We picked the forty best solutions and determined how frequent the functions and terminals were, and with them in order to identify which components are useful for classification and which are not. The functions and terminals in the EVOspatial and EVOspatial operators are counted only if their respective CMs are used by the EFL.

The experimental results about EVOspatial show that 18 solutions (45%) utilize ω in 31 occasions, follow by: 111 was used 33 times by 14 solutions (35%); (ω) was used 27 times in 13 solutions (32.5%); |−| and (ω) were used 24, 23, and 41 times, respectively, by 10 individuals (25%).

All functions in Fspectral were utilized by at least, but only, two best solutions. On the other hand, the terminals in Tspatial is 1,719 wavelengths, conforming H, plus 22 spectral indexes, so only a small percentage was used by the forty best solutions. The most used reflectances were p_550, p_665, p_560, and p_705 in two solutions. Only three spectral index were used and were: WSI(p_500, p_705), WII(p_500, p_705), and EVI(p_450, p_705, p_905, p_905).

In the same way, we analyzed the EVOspatial, ErodeSet_s, function was the most used by 3 individuals in 6 occasions, followed by: BottomHat, |−|, Open(p_665, p_705), p_705, p_705, p_905, p_905. The most frequent CM was CMspatial, 38 of the solutions (95%) utilized it in 206 occasions. Cspatial was used in 31 solutions (77.5%) 144 times. This means that the spectral information has been more useful than spatial to solve the problem during the BP process.

In the next section we go into the analysis of the effectiveness of the best solution during the training and testing stages.

Effectiveness of Solutions

The well solution was selected based first, on the fitness obtained during the training and, second, on the one obtained during the testing. The most eligible individual had a value of $\kappa = 0.3621$ in training stage, and a value of $\kappa = 0.3047$ in testing stage and utilizes spectral and spatial information, see Table 7. It is based on p_550, p_665, p_705, p_905, p_905, and p_905, wavelengths. Figure 5 shows the error matrices obtained by the best individual applied to each one of the HIs. Matrices are ordered by sampling date, from oldest to most recent one. Three of the images were evaluated during test stage, see Figure 5 (a), (b), and (f), with bold captions. Below each matrix, the
values of κ and of Overall Accuracy (OA) rate are shown. The best κ and OA classifications were made in the image 050616, with 96.4% and 052016 image, with 96.12%, (g) and (h) respectively, during May month. The worst accurate classifications were obtained in 022616 image, with 93.35%, and 042216 image, with 93.9%, (b) and (f) respectively, both from testing stage. The decreasing of performance is expected in testing stage. On the other hand, the text colors allow to relate the values of the error matrices with Figure 6.

Figure 5 depicts the spatial distribution of errors/successes detection for each image. These images are sorted according to the date they were taken, starting with the oldest and ending with the most recent. Regarding the pixel’s color, green pixels indicate graves that were correctly classified as graves (true positives); gray pixels indicate graves that were incorrectly marked as no-graves (false negatives); red pixels indicate no-graves incorrectly classified as graves (false positives); blue pixels indicate no-graves correctly classified as no-graves (true negatives). These colors allow us to better analyze the classified images and infer information from the spatial arrangement of the results. For example, the pixel corresponding to the lower left corner of grave F4, see Figure 1 for graves labeling, was detected in ten of the twelve images. The most correctly detected grave was F2, it was fully detected in seven images and partially detected in one. Grave F6 was totally detected in three of the images. Finally, grave F7 was completely detected in one image and partially detected in three.
We also analyzed the reflectance characteristics of the bands used by the best solution. We started with the bands used by EVO\textsubscript{spectral}. First 969 nm (ρ\textsubscript{969}), usually the 970 spectral region is used to estimate plant water concentration, e.g., water band index (WBI) and water index (WI), see Table 3. There is a water absorption peak near 970 nm (Danson \textit{et al.}, 1992), so that the lower the leaf water content the higher reflectance is in the 970 nm region, because there is lower spectral absorption. In the same way, the band 1546 nm (ρ\textsubscript{1546}) belongs to the Middle-Infrared spectral region (MIR, 1300-2500 nm), the interaction of vegetation with radiation in this region depends partly on the volume of water content in leaf cells. For this reason, the use of MIR wavelengths is suggested for remotely sensing leaf water contents, particularly from 1550 to 1750 nm (Tucker, 1980). Considering the reflectance characteristics at wavelength 2216 nm (ρ\textsubscript{2216}), Olson and Zhu (1985) noted that the 2080 to 2350 nm band group (centered at 2215 nm) provides significant information about trees species, particularly near 2220 nm. Moreover, the bands around 2020 to 2220 nm are considered important for cellulose or lignin detection using remote sensing data as they are distinctive from soil minerals and less affected by atmospheric gases (Zandler \textit{et al.}, 2015). The last wavelength used by the EVO\textsubscript{spectral} of the best solution is 2368 nm (ρ\textsubscript{2368}). It has been said in (Herold, 2006) that the presence of calcium carbonates is suggested by features near 2330 nm, specifically for calcite (2300 nm) and dolomite (2370 nm). In this same sense, recently Roberts \textit{et al.} (2013) discovered that natural surfaces, including organic matter and microbial biomass, possessing a high density of carboxyl groups may be a mechanism by which ordered dolomite nuclei form. With regard to EVO\textsubscript{spatial}, it utilizes the 1271 nm wavelength (ρ\textsubscript{1271}). This may be explained because the spatial information can be best extracted from a high-contrast band and the wavelength of 1271 nm falls within the NIR waveband, which is a spectral region of high contrast during early vegetation regrowth. According to Dehaan and Taylor (2002), Samphire (\textit{Halosarcia pergranulata}), a dense succulent shrub, is characterized by having a distinctive slope to the infrared reflectance plateau between 1250 and 1400 nm, and the spectra of cultivated mixed rye and clover pasture vegetation, show a nearly at infrared reflectance high between 1250 and 1450 nm.

Experiments with Weighted-Kappa

In light of observations from the first experiment with Kappa coefficient as fitness function, we decided to do another experiment by changing the fitness function by the weighted Kappa (κ\textsubscript{w}) statistic to try to improve the results. The approach of κ\textsubscript{w} seems useful in two different situations: first, when categories of thematic maps are ordered. For ordered categories it can be easily recognized that relative disagreement may vary, even though κ\textsubscript{w} was introduced in the title of the article by Cohen (1968) as a measure for ‘nominal scale agreement’; second, when the categories are classified according to nominal or ordinal scales, and the seriousness of erroneous map data could be quantified for a certain map user or group of map users (Naesset, 2014). It is usually unrealistic to assume that all the errors are of identical importance because it is a human subjective decision.

In order to calculate κ\textsubscript{w} is necessary to establish a table of weights which measure the degree of disagreement between

![Figure 6. Binary saliency maps and classification of best individual over the twelve images.](image-url)
the two type of errors, false positive and false negative, the higher the disagreement the higher the weight. The table of weights should be a matrix with zeros in the main diagonal and positive values off the main diagonal. Table 8 provides the weights used in this experiment, which can be interpreted in the following way: the error of classifying a grave as no-grave is four times less desirable than the error of classifying a no-grave as grave. Hence the H-VAM will tend to find more graves at the cost of increasing false detections.

Table 8. Disagreement weight matrix used during experiment with weighted Kappa.

<table>
<thead>
<tr>
<th></th>
<th>G</th>
<th>NG</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>NG</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Although there are techniques for assigning values to the weights of disagreement, as proposed by Cicchetti and Allison (1971), in this study we based the selection of values on results obtained from several trials with different weight matrices.

Another change that was introduced to the fitness function is the binarization method of the image. While in the first experiments a binarization with threshold equal to zero was used, during these experiments the method proposed by Otsu (1979) was used to determine the binarization threshold dynamically. Below, we show the analysis of the results of this second experiment using kₙ as fitness function.

Structural Analysis of Solutions

In this section we analyze the structure of the best forty solutions by calculating the use frequency of functions and terminals. These components that are useful for classification and which are not. The functions and terminals in the EVO_spectral and EVO_spatial operators are only counted if their respective CMs are used by the EFI.

First, we analyzed EVO_spectral operator. Regarding $F_{spectral}$ set, 14 solutions (35%) used + in 30 occasions, followed by the NDSI and H-VAM, used by 11 solutions (27.5%), in 20 and 31 occasions, respectively. The 1–1 function was used by 8 solutions (20%) in 11 occasions, ×, (0) and ln were used by 6 solutions (15%) in 23, 40, and 17 occasions, respectively. This + function were also, the most utilized by the best solutions obtained in the experiments with $\kappa$. On the contrary, +, × and supremum functions were the less used. Therefore, based on the results of both experiments, it is logical to think that $\kappa$, H-VAM, ×, and the NDSI are useful for grave classification. About the terminals set $F_{spectral}$, the most used wavelengths were $\rho_{861}$, $\rho_{841}$, and $\rho_{2243}$ in 052016 image with $\kappa_1=0.5404$, $\kappa_2=0.4719$ and OA=94.46% and in 070816 image with $\kappa_1=0.4963$, $\kappa_2=0.4001$ and OA=92.8%, both are shown in Figure 7h) and 7k). The worst accurate classifications were obtained in the 030716 image with $\kappa_1=0.2052$, $\kappa_2=0.1976$ and OA=92.8%, and in 060316 image with $\kappa_1=0.3038$, $\kappa_2=0.353$ and OA=92.52%, both are shown in Figure 7c and 7i.

Effectiveness of Solutions

In the same way as with $\kappa$ experiments, we have selected the best individual, shown in Table 9, it got an average value of $\kappa_1=0.4285$ in the training stage and of $\kappa_1=0.3662$ in the testing stage. This solution is based on spectral and spatial information and is based on $\rho_{901}$ and $\rho_{741}$, and $\rho_{2243}$ terminal solutions. This solution has a peculiarity: most utilized wavelengths belong to the VNIR portion of the electromagnetic spectrum, except for $\rho_{2243}$ that belongs to the SWIR2 portion. The error matrices generated by the best solution on the twelve images, are shown in Figure 7. Three of the images were evaluated during testing stage: 030716, 052016, and 060316, their error matrices are shown in Figure 7c, 7h, and 7i, respectively, with bold captions. Below each error matrix are shown the values of $\kappa_1$, $\kappa_2$, and OA rate. The best $\kappa_2$ values were obtained in images 052016 with $\kappa_1=0.5404$, $\kappa_2=0.4719$ and OA=94.46% and in 070816 image with $\kappa_1=0.4963$, $\kappa_2=0.4001$ and OA=92.8%, both are shown in Figure 7h) and 7k).

In order to compare the performance of the best solution obtained with $\kappa$ against the one obtained with $\kappa_2$, we calculated the average precision and recall of the classification of the twelve Hs, and Table 10 depicts the results.

Considering that the purpose of using $\kappa_1$, was to guide evolution by penalizing more the classification of a “grave” as “no-grave” than a “no-grave” as a “grave”. The results shown in the Table 10 are consistent with that objective. Since penalizing one error more than the other, caused that the solutions tended to make the less penalized error without suffering a considerable decrease in its fitness, which in turn causes the decrease in its precision. But, by heavily penalizing the classification of a “grave” as a “no-grave”, the recalls improved considerably, relative to $\kappa$, which was what we were looking for when implementing $\kappa_2$, even at the cost of a lower accuracy.

Then, the average $\kappa$, precision, and recall measures, obtained by the best solutions of the $\kappa$ and $\kappa_2$ experiments were calculated for the first three months and for the last three months of the monitoring period, separately; see Table 11. These results suggest that, although some graves classification was possible in the images obtained during the first three months, the measures mentioned above improve notably in the images obtained after the three months. Figure 8 shows that, in the case of the best solution of $\kappa_2$ experiment, the classification of graves F4 and F6, the ones with less buried carcasses, two and one, respectively, improves in images acquired after three months. It seems logical to think that the number of buried bodies is related with the optimal window of time for image acquisition.

On the other hand, Figure 8 depicts the spatial distribution of detection errors/successes in each image. The most detected
grave was F2 (see Figure 1 for graves labeling); it was fully detected in five images and partially detected in seven, at last one pixel of F2 was detected in all images. The grave F7 was the most fully detected in eleven images. These results are consistent with the number of carcasses in the graves F2 and F7, since they are the ones with the most buried bodies, i.e., three and four respectively. The grave F6 was totally detected in one image and partially detected in two. Finally, the grave F4 was fully detected in one image and partially detected in another.

We also analyzed the properties of reflectance at the bands used by the best solution. First, the bands used by EVO spectral. Some studies link the use of spectral region near the 653 nm band ($\rho_{653}$) to high green biomass regions and to vegetation

Table 9. The best solution of experiment with weighted Kappa as fitness function.

<table>
<thead>
<tr>
<th>Best Individual</th>
<th>Fit_{test}</th>
<th>Fit_{val}</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVO_{spectral}: NDSI[Inf(((\rho_{641} - \rho_{2243})^3 + (\rho_{653})^{1/2} \times 1.90)) - 1.95)); (\rho_{641} - \rho_{2243})^4, \rho_{653})]</td>
<td>0.4285</td>
<td>0.3662</td>
</tr>
<tr>
<td>EVO_{spatial}: BTH (\epsilon_{653}, \rho_{653})</td>
<td>0.4339</td>
<td>0.3333</td>
</tr>
</tbody>
</table>

Table 10. Average precision, recall and $\kappa$ of the best solutions of $\kappa$, $\kappa_w$ and PLS-VIP (Silván-Cárdenas et al., 2017) experiments.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Average Precision</th>
<th>Average Recall</th>
<th>Average $\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>0.4339</td>
<td>0.3333</td>
<td>0.3477</td>
</tr>
<tr>
<td>$\kappa_w$</td>
<td>0.3383</td>
<td>0.4896</td>
<td>0.3645</td>
</tr>
<tr>
<td>PLS-VIP</td>
<td>0.2416</td>
<td>0.2813</td>
<td>0.2086</td>
</tr>
</tbody>
</table>

Figure 7. Error matrices for the twelve hyperspectral images obtained by the best solution during weighted Kappa experiments.
indices as EVI and SAVI (Peñuelas et al., 1995; Huete, 1988; Huete et al., 2002; Ionue et al., 2008). Considering the spectral characteristics of the band 796 nm, in (Peñuelas et al., 1995) the authors used the $\rho_{800}$ to characterize the physiology and phenology of a wide range of leaves species and conditions. Furthermore, $\rho_{796}$ and $\rho_{841}$ are in the NIR (700 to 900 nm) spectral region, NIR is only affected by leaf internal structure and leaf dry matter content but no by water content. A combination of SWIR and NIR reflectance information can remove the effect of leaf internal structure and leaf dry matter content and it can improve the accuracy of retrieving the vegetation water content (Fensholt and Sandholt, 2003). Regarding the $\rho_{2243}$, Sims and Gamon (2003) and Jin and Sader (2005) reported that the band centered at 2250 nm has water absorption features and it can be used for prediction of plant water status. Besides, as it had previously been seen in previous , there is significant information about trees species in spectral bands near 2220 nm (Olson and Zhu, 1985). With regard to EVO spatial, it utilizes the 861 nm wavelength ($\rho_{861}$) that also belongs to NIR spectral region.

About the complexity of the solutions obtained by the BP, the operations that consume more time are the MM, $|\cdot|^2$, $|\cdot|^k$, $|\cdot|^2$ and $|\cdot|^k$, and multiplication and division of images ($\times$, $\div$). In this study the HIs have a size of 19 x 19 pixels, so the execution time of a H-VAM algorithm in an H takes a fraction of a second in a computer Intel® Core™ 5-6500 at 3.20GHz, and 7.7GB in RAM. The best H-VAMs are simple and utilize a small amount of computational resources and time.

Finally, a comparison among the results of the experiments with $\kappa$ and $\kappa_w$, and the ones reported by Silván-Cárdenas et al. (2017) was conducted. In such study, the authors utilized a Partial Least Squares (PLS) method followed by a technique of variable selection, based on Variable Importance of Projection (VIP) method, called Bootstrap-VIP, to compute spectral separability for detection of graves. According to the study of Maitra and Yan (2008), the PLS technique is more efficient than the Principal Component Analysis (PCA) technique for dimension reduction due to the supervised nature of its algorithm. In light of this finding, it is not necessary to compare our results against those obtained through PCA. We obtained the error matrices from the images shown in Figure 8 in (Silván-Cárdenas et al., 2017) in order to compute the average $\kappa$, precision and recall. The results of such computation are shown in Table 10. Our two evolutionary approaches were better than PLS-VIP approach regarding these three criteria. This may be explained by the limited capability of a linear mapping of spectral information through PLS, unlike the highly non-linear HVAMs that put spectral and spatial information together. The best precision was obtained by our approach guided by $\kappa$. The best recall and detection accuracy, measured with $\kappa$ coefficient of agreement, were obtained by our evolutionary approach guided by $\kappa_w$. Thereby, the strategy of setting the $\kappa_w$ as fitness function of the BP optimization algorithm gave the expected results, improving the detection accuracy of BP compared with $\kappa$ and PLS-VIP.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>First three months average:</th>
<th>Last three months average:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>Precision</td>
<td>Recall</td>
</tr>
<tr>
<td>$\kappa_w$</td>
<td>0.37</td>
<td>0.3125</td>
</tr>
<tr>
<td></td>
<td>0.3276</td>
<td>0.4375</td>
</tr>
</tbody>
</table>

Table 11. Average precision, recall and $\kappa$, for the first and last three months, of the best solutions of $\kappa$ and $\kappa_w$ experiments.

Figure 8. Binary saliency maps and classification of best individual obtained with weighted Kappa over the twelve images.
The best H-VAM algorithms found by the BP showed to be simple and fast. In fact, the H-VAM with the best accuracy is the simplest and fastest of the two.

The research around the detection of clandestine graves through hyperspectral images began recently. For this reason, all the information provided by the few researches on this subject is of great importance for future research in hyperspectral imagery, remote sensing, and other areas as forensic geology. In further work, it is possible to incorporate the information gathered in this and other research to accelerate the evolution process and to create new, more efficient hyperspectral models of attention aiming graves detection. With this goal in mind, more research should be done to create graves detectors in different scenarios, since the environment where burials take place may have very different characteristics. The incorporation of these characteristics into the method could improve their accuracy.

BP and H-VAMs can be used in any hyperspectral image classification problem e.g., urban soil classification, vegetation classification. In addition, the H-VAMs can be modified based on the a priori knowledge we have of the problem and thus adapt it to extract the necessary characteristics to solve it. On the other hand, there is great difficulty in obtaining a large amount of hyperspectral images that better reflect the enormous variety of soils and vegetation in which the graves can be found. This small number of images can lead to the BP overfitting the H-VAMs and its accuracy decreasing in the presence of any change in the characteristics of the scenario. In this sense, another disadvantage is that if we increase the number of images of the training set and the size of those images, the execution time of the BP can be increased to the order of weeks. Although the obtained H-VAMs would be more robust to changes in the environment.

Among the contributions of this work, we enlisted the following. First, the resulting H-VAMs for graves detection performed better compared with previous work. Second, the BP allows the automatic design of visual attention models for hyperspectral images; this helps researchers to solve the problem of moving from the classic model of visual attention to the hyperspectral model by hand. Third, a comparison of the results of using the $\kappa$ and $\kappa_i$ coefficients demonstrated that guiding the search of H-VAMs with $\kappa_i$ yields better accuracy results due to the introduction of the degree of disagreement. And finally, the new VAM, called H-VAM, for the fusion of spatial and spectral features of hyperspectral images is proposed as the base of task-specific evolved VAMs.

Acknowledgments
This research and data used in this paper was acquired thanks to the support provided by the institutional funds (FOINS) of the Mexican National Science and Technology Council (CONACyT), Grant No. 249205 “Viabilidad de las imágenes hyperspecles para la detección de fosas clandestinas en México”. In addition, the authors acknowledge the valuable comments of the anonymous reviewers and editors of PE&RS.

References


Plaza, A., J.A. Benediktsson, J.W. Boardman, J. Barelle, L. Brusszone, G. Camps-Valls, J.


