INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Reconhecimento Automático de Objetos Megalíticos em Áreas de Interesse em Imagens de Satélite

Automatic Recognition of Megalithic Objects in Areas of Interest in Satellite Imagery

David Galvão Chambel Caçador

Mestrado em Engenharia Informática

Orientadora:
Doutora Ana Maria Carvalho de Almeida, Professora Associada,
Iscte Instituto Universitário de Lisboa

Co-Orientador:
Doutor João Pedro Afonso Oliveira da Silva, Professor Auxiliar,
Iscte Instituto Universitário de Lisboa

Novembro,2020

TECNOLOGIAS
E ARQUITETURA

Departamento de Ciências e Tecnologias da Informação

Reconhecimento Automático de Objetos Megalíticos em Áreas de Interesse em Imagens de Satélite

Automatic Recognition of Megalithic Objects in Areas of Interest in Satellite Imagery

David Galvão Chambel Caçador

Mestrado em Engenharia Informática

Orientadora:
Doutora Ana Maria Carvalho de Almeida, Professora Associada,
Iscte Instituto Universitário de Lisboa

Co-Orientador:
Doutor João Pedro Afonso Oliveira da Silva, Professor Auxiliar,
Iscte Instituto Universitário de Lisboa

Novembro,2020

Agradecimentos

Agradeço a ambos os meus orientadores, a Professora Ana Maria de Almeida e o Professor João Pedro Oliveira, pela sua orientação, disponibilização de tempo ao longo deste período turbulento e valiosas contribuições.

Agradeço também à *European Space Agency* por me cederem as imagens sem as quais não seria possível realizar este trabalho.

Agradeço à minha família por me terem suportado ao longo de todo o meu percurso escolar, oferecendo o seu suporte incondicional quer emocional quer físico.

Obrigado.

Resumo

A integração de imagens aéreas e satélite com metodologias do âmbito da Aprendizagem Automática (Machine Learning) e Visão Computacional (Computer Vision) providenciou a capacidade de cobrir terrenos amplos e permitiu a extração de novas características do terreno, fornecendo meios para localizar, monitorizar e proteger remotamente de destruição locais com património cultural. Este trabalho procura unir informações espectrais e espaciais, derivadas de imagens multiespectrais, hiperespectrais e pancromáticas cedidas e obtidas pelos satélites da ESA, para implementar um sistema de deteção automática de alto desempenho capaz de detetar dolmens enterrados ou cobertos por vegetação. Separadamente, implementaram-se métodos onde se tentou desenvolver um sistema baseado na assinatura espectral do material dos dolmens, para imagens hiperespectrais, e, para imagens pancromáticas e multiespectrais, um sistema para extrair índices espectrais, fundir todos os índices extraídos num só e aplicar deteção de círculos para identificar locais onde haja grande probabilidade de existir um dólmen, após eliminação de falsos positivos através de uma técnica supervisionada de Aprendizagem Automática. As imagens hiperespectrais não demonstraram capacidade de definir uma assinatura de material de dolmens e, por extensão, não aptas a delinear automaticamente regiões com alta probabilidade de presença de dólmen, devido à grande dimensão dos pixels em comparação com os pixels dos dolmens conhecidos e os arredores dos dolmens serem demasiado semelhantes. O sistema criado da fusão de imagens pancromáticas e multiespectrais mostrou-se capaz de detetar localizações de dólmen, provando, simultaneamente, que parte da informação na ontologia existente para dólmens em Portugal (nomeadamente, a inserção usual próximo de fontes de água) pode ser usada para delinear áreas de alta probabilidade de presença de dólmen e que o uso de métodos de aprendizagem supervisionada permitiu eliminar cerca de 87.2% falsos positivos.

Palavras-Chave: Imagens Multiespectrais; Imagens Pancromáticas; Imagens Hiperespectrais; Transformação de Hough Circular; Índices Espectrais; Aprendizagem Automática.

Abstract

The integration of airborne and satellite imagery with Computer Vision and Machine Learning methodologies provided the ability of covering ample ground and enabled the detection of new terrain features, providing means to remotely locate, monitor and protect from destruction sites of cultural heritage. This work seeks to fuse spectral information obtainable from multispectral or hyperspectral images with spatial information derived from panchromatic images, provided by ESA and obtained from its satellites, to implement a high-performance automatic detection system capable of detecting buried or covered by vegetation dolmens. Separate methods were implemented, where for hyperspectral images a system was attempted based on the dolmens respective spectral material signature, and for panchromatic and multispectral images a system that extracted spectral indices, fused all into one and applied circle detection to identify dolmen locations, eliminating false positives through supervised machine learning. The hyperspectral images could not be used for the creation of a dolmens' material signature, and by extension cannot automatically delineate regions of high likelihood of dolmen presence in images, due to the size of each pixel being of much higher dimensions than the known dolmens and the dolmens surrounding environments being too similar. The system created through the fusion of panchromatic and multispectral images proved capable of detecting dolmen locations, while simultaneously proving that part of the existent defined ontology of the dolmen (their insertion near water sources) can be used to delineate areas of high probability of dolmen presence, and that using supervised learning methods can enable the elimination of around 87.2% of false positives.

Keywords: Multispectral Images; Panchromatic Images; Hyperspectral Images; Circular Hough Transform; Spectral Indices; Supervised Machine Learning.

Index

Index of Tables

Index of Figures

List of Abbreviations and Acronyms

3D – Three-Dimensional

BAI – Built-up Areas Index

BoVW – Bag-of-Visual Words

CDF – Cumulative Density Function

CHRIS – Compact High Resolution Imaging Spectrometer

CHT – Circular Hough Transform

CLAHE – Contrast-Limited Adaptive Histogram Equalization

CNNs – Convolutional Neural Networks

CS – Contrast Stretching

DEM – Digital Elevation Models

DS – Decorrelation Stretch

DSR – Design Science Research process

DTM – Digital Terrain Models

ESA – European Space Agency

GC – Gama Correction

GEOBIA – Geographic Object-Based Image Analysis

HE – Histogram Equalization

HS – Hyperspectral

HSV – Hue, Saturation and Value

IHS – Intensity, Hue, Saturation

LiDAR – Light Detection and Ranging

MS – Multispectral

NDSI – Normalized Difference Soil Index

NDVI – Normalized Difference Vegetation Index

NDWI – Normalized Difference Water Index

NHFD – Non-Homogeneous Feature Difference

NIR – Near Infrared

NIR1 – Near Infrared 1

NIR2 – Near Infrared 2

NLPCA – Nonlinear Principal Component Analysis

OBIA – Object-Based Image Analysis

ORS2A – Ortho Ready Standard Level 2A

PAN – Panchromatic

PCA – Principal Components Analysis

RGB – Red, Green, Blue

RS – Remote Sensing

SDI – Shadow Detection Index

TOA – Top-of-Atmosphere

WV-2 – WorldView-2

Chapter 1 – Introduction

1.1. Context

With the integration of airborne and satellite imagery alongside Computer Vision and Machine Learning methodologies in the field of archaeology, the way that archaeological research is performed is changing. With those systems' ability, via airborne and satellite imagery that provides a more ample ground coverage, the detection of new terrain features is facilitated [1], in a scale unfeasible when relying in human ground coverage alone. Moreover, we now have the means to remotely locate and monitor sites of cultural heritage and protect it from destruction, caused by both natural and anthropogenic factors [2]–[6] and while Portugal is rich in megalithic monuments, there still exist monuments left unknown, namely in the area of Alentejo, which further facilitates their degradation and destruction through natural means or by human development of the territory [4]. It stands to notice that there are key benefits for this remote identification in employing satellite over airborne imagery such as the greater spectral range and the higher spatial level they offer [7].

1.2. Motivation

Within this context, and particularly in the archaeological field, we have been witnessing an increased use of satellite imagery, which itself led to a need to optimize the traditional manual methods employed in image analysis (e.g. manual classification of an image by an expert), with investigations and case studies occurring at various locations and over various differing object types, as evidenced in multiple studies [1], [2], [14], [3], [5], [8]–[13]. However, both due to advances made in satellite sensors, restrictions in data accessibility and the variation in archaeological structures from region to region, it has resulted in various approaches being undertaken, either for the same or for different, data types [6].

Nevertheless, the still small number of existing studies on the field of archaeology has already shown both the need to automate the process of remote identification of ground zones presenting higher likelihood of possessing monuments, and that, from the existent

approaches, it is hard to extract useful knowledge applicable to different archaeological work as the existing works tailor their methods for the particular monuments, image types, and environment specifications.

Finally, the awareness of the existence and easiness of access to domain knowledge expertise for dolmens and the area they are mostly inserted in Portugal [15], as well as the possibility of acquisition of different types of satellite images for the geographical area in question, captured and shared by the European Space Agency (ESA), makes this area suitable for the investigation on the satisfiability of the development of automated methods for dolmen's remote identification in the Alentejo district.

1.3. Research Questions

Taking into consideration the context and motivation behind it, this dissertation focuses mainly in the creation of a system capable of automatically recognizing megalithic objects (specifically dolmens) in areas of interest (namely in Alentejo) using satellite imagery. The system will investigate the use of different image types and methods and will use existing domain knowledge for informed classification decision. Such a system has the potential of easing the detection of currently unknown monuments, by aiding archaeologists through the delineation of areas of higher probability of dolmen presence. To achieve the successful completion of this purpose, we intend to provide answers for the following three research questions.

Can the use of satellite spectral analysis of the image's ground allow for the extraction of spectral information that enables the automatic detection of megalithic monuments, usually buried or covered by vegetation, and their detection?

In what way can the fusion of the various sources of satellite image types (panchromatic, multispectral, or hyperspectral) be achieved in order to minimize the loss of either spectral or spatial information incurred during the process and still gain in features for identification of relevant areas of interest in ground images?

For the various types of satellite images used, and in view of the answers to previous questions, what is the most effective features extraction strategy for improving the performance of a dolmen's automatic detection system?

1.4. Objectives

Under these optics, this work seeks to make use of known indices and other spectral information, obtainable from either multispectral or hyperspectral satellite images, to allow for the implementation of image pattern recognition techniques and detect buried or covered by vegetation archaeological remains, namely dolmens.

In addition, we shall investigate the best approach for the implementation of the intended automatic detection system using the different image sources and types next described, fusing the image with the highest spatial information available with the extracted spectral features, that allows for the detection of the area's archaeological remains with the minimal possible false dolmen detection.

1.5. Research Methods

The research method used was the Design Science Research process (DSR), since this research is undertaken in order to design an artefact: an automated image object identification system. According to [16], the DSR is subdivided into six main steps, that can generally be described as: (1) an identification of the problem and respective motivation, where the research problem is defined and the value of the solution justified; (2) the objectives for the solution are defined; (3) relates to the design and development, that is, the creation of an artefact, such as models, instantiations, methods, or constructs; (4) the demonstration of the use of the developed artefact to solve one or more instances of the problem, either through experimentation, simulation, case study, proof, or other appropriate activity; (5) the evaluation of the solution, comparing the obtained results from using the artefact with the stated objectives; finally, (6) the communication of both the problem and its importance, the developed artefact and both its utility and novelty, the rigor of its design and its effectiveness to other researchers and relevant audiences.

The above described method can be assigned to four main phases: the problem identification phase, that encompasses both the problem and motivation identification and definition of objectives; the solution design phase, comprising the third step of design and development of the artefact; the evaluation phase, covering the demonstration of the

established artefact and evaluation of the solution; the final phase, the summarization of the results and respective communication through their publishing [17].

Accordingly, both the first and second steps are described in the introductory chapter. The next steps are primarily defined through the literary review, conducted in the second chapter of this thesis. The third step id integrated in the third chapter of the dissertation. Both the fourth and fifth steps stand as the fourth chapter, discussing the implementation and results and, finally, the sixth step is found in the concluding chapter, where the main conclusions, contributions, limitations, and future research are communicated

1.6. Structure and Organization of the Dissertation

The current thesis is divided in five chapters that reflect the phases undergone until its conclusion.

The first chapter introduces the investigation's context and motivation, as well as its theme, objectives, and brief structure description.

The second chapter divulges a revision of the related literature, the theoretical foundations beyond the conception of the work that provide the theoretical bases needed to understand it.

The third chapter describes the methodological development, revealing the needed background for the system's implementation.

The fourth chapter presents the implementation used and an analysis and discussion of the results obtained.

Finally, in the fifth chapter, the conclusions of this study and recommendations, limitations, and future research are presented.

Chapter 2 – Literature Review

This chapter describes the Literature Review and the context upon which this thesis is inserted that is derived from existent scientific works, being divided into nine sections: the first introducing the archaeological context of the objects to be detected; the second, presenting a previously ontology of the monuments, created by an expert; the third, expands on the automatic remote sensing methodologies; the fourth, introduces the satellite imagery types and image fusion methods; the fifth, presents existing classification based approaches and related works; the sixth, describes the methodologies used in related works; and the seventh, presenting some conclusions obtained with this investigation.

2.1. Archaeological Context

In the region of Pavia, the archaeological monuments investigated, i.e., the dolmens, stem from a currently unknown period over 4000 years ago and it is currently unknown whether they were a target of reuse of an older, unused or vacant, monument that was already present or were built from the ground up, with only their current shape being known [4], [18]. Since they are present through the rise and fall of several civilizations, they may have been reused, buried, destroyed or annexed to other structures, and be found in rural or city areas [4], [15].

The dolmens can be found dispersed throughout the region's territory, majorly distributed near the main riversides and areas near rocky outcrops[1] [4], [6], [12]. They are commonly found grouped up to a dozen, with distances in-between being relatively small. However for the rare isolated one's, the distance between it and the nearest monument doesn't exceed a few kilometres [4]. Based on data obtained by [4] from the database of Portal do Arqueólogo da Direção Geral do Património Cultural, a total of sixty eight identified dolmens are present in the study area. From this data, an approximate

[1] Rocky outcrops are geological features, usually steep, that support ecological influence beyond their area [93].

geographic location, of which most of them are represented with a coordinate accuracy ranging from 20m to 200m, of the dolmens in the area can be seen in Figure 1.

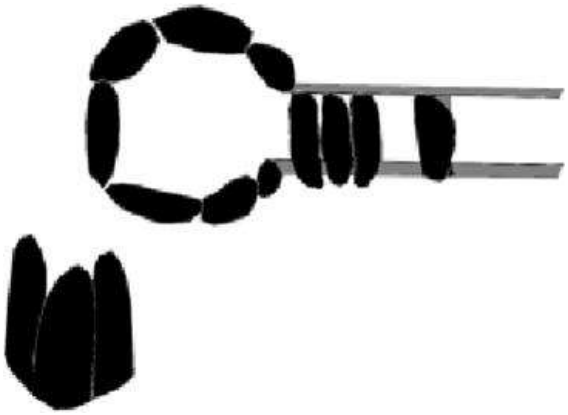


Figure 1:Example of a dolmens' schematic aerial view survey Source: [6].

These monuments have been identified to possess, in their majority, a polygonal shape, with seven pillars and a corridor of varying dimensions [4], [15], [18] that leads towards the polygonal shaped chamber, ranging from 2 to 5 meters and that is usually covered by a slab, in the format of a table or a hat [4], [6], [15]. While in their original form, these monumental structures were not visible, being covered by successive layers of earth and stone, called barrows or tumulus, the presence of the tumulus may have disappeared over the long period of time since its construction [4], [15].

Although the monuments already investigated in the case study's region, introduce certain unique polymorphisms to the dolmens structure, it's still possible to find similarities between them and thus allow for their identification through ground or vertical (aerial and satellite) images [4], [6], [19].

Consequently, while dolmens have been observed with or without several of their features depending on their state of degradation and destruction, throughout the entire region all share a common structure: a chamber, the primary feature that allows dolmens visualization from aerial and satellite imageries [4], [15], [19]. Considering the natural environment in which rocks are placed irregularly, the geometric shape present in the chamber makes it stand out from the surroundings, as even if such regular forms are found in nature they are usually smaller than the 2 to 5 meters diameter of the dolmens' chamber [4], [15].

2.2. Dolmen Ontology

An ontology defines what an object is and its composing features. As such, ontologies enable the logical organization of current existing knowledge on diverse themes, working as a methodology for the integration and representation of current information available from different domains [15], [20].

It facilitates the sharing and reuse of information, and increases the common understanding of knowledge of a domain between a machine and a man, allowing machines to process and collect resources intelligently, while simultaneously facilitating communication in-between the various devices present in the network [15], [20].

According to [15, p. 2], [21], an ontology is "*composed of several entities: a set of C concepts; a set of R relations; and a set of A axioms*". It allows to describe an object, based on its attributes, and use of these descriptions to recognise others of the same class [15].

As stated by [15], [22], ontologies are created to be easily reused and integrated with other ontologies and software systems, like for this case study that seeks to integrate the developed ontology of dolmens present in Pavia into an automatic classification system in order to detect and classify similar monuments.

The knowledge graph representing the dolmens' ontology, seen in Figure 2, provide the necessary context for the improvement of the machine learning algorithms capabilities and identification of hidden patterns [15].

Figure 2: Ontology model of dolmens of the Mora region Source: [15].

2.3. Automatic Remote Sensing Methodologies

Since the beginning of the 20th century, Remote Sensing (RS) techniques have been applied to aerial photographs with the intuit of identifying features not visible at ground level and detecting objects of interest [2], [6].

However, with the appearance of satellite images, the use of RS methods applied to aerial photography have seen their use extended towards their application on satellite imagery, whose large scale datasets created new trials for their respective image analysis [23].

2.3.1 Benefits of Satellite Imagery

With the continuous technological development of satellites, RS techniques became able to be used to cover any area present in the world in a relatively short period [6], [14], [23], as exemplified by the WorldView-2 (WV-2) satellite that offers fast retargeting, having an average revisit time of any site around the world of 1.1 days, and possesses the capability of collecting data of vast areas of over 10,000km² [24], [25]. Further, as sensors

continued to improve, not only has it allowed to capture imagery of higher spatial resolutions, giving rise to panchromatic (PAN) and Light Detection and Ranging (LiDAR) imagery types, but it has also allowed to capture a greater number of spectral bands, which in turn has increased the usage of multispectral (MS) and hyperspectral (HS) images [26].

The main advantage of using satellite imagery over aerial photography is the possibility of using a wider number of spectral bands to extract spectral information of an object [2], [8], [9]. This has turned satellite imagery into a valuable source of data for RS methods [2].

As such, with an increased capacity of detecting and capturing features at ever higher spatial, geometrical, spectral and radiometric resolutions, of objects present at ground level through satellite imagery, RS techniques applied to the field of archaeology have spiked, as shown by the multitude of studies performed [1]–[3], [5], [8], [11]–[14].

2.3.2 Pattern Recognition Methodologies

Due to the rise of RS methods and the ever-increasing quantity of data collected from more advanced sensors, a need to facilitate and optimise the detection of various objects in images was created as its interpretation becomes increasingly more problematic to be done manually [6], [13].

The means by which it was achieved, was by utilizing Machine Learning and Computer Vision methodologies, thereby allowing for the recognition of patterns present in both aerial and satellite imagery, from which it becomes possible to automatically extract information [27] that will allow the identification and classification of megalithic structures in existing images [6].

As such, in the field of archaeology, these methods have come to represent a "*non-intrusive restrictive search*" [6, p. 1], serving as a tool of optimization of traditional fieldwork [6], [14], [27], allowing for diminishing its high monetary cost and time-consuming tasks [6], [14]. Further, such methods represent the capacity to protect the megalithic structures cultural heritage and avoid their destruction [2], [5], [6].

2.4. Satellite Imagery Types

Due to the different sensors available to capture satellite data, different authors approach the field with different types of imagery. Some put greater emphasis in the spectral differences present [2], [5], [8], while others prefer to capture and use spatial features with greater clarity [1], [3], [12], [13]. Finally, some attempt to get the best of both approaches by recurring to the fusion of different types, one spectral and one spatial, of imagery [7], [9]–[11], [14].

2.4.1 LiDAR and Panchromatic Imageries

PAN and LiDAR sensors grant the capture of imagery with greater spatial features and higher resolution. PAN sensors allow for the capture of a large span of wavelengths, present in the visible spectrum of light, in a single band [8], [14]. This type of sensors produces imagery of higher spatial resolution [2], [8], which enable the extraction of more spatial features and thus the visualization of variances [9].

LiDAR sensors possess the capability of generating high resolution terrain models [1], [3], [12], [13], for the visualization of high-resolution terrain spatial features. To do so, they use *"pulsed laser beams to obtain precise three-dimensional (3D) information about the terrain"* [12, p. 3] and recorded in point clouds [3], [11]. The point clouds are then processed in order to generate Digital Elevation Models (DEM) [1] or Digital Terrain Models (DTM) [1], [3], [10], [12], [13], which allow to use highly accurate levels of elevation to discover and visualize different topographic features [1], [10], [11]. As stated, *"by incorporating 3D data, the detection of archaeological features becomes easier"* [28, p. 5].

Thus, LiDAR data has become one of the most frequently used data in archaeological studies of recent times, to perform surveys and prospection of new sites [1], [3], [12], [28].

2.4.2 Multispectral and Hyperspectral Imageries

MS and HS sensors grant the opportunity of detecting higher spectral differences in captured imagery.

MS sensors are equipped to be able to *"capture data within the visible and non-visible spectrum, encompassing a portion of the ultraviolet region, the visible, and the IR region"*

[2, p. 2], presenting the captured information divided by bands, each representing a small part of the total wavelength of the light spectrum [2], [8], [14]. Over time, it has come to be considered the standard means by which ground cover and soil types should be classified [2], by making use of specific spectral bands or a combination (an Index) of these spectral bands [8], corroborated by spectral features presenting a higher degree of separability [9]. However, according to [5, p. 659], this type of sensors "*usually lack in spatial and spectral resolutions necessary in identifying buried archaeological structures*", whose low spectral resolution stems from the lack of possible captured bands from their limited spectrum while the low spatial resolution stems from its sensors focusing on the capture of spectral data.

HS sensors obtain information relative to the optical, physical and chemical properties of the materials and surface of the earth [29]. The main differences between MS and HS stem from the higher number of bands captured and represented in HS with respect to MS [2], [10], [26], [29]. An MS sensor presents up to around 20 spectral bands [2] while a HS sensor presents usually from 100 to 200 bands or higher [2], [29], [30], [2], [29], [30]. Moreover, the MS spectral bands are spread out in the lights spectrum [10], [26], whereas, for the HS, spectral bands are adjacent to each other [5], [26], [29], which gives the HS sensors the capacity to obtain useful information by creating dense spectral signatures [10], where each pixel represents the complete spectrum [5] and effectively overcomes the resolution drawback of MS sensors [5].

Both sensors can be used to create composite features, such as indexes for vegetation and water, being helpful for the discovery of vegetation and nearby water bodies [10], [11], [29]. Additionally, HS sensors are more suited to map mineralogy of land materials [29].

2.4.3 Image Fusion

As stated, some of the works have pursued the benefits of several types of imagery simultaneously, by means of a combination of imagery types, in which one type focuses on the capture of spectral features and the other focuses on capturing detailed spatial features, using the previously described sensors.

This technique, called image fusion or pan-sharpening, is a way to integrate the spatial detail of a high-resolution PAN or LiDAR image and the colour information of a low-

resolution MS or HS image to produce and make use of a high-resolution MS or HS image [6], [31], [32], allowing for an increase in interpretation capabilities and, overall, resulting in more reliable information [7].

There have been many image fusion methods created and applied to fuse two differing images [7], [32], with some stating that there exists over a hundred variations of this technique [31]. Overall, the most effective and popular are: Intensity, Hue, Saturation (IHS) [7], [31], [32]; Principal Components Analysis (PCA) [7], [31], [32]; wavelet base fusion [31], [32]; and arithmetic combinations [31].

However, each of these methods have limitations, which are further exacerbated by the different sensors of origin and respective imagery characteristics [7], [31]. The main problem is colour distortion [7], [8], [31], that is, the distortion of "*the spectral characteristics of the original multispectral images to different extents*" [7, p. 1]. To solve this problem, various new methods, such as: synthetic variable ratio, Gram-Schmidt transform and smoothing filter-based intensity modulation [7], and new strategies of applying existing and tested methods, such as the case of the IHS method stretching each I, H and S band in accordance with the individual dataset [31], of image fusion were developed.

2.5. Image Classification Main Approaches

As mentioned, the explosive growth of available data lead to the need for the creation of automatic ways to extract information from it. Currently, for classification, the most commonly used approaches in the field of image analysis are: classification based in pixel analysis and based in objects [2], [13]. Nevertheless, a new approach has recently gained the researchers attention: usage of convolutional neural networks (CNNs) to automatically extract information from images [23]. These three approaches are next described.

2.5.1 Pixel-Based Classification

As implied by the name, this technique makes use of pixels, the smallest element of an image [13], [23]. In it, the individual pixel attributes serve as a basis for classification by assembling similar values of adjacent pixel attributes to recognize (or not) an archaeological feature [13], [23], [28]. The aim of this approach is "*to establish a per-*

pixel relationship with a certain class category based on the attribute information of itself and its spatial neighbourhood" [13, p. 3].

To achieve its goal, the approach must rely on "*the separability of the different classes*" [23, p. 1], usually measured though the computation of a distance metric or a separability index [5], [13].

2.5.2 Object-Based Classification

This method is a two-step process, with one being object-based image analysis (OBIA) and the other being classification [1], [2], [13].

OBIA, as the name implies, encompasses techniques that, when applied, result in either an image being split into meaningful and homogeneous segments of non-overlapping objects based on a specific criteria (shape, scale, etc.) [1], [2], [13], [28], [33], or in the definition of specific multiscale characteristics from which it may base the segmentation of the image [28], [34]. As such, its goal is to "*subdivide the image into homogeneous segments that describe the target features (pit, burial mound, etc.) as correctly as possible*" [13, p. 4]. The segmentation parameters are obtained through a trial-and-error process on the user part [1], through use of point-based, edge-based, region-based, or combinations of these, segmentation techniques [13], [33].

In fact, it has given rise to another discipline, the Geographic OBIA (GEOBIA), whose purpose is essentially the same as OBIA but for its application with Earth's geographic components [13], [28]. This has caused the term to cover the majority of OBIA's applications in the field of archaeology [28].

The classification technique is based on the segments obtained, where each segment is allocated a corresponding class based on the target object characteristics, e.g.: geometry, spectral values, neighbourhood relationships or semantic groupings [1], [2], [13]. The idea is *"to pass the segments through a series of decision rules, which serve to merge or split segments, and assign class labels to segments, according to user-specified rules"* [1, p. 4], where the rules are the objects properties of the employed target [13].

2.5.3 Convolutional Neural Network Classifiers

This method works through a "*system of interconnected neurons that pass messages to each other*" [23, p. 3] forming an artificial neural network capable of modelling

complex functions and being used as frameworks to classify input data [23]. In order to categorize and classify images, the input data is either an image, interpreted as a numerical grid of pixel values ranging from 0 to 255 depending on their colour intensity, or a derived set of features from the image with its goal being predicting the correct class of the associated image [23], [35].

They work through the junction of three components, the convolutional layer responsible for analysing the inputs, the pooling layer capable of summarizing the information obtained from a convolutional layer and the fully connected layer that makes use of the summarizations as inputs to classify the input image [35], [36]. Through this approach they learn to generalize and draw features from a large set of labelled images, without relying on handcrafted rules or parameter sets, and whose excellent results have consistently outperformed humans in visual object detection rates [36].

This approach's main advantage is the use of transfer learning that allows to extend pre-trained CNNs (e.g.: GoogleNet, AlexNet and OverFeat) on extensive datasets of labelled images, that have been proved to possess enough representation power to perform recognition tasks on different types of target images and thus make use of them in fields originally restricted by the small size of labelled data to learn from, such as archaeology, with extracted features being proven to be useful in image classification in RS images [36], [37].

2.6. Archaeological Methodologies

Applications for the extraction, detection, and classification of objects in images have been used in a variety of different fields of study successfully. Some of these fields include archaeology [1]–[3], [5], [8], [12]–[14], [28], forestry surveillance and classification [9], [10], urban building and greenery identification [11], cartography, surveillance, reconstruction and location [6].

When applied to archaeology, these applications come with a major drawback: each application is created based on the specific features of a particular region and for a certain structural monumental typology, which results in a non-generalizable solution for other applications in different geographical or cultural contexts. This stems from the very high

structural polymorphism present in archaeological remains, originating from differences in culture and available materials in the territory, presenting highly divergent characteristics or environmental insertions [6].

Currently, even at a more abstract level, where replication of methodologies could be possible, no approach is consensual on which type of images to use or even if a combination of certain types of images should be used, or which of the previous classification approaches is more appropriate to use.

In the context of using PAN and MS type images, which are usually fused in order to obtain both spectral and spatial features, the use of vegetation indexes proved to be a crucial factor for archaeological detection through enhancement of crop marks created due to near-surface monuments [2], [8], [38]. The results being further improved by the utilization of edge extraction algorithms, which emphasize, enrich and extract superficial anomalies [8].

On the other hand, on the usage of only PAN imagery, it is imperative to speak of the work of [14], who explored and proposed a method to detect circular patterns in agricultural lands through template matching methodologies. Other authors have focused on using only MS image type through the use of linear combinations applied to the different MS visible-near infrared bands to derive a set of components allowing to distinguish between the crop marks, the background soil and the remaining healthy vegetation, depending on the vegetations phenological cycle [39]–[41]. On the last case, the crop mark component was proven to be able to be used separately as a detection tool while others have instead made use of the separability index to detect the buried monuments.

When using HS type images, these are commonly fused with high-resolution PAN or LiDAR image type, that increase the capacity of detecting and the outlining of features [2]. This type of images is highly used for crop mark detection and detection of buried archaeological monuments, due to its capacity of in-depth spectrum analysis of land soil and vegetation. When alone, Nonlinear Principal Component Analysis (NLPCA) to obtain land cover types and biophysical information, followed by use of Separability Index to measure difference between tonal anomalies detected by NLPCA and surrounding areas, has been used to great effect to detect buried structures [5].

Alternatively, the recent use of LiDAR imageries that make use of DTMs and from them perform a multiscale topographic analysis, by calculating an integral image and respective topographic deviation in order to obtain a Multiscale Topographic Position, has proved to be advantageous to semi-automatic feature detection and visual interpretation [3]. Other authors have opted to use inverted pit detection methods to detect buried monuments, in the form of mounds, by inverting LiDAR DTMs before applying the detection method [1]. And some works have made use of LiDAR data for identification based on archaeological morphology in the form of concentric circles patterns [12].

On a final note, recently developed environments such as Google Earth Engine have allowed the creation of applications in various disciplines by easing data access and permitting geospatial analysis at planetary scale [42], [43], including for archaeology. According to [6], images from these engines have, in some cases, been successfully applied to identify monuments [43]–[45]).

As it can be observed in Table 1, which constitutes a summary of the works reviewed, in archaeology, no unique approach to either extract information or to classify objects in an image exists.

Table 1: Used methodologies in applications created for automated identification of archaeological areas of interest.

Year	Country	Bibliography	Data Type	Detection Method	Approach
2006	Italy	[8]	PAN / MS	Edge Detection	Object-based
2007	Italy	[38]	PAN / MS	Edge Detection	Object-based
2009	Greece	[2]	PAN / MS / HS / Aerial	Predictive Modelling	Object-based
2009	Norway	[14]	PAN / MS	Decision Trees	Object-based
2013	Italy	[5]	HS	Spectral Separability Index (SSI)	Pixel-based
2013	Greece	[40]	MS	SSI & Visual Interpretation	Pixel-based
2013	Hungary	[39]	MS / Ground HS / Aerial	Edge Detection / Photointerpretation	Object-based
2014	China	[44]	Google Earth VHR	Edge Detection: Canny Edge Detector	Object-based
2016	Tonga	[1]	LiDAR	Inverted Mound Algorithm / Homogeneity	Object-based
2016	Sweden	[13]	LiDAR	Minimum Distance: Euclidean / Mahalanobis	Pixel-based
2016	Austria	[13]	LiDAR	Homogeneity	Object-based
2016	Italy	[41]	MS / PAN	SSI / Minimum Distance	Pixel-based
2017	Greece	[43]	MS / PAN	SSI / Visual Interpretation	Pixel-based
2018	France	[3]	LiDAR	Random Forest	Object-based
2018	Austria/Italy	[46]	Airborne HS	Ranked SSI / Visual Interpretation	Pixel-based
2018	Greece	[46]	MS	Ranked SSI / Visual Interpretation	Pixel-based
2019	Spain	[12]	LiDAR	Landform Classification	Object-based

2.7. Conclusions

While the results presented in the related literature show high levels of false positives, the automatic methodologies for monument's detection in images discussed enable a quicker and easier method of detecting new archaeological monuments, allowing to restrict human ground prospection to zones of higher likelihood of bearing buried monuments and surveillance in areas difficult to reach.

Different types of monuments and ground may need different methodologies to detect and classify the archaeological structures. These should be based in the diverse characteristics, whether the monument is buried or exposed, and on the surrounding environment. The utilization of image fusion methodologies, that fuse spectral features with detailed spatial features of a different image type, has showed to be, in this aim, the better approach to analyse an image.

Over the recent years, there has been a shift in the types of images used, from PAN to LiDAR as it allows for greater and more accurate morphological characteristics by changing the paradigm from a two-dimensional to 3D data, and from MS to HS that are capable of capturing far more distinct and accurate environmental characteristics. When it comes to the choice of image type to be used, the preference depends on the archaeological object to be detected. Usually, for buried structures, the preferences go for HS to detect through the buried object's spectral information or, if mounds exist, LiDAR to construct a 3D model of the mound's presence and shape. Meanwhile, when the evidence of archaeological presence involves vegetation, there is a preference for MS and it's respective spectral indices, i.e., for crop marks identification, and for a fusion of MS and PAN types when exposed structures are intended due to their mixing of general shape with identification of basic spectral material identification (vegetation, soil, water, …).

Finally, no studies of this type have been undertaken in Portugal nor over this specific archaeological monument.

Chapter 3 – Methodology

Accordingly with the DSR methodology, after the initial phases where the problem has been identified, the respective motivation described, and the objectives for the intended solution defined, and after the description of the state of the art, this chapter will describe the methodological development and main research stages that have been performed. Beginning by the extraction of spectral and spatial features from the various image types through the development of the required spectral indices, spectral material signatures and spatial analysis of pixels groups, experiments in the creation of a supervised machine learning method capable of detecting whether certain areas in the satellite image can be identified as presenting high likelihood for a dolmen's being present.

To ensure that the objectives are achieved with optimal results, several techniques were tried and tested to ensure peak performance of the developed algorithms. The following steps will explain the theoretical basis for the techniques used, as described upon the conceptual system's architecture in Figure 3. All the needed implementations used MATLAB.

This chapter is divided into eight sections. Starting by a conceptual architecture for the developed recognition system, the next section describes the prior processing of images. Then, a discussion of the used image's colour spaces follows, succeeded by the basis of the image enhancements techniques and by the spectral band combinations used to form the various spectral indices. Next, the testing of image fusion methods is described, followed by a section describing feature extraction methods. Finally, the classifier models used are detailed.

3.1. Conceptual System's Architecture

Based on the related literature (Chapter 2), two conceptual models, capable of automatically detecting areas of archaeological interest with a high likelihood of containing monuments were created Figure 3: one using HS imagery and another using a fusion of MS and PAN imagery.

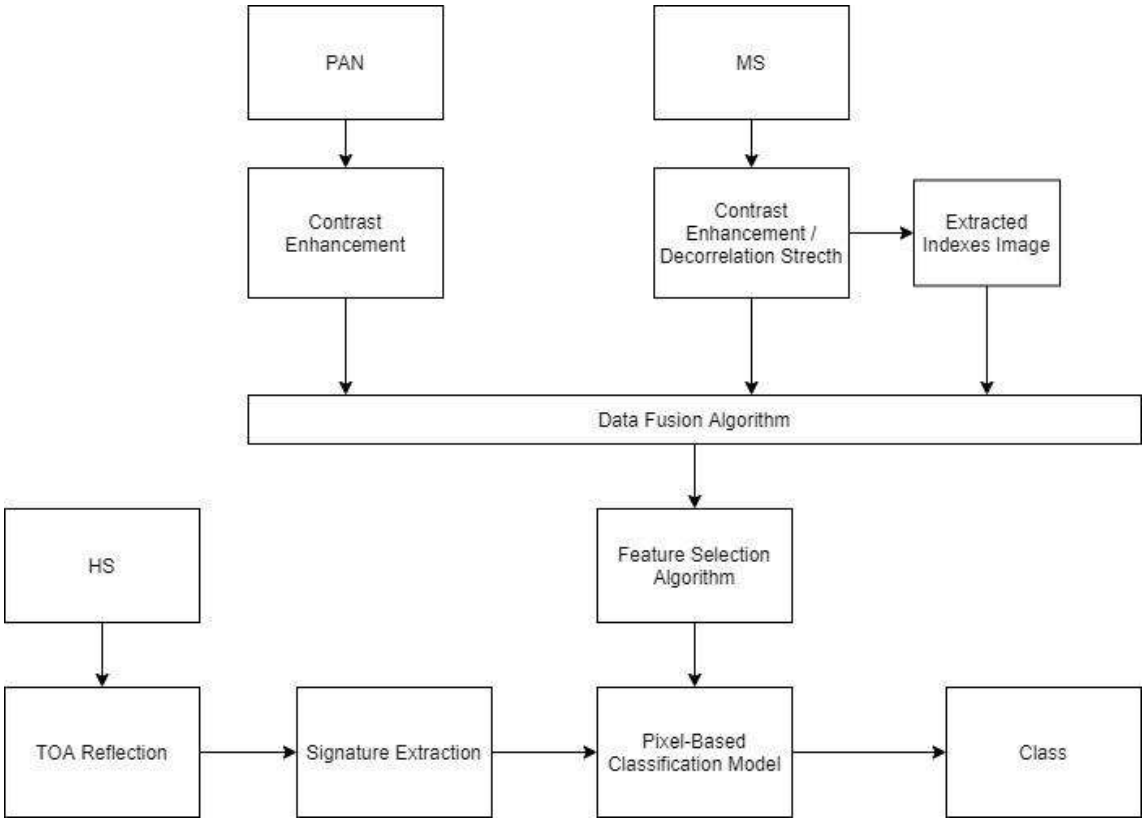


Figure 3: Proposed workflow and system's architecture.

3.2. Prior Processing

To facilitate their use by the end-user, when the satellite acquires each image of the differing types, they undergo several processes before their delivery up to specific standards of quality requested for the current project, the "Ortho Ready Standard Level 2A" (ORS2A) for the PAN and MS types and level 1A for the HS type. These standards of image quality allow for images to be in correctly mapped coordinate spaces, to improve feature classification and identification as well as performing image enhancements more flexibly, and for end-users to get any project running far more easily [47].

Any products obtained from the Proba-1 satellite are level 1A and undertake no additional pre-processing before delivery. This means that the received products (images) are the raw representation of the captured Top-of-Atmosphere (TOA) radiance by the on-board instruments.

For the WV-2 satellite, all images received were ORS2A, where radiometric, geometric and sensor corrections (like the ones next described) have been applied and

images have been mapped into a cartographic projection at a constant base elevation in order to allow for orthorectification [47].

3.2.1 Masking

A process commonly used in image processing, where all the pixels outside a fixed boundary are set to intensity value 1, effectively turning them black. It is used on occasions where those pixels are irrelevant to what is being considered. It possesses the side effect of boosting computation as those pixels stop being considered.

3.2.2 Radiometric Correction

As stated by [48, p. 16], the raw (digital numbers) WV-2 imagery "*undergoes a radiometric correction process to reduce visible banding and streaking in WorldView-2 products*". This is necessary since any change in detector response, in gain or offset, lens falloff and contamination on the focal plane manifests as either streaks or banding in the raw image. Thus, for any products of the WV-2 satellite, a *Relative Radiometric Correction* method is applied, responsible for minimizing these image artefacts that include a non-uniformity correction and a dark offset subtraction, in accordance with formula *(1)*:

$$q_{Det,Band} = \frac{p_{Det,Band} - A_{Det,Band}}{B_{Det,Band}} \quad (1)$$

where q is the radiometrically corrected detector data for specific detector Det and band $Band$, p the raw detector data, A the dark offset for a specific image acquisition, and B the detector relative gain.

Finally, the outputted detector data q is spatially resampled to create the delivered image with radiometrically corrected image pixels.

3.2.3 Geometric and Sensor Correction

The WV-2 ORS2A products are projected at a constant base elevation to a plane through use of map projections in order to allow for orthorectification, normalizing the

topographic relief present through unrefined elevation models, either by averaging the elevations obtained or using elevation values supplied by end-users [25], [47]. Simultaneously, the ORS2A products have their sensors corrected in such a way that the "correction blends all pixels from all detectors into the synthetic array to form a single image" [25, p. 427], that is, blends the pixels of overlapped areas from different sensors such that it returns a single synthetic array capable of forming an image.

3.2.4 Colour Space

Applications like computer graphics, image processing, and computer vision make use of different colour space models, that is, mathematical models capable of representing colour information. Usually, this information is represented using either three or four different colour components, where the displayed colour produced from such models depends on the respective parameters and equipment's used [49]. Thus, the same feature extraction method applied to differing colour spaces extracts separate significant features to train different classification models, as each colour space represents distinct characteristics of the image.

In essence, the different colour space representation permits to take full advantage of the inherent multiple visual information captured through the spectral bands on either the HS or MS image type.

3.2.4.1 RGB Colour Space

The TrueColor space, commonly called the RGB colour space, is composed by a 3D cartesian coordinate system of the values present in the Red, Green and Blue bands in any image, showcasing all possible colours by adding the values in each of three primary colours. White is the sum of the maximum of all three values and black the absence (null value) of all of them. Chrominance and luminance components are intermixed with the three bands, making it inefficient for colour analysis and colour based segmentation algorithms [49]. From this colour space, any other possible colour space is obtainable through the application of linear or non-linear transformations [49] as it can be next seen.

3.2.4.2 HSV Colour Space

The Hue, Saturation and Value (HSV) colour space is, similarly to the RGB colour space, a 3D cartesian coordinate system representing: the hue, capable of representing the type or shade of all possible colours and thus can be used for distinguishing colours; the saturation, referring to the percentage of white light mixed with a pure colour/hue; and the value, embodying the perceived light intensity (brightness) of the colour, that is the amount of light illuminating a colour, being similar to the intensity or luminance characteristics [49], [50]. Further, this colour space possesses the advantage of describing the colours similarly to how the human eye perceives them and possessing the capability of separating chromatic and achromatic components from the image [49], [50].

To transform from the RGB to the HSV colour space the following equations *(2)-(4)* must be applied, where the variable R represents the Red Band, the G the Green Band, and the variable B the Blue Band:

$$H = arcos \frac{\frac{1}{2} * (2 * R - G - B)}{\sqrt{(R-G)^2 - (R-B) * (G-B)}} \quad (2)$$

$$S = \frac{\max(R,G,B) - min(R,G,B)}{\max(R,G,B)} \quad (3)$$

$$V = \max(R,G,B) \quad (4)$$

3.3. Theoretical Basis for Image Enhancements

Following its collection, a series of experiments to individually enhance each type of image was implemented, with the intuit of increasing their respective quality to allow for a better extraction of points of interest and thus, features.

This type of processing focused mainly on transforming the received radiance data to reflectance data and in reducing the noise level present in the HS images, enhancing the spectral capabilities of the MS images to obtain higher returns on developed indexes, as well as the enhancement of spatial information through fine-tuning PAN images for easier

detection. Further, spatial resolution enhancement is done through image fusion methods, described below, between the MS and PAN type images.

3.3.1 MS and PAN Image Enhancements

Throughout the years, this type of techniques has been focused on an effort to improve both an image's appearance and, in images of low luminance taken in environments lacking in natural light sources, its details [51]–[53]. It serves to increase the interpretability of the information contained in an image for human viewers or, in the case of automated processes, deliver higher quality feature input by increasing the distinctions between an image's features. Thus, its goal is to manipulate images to produce an image more appropriate for a specific purpose [54], [55]. Several techniques have been developed to enhance certain features of an image, such as: contrast, noise level, boundaries, sharpness and edges [53], [54].

According to [56], [57], one of the most frequent problems associated with satellite imagery is their contrast, that is, the difference in reflected luminance (the colour and brightness values) between two adjacent surfaces in the captured image, stemming from the possibility of being captured during periods of either dense darkness or brightness, and thus causing loss of information in areas that are uniformly dark or light, respectively. Therefore, this is a technique that enhances an image directly on its pixels' intensity level, enabling the production of images with subjectively better perception to end-users than the originals [51], [53], [58].

Variations of this technique have thus become one of the most widely required and used processing methods to improve feature visualization [57]. Some of the most popular variations used have been Decorrelation Stretching, Linear Contrast Stretching, Gamma Correction and General Histogram Equalization [56, p. 81].

Considering the variability of the degree of sun incidence on the area under study, depending on the time of day that the images were taken, and thus the variability of the luminance that may be present in the different areas of an image, it was found to be pertinent to implement and test several of the above described Contrast Adjustment techniques to the received images.

3.3.1.1 Histogram Equalization

For the present work, two methods for implementing a histogram equalization (HE) technique were tested: a general HE and a contrast-limited adaptive histogram equalization (CLAHE), that, at its essence, is a fusion of an adaptive and a contrast-limited HE approaches.

The general HE is a widely adopted technique for its simplicity, ease of implementation and good performance in enhancing low contrast among a variety of images, through making an image histogram as uniform as possible [51], [53], [58]. It works by remapping the intensity values through a transform function adaptively obtained from a cumulative density function (CDF), by means of flattening the density distribution and stretching the images' grey levels dynamic range [51]–[53]. This technique assumes an uniform quality level across the entire image, as it produces only one grey level map that provides the same enhancement across the image [59], [60]. However, its main drawback stems from errors in the shifting of the image mean brightness that cause enhanced images to develop unnatural enhancements in certain areas and artefacts originating from intensity saturation effects [51], [52], [58].

According to [61], the algorithm used by MATLAB to apply the standard HE is *histeq*, that minimizes the grayscale transformation T by using

$$histeq = \left|C_1\big(T(k)\big) - C_0(k)\right| \tag{5}$$

where C_0 is the cumulative histogram of the image and C_1 is the histograms cumulative sum for all intensities k, with k varying in intensity value ranges depending on the respective image class from [0, 1], [0, 255], [0, 65535] to [-32768, 32767]. This transformation is applied under the constraints that T must be monotonic and $C_1\big(T(k)\big)$ cannot exceed $C_0(k)$ by more than half the distance between the histogram counts at k. To map the image's grey levels to the new respective values it makes use of an inbuilt grayscale transformation.

Another method, CLAHE, is an adaptive approach that works under the opposite assumption: that the distribution of the grey level changes from an image area to another [59], [60]. Thus, the adaptive HE approach works by computing several histograms, each

of a section of the image, and by mapping each pixel based on the local section grey level distribution, where the transformation function applied is proportional to the CDF obtained from the pixel neighbours' intensity values [59], [60]. Thus, while this technique allows the improvement of local contrasts and edge definitions in the different images sections, it possesses the drawback that it tends to magnify noise in homogeneous sections of the image [60].

Conversely, the contrast limiting approach allows to solve issues surrounding the general HE application on cases where the distribution is highly localized, as these result in mapping transformation curves having slopes that cause similar grey levels to highly differ. When applied simultaneously with the adaptive method, allows for the limitation of the drawback of noise amplification [59], [60]. Regarding the contrast amplification of the adaptive HE method, it stems from the slope of the transform function that it is proportional to the neighbourhood CDF [60]. Thus, when adding this function to the contrast limiting HE in order to create the CLAHE method, it works by limiting the amplification through clipping of the histogram at certain values before computing the local neighbourhood CDF, thus limiting the slopes present in the CDF and therefore on the transformation function, where the clip limit (value for the clipping of the histogram) depends on the histogram normalization and thus on the size of the neighbourhood region [54], [59], [60].

Thus, in accordance with [61], the MATLAB's function that applies CLAHE is *adapthisteq* that operates on *tiles* (composed of small regions of the image that are clipped) and calculates the contrast transform function described above (*histeq* function) for each tile individually, after which, the neighbouring *tiles* are combined using bilinear interpolation to eliminate artificially induced boundaries.

3.3.1.2 Decorrelation Stretch

In addition to the above described methods, a Decorrelation Stretch (DS) technique was also tested. It serves as a method of decreasing high correlations between the different bands present in the images, a problem often found in MS images sets, thereby producing more colourful images through enhancing the colour separation, with it highlighting subtle differences as it improves visual interpretation and facilitates feature

discrimination [54], [55]. During this process, the "*original color values of the image are mapped to a new set of color values with a wider range*" [55, p. 115].

3.3.1.3 Linear Contrast Stretch

Linear Contrast Stretch or simply Contrast Stretching (CS), adjusts local contrast, highlighting the darker and lighter areas with finer details permitting the detection of variations in the data more easily, and displays a contrast improved image, through an expansion of the original range of intensity values of the image to span the range of intensity of the device, darkening pixels bellow certain intensity level while brightening pixels above certain intensity level [54], [55]. Among the variety of its types, the focus of this work was on the linear method of Minimum-Maximum Linear Contrast Stretch, that in accordance with [54, p. 14], the "*original minimum and maximum values of the data are assigned to a newly specified set of values that utilize the full range of available brightness values*".

According to [55, p. 115], this type of CS originates from using a linear thresholding function, that can be described as:

$$s = T * (r) \tag{6}$$

where s and r denote the intensity values of output and input at any image position (x, y), respectively, with T representing the thresholding function to be applied.

3.3.1.4 Gama Correction

Another example of image enhancement is through gamma correction (GC) (or Power Law Transformation). It's a non-linear adjustment method, making use of logarithmic transformations, that is more effective the greater the possible range of intensity values to brighten [62], [63], being frequently used to improve either the detail or the contrast of lower intensity values [63]. In cases where the obtained gamma value is 1, then it will map new intensity values linearly, for cases lower than 1, it will map them with greater weight towards brighter (higher) intensity values, and for gamma values higher than 1, it will instead map them with greater weight towards darker (lower) intensity values [63].

According to [63], the GC methods possess the following basis form:

$$S = C * R^{\gamma} \tag{7}$$

where S represents the resultant gamma corrected intensity level, R represents the input intensity level and C and γ being positive constants. Further, as seen in formula *(8)*, ε can be added, in case of an input being zero, to offset and obtain a measurable output.

$$S = C * (R + \varepsilon)^{\gamma} \tag{8}$$

3.3.2 HS Image Enhancements

On the HS type images front, the enhancements were directed for the transformation and extraction of reflectance information on individual pixels, which necessitated the calculation of TOA reflectance values from the original radiance values of the image. All algorithms and software used for these transformations are based on the open source developed tool BEAM and respective CHRIS/Proba Toolbox, a software that provides extensions to the base BEAM platform allowing to perform the following tasks: noise reduction, cloud screening, atmospheric correction, geometric correction, TOA reflectance calculation and feature extraction [64].

From the possible methods, only noise reduction and TOA reflectance calculation were able to be applied to the received images, with the remaining methods erroring due to missing metadata information.

3.3.2.1 Noise Reduction

In accordance with [64], this tool was developed with the express purpose of removing and/or correcting coherent noises commonly known as vertical stripping and drop-outs, two kinds of noise that frequently affect the HS images captured by remote sensing instruments. Broadly speaking, drop-outs occur when the instrument randomly fails and produces anomalous values, usually zero or negative values, in some odd pixels which disrupt future operations as they are heavily affected by these pixels. On the other hand, vertical stripping is an error usually found in push-broom sensors types, stemming from

irregularities on the entrance of the spectrometer originating from temperature dilation, that results in images gaining complex vertical patterns.

In order to correct and remove these two types of problems, [65] described the implemented algorithms on the software in detail. In summary, it's a four step process: first it detects and corrects drop-outs by detecting the negative or zero anomalous values and replacing them with a weighted, by its similarity (inverse of the Euclidean Distance) to the corrected pixel, average of a 3x3 neighbourhood pixels values; the second step is a preliminary correction of vertical stripping through an estimation of the cause, the slit, gained from a prior characterization of vertical stripping pattern kept on a look-up-table and the sensors temperature at the given moment of acquisition; in the third step, a robust correction method is further applied to the vertical stripping problem, that is capable of estimating the remaining vertical stripping for each band; and finally, on the fourth step, the estimated coefficients to correct vertical stripping are applied to correct the column values [64], [65].

3.3.2.2 TOA Reflection Calculation

As stated above, products delivered by the Proba-1 satellite return the captured TOA radiance data. This type of data was not useful for the current classification test. As such, it needed to be converted to TOA reflection before it could be used.

According to [64], the estimation for TOA apparent reflection is:

$$\rho(x, y, \lambda_i) = \frac{\pi * L(x, y, \lambda_i)}{\cos\big(\theta(x, y)\big) * I(\lambda_i)} \tag{9}$$

where $L(x, y, \lambda_i)$ is the sensor captured radiance at the (x, y) location of the image, $I(\lambda_i)$ represents the extra-terrestrial solar irradiance corrected for acquisition date and convolved with CHRIS sensor spectral bands, and $\theta(x, y)$ the angle between illumination direction and the perpendicular vector to the surface represented by the 'Solar Zenith Angle' attribute provided in the CHRIS metadata.

The solar irradiance $I(\lambda)$ is provided in mW/m^2/nm from 200 to 2400 nm, in accordance with the work of [66], then corrected to the Julian day of year of acquisition (based on received image metadata), J, in accordance with formula *(10)*:

$$I(\lambda) = \frac{1}{(1 - 0.01673 * \cos(0.9856 * (J - 4) * \pi/180))^2} * I(\lambda) \quad (10)$$

Then it is resampled to the CHRIS sensor spectral channels due to the referenced solar irradiance displaying a different spectral sampling, forming $I(\lambda_i)$, where i represents a specific CHRIS band, and thus representing the mean solar irradiance for a given band obtained through the integration of solar irradiation and spectral response, following:

$$I(\lambda_i) = \frac{\int_0^\infty S_i(\lambda) * I(\lambda) * d * \lambda}{\int_0^\infty S_i(\lambda) * d * \lambda} \quad (11)$$

where, for simplicity, the spectral response of a CHRIS band, $S_i(\lambda)$, is a bell-shaped function that hinges on the mid-wavelength, λ_i, and the band width, $\Delta\lambda_i$, of the band (with both values being present in the band metadata file) derived through the function:

$$S_i(\lambda) = \frac{1}{1+|2*(\lambda-\lambda_i)/\Delta\lambda_i|^4} \quad \lambda_i - \Delta\lambda_i < \lambda < \lambda_i + \Delta\lambda_i \quad (12)$$

3.4. Basis of Spectral Index Development

To make use of the spectral information obtained, the proceeding step was to create indices that allowed to categorize the presence of different features in the region, in accordance with the expert defined ontology of this subject. Thus, the focus lied in developing indices for the detection of vegetation, water, and ground material types.

In line with [46], [67], the efficient use of spectral reflectance measurements for any detection lies with identifying the most highly correlated spectral wavelength, usually a ratio of two or more spectral bands, with the specified target physical parameter on an acquired image, which severely limits the regions of interest on the spectrum. Spectral Indices may be derived both from MS and HS type images, where HS derives them through use of narrow spectral bands while MS can only derive them if the target spectral feature encompasses a great range of the spectrum [46].

The literature indicates the existence of hundreds of indices, as can be seen in the following index creation database [68] that tracks possible indices by satellite sensor and index application, of which several were chosen to be created and tested. The following points describe the indices that were used in this work.

3.4.1 Vegetation Indices

Indices that quantitively measure biomass and vegetative growth through use of unique vegetational spectral information, derived mainly from the range where the transition of the strong chlorophyll pigment absorption identified in the visible red band and the high reflectance of the leaf's mesophyll in the near infrared band occurs, with the final purpose of estimating the image pixels that represent alive and green vegetation [8], [38], [46].

The Normalized Difference Vegetation Index (NDVI) is the most widely used and well known index, indicating photosynthetic activities in plants and thus where variation in its values indicate variations in the accessibility of water and nutrients, being found to be related to the green leaf area index [8], [38]. Its normalization stems from the need to lessen the effects of possible variations induced by atmospheric contamination, with its values, if high, allowing to classify pixels as being covered by healthy vegetation and, if low, categorizing pixels as being covered by stressed or diseased vegetation [38].

According to [68], the NDVI can be obtained by using the following formula, where NIR represents a near infrared spectral band encompassing the wavelengths around 800nm (790-810nm) and the RED represents a visible red spectral band encompassing from 620nm to 700nm:

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{RED}}{\mathrm{NIR} + \mathrm{RED}} \tag{13}$$

In this work, this index can be calculated, using the WV-2 satellite images, by substituting the NIR variable by either of the satellite's NIR1 or NIR2 band alongside the RED variable being the detected Red band and, on the other hand, from the same formula it's also possible to calculate the Red Edge NDVI by substituting the RED variable with the Red Edge band instead [68].

3.4.2 Water Index

These indices are designed to highlight and map surface water features while simultaneously suppressing information unrelated to water and can thus separate water from non-water according to the selected threshold value [69], [70]. Many of the created indices require the use of a shortwave-infrared band not present in the received images, however the used WV-2 sensor, as asserted above, captures an eight-banded image of

which one of them is the coastal band that is sensitive to water areas and has the potential for water research and bathymetric studies [69].

The Normalized Difference Water Index (NDWI) has been extensively used to delineate water areas, being originally designed to account for the differences between green and NIR spectral bands, by taking into account the strong absorption by water features and the high reflectance by vegetation and soil features detected in NIR band [69]. Much like it happens with NDVI, its normalization serves the purpose of lessening the effects of potential variations generated by atmospheric contamination. However, and as stated above, adapting this index to the WV-2 satellite implies not only changing the original green band to the more water sensitive coastal band but the original NIR band with the NIR2 band.

Thus, according to [70], the NDWI adapted to a WV-2 satellite image can be obtained by using the following formula:

$$\mathrm{NDWI} = \frac{\mathrm{COASTAL} - \mathrm{NIR2}}{\mathrm{COASTAL} + \mathrm{NIR2}} \tag{14}$$

where the both the NIR2 and the COASTAL variables are substituted by the respective bands of the same name.

3.4.3 Shadow Index

In remote sensing, shadows are considered to be mutual features that can be divided in two classes, a self-shadow (direct light not illuminated as part of the object) and cast shadow (the direction of the light source in the object), and three categories, being either cloud shadows, shadows by natural features or shadows by urban features [71]. Thus, this type of index is developed to automatically detect shadows though highlighting and mapping shadowed areas while suppressing the remaining information available.

The Shadow Detection Index (SDI) was developed to be used with the WV-2 satellite spectral bands in mind and is thus optimized to make use of the Blue, NIR1 and NIR2 bands, that have been found to be effective and capable of differentiating, and thus detecting, between areas of shadow and non-shadow, with the index having been proven to be able to distinguish between dark objects and shadows [71].

Thus, according to [71], the SDI developed is obtained through the following formula, where all variables are substituted by the respective bands of the same name:

$$\text{SDI} = \frac{\text{NIR2} - \text{BLUE}}{\text{NIR2} + \text{BLUE}} - NIR1 \tag{15}$$

3.4.4 Soil Index

Much like the aforementioned indices, this type of indices focus on their capacity to identify and delineate areas where soil is the material that either dominates the background or the foreground of satellite images [72]. However, in remote sensing, soil is difficult to detect due to its complex physical and chemical compositions, its regional differences and the lack of a direct connection between soil abundance and its spectral signatures [73].

The Normalized Difference Soil Index (NDSI), according to [72], while normally making use of SWIR and NIR bands to represent the difference in reflectance values in soil areas, can instead delineate said areas by taking into account the unique differences present in soil response values between the Green and Yellow spectral bands due to the lack of SWIR bands in the WV-2 satellite. Much like it happens with the former Normalized Difference indices, its normalization serves the purpose of lessening the effects of potential variations generated by atmospheric contamination.

According to [72], the NDSI applied to the WV-2 satellite is obtained through the following formula, where the two variables are substituted by the respective spectral bands of the same name:

$$\text{NDSI} = \frac{\text{GREEN} - \text{YELLOW}}{\text{GREEN} + \text{YELLOW}} \tag{16}$$

3.4.5 Feature Difference Index

This indices display their importance when attempting to detect and map areas that are in direct contrast to the presented background information in which, in the case of remote sensing technologies, the background information can be assumed to be the natural occurrences and the direct contrast the areas that can be identified as being opposed to natural, and thus man-made [72].

The Non-Homogeneous Feature Difference (NHFD) classifies areas which contrast against the background owing to non-homogeneous features generally responding and standing out brighter than the background, through the differences present in the Red Edge and Coastal spectral bands affording the ability to segregate them [72].

In accordance with [72], the NHDF for use in WV-2 satellite outputs can be obtained through the following formula, where the variables are substituted with the respective spectral band of the same name:

$$\text{NHFD} = \frac{\text{RED EDGE} - \text{COASTAL}}{\text{RED EDGE} + \text{COASTAL}} \tag{17}$$

3.4.6 Built-up Presence Index

This type of index are generally applied for their usefulness in detecting and mapping asphalt and concrete surfaces, with emphasis in road extraction, both in urban and non-urban areas [74], [75]. In the available WV-2 spectral bands, the Blue and NIR bands are the closest in terms of wavelength to the available and displayed feature that best describes asphalt roads, the appearance of iron-oxide stemming from the oxidation process and exposure of rocky components, at the 520nm to 870nm [75].

While the Built-up Areas Index (BAI) makes full use of the above described phenomenon of capturing the presence of iron-oxide in said spectral bands to segregate and classify the individual pixels as either belonging to a road network or not, it has also been proven that while accurately extracting the roads it simultaneously misclassifies major water bodies as roads [75].

As stated by [74], [75], the BAI applied to the WV-2 satellite sensors is applied according to the following formula, where each of the variables is substituted with the respective spectral band of the same name:

$$\text{BAI} = \frac{\text{BLUE} - \text{NIR}}{\text{BLUE} + \text{NIR}} \tag{18}$$

3.5. Image Fusion Methods

Considering the results obtained from the pre-processing of the MS and PAN images and from the research performed on the object to detect, an understanding that neither

image would be able to detect the object alone arose, as they are usually covered by vegetation, which implies the need to detect it through the MS image, and are in possession of a diameter ranging from 2 to 5 meters, implying the need of PAN image for more detailed spatial information.

As defined above, image fusion or pan-sharpening is a process to make optimal use of the benefits present in both the Pan and MS images, and respective spectral indices, through the integration of high-resolution spatial data of the PAN image and the spectral information of the MS image to generate and make use of a high-resolution MS or HS image, with the goal of increasing the interpretation capabilities through the production of more reliable information while losing the least amount of either spectral or spatial information possible (Rahmani, Strait, Merkurjev, Moeller, & Wittman, 2010; Zeng et al., 2010; Zhang, 2004).

Of the existing methods of image fusion, two different types were tested. They are explained in the following sections.

3.5.1 IHS Transform

The IHS (Intensity, Hue, Saturation) transformation method is one of the most widely used and effective image fusion techniques, as stated above, owing to its basic computation, efficiency and producing fused images of high spatial resolution with low spectral resolution [32]. According to both [31], [32], [76], the IHS generally works through resizing the MS image to the same size as the PAN image and converting the MS colour image from the RGB (Red, Green, Blue) space, where the individual bands of the RGB can be any combination of the possible 8 MS bands available, into a IHS colour space, where the I component resembles, and is thus replaced by, the PAN image, after which the reverse IHS transformation is applied to transform it into the fused RGB image.

For each pixel, the IHS fusion, according to [76], is calculated according to the described three-step procedure:

First:

$$\begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-\sqrt[2]{2}}{6} & \frac{-\sqrt[2]{2}}{6} & \frac{2\sqrt[2]{2}}{6} \\ \frac{1}{\sqrt[2]{2}} & \frac{-1}{\sqrt[2]{2}} & 0 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (19)$$

Second: The I component is replaced with the PAN image, needing no calculation.

Third:

$$\begin{bmatrix} F(R) \\ F(G) \\ F(B) \end{bmatrix} = \begin{bmatrix} 1 & \frac{-1}{\sqrt[2]{2}} & \frac{1}{\sqrt[2]{2}} \\ 1 & \frac{-1}{\sqrt[2]{2}} & \frac{-1}{\sqrt[2]{2}} \\ 1 & \sqrt[2]{2} & 0 \end{bmatrix} * \begin{bmatrix} PAN \\ v_1 \\ v_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{-1}{\sqrt[2]{2}} & \frac{1}{\sqrt[2]{2}} \\ 1 & \frac{-1}{\sqrt[2]{2}} & \frac{-1}{\sqrt[2]{2}} \\ 1 & \sqrt[2]{2} & 0 \end{bmatrix} * \begin{bmatrix} I + (PAN - I) \\ v_1 \\ v_2 \end{bmatrix} \quad (20)$$

$$= \begin{bmatrix} R + (PAN - I) \\ G + (PAN - I) \\ B + (PAN - I) \end{bmatrix}$$

Where $F(X)$ is the fused image of the X band, for $X = R, G, B$.

3.5.2 BROVEY Transform

Like its IHS counterpart, the Brovey Transform is also one of the most popular and effective image fusion methods, though based instead on arithmetic combinations [31], [32]. In accordance to [31], in essence it works by multiplying each MS band by the PAN band and then dividing each result by the sum of the MS bands. The resultant image of this transformation emphasizes spatial information in degradation of spectral information, leading to colour distortion when the spectral range of the input images are different or they have significant long-term temporal changes [33], [77]. The following formula defines how to compute each individual RGB band to produce the Brovey fused image, according to [77]:

$$\mathrm{R}_{new} = \frac{R}{(R + G + B)} * PAN \quad (21)$$

$$\mathrm{G}_{new} = \frac{G}{(R+G+B)} * PAN \quad (22)$$

$$\mathrm{B}_{new} = \frac{B}{(R+G+B)} * PAN \quad (23)$$

3.6. Feature Extraction

Taking into consideration the resultant fused images, some methods of feature extraction were tested and applied to delineate desirable areas of high likelihood where the dolmens may be present. These methods will thus attempt to define and delineate the zones of interest where archaeologists should focus their efforts.

3.6.1 Circular Hough Transform

The Circular Hough Transform (CHT) technique, a variation of the Hough Transform, is capable of isolating, and thus detect, feature shapes (i.e.: lines, ellipses, circles, etc) in an image, in this case circles, that requires said features being specified in parametric form [78]. This method is often used due to having been proven to be tolerant to gaps in the features descriptions while remaining relatively unaffected by noise, occlusion and varying levels of illumination present in the image [61], [78].

According to [61], [78], the algorithm that calculates the circular shape follows the equation (24):

$$(x - x_0)^2 + (y - y_0)^2 = r^2 \quad (24)$$

where r represents the circle radius and (x_0, y_0) represents the coordinates of the circle centre. Further, the MATLAB's function (*imfindcircles*) employed CHT algorithm possesses three essential steps:

- Computes the accumulator array, where foreground pixels of high gradient are designated candidate pixels which cast votes, in a pattern that forms a complete circle of fixed radius, in the accumulator array.

- Estimates the centre from the accumulated votes through detecting the peaks in the accumulator array where the candidate's pixels vote patterns coincide.
- Estimate the radius of the circle, either through radial histograms or through encoding radius information when forming the accumulator array alongside the possible centre locations.

3.6.2 Bag-of-Visual Words

The Bag-of-Visual Words (BoVW) feature extraction method was made to manage both the variability of spectral and spatial content in RS images and has attracted attention for its capability of constructing midlevel representation of features instead of low-level ones, therefore permitting the visual vocabulary to possess semantic understanding of geospatial objects and low computation complexity through employment of patch-level detection and description that divide images into subregions of objects comprised of multiple homogeneous components [79], [80].

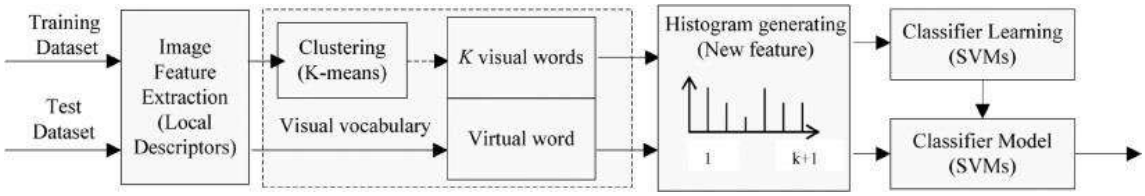


Figure 4: BoVW Framework Source: [80].

As observed in Figure 4, typically this method's framework comprises of the following steps, after which the resultant global feature representation is fed into the chosen classifier model:

- Feature Extraction, using patch detection and description to form feature vectors from the objects.
- Feature Pre-Processing, due to extracted features usually possessing high dimensionality and strong correlation between them, it is normally necessary to apply the PCA method, a statistical procedure using orthogonal transform where, usually, the resultant number of variables is less than the original number (leading to dimensional reduction), to map them into a set of linearly uncorrelated variables called principal components.

- Codebook Generation, executed during the training phase of the method using a method of unsupervised learning (either K-means Clustering or Gaussian Mixture Models), where each cluster centres is assigned as a visual word, and thus creating the visual vocabulary that describes the object.
- Feature Encoding, where each extracted patch is assigned to the closest visual word through Euclidean Distance and a histogram, that represents an object, is created by counting the occurrences of each visual word [79], [80].

3.7. Classifiers

In an effort to reduce the number of areas of high likelihood of dolmen presence following the automatic identification of circles in the images, an ensemble of several different classifiers was trained, tested and implemented to automatically classify each obtained circle as either showing dolmen presence or not. The following sub-sections will expand on the theoretical basis of the classifiers used.

3.7.1 Logistic Regression

The Logistic Regression classification approach considers the existence of a nonlinear relationship between the independent variables (e.g.: nominal, ordinal, interval or ratio scale measurements), considered as predictors of the dependent ones, and the dependent variables, present in binary format, that takes the form of a Bernoulli Distribution owing to the binary format of the dependent variable [81]. Accordingly, this relationship is expressed through the following equation (25) [81]:

$$\mathrm{p} = \frac{1}{1 + e^{-z}} \tag{25}$$

where p is the probability, ranging from 0 to 1, of the occurrence of object detection (in this case, the dolmen) and z is the linear combination of related independent variables that allows to determine the probability of occurrence.

3.7.2 K-Nearest Neighbours

The K-Nearest Neighbours classification method has been widely used owing to its simplicity, flexibility, and consistently high performance, that classifies new data through storing of all the data used to train the model and then calculating a similarity measure (distance function) between the new and stored data, setting the new data as the same class as the majority of its k-nearest neighbours classes [82]–[84]. In this method, K is commonly calculated through cross-validation, however small and odd K values are usually selected to break ties while larger K values have better precision due to reducing noise [82], [83].

According to [82], there are three possible distance functions used to calculate distance for continuous variables, the Euclidean, Manhattan, and Minkowski (equations (26)-(28)) and one for categorical variables, the Hamming Distance (equation (29)):

$$\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{26}$$

$$\sum_{i=1}^{k}|x_i - y_i| \tag{27}$$

$$\left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{1/q} \tag{28}$$

$$D_h = \sum_{i=1}^{k}|x_i - y_i| \tag{29}$$

3.7.3 Support Vector Machines

The Support Vector Machine classification technique is a non-parametric supervised learning method whose goal is the production of a model capable of predicting the target class value for given data attributes, training the model through mapping feature vectors into a higher dimensional space and finding a linear hyperplane that is capable of separating this higher dimensional space with maximal margin [82], [84]. To do so, it trains the model according to the following equation (30) [82]:

$$\min_{w,b,\xi} \frac{1}{2} * w^t * w + C \sum_{i=1}^{l} \xi_i \quad (30)$$

$$\text{subject to } y_i * (w^T * \Phi * (x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0$$

where Φ maps feature vectors to higher dimensional space, $C > 0$ is the penalty parameter of the error term and $K(x_i, x_j) \equiv \Phi * (x_i)^t * \Phi * (x_j)$ is a possible kernel function (e.g.: Linear, Polynomial, etc).

3.7.4 Naïve Bayes

The Naïve Bayes classification method is one of the most competitive learning algorithms that as seen wide usage in RS classifications owing to its highly scalable learning, it estimates the probability of an observation belonging to a predefined category through defining a model based on the Bayes theory where, based on the variables present in the used training data, it defines the prior probability of each class originating from a conditional probability estimation where the independence of the predictor variables is assumed [81], [83], [84].

According to [81], [83], [84], the estimated probability of being a certain class is obtained based on the following equation (31):

$$P(h|D) = \frac{P(D|h) * P(h)}{P(D)} \quad (31)$$

where $P(h|D)$ is the estimated probability of belonging to class h given the variables D (also denominated posterior probability), $P(D|h)$ is the probability of D given h (also called likelihood), $P(h)$ is the prior probability of categorical hypothesis h (denominated prior) and $P(D)$ is the prior probability of training data variables D (also called evidence), where due to this method assuming the conditional independence of predictor variables, $P(D|h)$ can also be calculated through equation (32):

$$P(D|h) = \prod_{i=1}^{k} P(\frac{D_i}{h}) \quad (32)$$

3.7.5 Linear Discriminant Analysis

The Linear Discriminant Analysis classification technique is used to discriminate between different groups according to the individual related features of each possible

class while identifying the features with the greatest contribution towards class separation and creating a predictive model, composed of a discriminant function based on a linear combination of said strongest variables, that is capable of predicting a class for new data by calculating individual class centroids, i.e.: the mean value of the discriminant scores, from the training data and , finally, predicting the most likely class for new data by calculating as many linear equations functions as there are classes and from them compute and choose the highest classification score, taking it as the predicted class [82], [85].

The discriminant function, according to [82], [85], takes the form of the following formula:

$$D = V_1 * X_1 + V_2 * X_2 + V_3 * X_3 + \cdots + V_i * X_i + \alpha \quad (33)$$

where D is the discriminant function, V the discriminant coefficient or weight for that attribute, X the features of the attribute, α is a constant and i is equivalent to the number of predictor variables.

3.7.6 Decision Trees

The Decision Trees classification method has become increasingly important owing to their computation efficiency and conceptual simplicity where, unlike in other approaches, instead of simultaneously using a set of features to classify in a single step, it instead performs the classification based on a multistage or hierarchical decision scheme or a tree like structure that are composed of a root node, that holds all the data, a set of internal nodes, or splits, and a set of terminal nodes, called leaves, where each node makes a binary decision that splits a class from the remaining ones until no more decision splitting may be made, at which point the terminal leaf node is reached, in a process known as top-down approach [86]–[88].

Alongside the DT possessing the ability to automatically execute feature selection and complexity reduction, its basic concept of splitting complex decisions into several simpler ones allows for a more interpretable solution and of both the predictive and generalization capability of the classification [86]–[88]. Finally, if the target variable is a discrete value this process is known as decision tree classification, while if it's a continuous value it is known as decision tree regression [88].

3.7.7 Ensemble

Ensemble methods work by combining the predictions of several, usually weak, classifiers into a single composite classifier that has generally been found to be more accurate than any of its individual parts and capable of being used to reduce the error of its weaker parts [86], [87]. Another example of an ensemble method is boosting, in which rather than making use of different classifiers, it instead generates a series of the same classifier iteratively where the training set chosen emphasizes the selection of incorrect classifications more times than correctly classified samples of the previous classifier, becoming a method capable of improving the algorithms performance and reducing the errors of its weaker classifiers by aggregating the generated classifiers into a composite classifier and having them vote on the same sample [86], [87].

Chapter 4 – Geographical Area, Images, and Image Enhancement

The following chapter will expand on the geographical region that was studied, alongside the received and used images, ending with the steps taken to enhance the received images and demonstrating the obtained enhanced final products.

Thus, this chapter is divided into three sections, one presenting the geographic information of the researched region, another presenting the technical information of the PAN, MS and HS images, and ending with one section that presents the various techniques attempted to enhance the utilized images.

4.1. Geographical area of the case study

The area being studied is the region of Pavia in Mora, Alentejo, Portugal, since it is the one for which we have expert data and knowledge to assess the success of the classification. It encompasses an area of 185 km^2.

In accordance with [4], it's a territory composed mainly of alkaline granites, granodiorites, tonalites and trondhjemite. According to different authors, this territory houses the most extensive plateaus of Portugal, with local topography curves averaging an altitude rounding the 200m without a high percentage of variation for either declivity or relief [4], [89].

This area is situated in central Alentejo, a region comprised of the three major hydrographic basins in the south of Portugal. More precisely, the region of Mora is in the tertiary basin of the river Tejo [4]. According to [4], [18]), this type of landscape.

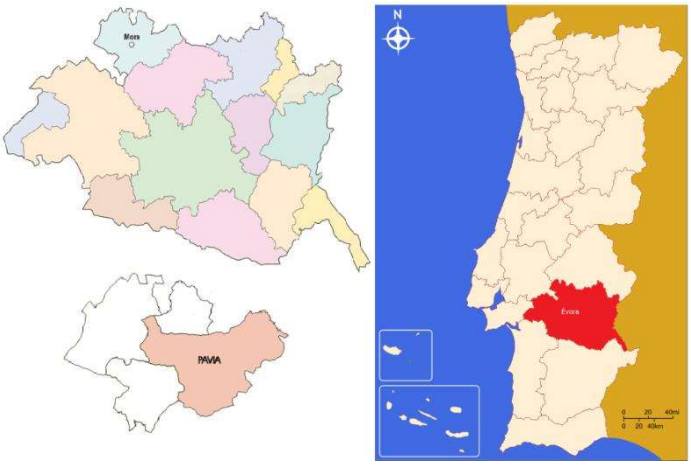


Figure 5: Area of study (Pavia, Mora) within the map of Portugal.

4.2. Satellite Images used

To study this region, three different types of satellite images were used: PAN, MS, and HS.

4.2.1 Panchromatic and Multispectral Images

Regarding the PAN and MS types, due to the unavailability of satellite images containing Pavia's total area, it was necessary to use two tiles of separate images for both imagery types and, as such, this work makes use of two PAN and two MS images.

With one PAN and one MS image captured by the WV-2 satellite, collected on separate occasions (30 of August, 2018 and 14 of March, 2019), with both sets possessing 0% of cloud coverage and the most recent being 19.6° off-nadir and the remaining being from 26° to 26.4° off-nadir viewing angle. Both sets present individual images that are ORS2A, in which radiometric, geometric and sensor corrections are applied and mapped to a cartographic projection at a constant base elevation. The originally received image sets alongside the expertly identified visible dolmen locations can be seen in Figure 6.

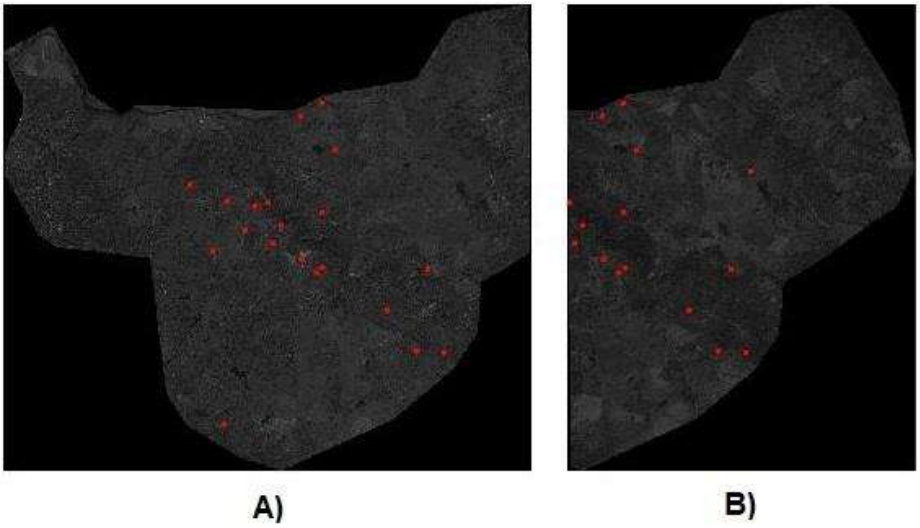


Figure 6: Images received on the first set, B), and on the second set, A), where red points indicate approximate locations of expert identified visible dolmens.

The PAN images present a pixel spatial resolution of 50cm, presenting an average collected ground sample distance of 0.517m and 0.564m, while the MS images have a pixel spatial resolution of 2m, displaying an average collected ground sample distance of 2.067m and 2.257m. Additionally, the captured MS images possess eight bands that allow to showcase spectral differences, these being: Coastal, Blue, Green, Yellow, Red, Red Edge, Near Infrared (NIR) 1 and NIR2. In Figure 7 depicts a plot with each bands' relative range of wavelengths alongside their relative response level, while the following Table 2, states the comprehensive full range of captured wavelengths and respective spatial resolutions possible.

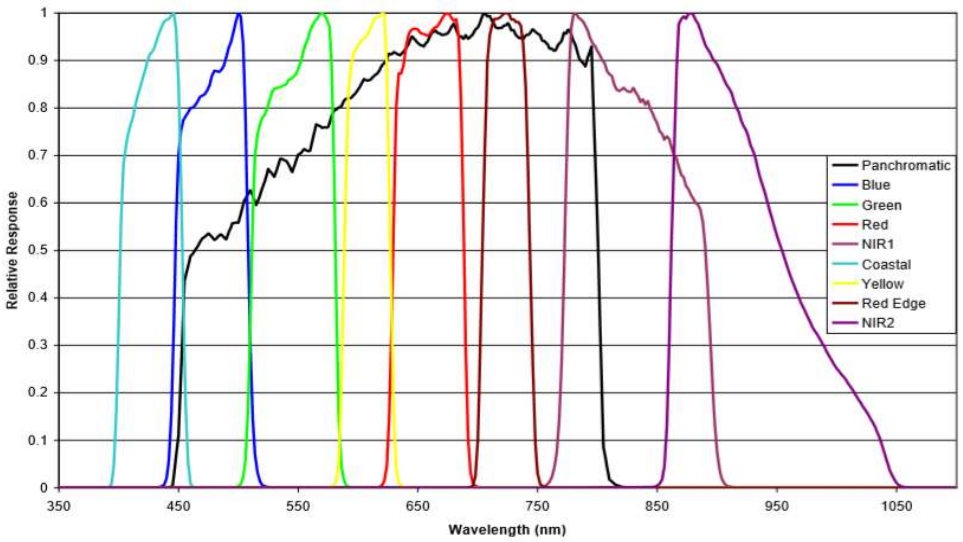


Figure 7:WorldView-2 Relative Spectral Response (nm) Source: [48].

Table 2: WorldView-2 satellite characteristics of the Visible-Near Infrared Spectrum Source: [47].

	Coastal Blue (nm)	Blue (nm)	Green (nm)	Yellow (nm)	Red (nm)	Red Edge (nm)	Near IR1 (nm)	Near IR2 (nm)	Pan (nm)	Spatial Resolution
WorldView-2	396 – 458	442 – 515	506 – 586	584 – 632	624 – 694	699 – 749	765 – 901	856 – 1043	447 – 808	PAN: 0.50m Others: 2m

This different spatial resolution results in the images respective pixel resolutions diverging widely. On the first collected set of images (30 of August, 2018), the MS pixel resolution is 8335x6209x8, while in the PAN it is 33340x24836. On the second set (14 of March, 2019), the MS resolution is 8335x9428x8 with the PAN being 33340x37712. All of them are represented in 16 bits per pixel.

According to [24], the role of each MS bands, in regards to how they generally affect and are affected by, can be summarized as:

- **Coastal Blue:** While aiding in vegetative analysis due to being absorbed by healthy plants, its main purpose is for bathymetric experiments. It is considerably affected by atmospheric scattering and can prove useful for enhancements to atmospheric correction techniques.
- **Blue:** It is absorbed by the chlorophyll present in plants, provides good water penetration and is less affected by atmospheric scattering and absorption than Coastal Blue.
- **Green:** Puts more focus on the peak reflectance values of healthy vegetation, making it an ideal for plant vigour estimates. Alongside the Yellow band, it makes them useful in differentiating between types of plants.
- **Yellow:** important for feature classification, it serves as a method of detection for “Yellowness”, the unhealthiest, of plants in water or land.
- **Red:** Is absorbed by healthy plants materials, thus serving as one of the key bands for vegetation discrimination, and as proved its usefulness in classifying bare soils, roads, and other geographical features.

- **Red Edge:** Strategically adjusted to put the focus on the high reflectivity of vegetation response, it is effective in evaluating plant health and classifying vegetation.
- **Near Infrared 1 (NIR1):** Effective in estimating a plants biomass and the moisture contents present, it serves mainly for splitting vegetation from water bodies, detecting vegetation types and distinguishing between soil types.
- **Near Infrared 2 (NIR2):** While partially overlapping with the NIR1 Band, it is less affected by the atmosphere, enabling a larger analysis of the vegetation and execution of studies over biomass.

4.2.2 Hyperspectral Images

Regarding the collected HS image, it is composed of a set of five images individually taken and captured by the Proba-1 satellite using the aboard instrument Compact High Resolution Imaging Spectrometer (CHRIS). This tool captures one image as closely as possible to the +/- 55°, +/- 36° and 0° observation angles for an established target [90], [91]. The HS image was collected on the 8 of January 2020, possessing thirty-seven spectral bands related to land channels and a resolution of 748 rows and 766 columns and a nadir ground sampling distance of 17 meters. This image and respective set of images that compose it, are then processed until achieving the Level 1A, by applying radiometric and wavelength calibrations to ensure accurate image representations, returning the captured TOA radiance levels [64], [90], [92].

In accordance with the satellite's technical, Annex A shows the minimum, middle and maximum values possible for the wavelength in each bands', in nanometres, alongside the pixels respective width for the thirty-seven and eighteen band image respectively. An example of a band of this image is present in Figure 8.

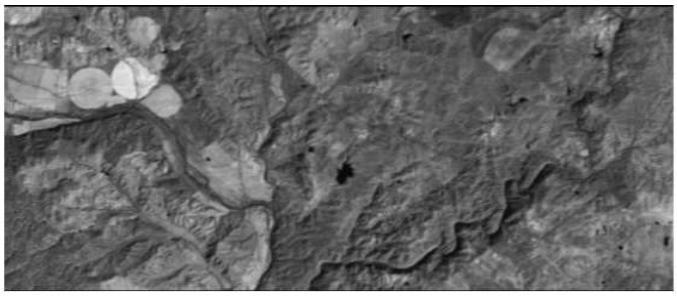


Figure 8: Band 27 of the original HS image received.

4.3. Image Enhancements Techniques

Before being able to implement any pattern recognition techniques or spectral information extraction, the received images needed to undergo certain procedures to both pre-process and enhance them. As such, and in accordance with the proposed workflow in Figure 3, each image underwent several image enhancements techniques.

4.3.1 Hyperspectral Imagery Enhancements

As stated above, the captured HS images return not the needed reflectance values but instead the TOA radiance levels, thus two methods were experimented to both enhance and convert the original radiance values to the desired reflectance values.

Regarding the experiment of implementing Noise Reduction through the previously described algorithm (see Noise Reduction), instead of removing and correcting the noises commonly known as *vertical stripping* and *drop-out pixels*, it added vertical stripping to the image, as it can be seen in Figure 9 on the right. Due to the obtained results proving that the implementation of this algorithm worsens the images instead of improving them, it was decided to not use it to process the HS images moving forward.

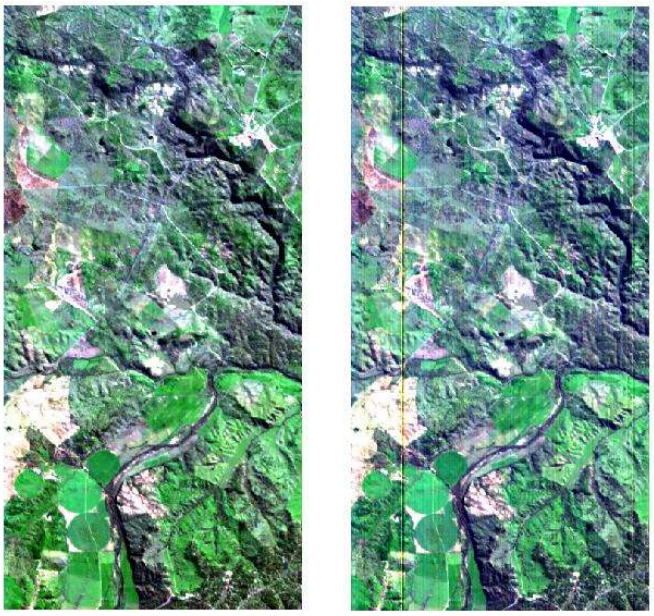


Figure 9: Visible differences between the original HS image received (left) and HS image after undergoing Noise Reduction (right) that now presents vertical stripping.

The following step was the conversion of TOA radiance pixel values to TOA reflectance pixel values through the TOA Reflectance Calculation algorithm previously described. While it did not have any outward effect in the way pixel values can be seen in an image, it rescaled and altered the resultant pixel values graphs. In Figure 10, it is possible to see an example of comparison between the radiance graph and the reflectance graph produced for a randomly assigned pixel.

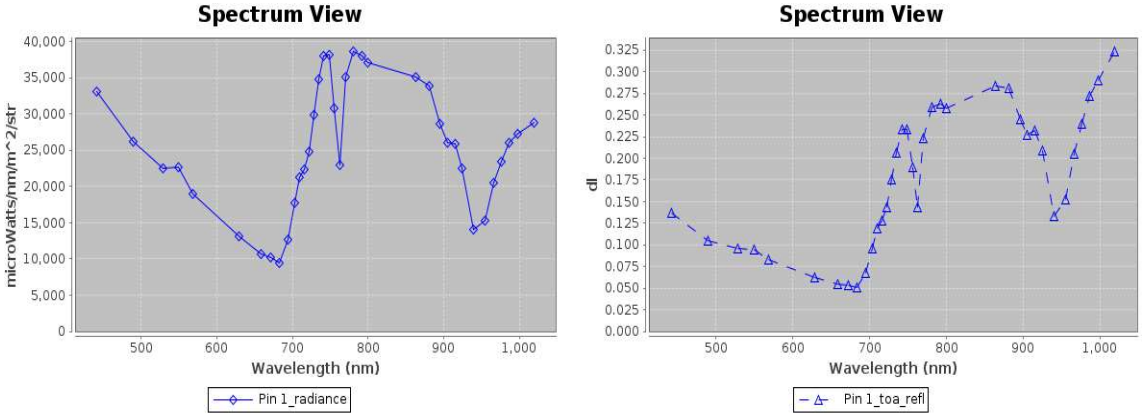


Figure 10: Comparison between original HS image radiance values received (left) and HS image after undergoing TOA Reflection Calculation method (right) presenting reflection values, in the y axis.

4.3.2 Panchromatic Imagery Enhancements

To enhance PAN imagery, the focus lied in checking and applying different contrast enhancing techniques. The first step was to extract from each image seen in Figure 11 a histogram, which revealed that the data concentrates itself within an extremely small part of the available dynamic range for both images. To compensate for this issue, different contrast enhancement techniques were tested to obtain the most fitted PAN image possible, including: HE, CLAHE, GC and CS. While Figure 12 shows the comparison between the original and the corrected histograms, Figure 13 and Figure 14 display the influence of the application of the different methods to each of the original images (where A) HE; B) CLAHE; C) CS; D) CS followed by CLAHE), Figure 15 showcases the best obtained results from applying the GC to the images.

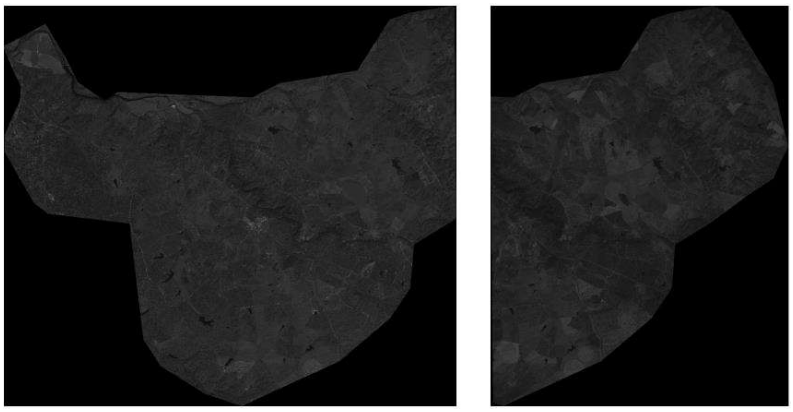


Figure 11: Original received PAN images, with the left being the image of the second set and the right being of the first set.

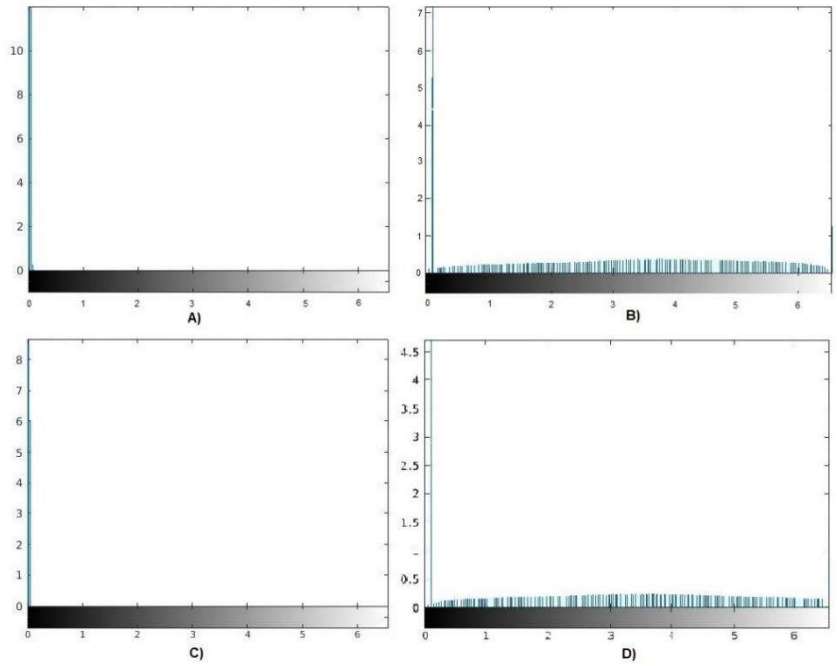


Figure 12: Comparison between original images Histograms (Left) and resultant corrected Histograms (Right).

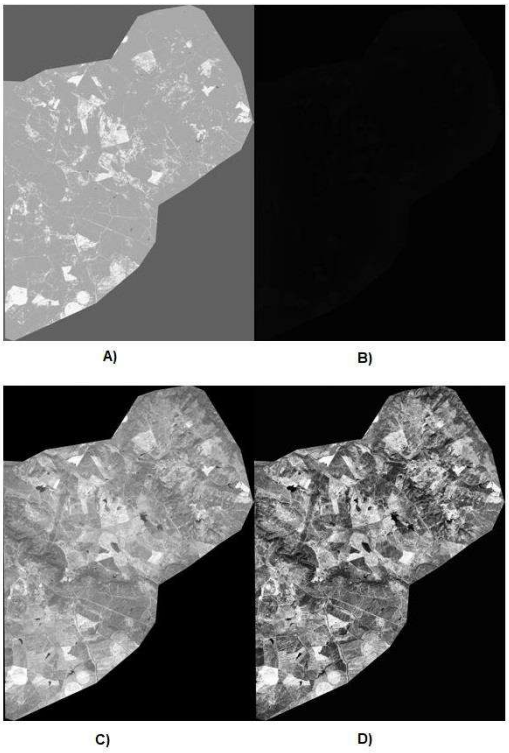


Figure 13: Comparison between contrast enhancement methods on the PAN image received on the first set.

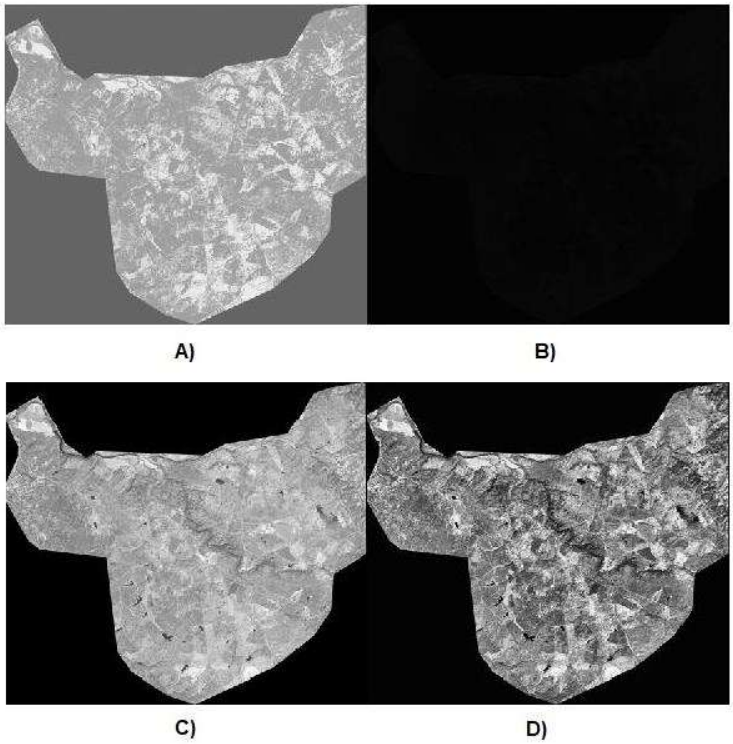


Figure 14: Comparison between contrast enhancement methods on the PAN image received on the second set.

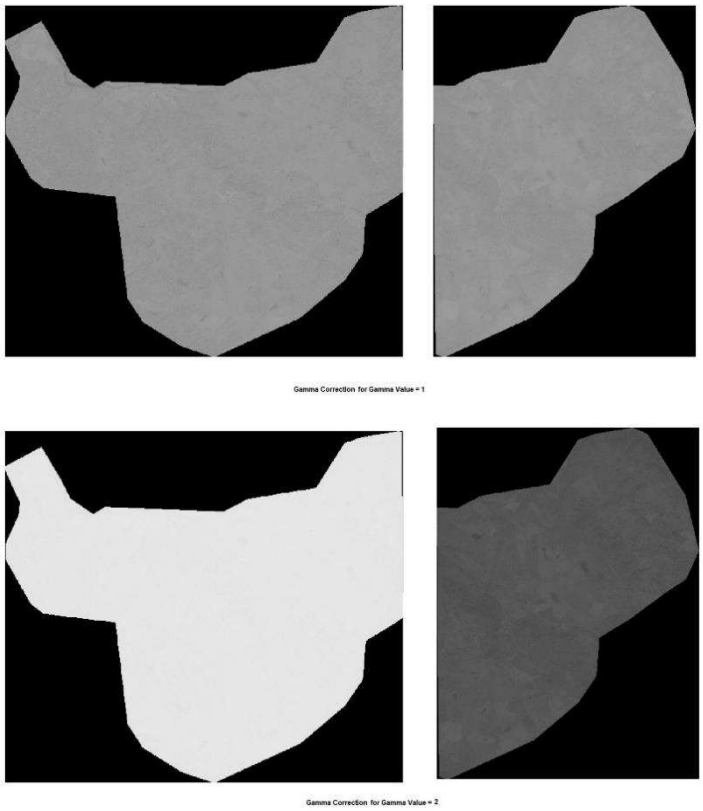


Figure 15: Best obtained results when applying Gamma Correction method to the received PAN images.

From the results obtained in Figure 13 and Figure 14, some conclusions can be reached: first, that neither the application of a single HE method or CLAHE method produce any usable results, with HE merely whitening the image and CLAHE losing nearly all visual information; and secondly, that only with the application of a CS method can any favourable result be achieved, as a single CS method allowed to extract an image with contrast but still relatively dull, and the application of a CLAHE method following it corrected this limitation and permitted to acquire an image with striking contrasts and enhanced edge definitions. Finally, Figure 15 demonstrates that the application of GC to the received images produce results that lose most definition and thus, should also not be considered.

4.3.3 Multispectral Imagery Enhancements

On enhancing MS imagery, the focus lies in checking and applying different contrast enhancing techniques as well as checking the correlation of the bands used to form the differing indexes and, if necessary, enhance it through decorrelation techniques.

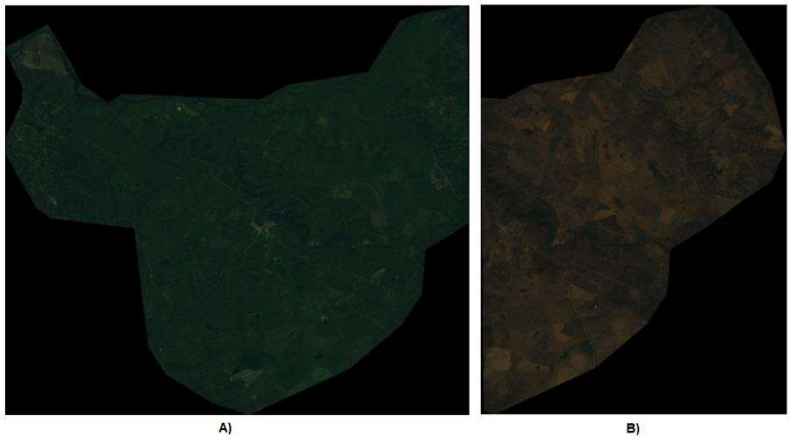


Figure 16: MS images received on the first set, B), and on the second set, A).

As depicted by Figure 16, the originally obtained images are dull, presenting very light contrast. To understand the reason for the said dullness, an exploration of their individual band histograms and respective correlations was made, whose results can be seen in the

standard band histogram shown in Figure 18 (for each individual band histogram, see Appendix A) and the RGB band correlation in Figure 17.

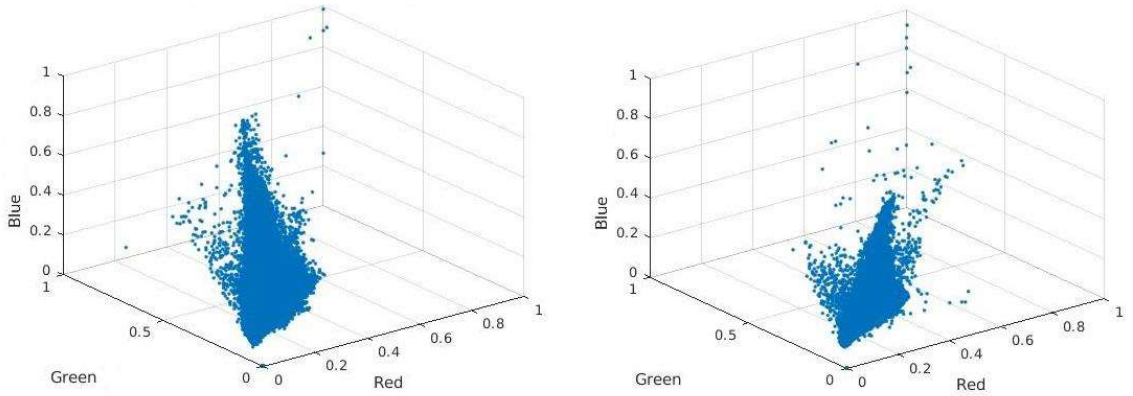


Figure 17: RGB band correlations for each of the MS image, on the left being for the second set while the right being for the first set of images.

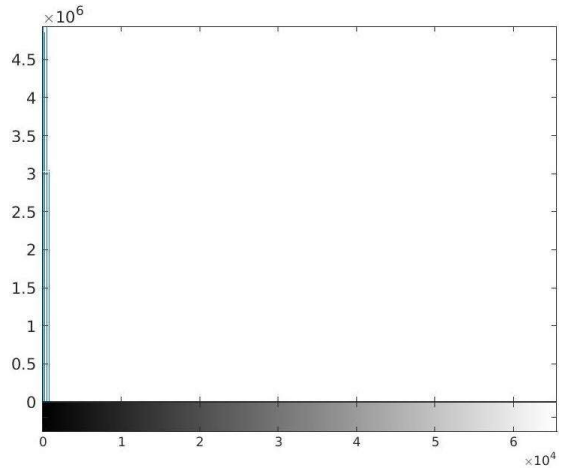


Figure 18: Standard example of band Histogram obtained for each band of each MS image.

Additionally, Figure 19 explores the band correlations between both the spectral band 7 and 8, representing the NIR bands, with the bands 3 and 2, the green and blue bands, necessary for any calculation of the Vegetation Indices.

From the histogram that was extracted (Figure 18), the data can be seen to concentrate itself within an extremely small part of the available dynamic range and thus, producing a dull image. Furthermore, the correlations extracted and displayed in Figure 17 that describe the degree of relationship between the colour component bands of each pixel, show high levels of correlation between bands lending to monochromatic dull images.

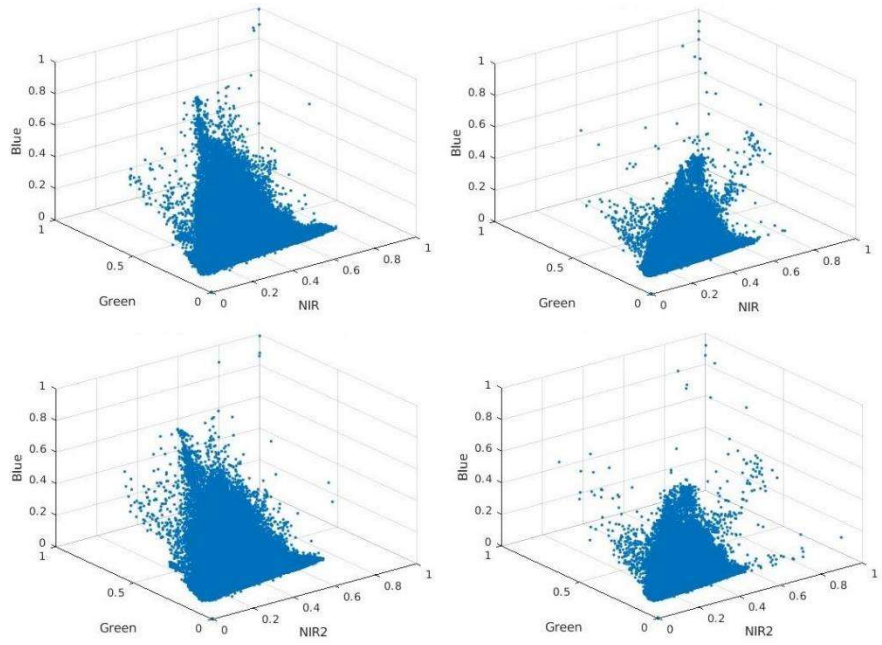


Figure 19: NIR1 and NIR2 band correlations with Green and Blue bands for each of the MS image, on the left being for the second set while the right being for the first set of images.

To correct this, similarly to the processing of the PAN images, different contrast enhancement techniques were tested to obtain the best possible MS image, including HE, CLAHE, GC and CS., Figure 20 displays the resultant corrected band histograms while Figure 21 shows the corrected band correlations.

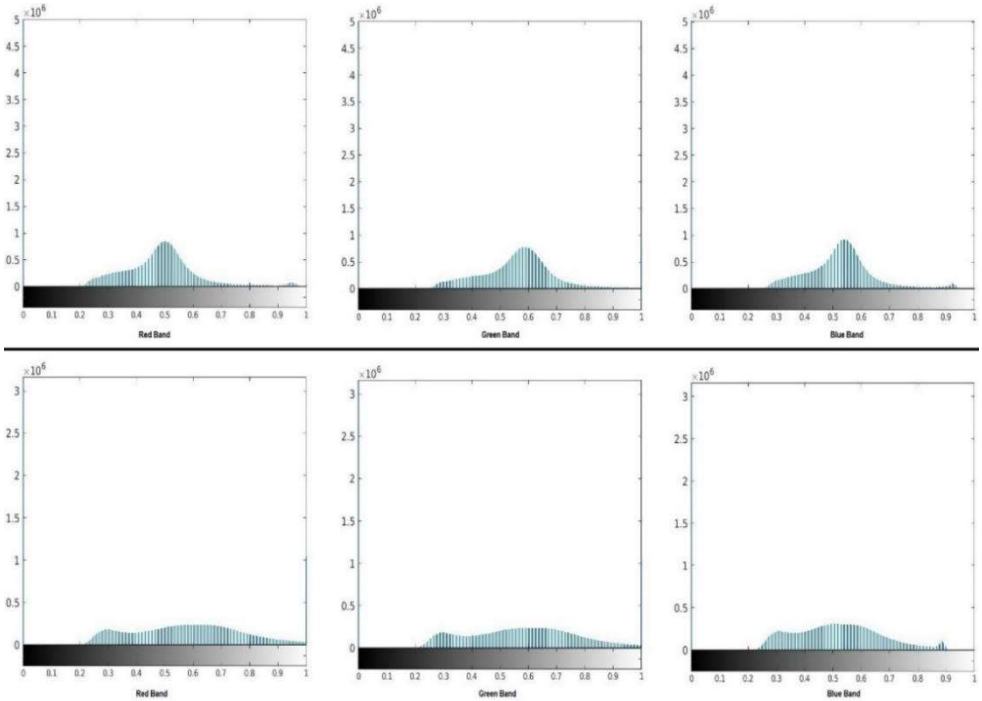


Figure 20: Corrected histograms representation of the Red, Green and Blue bands, with the upper three histograms belonging to the second set and the lower three to the first set of MS images.

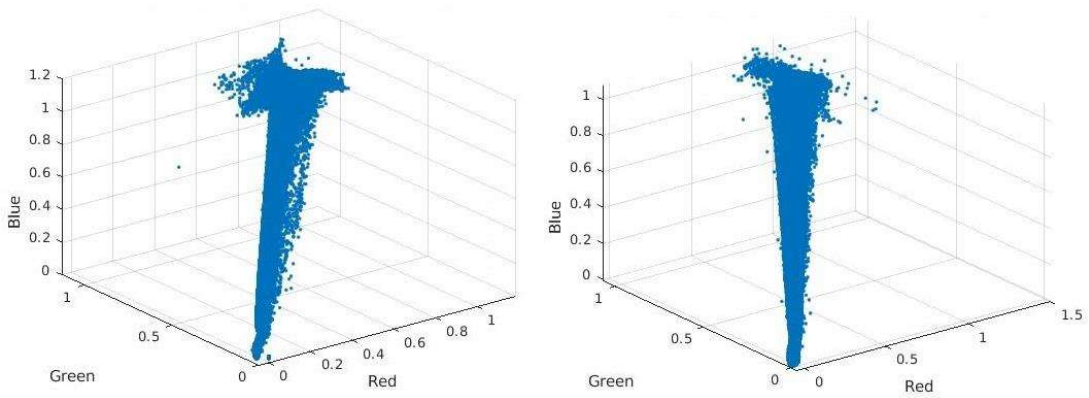


Figure 21: Corrected RGB band correlations for each of the MS image, on the left being for the second set while the right being for the first set of images.

However, while in both Figure 17 and Figure 19 the shown correlation between bands is somewhat sparse, after the corrections performed to enhance the images, Figure 21 continues to demonstrate an extremely high correlation between the visible bands, maintaining a monochromatic dull image. To remedy this situation, a DS method was applied to attempt to further segment the spectral information of each band, allowing to obtain the band correlations seen in Figure 22.

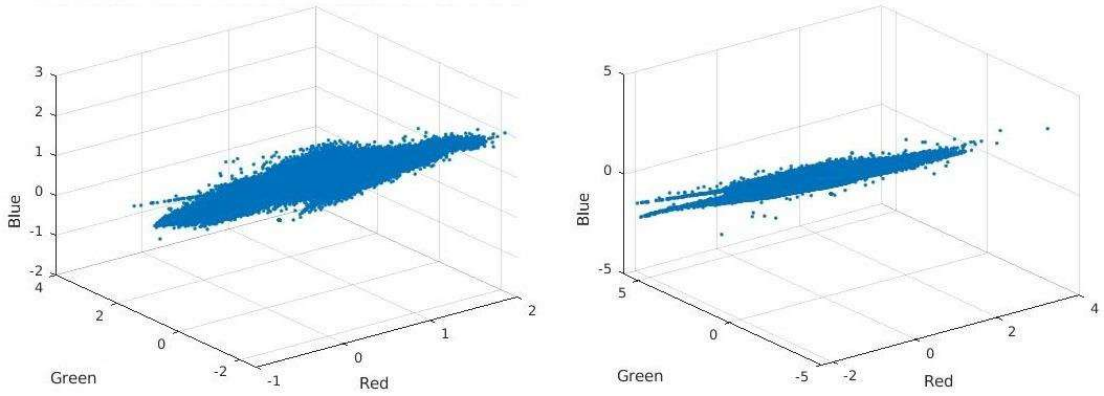


Figure 22: Corrected and Decorrelated RGB band correlation for each of the MS image, on the left being for the second set while the right being for the first set of images.

The decorrelated RGB bands in Figure 22 show a marked improvement in transforming the monochromatic image into an enhanced colour image, where each pixel main colour is easily visible. While linear correlation is still present, it can only be seen for higher ranges of pixel values in comparison to the low ranges present in Figure 17 and Figure 21.

Figure 23 and Figure 24 were the derived showcasing of the obtained images after applying different methods to the original ones, where: A) represents CS; B) represents

HE; C) represents CLAHE; D) represents CS followed by CLAHE; E) represents DS; F) represents CS followed by DS; G) represents CS followed by CLAHE and subsequently DS.

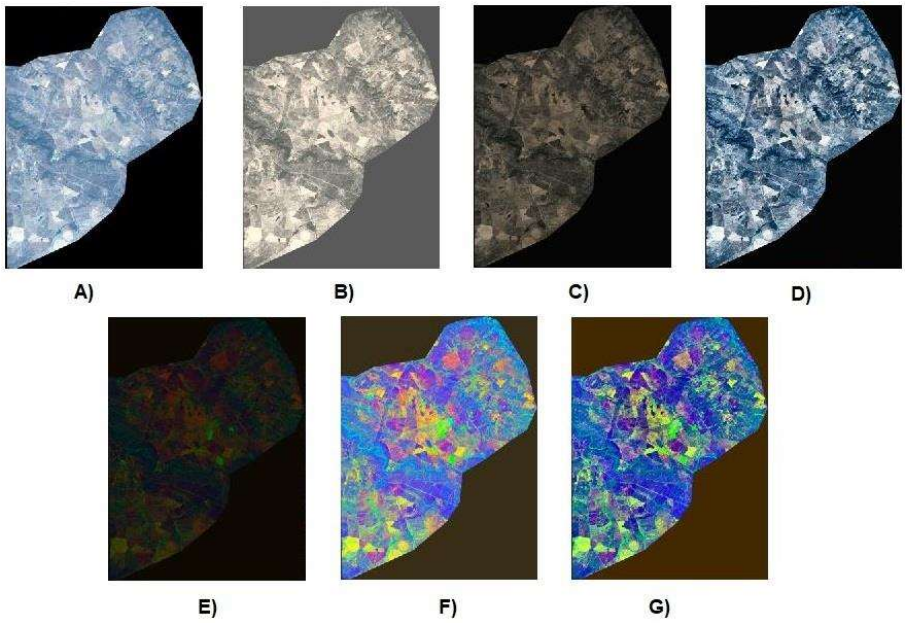


Figure 23: Derived images after application of contrast enhancement methods on the MS image received on the second set.

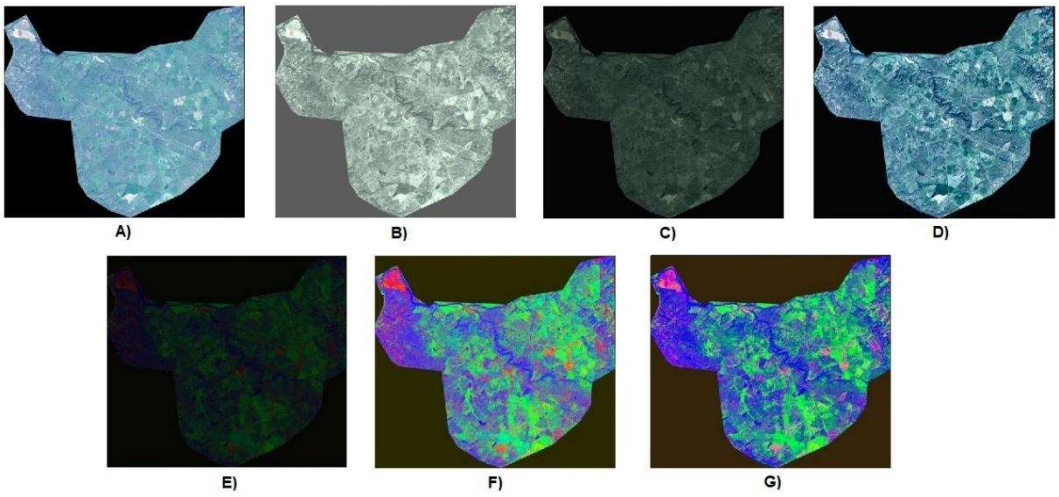


Figure 24: Derived images after application of contrast enhancement methods on the MS image received on the second set.

From this procedure, we can conclude the potential processes to be used, depending on the final objective. For the extraction of Spectral Indices, the process that uses DS and contrast enhancement methods should be applied to cases where the most essential part was the spectral information itself and thus, was applied only to extract Spectral Indices,

while the process that utilizes only contrast enhancement methods was to be applied for general purposes, such as image fusion.

Chapter 5 – Implementation, Results Analyses and Discussion

The following steps will expand on the experimented and implemented techniques ending with a description of the flowchart final developed algorithm, showcasing the products obtained by each of the individual steps and, finally, displaying the results obtained with its use.

This chapter is divided in three sections, one presenting the obtained results for each of the conceptualized steps in the system alongside conclusion taken from them, and one presenting the final and restructured system architecture.

5.1. Implementation Results

In accordance with the previously proposed system workflow in Figure 3, two different methods of dolmen detection were experimented with: one making use of MS and Pan type imagery and one using exclusively the HS images. In this chapter, the implementations using both imageries are explained.

5.1.1 HS Material Signature Method

This system's original objective is, essentially, to extract the reflectance values of the image's pixels for the range of wavelength present in the spectral bands of the HS image and individually compare them to the reflectance values of the isolated dolmen's material obtained in laboratory and publicized in an online spectral library, with the intuit of determining areas of high likelihood of dolmen presence through detecting the monuments composing material's presence in certain pixels.

To prove this systems viability, the average spectral reflectance of all the pixels with dolmen's expert identification was compared to the average spectral reflectance of several areas of surrounding pixels where no dolmens are known to be present. This comparison serves as a test to see if the presence of the dolmen has any effect in the extracted spectral reflectance values along the different bands when compared to areas where no dolmens are known to exist, such as the pixels immediately surrounding the various pixels with dolmen presence and areas randomly extracted from the region, and from this difference assign the continuous bands spectral reflectance values as signalling the presence of the

dolmens' material composition. Unfortunately, nearly all pixels found are inserted in similar backgrounds of natural areas and thus present similar reflectance values along the spectral bands, with the only difference being one dolmen found in the middle of the town of Pavia, whose unique case among dolmens hindered its impact in extracting the unique material spectral signature of the dolmens.

For this project to be viable, there should be a visible difference between the presence and absence of a dolmen in the pixels to make it possible to assign the difference to the presence of a unique material. The results of this test can be seen in Figure 25 where the pixel reflectance percentage values are shown for several NxN grids of adjacent pixels without dolmen presence, centred on expertly identified dolmen pixels.

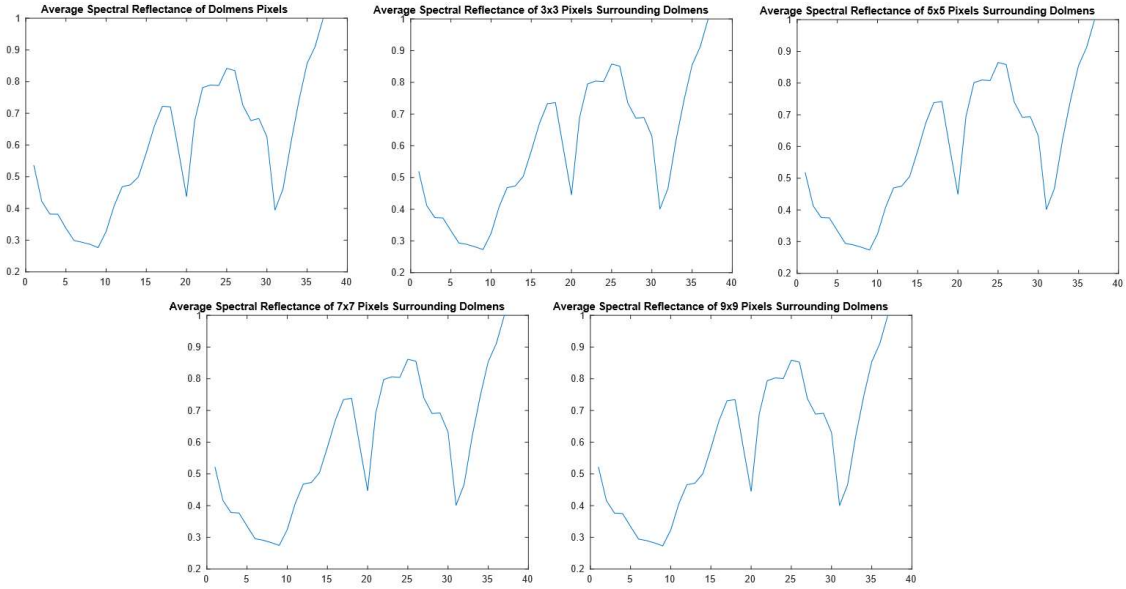


Figure 25: Comparison between pixel reflectance percentage values (y axis) and spectral wavelength in bands (x axis) for a NxN grid of surrounding pixels centred on an identified dolmen pixel.

Figure 25 shows that the average spectral reflectance values are maintained independently of the considered area, indicating that the spectral information of the region where the dolmens are inserted is the same as the ones outside the location of the monument permitting to conclude that this method is incapable of recognizing the dolmens material signature.

While we can conclude that achieving a direct method of object recognition will be virtually impossible, these images possess great levels of spectral information of far more detailed levels than the MS images. However no experiments were carried out in their

use, such as spectral indices extraction, as the total geographic area they covered was too small to be considered worthy of being studied in regards to the creation of an automatic machine learning method, further limiting the already small number of available identified dolmen locations.

5.1.2 PAN and MS Data Fusion Method

In this system, following the first step of the implementation to pre-process and enhance the received images, the extraction of spectral indices was attempted. Afterwards, the resultant indices and images were fused into a single image on which feature extraction methods were applied.

5.1.2.1 Image Fusion Methods

Following both the procedures of image enhancement and index extraction, a method to tie the information together was applied, this being the image fusion techniques. As previously described, two different techniques were applied to fuse the PAN and MS images: IHS and Brovey transformation method. The following Figure 26 and Figure 27 show the results of the application of these methods to a single part of the full image of the first set. chosen due to the previously identified presence of three dolmens by an expert, encompassing one tenth of both the length and width of the image and applied to the MS RGB bands, based on an image enhancement without the application of the DS method.

Figure 26: IHS Transformation applied to the contrast enhanced PAN and RGB MS partial image.

Figure 27: Brovey Transformation applied to the contrast enhanced PAN and RGB MS partial image.

5.1.2.2 Spectral Indices Extraction

As stated already, several vegetational, water, shadows, soil, man-made and road indices were extracted, according to the previously described formulas, to search for the presence of these elements in the image's area as an indicator for a dolmen's presence. According to the experts opinion, these are commonly found, either covered by vegetation or their presence causing vegetational growth outside of the norm, and are usually located in areas close to water sources and standing out from the background soil. The images showcasing the results obtained from the various extracted spectral indices can be seen in Appendix B.

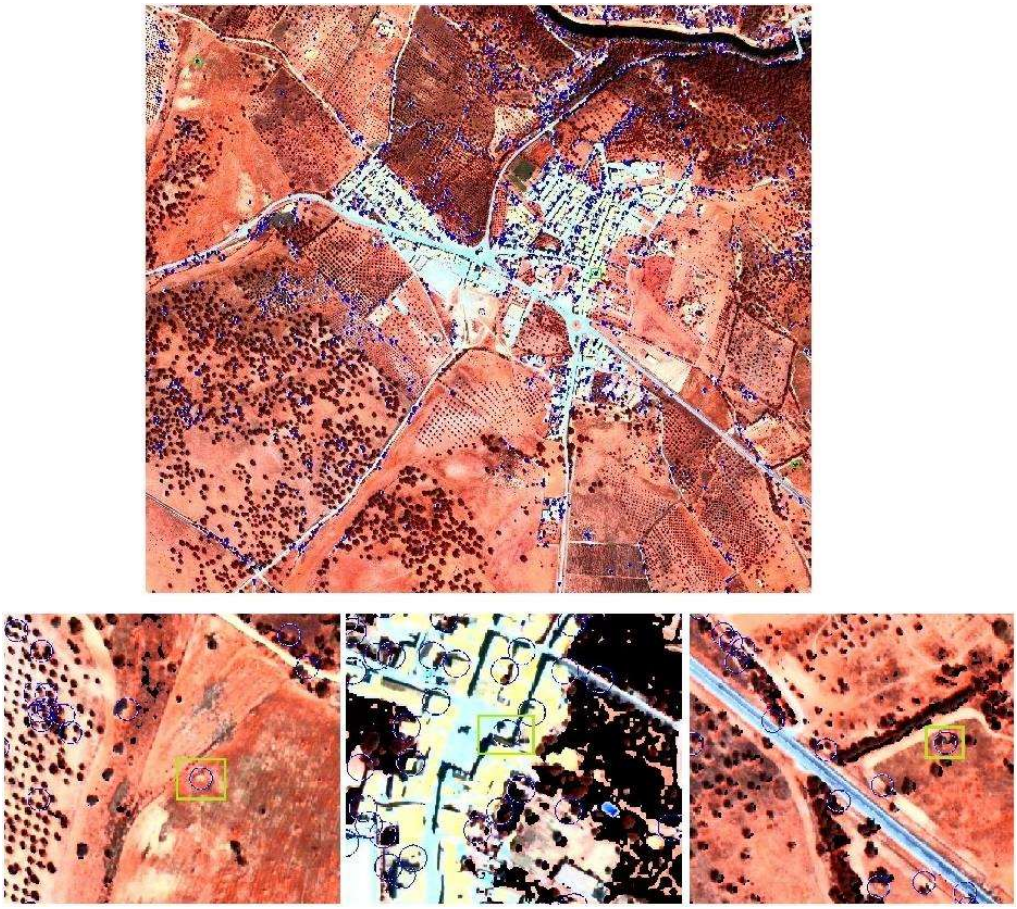


Figure 28: Example of CHT classification experiment Simple Detection on a partial False RGB image of the second set.

5.1.2.3 Circle Detection Method

To detect any identifiable spatial circular feature, either a full circle or a partial circle, associated to the dolmens present in the images, in an attempt to detect both the monuments identified by an expert and any possible new dolmen in open air, a CHT method was applied in three possible ways, in search of the best possible result of automatically detecting expertly identified dolmens.

The first experiment, hereafter described as experiment Simple Detection, is based on applying the CHT directly on the fused image (i.e.: Figure 26), followed by the application of one of each of the spectral indices developed (vegetation, water, soil, shadow, man-made structures and road) to mask the said image. Then, any circle whose radius is below 1.15m and above 3.25m and in whose isolated image, irrespective of the index detected, the original image occupies at least 20 percent of it is cut out. Figure 28

showcases the results on a small False RGB (formed by substituting the Blue band in RGB colour space with the NIR1 band, forming the NRG colour space) partial image where the blue dots signify possible monument locations, and the green squares identify the true and expertly defined location of the dolmens.

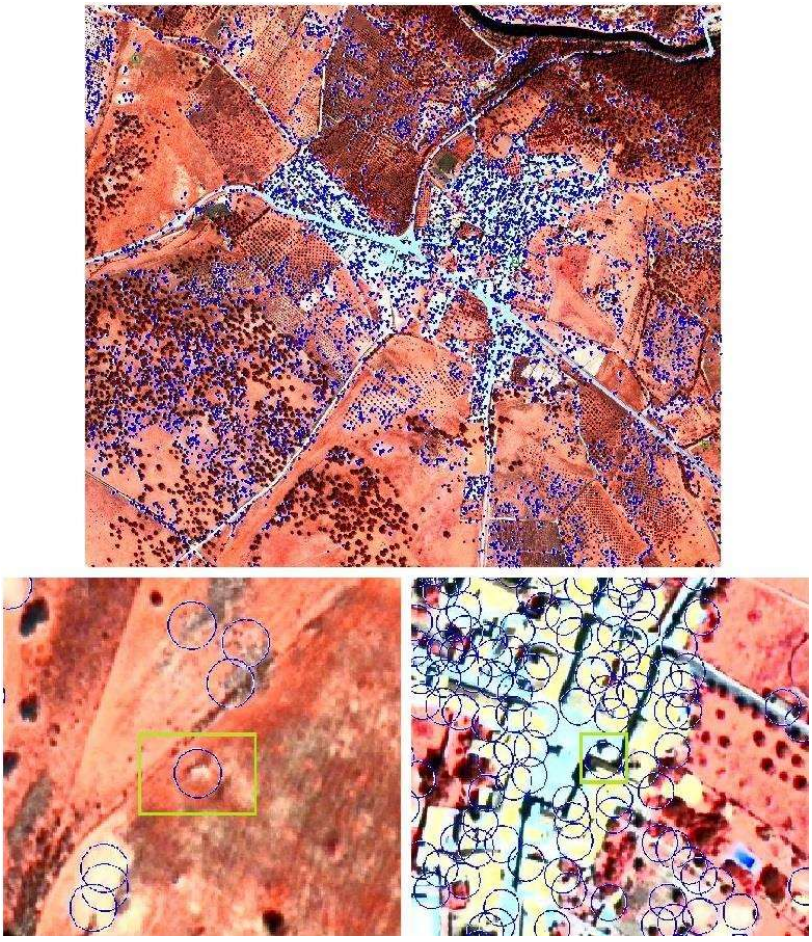


Figure 29: Example of CHT classification experiment Masked Detection on a partial False RGB image of the second set.

The second experiment, hereafter described as experiment Masked Detection, applies the CHT on the resultant image after applying and masking the original image with one of each spectral index developed and from its results cut any whose radius is below 1.15m and above 3.25m. Figure 29 showcases the results on a small False RGB partial image where, once more, the blue dots signify possible monument locations and the green squares identify the true and expertly defined location of the dolmens.

The third experiment, hereafter described as Simultaneous Detection, is the simultaneous application of both the above described Simple Detection and Masked

Detection approaches, while maintaining the exclusion of any result whose radius is below 1.15m and above 3.25m and in whose isolated image, irrespective of the index detected, the original image occupies at least 20 percent of its present in approach A (Figure 30).

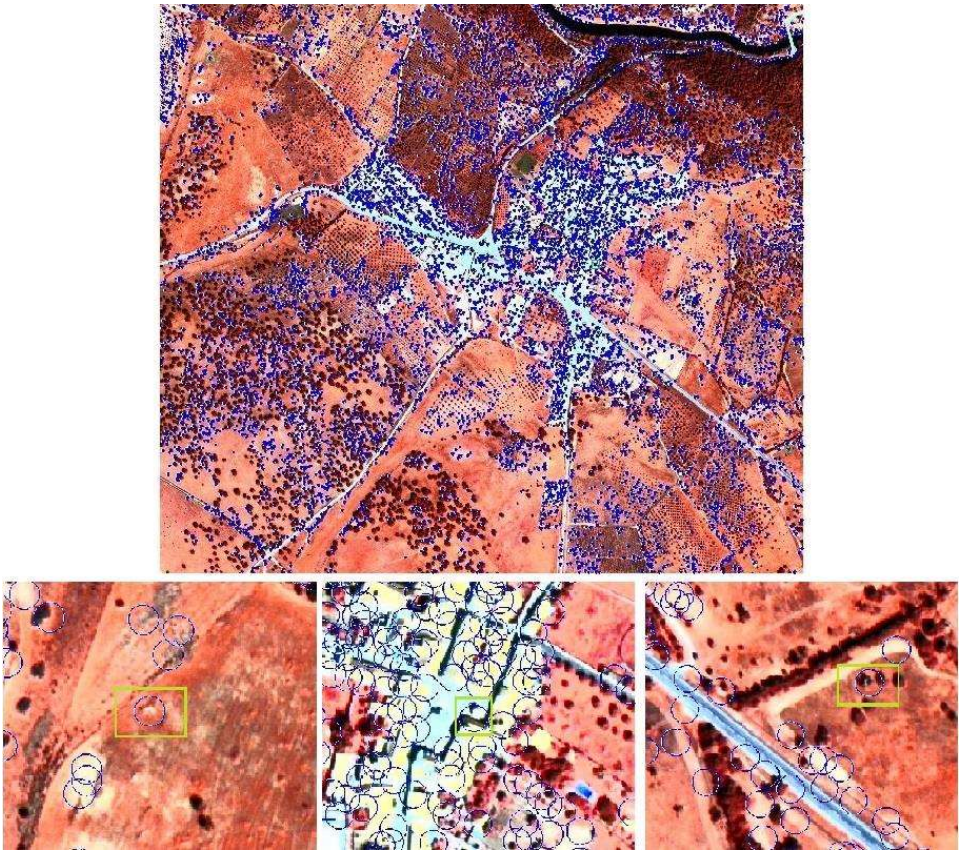


Figure 30: Example of CHT classification experiment Simultaneous Detection on a partial False RGB image of the second set.

Table 3 demonstrates each methods' capability in terms of detecting circles where the presence of dolmens has been previously identified by an expert both the three experiments applied to both images, while Table 4 was created to showcase and compare the total number of detected circles for the complete image.

Table 3: Expertly detected identified visible dolmens in circles for each experiment and image.

	Simple Detection	*Masked Detection*	*Simultaneous Detection*	*Expert Defined Dolmens (max.)*	*Best Total Percentage*
First Set Image	6	4	7	13	±54 %
Second Set Image	8	11	15	18	±83 %
Independent Dolmens	10	13	17	18	±94 %

Table 4: Detected circles from CHT where there exists no previously identified dolmen, for each experiment and image.

	Simple Detection	*Masked Detection*	*Simultaneous Detection*
First Set Image	122.937	184.292	275.721
Second Set Image	180.505	379.932	496.972
Total	303.442	564.224	772.693

Analysing both tables, it is easy to note that experiment Simultaneous Detection, that integrates both precursory approaches of Simple Detection and Masked Detection and elimination of overlapped detected circles, is the one that returns the highest number of identified dolmens, while simultaneously returning the highest number of detected circles., If both the detected visible dolmens and detected dolmens in each image are added together, the experiment Masked Detection attains a value of 14 dolmens and 564.225 circles whereas the experiment Simple Detection detects 13 dolmens and 303.442 circles. Thus, it can be concluded that Masked Detection is slightly better at detecting dolmens but suffers from a higher number of false positive circles (260.783) when compared to Simple Detection.

Finally, comparing Table 3 and Table 4 and taking into consideration the objective of detecting the highest possible number of visible dolmens, experiment Simultaneous Detection was chosen as the best tested method to consider for the remaining of the research. Further, taking into regard the extremely high number of false positives circles detected in comparison to the small value of true dolmens discovered, several steps were taken to try and reduce false circle dolmens while maintaining the maximum possible

number of true positive dolmens. These steps and results will be explained in the following sub-sections.

5.1.3 Delimitation of Areas Using Expert Knowledge

To delimitate the considered area to classify as possible dolmen locations, the knowledge of an expertly defined ontology for the dolmen insertion in the territory under study was used. Based on this ontology, such monuments are usually inserted on rural and elevated terrain zones, near rock outcrops and close (around less than 2km) to natural water sources, while defining both urban and agricultural plantation zones as not presenting said monuments due to them being destroyed for construction or agricultural activities.

From the available information, no defined urban/rural, agricultural, of elevated terrain or rock outcrops zones could be defined, leaving only possibility of defining regions based on water source proximity. Thus, areas found to be further away than 1.2km from a relatively large water source were masked and presented as not feasible for dolmen locations. The following figures (Figure 31 and Figure 32) demonstrate the differences between the originally received set of images and the resultant regions based on water proximity in accordance with the expert knowledge. With this method's implementation the detection of identified dolmens was unchanged. However, the number of false circles detected has dropped due to the reduction in the area of interest thus obtained. The reduction of area of interest in the first set of images (Figure 32) is of around 8.61%, while for the second set of images (Figure 31) it is around 5.04%. Table 5 displays the resultant false circles for each of the three experiments and two images.

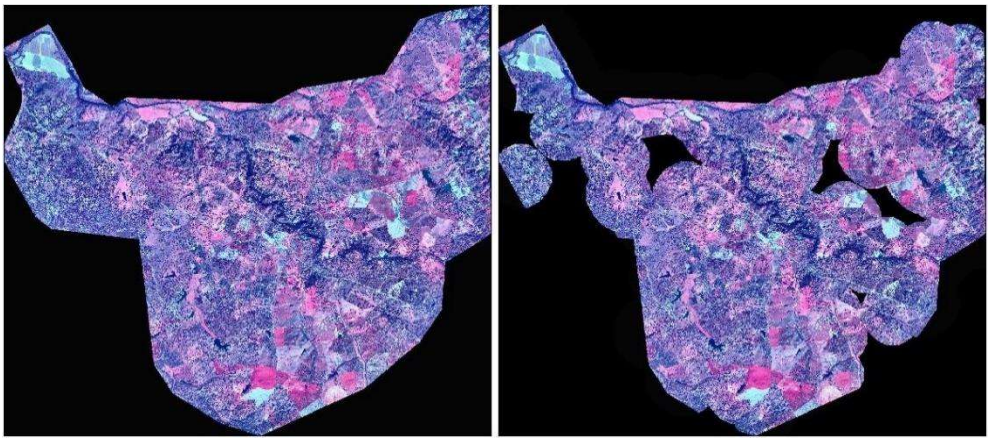


Figure 31: Difference between the received second set of images (left) and the ontology-based water regions (right).

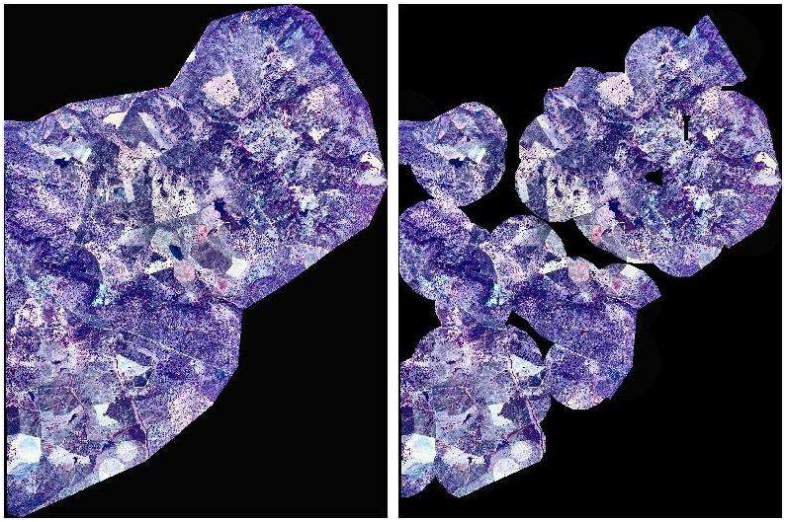


Figure 32: Difference between the received first set of images (left) and the ontology-based water regions (right).

Table 5: Detected circles from CHT where there exists no previously identified dolmen, for each experiment and ontology-based water defined region image.

	Simple Detection	*Masked Detection*	*Simultaneous Detection*
First Set Image	102.537	160.874	237.277
Second Set Image	159.076	346.615	450.471
Total	261.613	507.489	687.748

An analysis of Table 4 and Table 5 allow to conclude that the falsely detected dolmens by the CHT was effectively reduced: experiment Simple Detection eliminated ±13.8% of circles in both images, Masked Detection removed ±10.1% of circles in total and Simultaneous Detection removed ±11% of circles in both images when comparing

with the number of circles detected in the original images, therefore, proving the effectiveness of an ontology-based definition for restricting areas with higher likelihood of dolmen's presence, allowing for only searching for the objects in question within the defined boundaries.

5.1.4 Circle Classification Models

As analysed above, the application of the CHT circle detection technique retrieved an extremely high number of circles, independently of the type of experiment and of image, with the lowest number being around 120.000 and the highest around 500.000 (see Table 4). The application of ontology based defined regions, the number of circles for probable dolmen identification, although still high, dropped for a lowest of around 100.000 and a highest of around 450.000 (see Table 5). As such, to further reduce the erroneously detected circles, several classification models were tested, trained and implemented with different techniques, in such a way that the models effectively eliminated, or at least minimized, the loss of any of the circles detecting previously identified dolmens. To train these models, each detected circle was classified as either being a dolmen or not and then a square RGB image of 28 per 28 pixels centred on the circle was extracted and saved.

These models (see Classifiers) were trained recurring to the MATLAB Classification Learner app that allowed, based on the discovered visual features obtained through the application of BoVW method, to easily train a variety of cross-validation models of different types. It also chooses to apply (or not) a PCA when training and tunes the model by searching for the best hyperparameters for each model, allowing for their easy comparison and exporting. Finally, an ensemble of five of the best trained models was created and used to independently classify each detected circle as being a dolmen or not, deleting any circle that below a defined threshold, in this case three, of votes classifying as being dolmen.

Further, two different experiments were made for the training of the models: one that only considered the originally extracted dolmen circles and one that performed data augmentation by applying rotation to the images, in 90º degree angles, performing a full circle mapped rotations. However, when comparing the confidence of the two models for circles that encapsulate known dolmens, it was found that the model without rotation

affected the detection of previously identified dolmens when taking into consideration the majority-of-three voting system, by not being able to correctly classify one of the truly detected dolmens, while the model with rotation did accurately classify all the existent dolmens. Nevertheless, both schemes had clear effects on the number of falsely identified circles. Table 6 showcases those erroneously identified circles numbers obtained through method Simultaneous Detection for each image and experiment, with or without rotation, applied during training.

Table 6: Detected circles with more than three model votes through method Simultaneous Detection, where there exists no previously identified dolmen, for models trained with or without rotation.

	Without Rotation	*With Rotation*
First Set Image	114.979	117.925
Second Set Image	215.398	207.873
Total	330.377	325.798

From an analysis of the results described in Table 6, when taking into consideration that when applied with rotation was able to capture one more of the known dolmens, one that was partially inserted within a tree, in addition to the fact that the model with rotation enabled to obtain less misclassified circles allows to conclude that the developed model with rotation is overall better at detecting dolmens.

5.1.5 HSV Colour Space Detection Implementation

After the application of the CHT circle detection technique and retrieval of detected circles in False RGB images, it was found that this method retrieved an extremely high number of circles independently of the type of experiment and image it was applied on (see Table 4), while detecting a very low number of circles that truly possess dolmens. In efforts to both capturing more circles capable of detecting dolmen presence and diminishing even more the incorrectly classified circles, the best implemented process concluded from the previous sub-sections were directly applied to a converted HSV colour space of the original RGB image.

From this process, Table 7 was obtained were the number of identified dolmens captured before applying supervised classification is described. Table 8, on the other

hand, showcases the number of circles of known dolmens initially captured through a model with rotations of 90º degree angles applied to the dolmen dataset, alongside the respective misclassified number of circles.

Table 7: CHT applied to HSV colour space displaying truly detected dolmens in circles for experiment Simultaneous Detection with ontology defined regions.

	Dolmens Captured Through Simultaneous Detection	*Expert Defined Dolmens (max.)*	*Best Total Percentage*
First Set Image	7	13	±54 %
Second Set Image	16	18	±89 %
Independent Dolmens	17	18	±94 %

Table 8: CHT applied to HSV colour space displaying detected and false positive dolmens in circles for experiment Simultaneous Detection with ontology defined regions, after application of automatic classification model.

	Simultaneous Detection Classifier (Rotation)	*Expert Defined Dolmens (in Simultaneous Detection)*	*Best Percentage*	*False Positive Dolmens Classifier (Rotation)*
First Set Image	5	7	±71 %	107.400
Second Set Image	16	16	100 %	204.754
Independent Dolmens	16	17	±94 %	-
Total	-	-	-	312.154

Comparing the results displayed in Table 3 and Table 7, its noticeable that the application of CHT to the HSV colour space can detect one more of the known dolmens than its application to the False RGB colour space. Further, analysing Table 3, Table 6 and Table 8, it can be seen that while it couldn't classify two known dolmens in the first image set but could classify one more in the second image set, the misclassified dolmens of the model that applied the 90º degree angles rotation to the dolmen dataset are lower in the HSV colour space than the False RGB.

After analysing the obtained results, it was found that the reason the model couldn't identify the two missing dolmens was that both of them were not at least partially covered by vegetation, that is, what the model learned was not how to identify dolmens but instead how to identify trees with nearby rocks. This is also one of the main reasons for the great number of erroneously classified circles, as such features are common in the area. Finally,

taking into consideration that when this models equivalent was applied to the False RGB colour space, it was able to capture one extra dolmen due to the fact it was partially inserted within a tree, it can be concluded that the same behaviour is being observed on both model cases.

To solve this issue, another way was experimented to create and train the models for the classifiers through two steps: first, due to noting that the number of partially covered dolmens was much higher than the naked dolmens, several partial captures of said dolmens were extracted and added to the dataset; second, having proved that models with at least some rotation added were better able to isolate the dolmens and had proved beneficial in reducing incorrect classifications, for each of the images present in the dataset twenty-four rotations of 15º degree angles were applied to further augment the data. From this solution Table 9, demonstrates the number of circles of known dolmens captured through said model of partial dolmen images and 15º degree angle rotations, alongside the respective misclassified number of circles.

Table 9: Detected and misclassified dolmens in circles for experiment Simultaneous Detection in HSV colour space with ontology defined regions, after application of a data augmented automatic classification model.

	Simultaneous Detection Classifier (Rotation + Partials)	*Expert Defined Dolmens (in Simultaneous Detection)*	*Best Percentage*	*False Positive Dolmens Classifier (Rotation + Partials)*
First Set Image	7	7	100 %	32.763
Second Set Image	14	16	±88 %	57.211
Independent Dolmens	15	17	±88 %	-
Total	-	-	-	89.974

Comparing the results displayed in Table 8 and Table 9, although this new model did not prove capable of detecting all the initially verified dolmens through CHT it demonstrated the ability to cut the number of incorrectly classified circles in over ±71.2%, reducing the total number of false positives from 312.154 with the previous model to 89.974 in the new one. While the total number of true positives obtained in both image sets did not change, the total independently recognized dolmens did with the loss of two of the known dolmens present.

Although it can be seen in Table 7 that CHT applied to HSV had already proved incapable of detecting one of them, majorly due to the complete lack of any indication of its presence and near complete degradation (as can be seen in Figure 33, showing the HSV and the Google Earth image of the monument location side by side), in the case of the two misclassified dolmens, the CHT was capable of detecting some circularity being present either near the monument (e.g.: Figure 34) or on the monument's physical location (e.g.: Figure 35), resulting in obtaining an image with only an incomplete partial piece of the monument at their edge or due to the advanced levels of degradation of the monument in question the classification models not detecting them as dolmen locations, respectively.

The following Figure 34 and Figure 35 showcase the side by side comparison of the HSV and Google Earth images of the erroneously classified dolmen locations whilst Figure 36 and Figure 37 present the HSV and Google Earth image comparison of correctly classified dolmens.

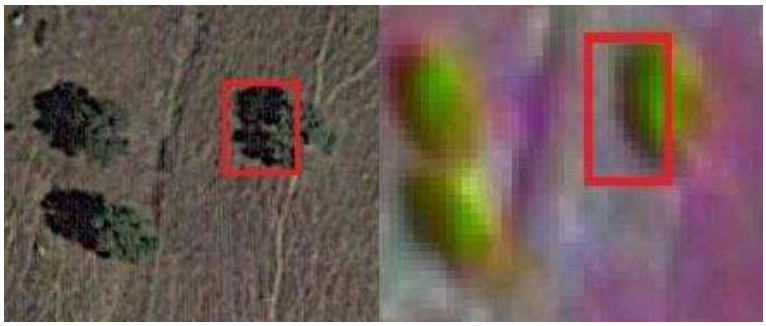


Figure 33: Google Earth and HSV image of the undetected CHT Gonçala 3 monument location, side by side.

Figure 34: Google Earth and HSV image of the erroneously classified Oliveira 1 monument location, side by side.

Figure 35: Google Earth and HSV image of the erroneously classified Antões 1 monument location, side by side.

Figure 36: Google Earth and HSV image of the correctly classified Gonçala 1 monument location, side by side.

Figure 37: Google Earth and HSV image of the correctly classified Adua 1 monument location, side by side.

Finally, taking into consideration what was learned through the Literary Review, this system results are great for an area of such size, as it has been found that the larger the area and the smaller the monument in question, the greater the number of false positives, with the inverse also applying.

5.2. System Architecture Restructure

Based on the results obtained during the implementation of the conceptual model, shown in Figure 3, said model can now be reconstructed after factoring in the discovered inability of use of HS images, the application of a knowledge-based method based on the defined ontology to delineate region of interest, the varied types of spectral indices

applied to discard unnecessary information and the multiple layers of classification applied to take into consideration varied obtained features. Thus, Figure 38 shows the reconstructed system architecture and workflow.

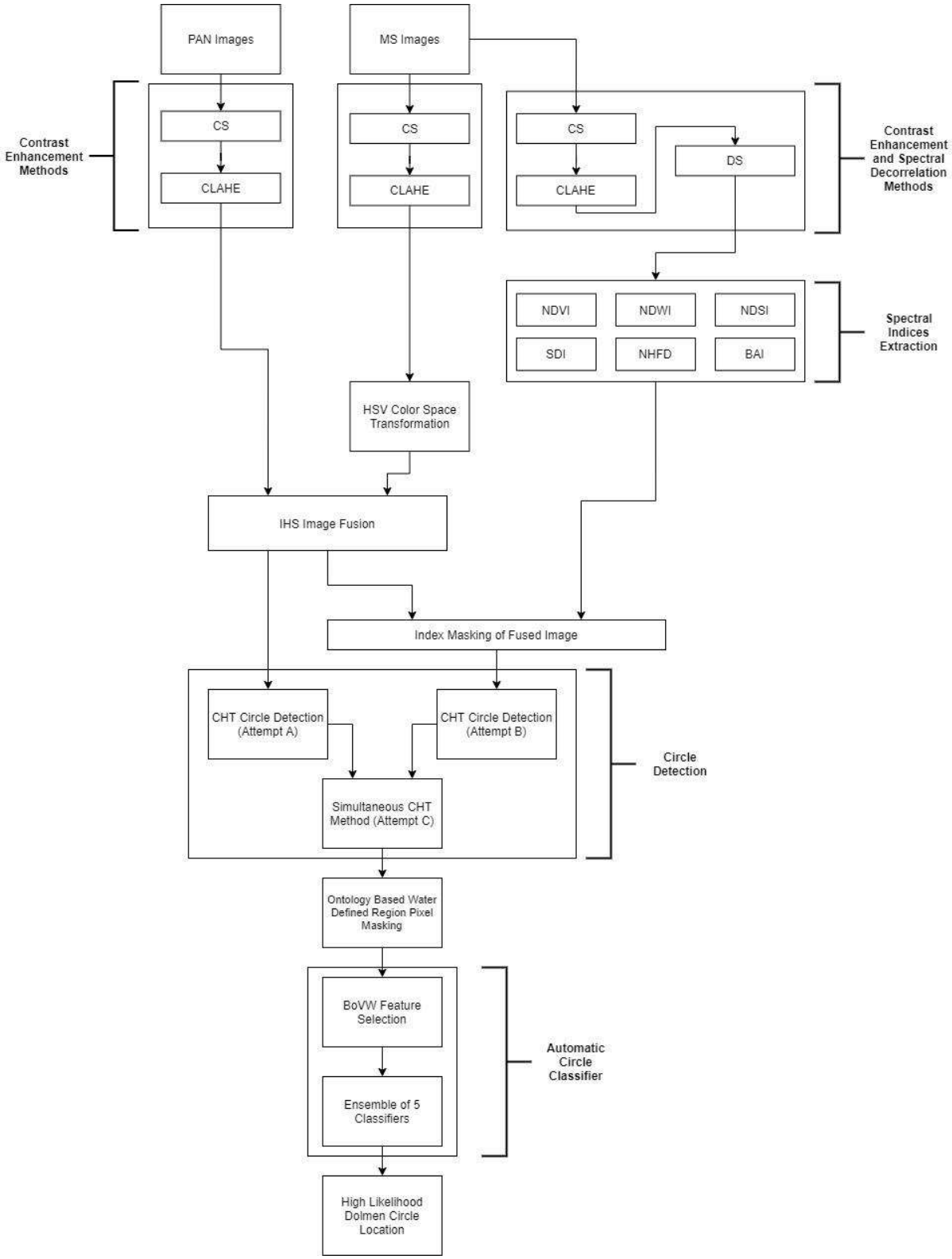


Figure 38: Reconstructed Workflow and System Architecture in use.

5.3. Discussion of Results

Based on the experiments carried out in all three image types, several results stand out.

First of all, we proved that the HS images provided were could not be used to extract a spectral material signature of the dolmens. This fact is due to a combination of (i) large pixel size resolution when compared to the dolmen' small dimension and (ii) the fact that. the monuments composing materials are in fact commonly found spread throughout the area of implementation. Further, the total area the HS image captures is a very small part of the whole area of study, limiting the usefulness of the spectral information extracted and eliminating a large part of the expertly positioned dolmens from consideration, leading to the HS images not being used alongside the other images type.

The experiments carried out in the creation of a system using PAN and MS images proved that dolmen detection is more accurate in the HSV colour space over traditional RGB colour space. Further, the system demonstrated that spectral indices can further define partially obstructed dolmens, allowing for the dolmens to show various degrees of circularity. Additionally, the developed system showed that expert knowledge based on ontologies can be, at least in part, used to delineate areas of high probability of dolmen presence and, finally, proved that an ensemble of several supervised classification models can classify detected circles (through CHT) such that it can capture most of the known dolmens for automated detection while eliminating the number of erroneously detected circles in over ±87.2%.

Chapter 6 – Conclusion and Future Work

This thesis investigated the possibility of creating an automatic system capable of using satellite imagery for identification of ground areas presenting high likelihood of the presence of archaeological monuments – dolmens. For this end, information captured in various types of image – MS, HS and PAN – was used in experiments to enable a system that, in the case study at our disposal, detected the greatest possible number of dolmens in a set of previously identified dolmens, knowledge provided by a human expert. The system was tunned so as to minimize the number of erroneously classified regions as showing high probability for dolmens implantations.

This chapter is divided in four sections presenting the conclusions, the study contributions to the academic and business-level, and, finally, the inherent limitations and direction for future research.

6.1. Main Conclusions

Taking this thesis objectives into consideration, it was possible to prove that the type of Hyperspectral image obtained from the Proba-1 CHRIS satellite sensor could not be used for the creation of a dolmens' material signature, mostly because of the fact that the surrounding environment was rich in the material used in the monument's fabric. This also happens, in part, due to the difference in size between the pixel's image resolution and the dolmen's physical size. The pixel in the HS image covers a ground area of 17m per 17m, while the dolmen, at best, encapsulates a diameter of around six meters. Thus, the HS images can neither be used to enable a system capable of automatically defining small regions of high likelihood of presence of dolmens within an image, nor for obtaining information to be fused into a system using MS and Pan images to delineate areas of interest where to apply circle detection on.

On the other hand, in successfully creating a system using Panchromatic and Multispectral images that proved capable of detecting dolmen locations of up to fifteen of eighteen of the previously identified dolmens (around 83% of accuracy), this system proved that, at least partly, the expertly defined ontology of the dolmen (the domain knowledge) is of added-value for increasing the performance of automated dolmen

detection systems. In this case, the knowledge that dolmens are usually inserted near relevant water sources was transferred and used to enhance the automatic classification of image areas of higher probability of dolmen's presence. Further, while it is possible to detect some of the more visible and striking dolmens in satellite imagery through direct application of a circle detection method, only with the use of spectral information and respective creation of spectral indices was it possible to eliminate pixels of known unwanted substances, allowing to define circular or partially circular regions presenting higher probabilities for the location of dolmens that demonstrated capacity of capturing areas with visual obstacles for clear dolmen visualization.

The developed system also allowed to conclude that dolmens are easier to detect when using the HSV colour space of the satellite imagery, returning both slightly better detection of verified dolmen locations and a lower number of erroneously classified dolmen sites. Finally, it. was possible to identify the possibility of using supervised machine learning methods to create models capable of automatic classification of individual locations as being possible dolmen locations based on similarity of visual features to previously expertly defined sites. The developed supervised model allowed to eliminate around 87.2% of false positives.

6.2. Contributions

6.2.1 Academic Implications

This thesis mainly contributes with a preliminary work on automated remote sensing classification systems in satellite images of archaeological monuments in Portugal, while simultaneously contributing towards the start of research on image automated classification for heritage detection.

Further, it offers proof of concept towards the effectiveness of usage of an existent ontology of environmental dolmen insertion for defining macro regions where dolmens can be expected to be present and thus, be used for the construction of a mixed knowledge-based classification system.

6.2.2 Business-Level Implications

Meanwhile, this proof of concept enables further research into an automatic tool for aiding archaeologists in the detection of new dolmens, without needing to recur to time-consuming ground cover of whole regions, and simultaneously aiding in identifying zones of cultural importance to be protected in rural and agricultural regions.

6.3. Study Limitations

This study main limitations are found in the collected Hyperspectral images. Several factors contributed for these limitations. First, the images that ESA sent us only covered a small part of the total region of interest. On the other hand, the ground sampling distances and, thus, the correspondent pixel size, were too big for the needed level of ground analysis required, leading to problems in analysing the much smaller monuments. Finally, the total number of spectral bands captured were too low: 37. Far more spectral bands were needed to develop spectral material signatures with high accuracy.

A big limitation was also found in the decayed state of most of the monuments in the case study area, having caused the loss of the former circular shape the monuments were built with, and leaving mostly irregular shapes to be detected. This fact led to some monuments simply being undetectable since the system was unable to distinguish between them and common rock formations present in the region. Other limitations involve the low number of previously identified dolmens located in the region leading to a low number of previously annotated images for the training phase of the classifier models, which necessarily hinders learning.

6.4. Future Research

Megalithic monuments present local and fragrant specificities, preventing the application of existent automated approaches to new geographical areas. While archaeological remains automatic detection systems using aerial and satellite images have already began to gain traction in recent years, such has been noted to happen mostly on foreign countries, and therefore leaving a significant lack of such systems being

developed and applied to Portuguese regions and monuments, leaving much room for growth.

While this dissertation proved successful in using an already developed ontology to aid in detection, it would be of interest to develop the system in order to encapsulate all of the regions where the archaeological monuments in question have been detected, and expand on the system by adding new domain knowledge, based on the ontology, capable of further refining areas of high probability of presence.

This thesis makes use of Panchromatic and Multispectral image types only. Different image types, such as Hyperspectral and LiDAR, could provide distinct and more interesting avenues for the automatic detection of dolmens, providing that the spatial resolution is fine enough for isolating smaller patches of ground and enable new paths for signature identification and discovery of new features related to dolmens' height.

6.5. Acknowledgements

We are thankful for the data provided by the European Space Agency.

Bibliography

[1] T. Freeland, B. Heung, D. V. Burley, G. Clark, and A. Knudby, "Automated feature extraction for prospection and analysis of monumental earthworks from aerial LiDAR in the Kingdom of Tonga," *J. Archaeol. Sci.*, vol. 69, pp. 64–74, 2016.

[2] D. Alexakis, A. Sarris, T. Astaras, and K. Albanakis, "Detection of neolithic settlements in thessaly (Greece) through multispectral and hyperspectral satellite imagery," *Sensors*, vol. 9, no. 2, pp. 1167–1187, 2009.

[3] A. Guyot, L. Hubert-Moy, and T. Lorho, "Detecting Neolithic burial mounds from LiDAR-derived elevation data using a multi-scale approach and machine learning techniques," *Remote Sens.*, vol. 10, no. 2, 2018.

[4] A. Câmara, L. Rocha, and T. Batista, "A fotointerpretação como recurso de prospeção arqueológica. Chaves para a identificação e interpretação de monumentos megalíticos no Alentejo: aplicação nos concelhos de Mora e Arraiolos," Universidade de Évora, 2017.

[5] R. M. Cavalli, G. A. Licciardi, and J. Chanussot, "Detection of anomalies produced by buried archaeological structures using nonlinear principal component analysis applied to airborne hyperspectral image," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 6, no. 2, pp. 659–669, 2013.

[6] A. Câmara, "Automatic Methods for the Identification of Archaeological Barrows in Satellite Images," 2017.

[7] Y. Zeng, W. Huang, M. Liu, H. Zhang, and B. Zou, "Fusion of satellite images in urban area: Assessing the quality of resulting images," *2010 18th Int. Conf. Geoinformatics, Geoinformatics 2010*, no. 40771198, pp. 1–4, 2010.

[8] R. Lasaponara and N. Masini, "Identification of archaeological buried remains based on the normalized difference vegetation index (NDVI) from quickbird satellite data," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 3, pp. 325–328, 2006.

[9] N. Kosaka, T. Akiyama, B. Tsai, and T. Kojima, "Forest type classification using data fusion of multispectral and panchromatic high-resolution satellite imageries," *Int. Geosci. Remote Sens. Symp.*, vol. 4, no. C, pp. 2980–2983, 2005.

[10] O. Kose and E. Ozbek, "Fusion of Hyperspectral and LIDAR Remote Sensing Data for Classification of Complex Forest Areas," *Andrologia*, vol. 47, no. 6, pp. 603–603, 2015.

[11] C. Debes *et al.*, "Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2405–2418, 2014.

[12] E. Cerrillo-Cuenca and P. Bueno-Ramírez, "Counting with the invisible record? The role of LiDAR in the interpretation of megalithic landscapes in south-western Iberia (Extremadura, Alentejo and Beira Baixa)," *Archaeol. Prospect.*, no. March, pp. 1–14, 2019.

[13] C. Sevara, M. Pregesbauer, M. Doneus, G. Verhoeven, and I. Trinks, "Pixel versus object - A comparison of strategies for the semi-automated mapping of archaeological features using airborne laser scanning data," *J. Archaeol. Sci. Reports*, vol. 5, pp. 485–498, 2016.

[14] Ø. D. Trier, S. Ø. Larsen, and R. Solberg, "Automatic detection of circular structures in high-resolution satellite images of agricultural land," *Archaeol. Prospect.*, vol. 16, no. 1, pp. 1–15, 2009.

[15] A. Câmara, A. de Almeida, J. Oliveira, and M. Silveira, "Photointerpretation as a Tool to Support the Creation of an Ontology for Dolmens.," in *XXVII Meeting of the Portuguese Association for Classification and Data Analysis, Lisboa, 22-24 Outubro.*, 2020.

[16] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007.

[17] P. Offermann, O. Levina, M. Schönherr, and U. Bub, "Outline of a design science research process," *Proc. 4th Int. Conf. Des. Sci. Res. Inf. Syst. Technol. DESRIST '09*, 2009.

[18] L. Rocha, "As origens do megalitismo funerário no Alentejo Central: a contribuição de Manuel Heleno," Universidade de Évora, 2005.

[19] A. Câmara and T. Batista, "Fotointerpretação e Sistemas de Informação Geográfica: Contributo para a Identificação de Dólmens em Portugal: O Caso de Mora e Arraiolos," *Iber. Conf. Inf. Syst. Technol. Cist.*, 2017.

[20] E. A. M. Morais and A. P. L. Ambrósio, "Ontologias: conceitos, usos, tipos, metodologias, ferramentas e linguagens," p. 21, 2007.

[21] N. Maillot, M. Thonnat, and A. Boucher, "Towards ontology based cognitive vision," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2626, pp. 44–53, 2003.

[22] D. Kirasić and D. Basch, "Ontology-based design pattern recognition," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5177 LNAI, no. PART 1, pp. 384–393, 2008.

[23] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Fully convolutional neural networks for remote sensing image classification," *Int. Geosci. Remote Sens. Symp.*, vol. 2016-Novem, no. 2, pp. 5071–5074, 2016.

[24] DigitalGlobe, "The Benefits of the Eight Spectral Bands of WorldView-2," 2013. [Online]. Available: https://dg-cms-uploads-production.s3.amazonaws.com/uploads/document/file/35/DG-8SPECTRAL-WP_0.pdf. [Accessed: 17-Dec-2019].

[25] P. Cheng, "Geometric correction and automatic DEM extraction of WorldView stereo data," *32nd Asian Conf. Remote Sens. 2011, ACRS 2011*, vol. 1, no. 1, pp. 426–431, 2011.

[26] N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," *Opt. Eng.*, vol. 52, no. 9, p. 090901, 2013.

[27] J. R. Parker, *Algorithms for image processing and computer vision*. John Wiley & Sons, 2010.

[28] D. S. Davis, "Object-based image analysis: a review of developments and future directions of automated feature detection in landscape archaeology," *Archaeol. Prospect.*, no. October 2018, 2018.

[29] M. V. Kumar and K. Yarrakula, "Comparison of efficient techniques of hyper-spectral image preprocessing for mineralogy and vegetation studies," *Indian J. geomarine Sci.*, vol. 46, no. May, pp. 1008–1021, 2017.

[30] H. K. Aggarwal and A. Majumdar, "Hyperspectral Image Denoising Using Spatio-Spectral Total Variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 442–446, 2016.

[31] Y. Zhang, "Understanding Image Fusion Existing image fusion techniques," *Photogramm. Eng. Remote Sens.*, no. June, pp. 657–661, 2004.

[32] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 746–750, 2010.

[33] J. Schiewe, "Segmentation of High-Resolution Remotely Sensed Data - Concepts , Applications and Problems," *Symp. Geospatial theory, Process. Appl.*, vol. XXXIV, p. 6, 2002.

[34] G. J. Hay, G. Castilla, M. A. Wulder, and J. R. Ruiz, "An automated object-based approach for the multiscale image segmentation of forest scenes," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 7, no. 4, pp. 339–359, 2005.

[35] G. Caspari and P. Crespo, "Convolutional neural networks for archaeological site detection – Finding 'princely' tombs," *J. Archaeol. Sci.*, 2019.

[36] K. Lambers, W. B. Verschoof-van der Vaart, and Q. P. J. Bourgeois, "Integrating remote sensing, machine learning, and citizen science in dutch archaeological prospection," *Remote Sens.*, vol. 11, no. 7, pp. 1–20, 2019.

[37] I. Zingman, D. Saupe, O. A. B. Penatti, and K. Lambers, "Detection of Fragmented Rectangular Enclosures in Very High Resolution Remote Sensing Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4580–4593, Aug. 2016.

[38] R. Lasaponara and N. Masini, "Detection of archaeological crop marks by using satellite QuickBird multispectral imagery," *J. Archaeol. Sci.*, vol. 34, no. 2, pp. 214–221, 2007.

[39] A. Sarris *et al.*, "Integration of geophysical surveys, ground hyperspectral measurements, aerial and satellite imagery for archaeological prospection of prehistoric sites: The case study of Vészto{doubleacute}-Mágor Tell, Hungary," *J. Archaeol. Sci.*, vol. 40, no. 3, pp. 1454–1470, 2013.

[40] A. Agapiou, D. Alexakis, A. Sarris, and D. G. Hadjimitsis, "Orthogonal equations of multi-spectral satellite imagery for the identification of un-excavated archaeological sites," *Remote Sens.*, vol. 5, no. 12, pp. 6560–6586, 2013.

[41] A. Agapiou, V. Lysandrou, R. Lasaponara, N. Masini, and D. G. Hadjimitsis, "Study of the variations of archaeological marks at neolithic site of lucera, Italy using high-resolution multispectral datasets," *Remote Sens.*, vol. 8, no. 9, 2016.

[42] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, 2017.

[43] A. Agapiou, "Remote sensing heritage in a petabyte-scale: satellite data and heritage Earth Engine© applications," *Int. J. Digit. Earth*, vol. 10, no. 1, pp. 85–102, 2017.

[44] L. Luo *et al.*, "Automated extraction of the archaeological tops of qanat shafts from VHR imagery in Google Earth," *Remote Sens.*, vol. 6, no. 12, pp. 11956–11976, 2014.

[45] S. Ghaffarian and S. Ghaffarian, "Automatic building detection based on supervised classification using high resolution Google Earth images," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 40, no. 3, pp. 101–106, 2014.

[46] D. Cerra, A. Agapiou, R. M. Cavalli, and A. Sarris, "An objective assessment of hyperspectral indicators for the detection of buried archaeological relics," *Remote Sens.*, vol. 10, no. 4, pp. 1–25, 2018.

[47] DigitalGlobe, "Standard Imagery." [Online]. Available: https://earth.esa.int/documents/10174/2525050/DigitalGlobe-Standard-Imagery.pdf. [Accessed: 17-Dec-2019].

[48] T. Updike and C. Comp, "Radiometric Use of WorldView-2 Imagery Technical Note," *DigitalGlobe*, no. November, pp. 1–17, 2010.

[49] P. Ganesan and V. Rajini, "Assessment of satellite image segmentation in RGB and HSV color space using image quality measures," *2014 Int. Conf. Adv. Electr. Eng. ICAEE 2014*, 2014.

[50] J. Q. Ma, "Content-based image retrieval with HSV color space and texture features," *2009 Int. Conf. Web Inf. Syst. Mining, WISM 2009*, no. 3, pp. 61–63, 2009.

[51] K. Singh and R. Kapoor, "Image enhancement using Exposure based Sub Image Histogram Equalization," *Pattern Recognit. Lett.*, vol. 36, no. 1, pp. 10–14, 2014.

[52] C. H. Ooi and N. A. M. Isa, "Quadrants dynamic histogram equalization for contrast enhancement," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2552–2559, 2010.

[53] C. Lee, C. Le, Y. Y. Lee, and C. S. Kim, "Power-constrained contrast enhancement for emissive displays based on histogram equalization," *IEEE*

Trans. Image Process., vol. 21, no. 1, pp. 80–93, 2012.

[54] A. Saxena, "An Exhaustive Analysis On Various Foggy Image Enhancement Techniques.," vol. 3, no. 1, pp. 11–17, 2014.

[55] S. Chib and M. S. Devi, "Performance Analysis of Enhancement Techniques for Satellite Images," *Int. J. Comput. Sci. Eng.*, vol. Vol.-4, no. 12, pp. 113–119, 2016.

[56] A. K. Bhandari, M. Gadde, A. Kumar, and G. K. Singh, "Comparative analysis of different wavelet filters for low contrast and brightness enhancement of multispectral remote sensing images," *2012 Int. Conf. Mach. Vis. Image Process. MVIP 2012*, no. i, pp. 81–86, 2012.

[57] A. K. Bhandari, A. Kumar, and G. K. Singh, "SVD based poor contrast improvement of blurred multispectral remote sensing satellite images," *Proc. 2012 3rd Int. Conf. Comput. Commun. Technol. ICCCT 2012*, pp. 156–159, 2012.

[58] D. Sheet, H. Garud, A. Suveer, M. Mahadevappa, and J. Chatterjee, "Brightness preserving dynamic fuzzy histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2475–2480, 2010.

[59] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *J. VLSI Signal Process. Syst. Signal Image. Video Technol.*, vol. 38, no. 1, pp. 35–44, 2004.

[60] V. R. Ganesh and H. Ramesh, "Effectiveness of contrast limited adaptive histogram equalization technique on multispectral satellite imagery," *ACM Int. Conf. Proceeding Ser.*, pp. 234–239, 2017.

[61] The MathWorks Inc., "Mathworks," 2020. [Online]. Available: https://www.mathworks.com/. [Accessed: 11-Mar-2020].

[62] N. Sharma and O. P. Verma, "Gamma correction based satellite image enhancement using singular value decomposition and discrete wavelet transform," *Proc. 2014 IEEE Int. Conf. Adv. Commun. Control Comput. Technol. ICACCCT 2014*, no. 978, pp. 1286–1289, 2015.

[63] A. K. Bhandari, A. Kumar, and G. K. Singh, "Improved knee transfer function and gamma correction based method for contrast and brightness enhancement of satellite image," *AEU - Int. J. Electron. Commun.*, vol. 69, no. 2, pp. 579–589, 2015.

[64] Brockmann Consult *et al.*, "CHRIS/Proba Toolbox for BEAM (CHRIS-Box)." [Online]. Available: http://www.brockmann-consult.de/beam/chris-box/theory-details.html. [Accessed: 03-Apr-2020].

[65] L. Gómez-Chova, L. Alonso, L. Guanter, C. V. Gustavo, J. Calpe, and J. Moreno, "Correction of systematic spatial noise in push-broom hyperspectral sensors: Application to CHRIS/PROBA images," *Appl. Opt.*, vol. 47, no. 28, 2008.

[66] G. Thuillier *et al.*, "THE SOLAR SPECTRAL IRRADIANCE FROM 200 TO

2400 nm AS MEASURED BY THE SOLSPEC SPECTROMETER FROM THE ATLAS AND EURECA MISSIONS," *Sol. Phys.*, vol. 214, pp. 1–22, 2003.

[67] A. K. Mahlein *et al.*, "Development of spectral indices for detecting and identifying plant diseases," *Remote Sens. Environ.*, vol. 128, pp. 21–30, 2013.

[68] V. Henrich, G. Krauss, G. Christian, and C. Sandow, "Index Database," 2012. [Online]. Available: https://www.indexdatabase.de/. [Accessed: 19-Feb-2020].

[69] C. Xie, X. Huang, W. Zeng, and X. Fang, "A novel water index for urban high-resolution eight-band WorldView-2 imagery," *Int. J. Digit. Earth*, vol. 9, no. 10, pp. 925–941, 2016.

[70] P. Maglione, C. Parente, and A. Vallario, "Coastline extraction using high resolution WorldView-2 satellite imagery," *Eur. J. Remote Sens.*, vol. 47, no. 1, pp. 685–699, 2014.

[71] K. Shahi, H. Z. M. Shafri, E. Taherzadeh, and A. S. Area, "A Novel Spectral Index for Automatic Shadow Detection in Urban Mapping Based On WorldView-2 Satellite Imagery," *Int. J. Comput. Electr. Autom. Control Inf. Eng.*, vol. 8, no. 10, pp. 1482–1485, 2014.

[72] A. Wolf, "Using WorldView 2 Vis-NIR MSI Imagery to Support Land Mapping and Feature Extraction Using Normalized Difference Index Ratios.," *Unpubl. report, Longmont, CO Digit.*, 2010.

[73] Y. Deng, C. Wu, M. Li, and R. Chen, "RNDSI: A ratio normalized difference soil index for remote sensing of urban/suburban environments," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 39, pp. 40–48, 2015.

[74] P. Mhangara, J. Odindi, L. Kleyn, and H. Remas, "Road extraction using object oriented classification."

[75] K. Shahi, H. Z. M. Shafri, E. Taherzadeh, S. Mansor, and R. Muniandy, "A novel spectral index to automatically extract road networks from WorldView-2 satellite imagery," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 18, no. 1, pp. 27–33, 2015.

[76] M. Choi, "A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1672–1682, 2006.

[77] R. Gharbia, A. H. El Baz, A. E. Hassanien, and M. F. Tolba, "Remote Sensing Image Fusion Approach Based on Brovey and Wavelets Transforms," in *Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014*, 2014, pp. 311–321.

[78] W. M. K. W. M. Khairosfaizal and A. J. Nor'aini, "Eyes detection in facial images using circular hough transform," *Proc. 2009 5th Int. Colloq. Signal Process. Its Appl. CSPA 2009*, pp. 238–242, 2009.

[79] X. Peng, L. Wang, X. Wang, and Y. Qiao, "Bag of visual words and fusion methods for action recognition: Comprehensive study and good practice," *Comput. Vis. Image Underst.*, vol. 150, pp. 109–125, 2016.

[80] S. Xu, T. Fang, D. Li, and S. Wang, "Object classification of aerial images with bag-of-visual words," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 366–370, 2010.

[81] P. Tsangaratos and I. Ilia, "Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size," *Catena*, vol. 145, pp. 164–179, 2016.

[82] R. A. Nugrahaeni and K. Mutijarsa, "Comparative Analysis of Machine Learning KNN, SVM, and Random Forests Algorithm for Facial Expression Classification," pp. 163–168, 2016.

[83] M. J. Islam, Q. M. J. Wu, M. Ahmadi, and M. A. Sid-Ahmed, "Investigating the Performance of Naive- Bayes Classifiers and K- Nearest Neighbor Classifiers," in *2007 International Conference on Convergence Information Technology (ICCIT 2007)*, 2007, pp. 1541–1546.

[84] M. Li, L. Ma, T. Blaschke, L. Cheng, and D. Tiede, "A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 49, pp. 87–98, 2016.

[85] E. Taherzadeh and H. Z. M. Shafri, "Development of a Generic Model for the Detection of Roof Materials Based on an Object-Based Approach Using WorldView-2 Satellite Imagery," *Adv. Remote Sens.*, vol. 02, no. 04, pp. 312–321, 2013.

[86] M. Pal and P. M. Mather, "Decision tree based classification of remotely sensed data," *22nd Asian Conf. Remote Sens.*, vol. 5, no. November, pp. 5–9, 2001.

[87] M. Pal, "Ensemble Learning with Decision Tree for Remote Sensing Classification," *Proc. World Acad. Sci. Eng. Technol. Vol 26 Parts 1 2 December 2007*, vol. 26, no. December, pp. 735–737, 2007.

[88] M. Xu, P. Watanachaturaporn, P. K. Varshney, and M. K. Arora, "Decision tree regression for soft classification of remote sensing data," *Remote Sens. Environ.*, vol. 97, no. 3, pp. 322–336, 2005.

[89] A. M. P. J. Ferreira, "Dados Geoquímicos de Base de Sedimentos Fluviais de Amostragem de Baixa Densidade de Portugal Continental: Estudo de Factores de Variação Regional," Universidade de Aveiro, 2000.

[90] M. Cutter and L. Johns, "Chris data products - Latest issue," *Eur. Sp. Agency, (Special Publ. ESA SP*, no. 593, pp. 1–6, 2005.

[91] Surrey Satellite Tecnhology Limited, "CHRIS Data Format," 2008. [Online]. Available: https://earth.esa.int/c/document_library/get_file?folderId=23844&name=DLFE-592.pdf. [Accessed: 16-Jan-2020].

[92] M. Cutter, "Review of aspects associated with the CHRIS calibration," *Eur. Sp. Agency, (Special Publ. ESA SP*, no. 578, pp. 1–5, 2004.

[93] J. A. Fitzsimons and D. R. Michael, “Rocky outcrops: A hard road in the conservation of critical habitats,” *Biol. Conserv.*, vol. 211, pp. 36–44, 2017.

Annexes e Appendixes

Annex A

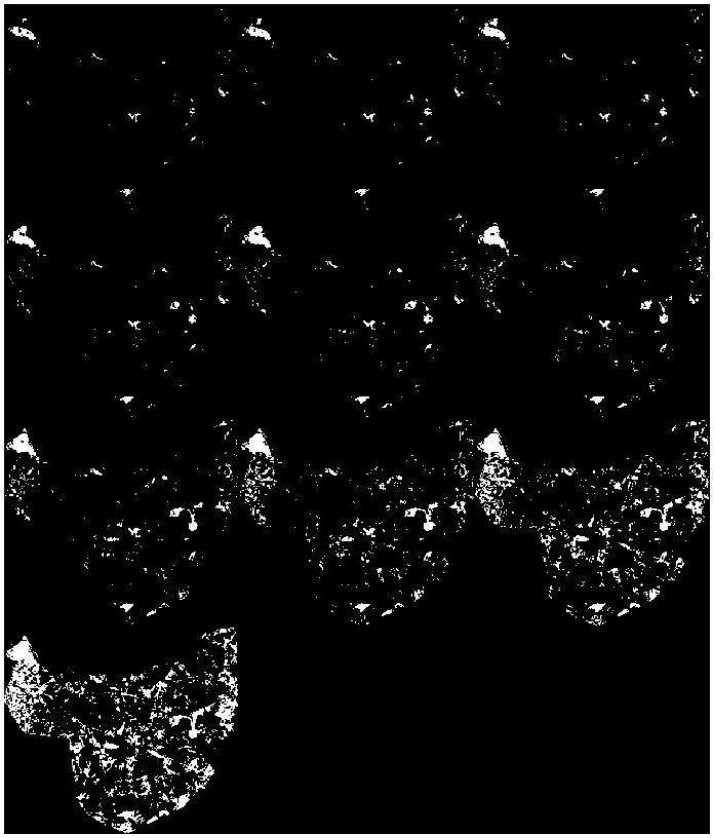
Proba-1 CHRIS satellite characteristics of the wavelengths of each band for the 37-band image Source: [91].

Band	Minimum (nm)	Middle (nm)	Maximum (nm)	Width (nm)
H1	438	442	447	9
H2	486	489	495	9
H3	526	530	534	9
H4	546	551	556	10
H5	566	570	573	8
H6	627	631	636	9
H7	656	661	666	11
H8	666	672	677	11
H9	677	683	689	11
H10	694	697	700	6
H11	700	703	706	6
H12	706	709	712	6
H13	712	716	719	6
H14	719	722	725	6
H15	725	728	732	7
H16	732	735	738	7
H17	738	742	745	7
H18	745	748	752	7
H19	752	755	759	7
H20	759	762	766	7
H21	766	770	773	7
H22	773	777	788	15
H23	788	792	796	8
H24	796	800	804	8
H25	863	872	881	18
H26	881	886	891	10
H27	891	895	900	10
H28	900	905	910	10
H29	910	915	920	10
H30	920	925	930	10
H31	930	940	950	20
H32	950	955	960	10
H33	960	965	971	11
H34	971	976	981	11
H35	981	987	992	11
H36	992	997	1003	11
H37	1003	1019	1036	33

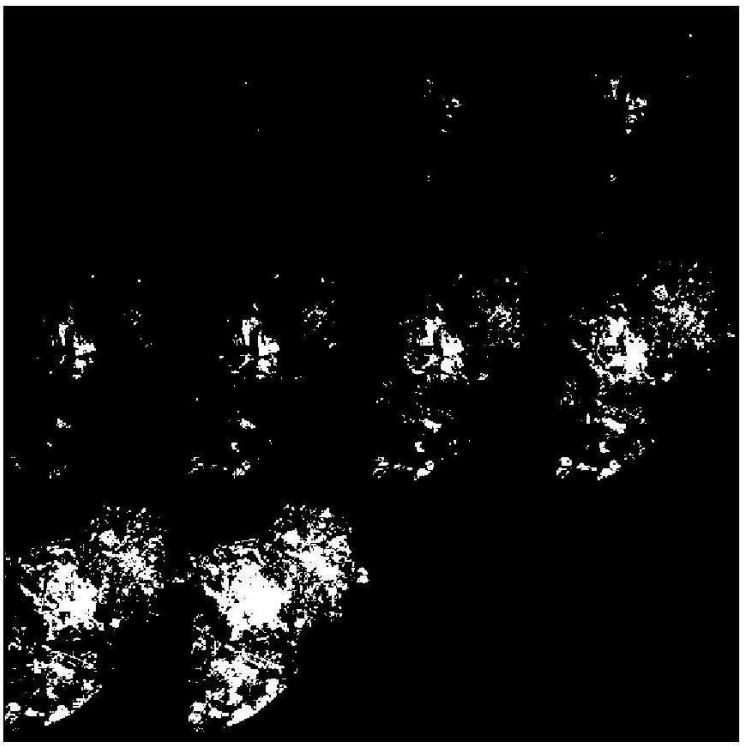
Appendix A

Band Histogram for each of the MS image, on the left being for the second set while the right being for the first set of images.

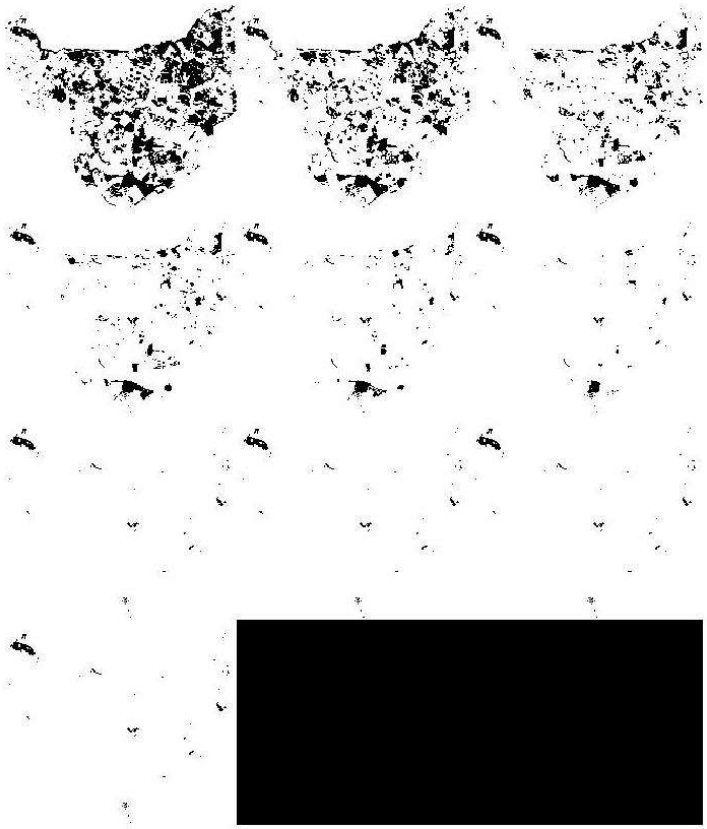
Appendix B



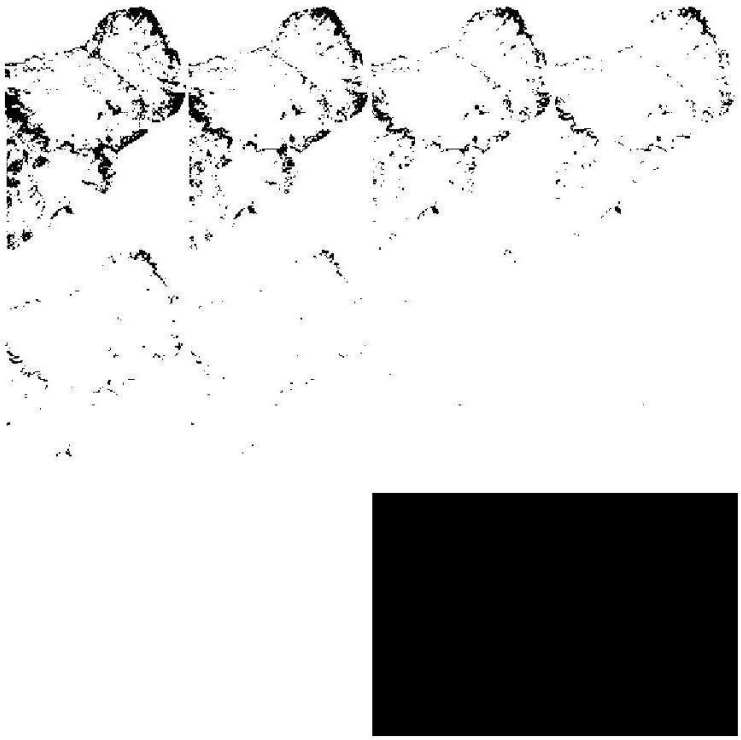
Atmospherically Resistant Vegetation Index 2 applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



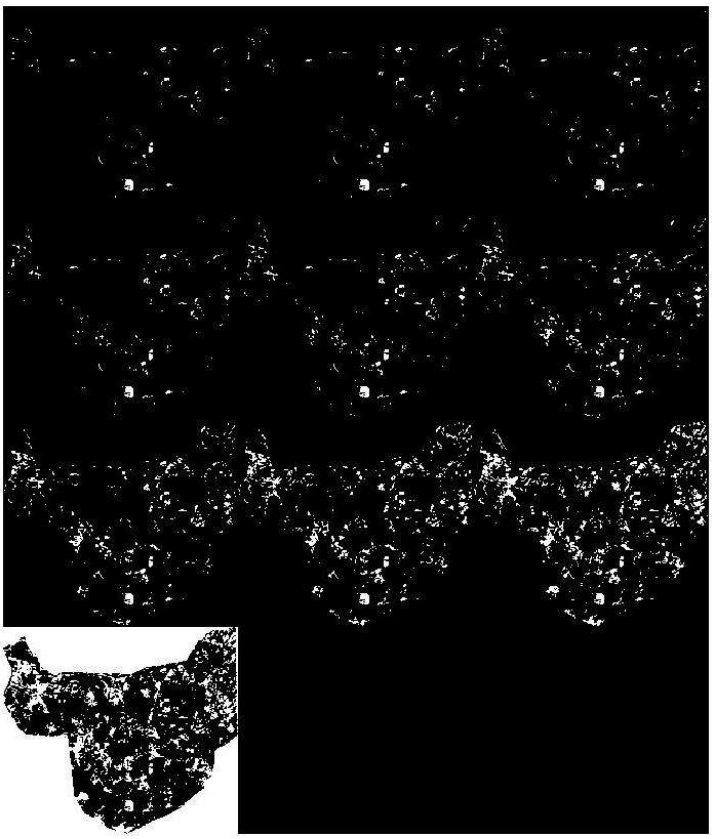
Atmospherically Resistant Vegetation Index 2 applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



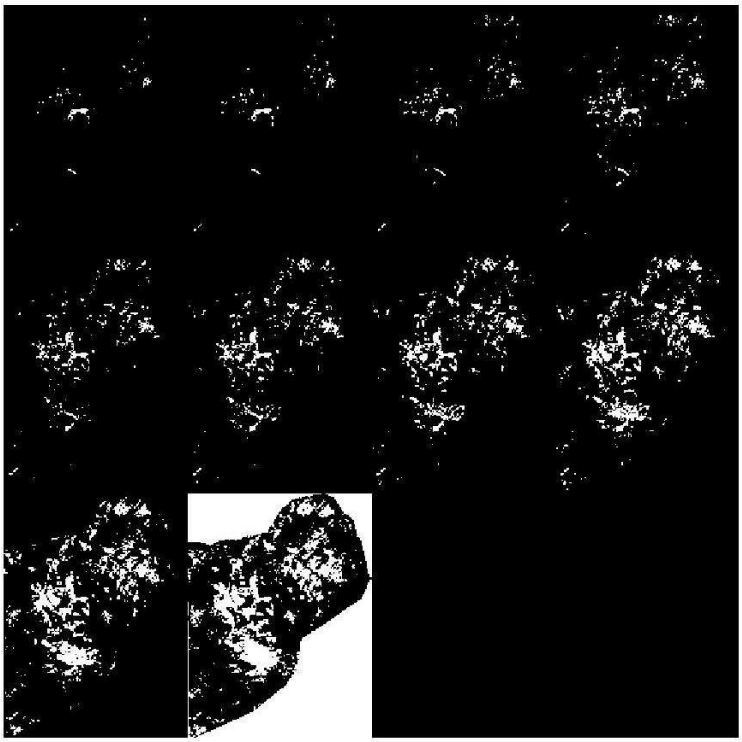
Difference Vegetation Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



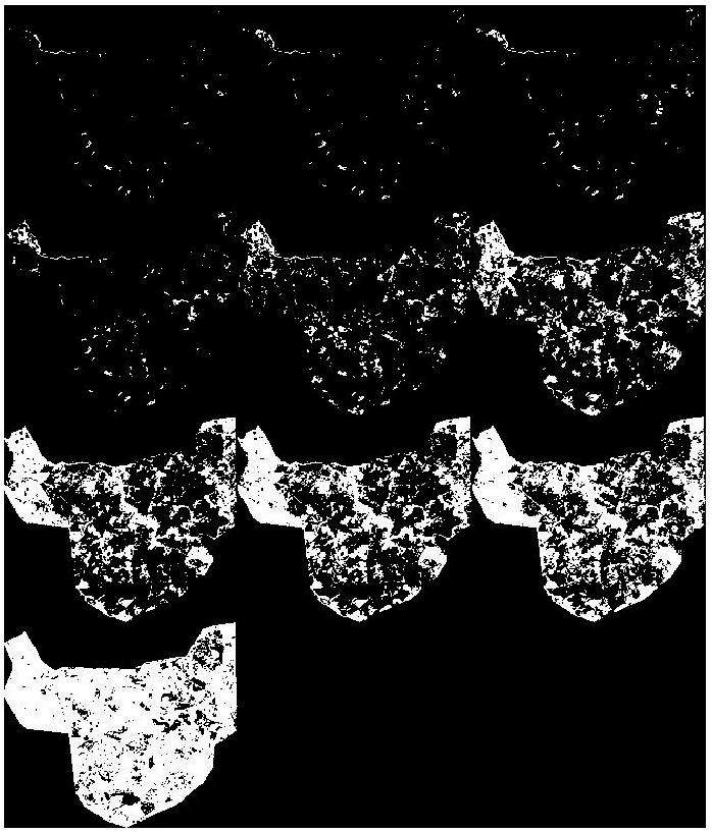
Difference Vegetation Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



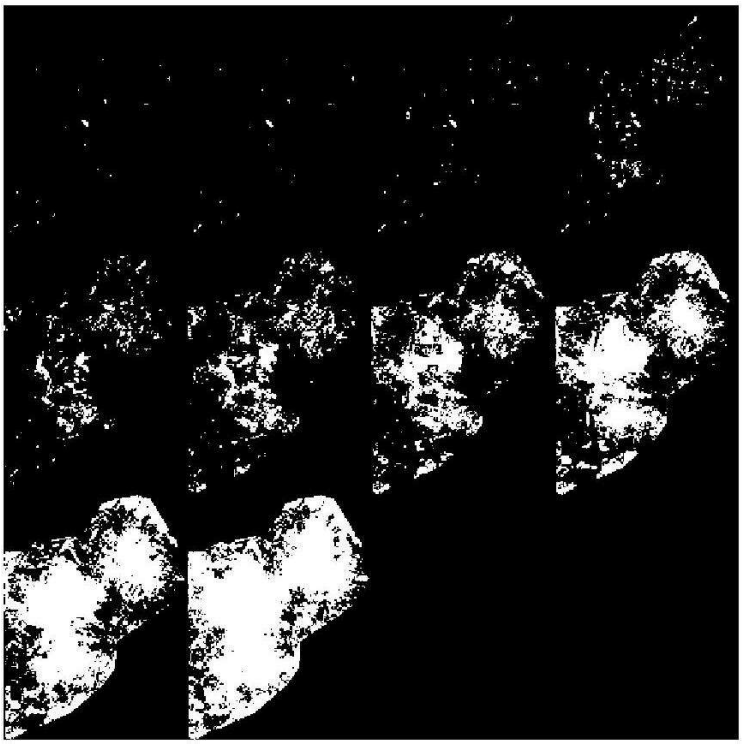
Enhanced Vegetation Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



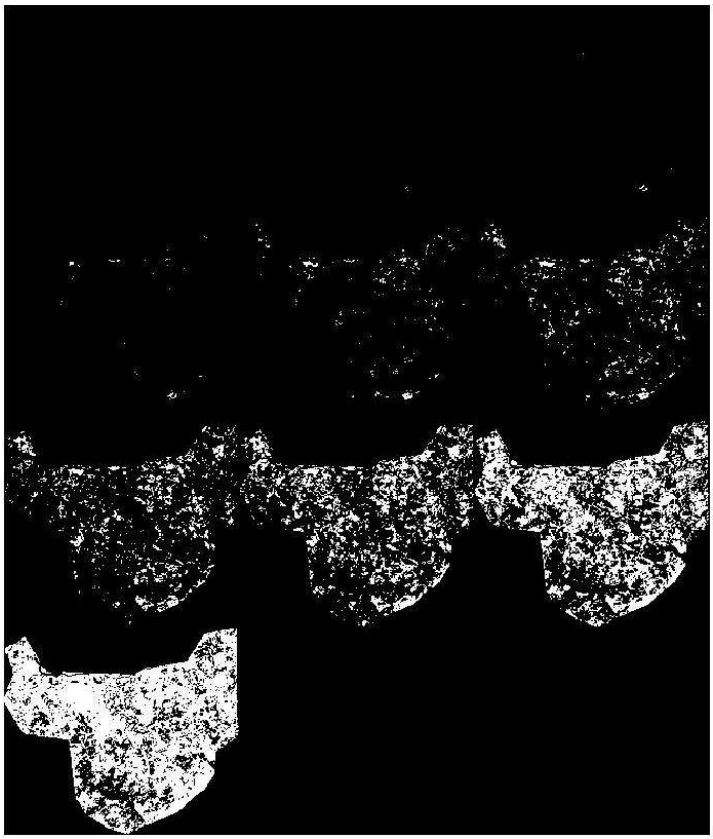
Enhanced Vegetation Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



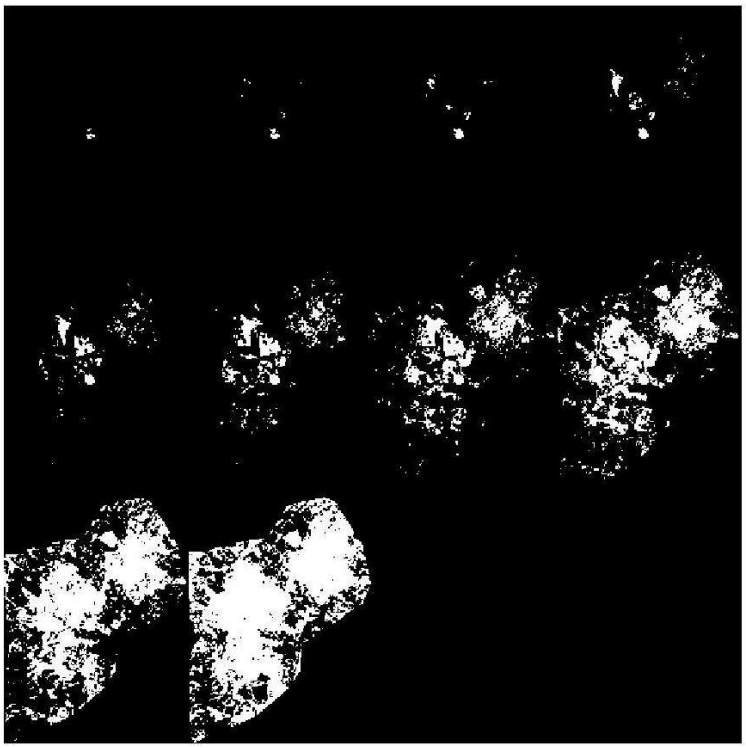
Green Normalized Difference Vegetation Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



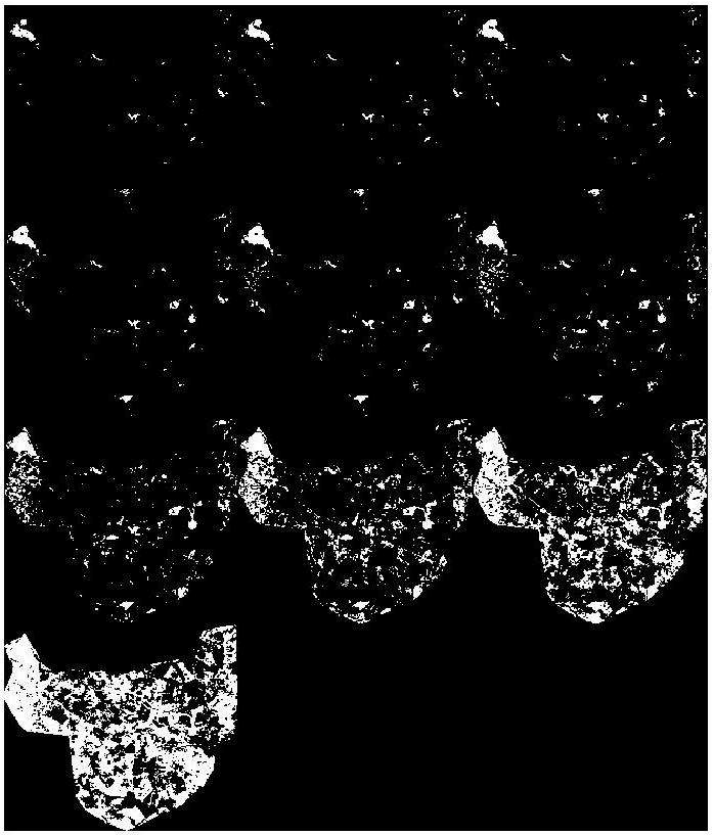
Green Normalized Difference Vegetation Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



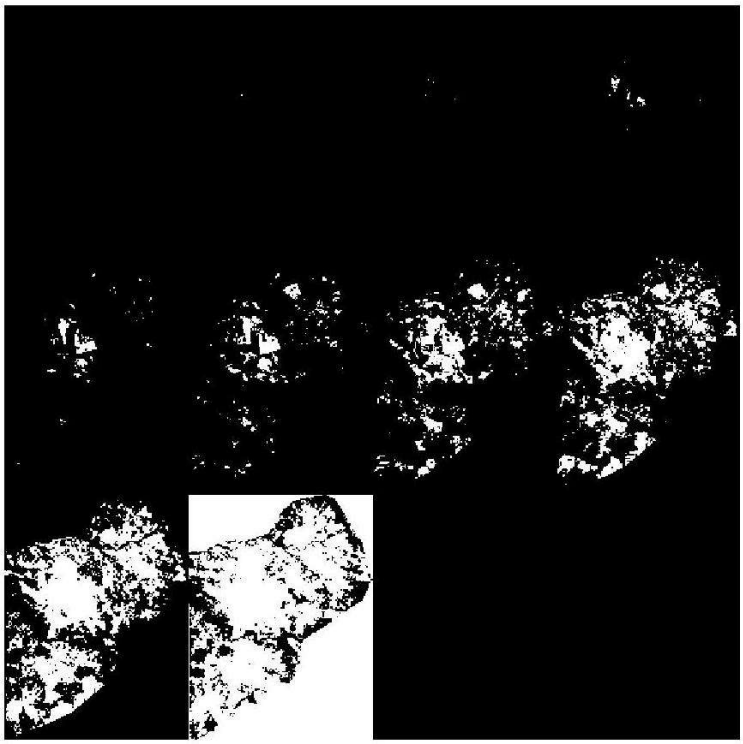
Normalized Difference Chlorophyll Vegetation Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



Normalized Difference Chlorophyll Vegetation Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.

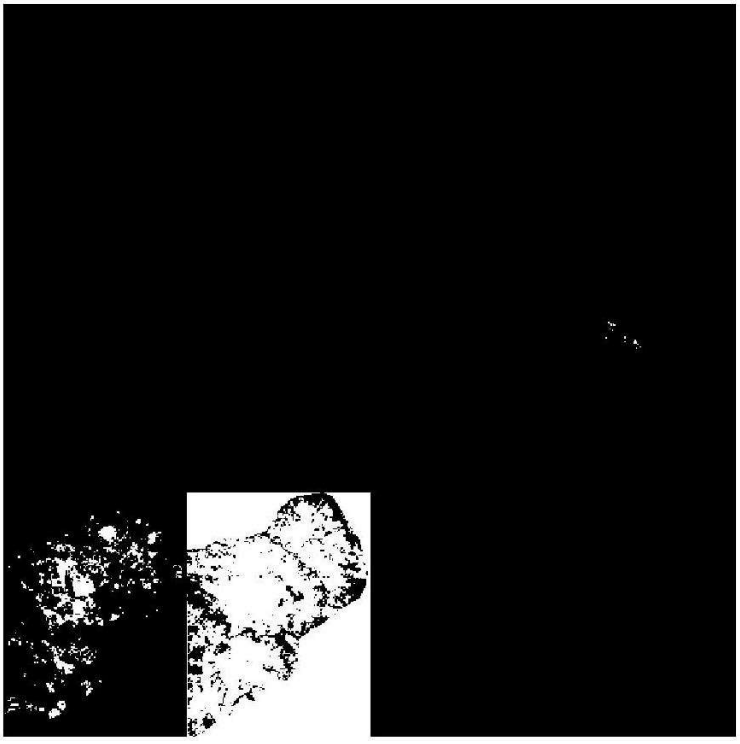


Red Edge Normalized Difference Vegetation Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.

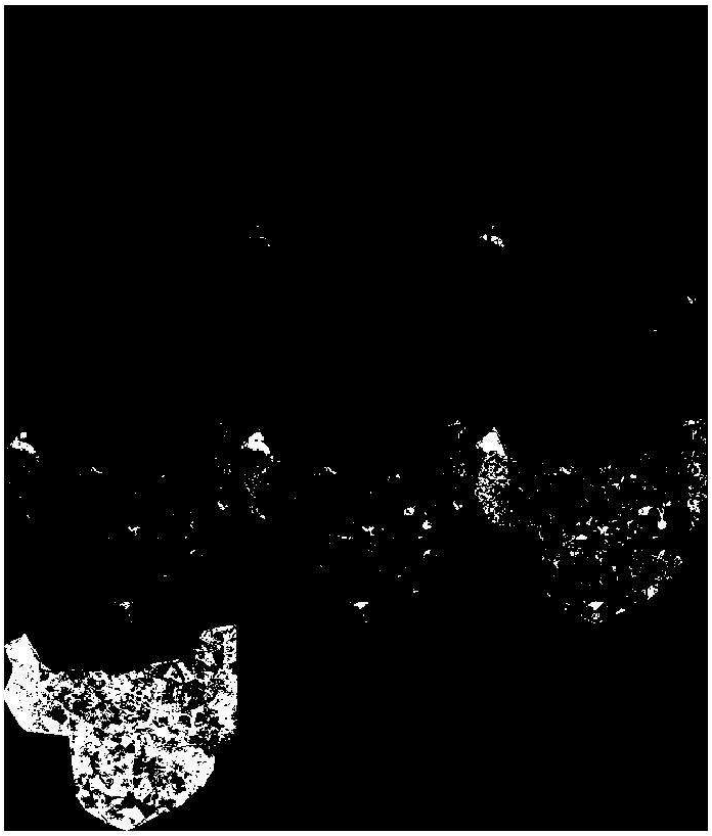


Red Edge Normalized Difference Vegetation Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.

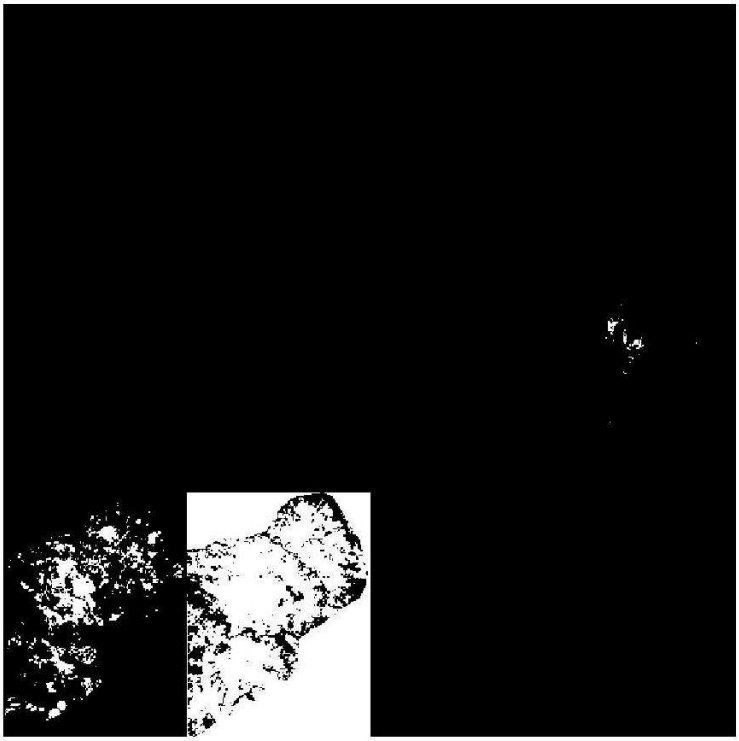
Renormalized Difference Vegetation Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



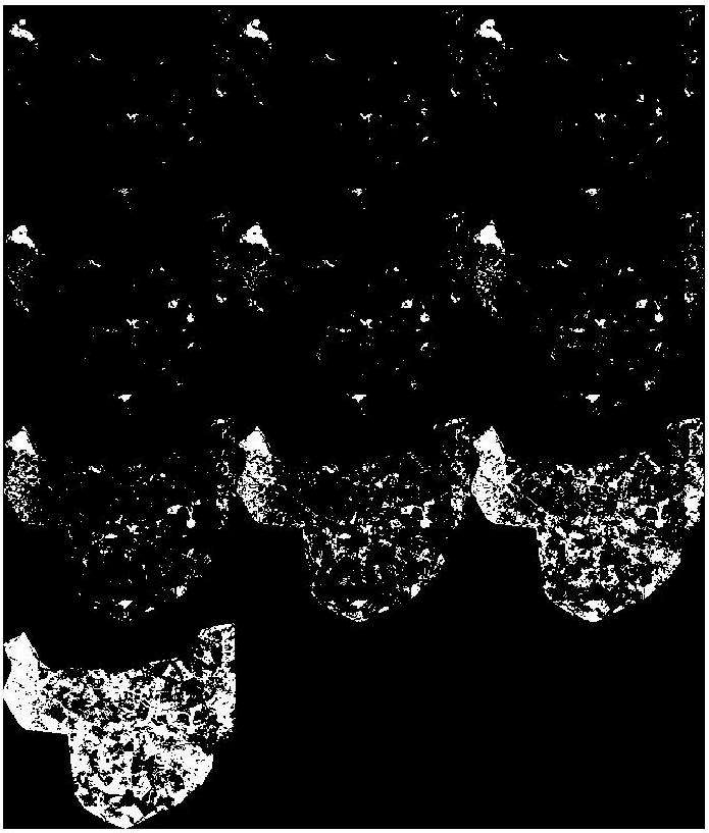
Renormalized Difference Vegetation Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



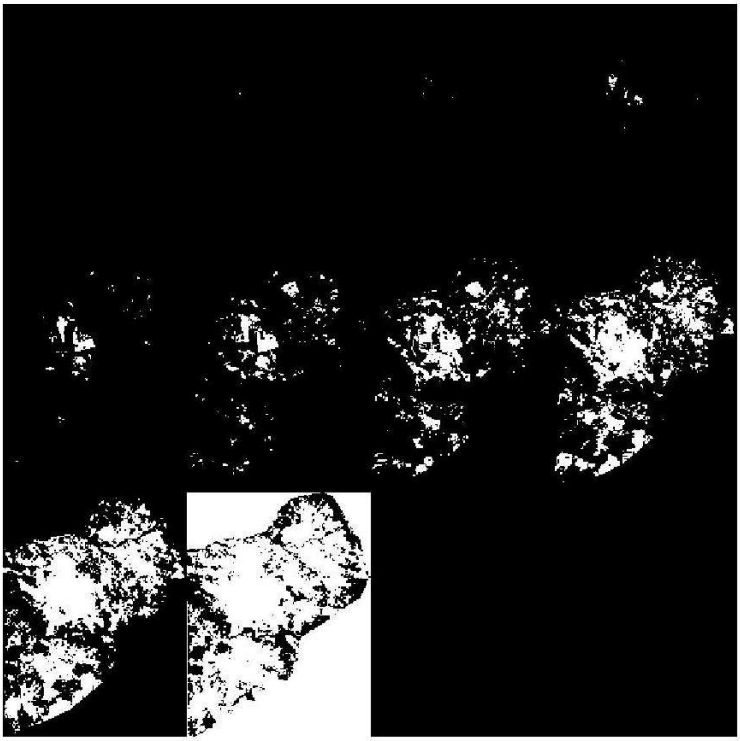
Soil Adjusted Vegetation Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



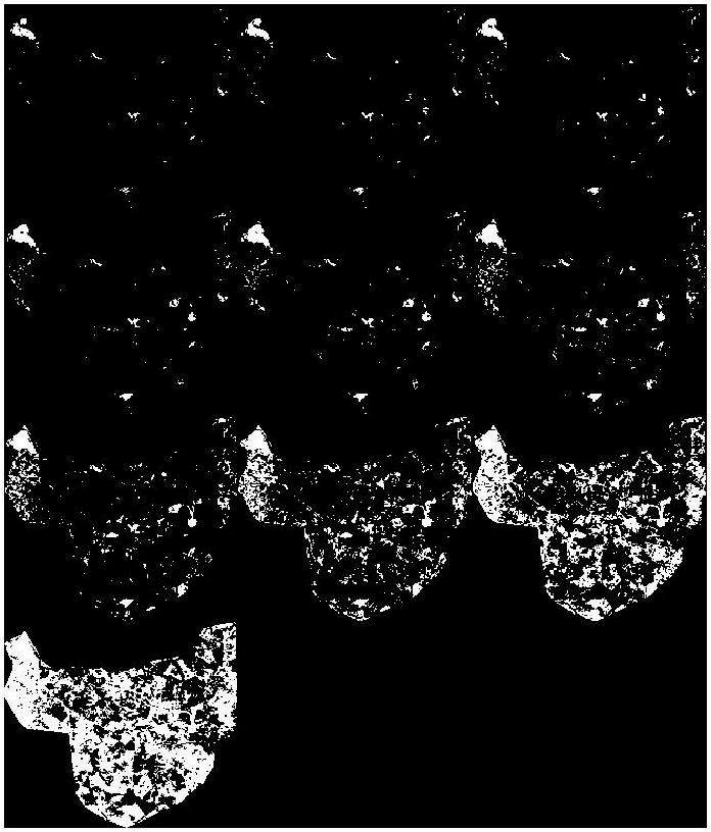
Soil Adjusted Vegetation Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



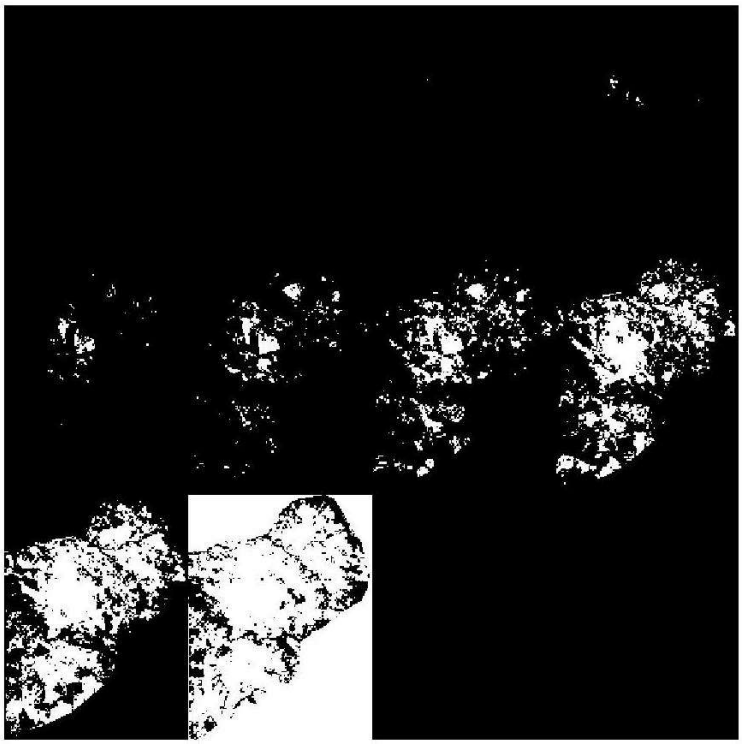
Wide Dynamic Range Vegetation Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



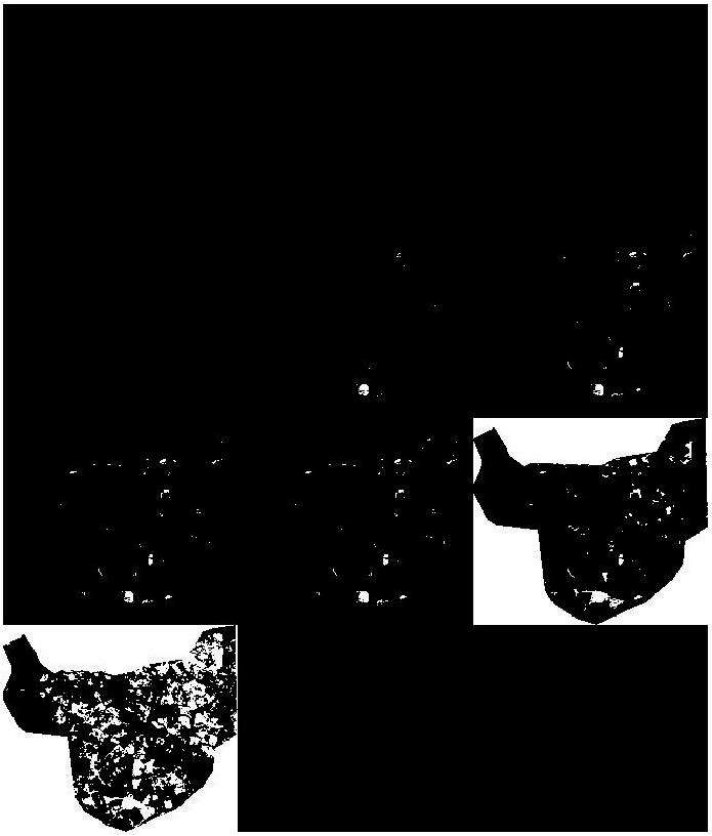
Wide Dynamic Range Vegetation Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1



NDVI Vegetation Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.

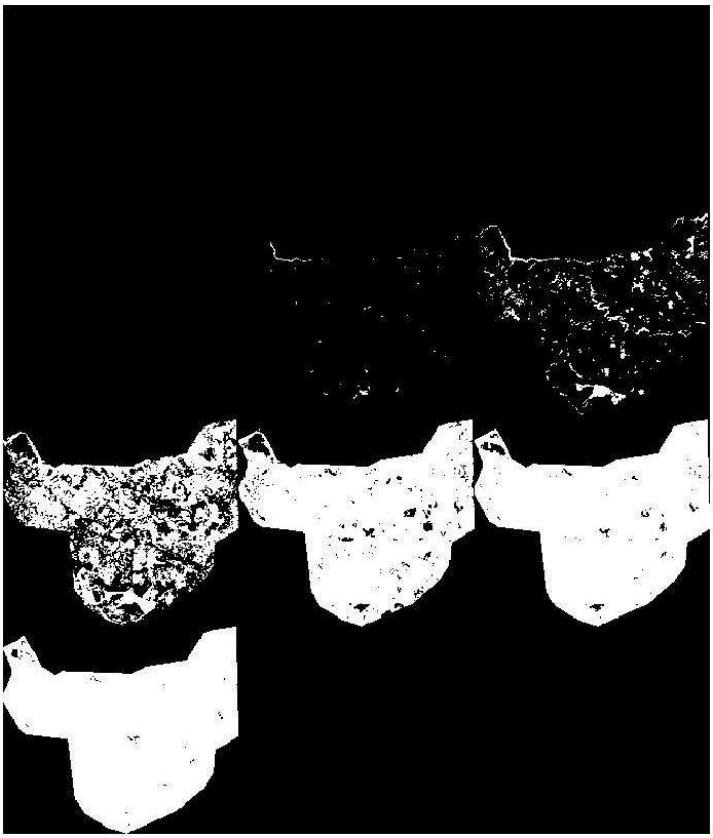


NDVI Vegetation Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.

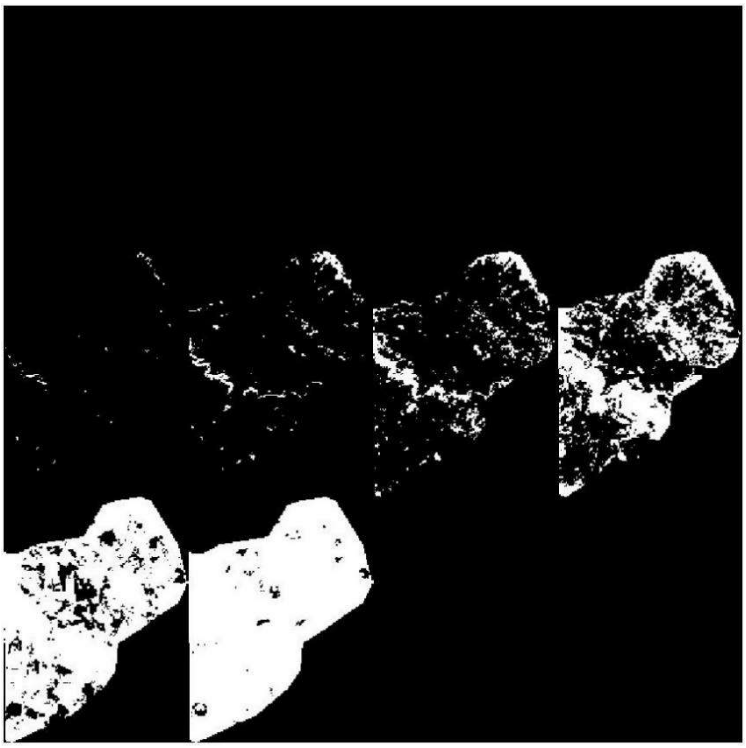


NDWI Water Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.

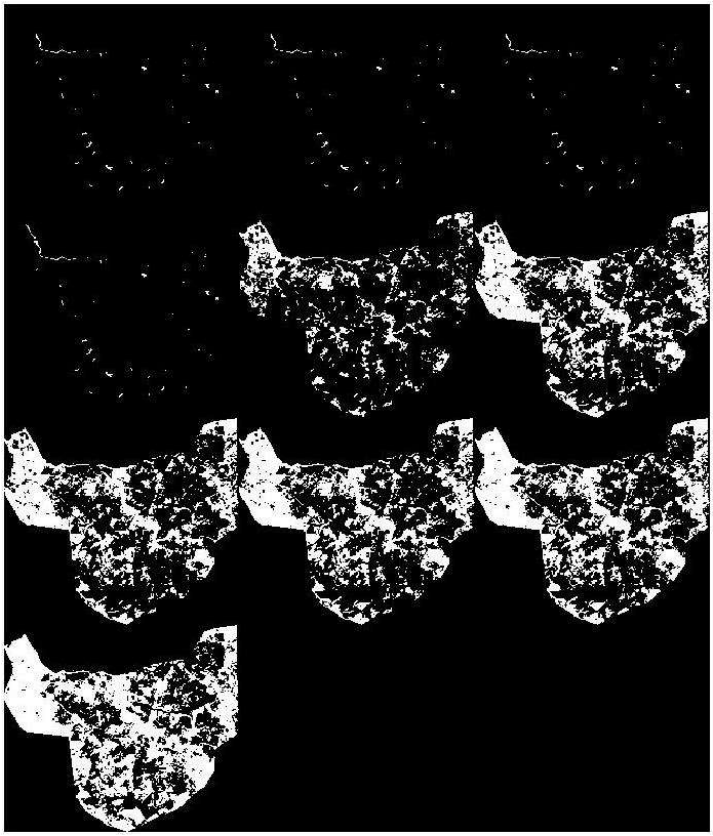
NDWI Water Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



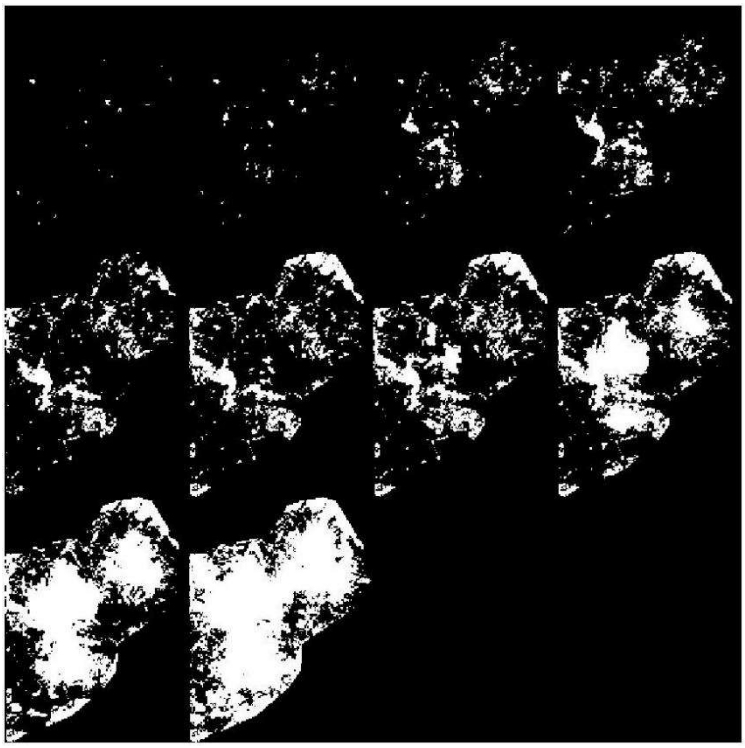
NDSI Soil Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



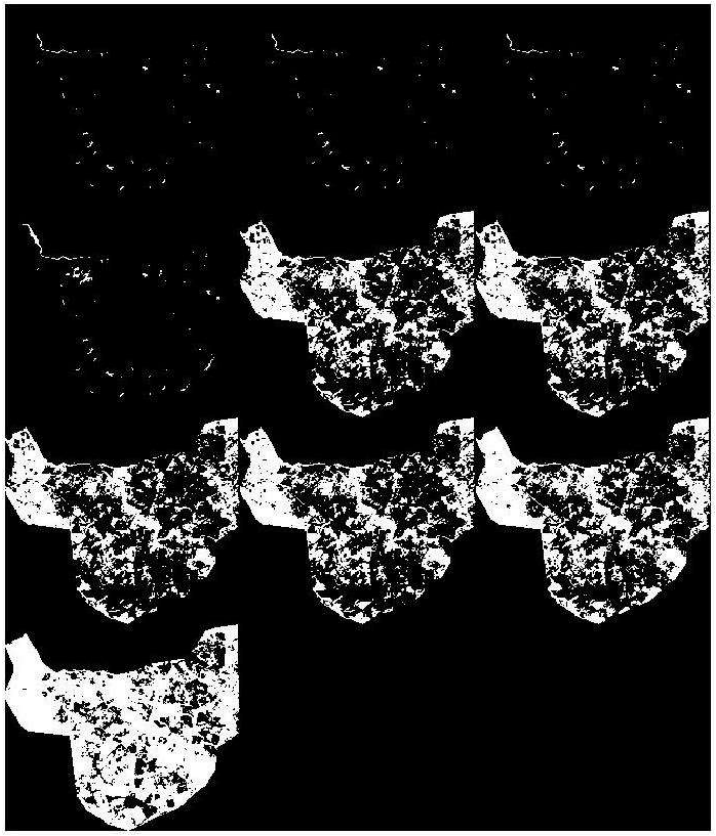
NDSI Soil Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



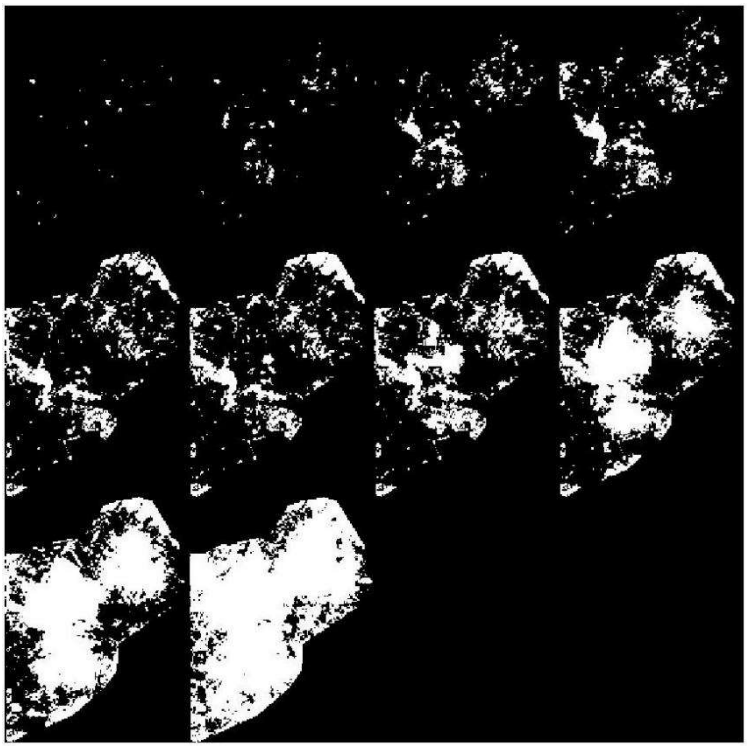
SDI Shadow Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



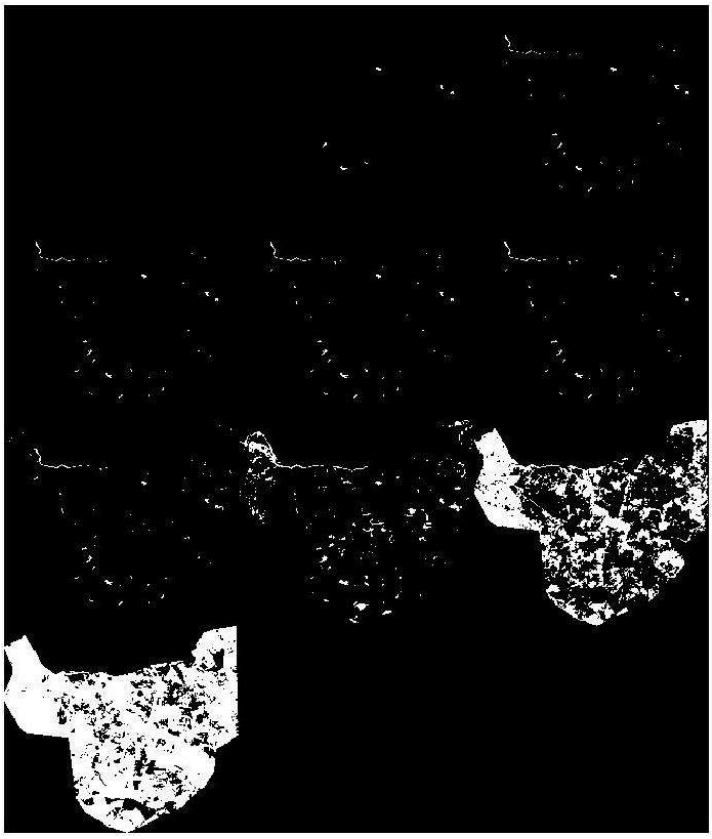
SDI Shadow Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



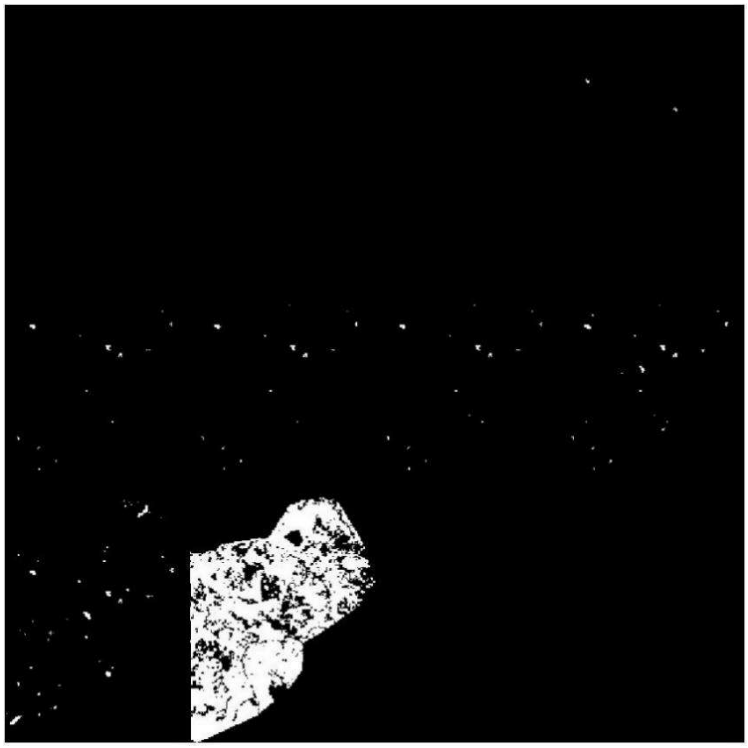
NHFD Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



NHFD Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



BAI Index applied to the image of the second set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.



BAI Index applied to the image of the first set, for values superior to a threshold ranging from 0.9 to 0 with decrements of 0.1.