



Article

Genetic Programming Approach for the Detection of Mistletoe Based on UAV Multispectral Imagery in the Conservation Area of Mexico City

Paola Andrea Mejia-Zuluaga ^{1,†} , León Dozal ^{2,†} and Juan C. Valdiviezo-N. ^{3,*}

¹ Centro de Investigación en Ciencias de Información Geoespacial (CentroGeo), Ciudad de México 14240, Mexico; paomejia23@gmail.com

² CONACYT—CentroGeo, Aguascalientes 20313, Mexico; leon.dozal@gmail.com

³ CONACYT—CentroGeo, Yucatán 97302, Mexico

* Correspondence: jcvaldiviezon@gmail.com

† These authors contributed equally to this work.

Abstract: The mistletoe *Phoradendron velutinum* (*P. velutinum*) is a pest that spreads rapidly and uncontrollably in Mexican forests, becoming a serious problem since it is a cause of the decline of 23.3 million hectares of conifers and broadleaves in the country. The lack of adequate phytosanitary control has negative social, economic, and environmental impacts. However, pest management is a challenging task due to the difficulty of early detection for proper control of mistletoe infestations. Automating the detection of this pest is important due to its rapid spread and the high costs of field identification tasks. This paper presents a Genetic Programming (GP) approach for the automatic design of an algorithm to detect mistletoe using multispectral aerial images. Our study area is located in a conservation area of Mexico City, in the San Bartolo Ameyalco community. Images of 148 hectares were acquired by means of an Unmanned Aerial Vehicle (UAV) carrying a sensor sensitive to the *R*, *G*, *B*, red edge, and near-infrared bands, and with an average spatial resolution of less than 10 cm per pixel. As a result, it was possible to obtain an algorithm capable of classifying mistletoe *P. velutinum* at its flowering stage for the specific case of the study area in conservation area with an Overall Accuracy (OA) of 96% and a value of fitness function based on weighted Cohen's Kappa (k_w) equal to 0.45 in the test data set. Additionally, our method's performance was compared with two traditional image classification methods; in the first, a classical spectral index, named Intensive Pigment Index of Structure 2 (SIPI2), was considered for the detection of *P. velutinum*. The second method considers the well-known Support Vector Machine classification algorithm (SVM). We also compare the accuracy of the best GP individual with two additional indices obtained during the solution analysis. According to our experimental results, our GP-based algorithm outperforms the results obtained by the aforementioned methods for the identification of *P. velutinum*.

Keywords: evolutionary computation; image detection; forest pest; supervised learning; vegetation index; computer vision; change detection



Citation: Mejia-Zuluaga, P.A.; Dozal, L.; Valdiviezo-N., J.C. Genetic Programming Approach for the Detection of Mistletoe Based on UAV Multispectral Imagery in the Conservation Area of Mexico City. *Remote Sens.* **2022**, *14*, 801. <https://doi.org/10.3390/rs14030801>

Academic Editor: Michael Sprintsin

Received: 1 December 2021

Accepted: 3 February 2022

Published: 8 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Phoradendron velutinum, also known as *P. velutinum* or true mistletoe, is a hemiparasitic shrub plant of aerial parts of trees and shrubs that usually attaches itself to broadleaf and mixed conifer forests. Since *P. velutinum* spreads itself by employing a ballistic propulsion method, the detection of new shoots is difficult, causing the decline of around 23.3 million hectares of conifers and broadleaves in Mexico. Hence, efficient identification of plant infestation usually requires manual inspection, becoming a labor-intensive task [1].

One of the regions affected by this pest is the conservation area located in the San Bartolo Ameyalco Community, south of Mexico City; there, the uncontrollable presence of true mistletoe has caused different social and environmental issues. For example, the

abundant presence of mistletoe in combination with other stress factors can reduce the vitality of trees and forest mortality could precede an ecological imbalance caused by the decrease of both carbon sequestration and the recharge of the aquifer that depends on this forest area [1]. This region is also socially and economically impacted as it is the conservation land that contributes the greatest amount of water and other environmental services to the Valley of Mexico [2]. In recent years, different methodologies focused on the detection of pests and diseases in plants have been proposed based on aerial images collected by Unmanned Aerial Vehicles (UAVs) and machine learning algorithms [3–5]. For example, the authors of [6] proposed a deep-learning-based method of leaf disease detection in a variety of plants. In the case of forest pests, studies exist on the detection of bark beetle in both mixed forests [7] and urban forest at individual tree level [8]. Additionally, the spectral detection of *Dendrolimus tabulaeformis* is presented in a work by Zhang, N. et al. [9].

In relation to mistletoe, a few studies have been carried out regarding its autonomous detection in different environments, some of which are described next. The first proposals aimed at mistletoe identification in fir trees made use of a hyperspectral line scanner and supervised classification techniques [10,11]. Later, based on the discovery that mistletoe species show a lower temperature compared to their hosts as they transpire more water, Maes W. et al. in [12] employed UAV-collected RGB and thermal image sets, then the temperature difference was used to indicate the presence of mistletoe in eucalyptus trees. Recently, authors in [13] presented preliminary results on the capabilities of multispectral aerial images in conjunction with different types of Convolutional Neural Networks (CNN) for the detection of mistletoe. Additionally, the automatic detection of dwarf mistletoe, also known as *Arceuthobium globesum*, present in the Nevado de Toluca volcano (Mexico) was reported in [14]; for that purpose, UAV-collected RGB images were analyzed by means of the k-nearest neighbors technique and colorimetry algorithms. The authors concluded that a spatial resolution greater than or equal to 10 cm/px is required for the automatic detection of forest pests based on aerial images and supervised learning. Recently, the authors of [15] employed an UAV to collect series of high-resolution RGB images, from which a canopy height model (CHM) was derived; individual tree crowns were delineated manually and automatically based on the CHM. Then, a set of vegetation indices were computed and a random forest classification scheme was used to delineate individual trees infected by mistletoe (*Viscum album*) in mixed broadleaf forests located in Iran. On the other hand, a first approach to the study of *P. velutinum* employing evolutionary computation was presented by the authors in [16], proposing the development of a spectral index to detect the parasitic plant, considering algebraic and neighborhood functions; In this work, the feasibility of using evolutionary methods for the detection of forest patterns in images was analyzed as providing 89% accuracy.

Although the aforementioned research provides partial solutions for automatic mistletoe detection and classification using remote sensing techniques, the following issues require special attention. (1) A thousand varieties of mistletoe exist, all negatively impacting on host tree physiology and health, soil nutrient cycling, and stand dynamics [1,13]; (2) some mistletoe species show very similar color and texture features compared with the leaves of the host plant [13]; (3) in most cases, detection at an early stage is difficult. Therefore, the development of efficient methodologies providing early detection of some mistletoe species for adequate control has become necessary.

This paper proposes the use of an Evolutionary Computing technique, known as Genetic Programming (GP), in conjunction with UAV multispectral imagery with the aim of determining an algorithm specifically designed for the detection of *P. velutinum*. Genetic Programming helps to automate the search for solutions to different user-defined problems, which has been implemented in different areas, such as regression analysis [17], prediction [18,19], classification [20], symbolic regression [21], optimization [22], among others. The diversity of applications is possible since the algorithms of GP perform an abstraction of the knowledge represented in mathematical expressions or systems based on rules, in a probabilistic search space inspired by the Darwinian theory of evolution. The

contributions of this research are the following. We propose a Genetic-Programming-based methodology, designed specifically for the detection of *P. velutinum* in the flowering stage from forest areas. This algorithm allowed us to derive a detector of this particular mistletoe species, showing a higher accuracy rate (around 96.6%) compared to other classification methods. Since the optimal solution found by GP involves algebraic operations among the input spectral bands, this solution can be easily implemented for mistletoe detection in other forest areas, different to the one reported here.

The organization of the document is as follows. In Section 2, we describe the main characteristics and location of the study area. Section 3 presents the materials and methods used in our study, including the UAV characteristics, multispectral image collection, and our developed GP model. In Section 4, the results achieved with our GP-based algorithm for *P. velutinum* identification are compared with two traditional methods. Section 5 outlines some important remarks concerning the implementation of our GP algorithm for pest detection, while the main conclusions and the direction for future work are stated in Section 6.

2. Study Area

The study area is located in the conservation site, registered as Community Conservation Area of the San Bartolo Ameyalco Community, located at coordinates 19°20'N and 99°16'W at a height of about 2420 m.a.s.l. in Alvaro Obregón Borough, south of Mexico City. San Bartolo Ameyalco was registered with the Community Areas for Ecological Conservation program in 2017, with the aim of protecting, improving, and conserving the natural resources and environmental services that this area provides to Mexico City.

The conservation area in San Bartolo Ameyalco is made up of 244 hectares, most of which are cataloged as forest land, due to the presence of the following species: oyamel (*Abies religiosa*) forests covering 23% of the surface, pine (*Pinus* spp) covering 38%; cypress (*Cupressus*) and oak (*Quercus* spp) forests are also present. In addition, 22% of the surface is used for rainfed agriculture. Figure 1 shows the location of the study area.

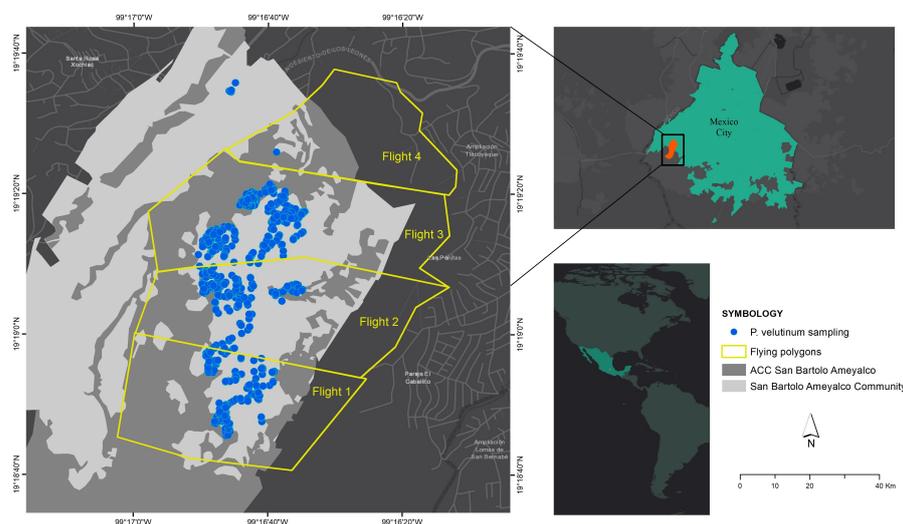


Figure 1. Location of the study area and flight polygon distribution for Unmanned Aerial Vehicle (UAV) image collection.

P. velutinum is a forest pest, also known as barbas, bungu, graft, evil eye, mistletoe or on stick [23], which is part of the *true mistletoe* species; these species are normally found in temperate zones of Mexico City [24] as in the case of San Bartolo Ameyalco Community.

P. velutinum is a hemiparasitic shrub with branched stems, opposite and decussate leaves, inflorescences in the form of articulated pedunculated spikes. It generally occurs in the aerial parts of trees or shrubs because the location in the higher parts allows it to obtain sufficient illumination for the photosynthesis process. Due to its hemiparasitic condition,

it is able to extract compounds from its host and photosynthesizes organic matter [25]. Physiologically, it is up to 80 cm long, with internodes of 8 cm long, without cataphiles, and with a petiole of 5–20 mm in length, the coloration of this parasite is green or yellowish with a velutine pubescence frequently yellow in the young parts and the mature parts glabrescent. It has male inflorescences of 2.5 cm in length, 2 to 5 segments of 6 to 40 flowers each in 6 longitudinal rows, the female inflorescence is 4 cm long with fruit, with 2 to 4 segments of 12 to 30 flowers per segment in 6 rows [26].

These species have an establishment stage, in which the seed with chlorophyll endosperm produces simple sugars as a source of energy before germination. Afterwards, the incubation stage follows, in which a radicle develops that penetrates the host cortex until it reaches the vascular tissues where it develops cortical haustoria (0.8–12 cm) [27]. Once established with aerial stems and flowers, a pollination process begins that can take around 4–6 weeks. Finally, the plant disperses the seed through ballistic propulsion, i.e., projecting the mature fruit around 15 m, which allows the seed to adhere to the branches of a new host [28].

3. Materials and Methods

3.1. Multispectral Image Collection

Multispectral image collection in the study area was carried out on 13 April 2021, at 12:30 p.m., with the support of the Commission of Natural Resources and Rural Development (CORENA) technical team, who provided information on the location of trees infested by *P. velutinum*.

At a solar elevation angle of 57.58° and an azimuth of 103.45°, aerial multispectral images of the affected regions were acquired using a P4 Multispectral UAV, frequently used in precision agriculture applications, environmental monitoring, and in-plant inspection and maintenance. This equipment has a camera with six 1/2.9" and 2 Megapixel CMOS sensors, composed by one RGB sensor ($R = 450 \text{ nm} \pm 16 \text{ nm}$, $G = 560 \text{ nm} \pm 16 \text{ nm}$, $B = 650 \text{ nm} \pm 16 \text{ nm}$) for the visible spectrum interval, and five monochrome sensors that include coverage in the red edge spectral band ($REG = 730 \text{ nm} \pm 16 \text{ nm}$) and near-infrared ($NIR = 840 \text{ nm} \pm 26 \text{ nm}$).

In order to cover the largest surface in the shortest flight time, the study area was divided into four polygons, as shown in Figure 1 and Table 1. The flying altitude was kept between 80 and 110 meters to achieve a Ground Sample Distance (GSD) lower than 10 cm, as suggested by [14]. The distribution of the flight polygons covered a total ground area of around 147.25 ha, recorded in 2565 images. Indeed, each image is composed of 3 individual files corresponding to RGB, REG, and NIR bands, respectively, with a dynamic range of 16 bits. The full estimated flight time was about 1 h and 40 s.

Table 1. Polygons distribution for Unmanned Aerial Vehicle (UAV) image collection.

Polygon	Area (ha)	Flying Height (m)	GSD (cm/px)	Number of Images
1	31.40	110	5.8	474
2	46.34	80	4.2	1244
3	49.14	90	4.8	522
4	20.37	110	5.8	325

3.2. Hardware and Software

The data preprocessing experiments that will be presented in subsequent sections were performed on a WorkStation Intel(R) Xeon(R) W-2102 CPU @ 2.90GHz, 16 RAM and Windows 10 as the operating system. In addition, the implementation of genetic programming to find an optimal solution to *P. velutinum* detection from UAV multispectral images was carried out using MATLAB R2019a software and GP-Lab toolbox [29]; these

experiments were run on a server Intel(R) Xeon(R) CPU. E5-2690 v2 @ 3.00GHz 250 RAM and Ubuntu 18.04.5 LTS operating system (GNU / Linux 4.15.0-122-generic x86-64).

3.3. Image Preprocessing

To carry out an adequate training of algorithms with GP, a preprocessing of the images was necessary, which consisted of the following steps: 1. The registration of the *REG* and *NIR* bands with respect to the *RGB* channels, since the sensor used by the P4 Multispectral equipment presents a shutter lag, causing displacements between the bands of each image; 2. The creation of masks that specify the pixels that show the presence of *P. velutinum*; 3. The creation of a data set with the necessary characteristics to train the GP model.

3.3.1. Image Registration

The registration of the spectral bands was implemented through the *Speeded up Robust Features* (SURF) algorithm [30], which is designed as a detector and descriptor of local characteristics using box filters for handling scales. The workflow of this registration algorithm is composed of the detection of points of interest, the description of the local neighborhood to find unique characteristics that can be representative of the orientation of the objects in the scene, and the coincidence of points from contrast comparison in areas of interest.

The SURF registration of the aforementioned spectral bands was implemented in the MATLAB programming platform. The registration algorithm is carried out employing the *detectSURFFeatures* function for the detection of objects, returning the points of interest, from which the characteristic vectors (descriptors) are extracted with the *extractFeatures* function together with their locations using the binarization of the image. The matching characteristics for both data sets (grayscale *RGB* – Ground truth and *REG* and *NIR* bands – moving image), are matched through the *matchFeatures* function.

The registration was carried out on 2565 images, together with a combination of parameters optimized for the study area; for this, an iteration of characteristics was carried out to maximize the registry in a forest environment with few buildings and roads. This previous study was carried out on 100 randomly selected images, considering the Structural Similarity Index (SSIM) [31] as a metric, due to its ability to evaluate the similarity between two images by weighting the luminance, contrast, and structure of the image without compressing or distorting it.

3.3.2. Mask and Data Set Creation

The process of creating the masks is based on the GPS sampling of *P. velutinum*, made up of 885 points captured in the conservation area of Mexico City by CORENA. This *InSitu* database was geographically located on the multispectral aerial images to identify infestation points with a photo-interpretation process, resulting in binary masks with the absence/presence of *P. velutinum* from ROI polygons, whose pixels with a value equal to 1 represent the parasitic plant and the pixels with a value of 0 represent the background (local trees, bare ground, and some buildings). Subsequently, a selection of 94 images showing the presence of mistletoe in different proportions was made. The criteria for this selection were: to avoid redundant spatial information within the set (overlapping) and all images should contain a clear presence of mistletoe.

Finally, for the creation of the data set with the presence/absence of *Phoradendron velutinum*, tiles of 300×300 pixels were made to increase the amount of training data and reduce computational time. Each tile was selected by maximizing the presence of the pest, obtaining a total of 250 tiles with a clear presence of *P. velutinum* in each section.

3.4. Algorithm to Detect Mistletoe

As we mentioned earlier, the objective of this research is to automatically design an algorithm capable of accurately detecting mistletoe in multispectral images by using GP. To this end, we propose using a transformation of the five spectral bands of an image, whose

response indicates the presence of mistletoe. The following step consists of an iterative non-linear normalization operator proposed by Itti and Koch [32]; they implemented a spatial competition scheme, inspired by physiological and psychological studies of corticocortical connections in early visual areas of the brain cortex, which results in a complex balance of excitation of highlighted regions of the image and inhibition of remaining regions. Finally, the normalized results are binarized by using a prominence threshold equal to zero. All values that are less than or equal to zero are assigned the value of 0 (no detection), while those greater than zero are assigned the value of 1 (detection). This binarization allows us to directly compare the results of the mistletoe detector algorithm with the binary image of ground truth. Figure 2 shows the operation diagram of the algorithm to detect mistletoe in multispectral images. The following Section 3.4.1 describes the details of the proposed GP approach for automatic design of a suitable image transformation for mistletoe detection.

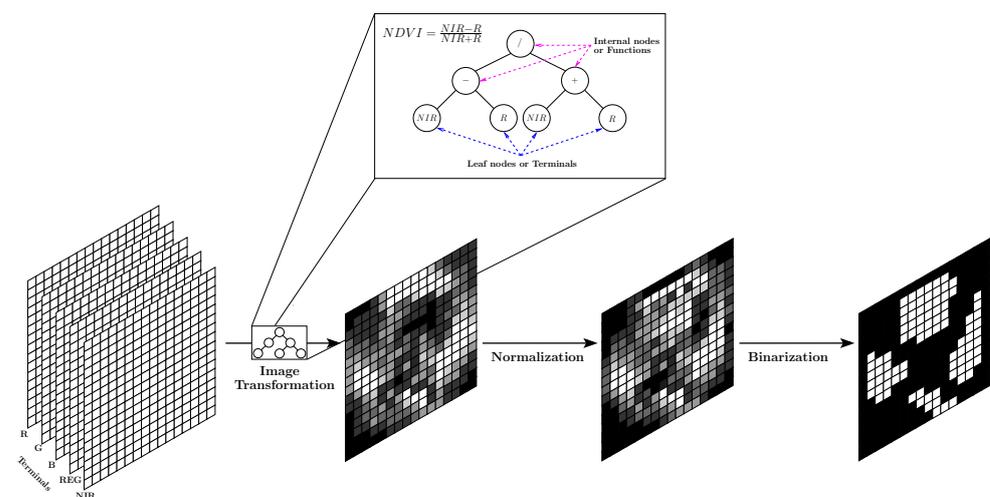


Figure 2. Complete diagram of the operation of the algorithm to detect mistletoe in multispectral images. In the Image Transformation stage, we illustrate, as an example, the internal nodes or functions and the leaf nodes or terminals required by the Genetic Programming (GP) algorithm to implement the NDVI index.

3.4.1. Genetic Programming for Feature Extraction

John Koza stated that “the problem of getting computers to learn to program themselves by providing a domain-independent way to search the space of possible computer programs for a program that solves a given problem” [33]. In 1992, he proposed the paradigm of Genetic Programming (GP) to adaptively and intelligently search for a single computer program that is good enough to solve a specific task. GP is capable of finding solutions close to the optimum, sometimes optimal solutions, in a treatable time. As we mentioned earlier, the goal of GP is to find an image transformation for the algorithm to accurately detect mistletoe.

Candidate solutions, i.e., image transformations, are represented as binary trees, called individuals. The root and internal nodes of the tree represent functions on the leaf nodes, called terminals, which represent the input data, e.g., the spectral bands of the images (see Figure 2). Functions and terminals should be selected in order to establish a suitable search space for the search for a suitable detector. For this reason, it is also necessary to adopt a precise criterion, called fitness function, to evaluate the performance of different individuals in carrying out mistletoe detection. Initially, GP generates a population of randomly created individuals, using the functions (protected functions are displayed in Table 2) and terminals shown in Table 3. These functions were chosen based on the algebraic operations used by different spectral indices and on image processing routines for image enhancement.

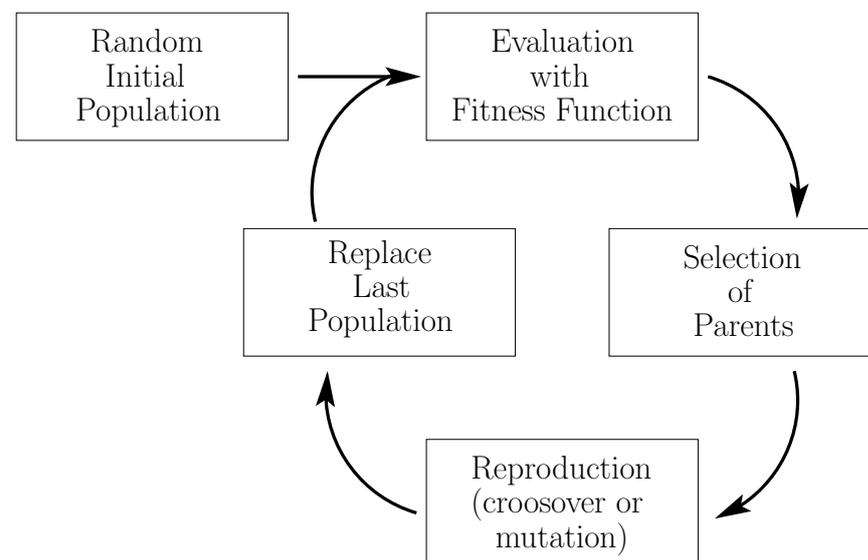
Table 2. Protected functions used for GP algorithm.

Protected Function	Input Arguments	Output Argument
Division	a, b	a (if b = 0) a/b (otherwise)
Square root	a	0 (if a ≤ 0) sqrt (a) (otherwise)
Base 2 Logarithm	a	0 (if a = 0) log ₂ (abs(a)) (otherwise)
Base 10 Logarithm	a	0 (if a = 0) log ₁₀ (abs(a)) (otherwise)

Table 3. Parameters used for GP algorithm.

Feature		Feature	
Generations	50	Selection	Lexictour
Population size	100	Elitism criteria	Keepbest
Initialization	Rampedinit	Functions set	+, + , - , ÷, ×, x ² , √x, e ^x , log ₂ , log ₁₀ , Ln, ¬, x , [x]
Crossover	Probability 0.7	Terminals set	R, G, B, REG, NIR
Mutation	Probability 0.3	Fitness function	Weighted Cohen's Kappa

Afterwards, the operation of GP can be divided into four steps that are executed iteratively; these steps are: evaluation, selection, reproduction, and replacement (see Figure 3). Therefore, the next step is to assess the accuracy of each individual when used to detect mistletoe.

**Figure 3.** Diagram of the main steps of Genetic Programming.

We propose using the weighted Kappa coefficient (k_w) [34] as a fitness function to evaluate the detection of individuals because k_w incorporates degrees of disagreement on a ratio scale for each of the possible results in a confusion matrix C , so that disagreements of varying severity are weighted accordingly. In other words, k_w can be used in situations where some errors are more serious than others, as in this study. We established the disagreement weights based on the opinion of experts in the detection and treatment of pests who say that because the resources assigned to these forestry tasks are limited, the cost of transportation and attention to false detections should be minimized. Therefore,

we conclude that it is worse to label a pixel as mistletoe, when it is not (false positive) than to label a pixel as not mistletoe when it is (false negative). Accordingly, we set the disagreement matrix of weights D (shown in Table 4), which represents the degree of disagreement between the two types of errors. The higher the disagreement, the greater the weight value, and it is convenient (but not necessary) to assign zero to the agreement diagonal, that is, no disagreement [34]. The disagreement matrix can be interpreted as follows: the error of misclassifying a pixel as a mistletoe-infested tree is twice as bad as the mistake of classifying an infested tree pixel as non-infested, as it is more expensive in terms of time, financial, and human resources to classify an area with mistletoe infestation, mobilize a commission to carry out the appropriate forest health activities, and find that there is indeed no presence of the parasitic plant. Such a fitness function was chosen to cause the GP search to target a search subspace where candidate mistletoe detectors improve their accuracy without too many false positives.

Table 4. Disagreement matrix D used during experiments with weighted Kappa.

		Predicted	
		Infested	Other
Real	Infested	0	1
	Other	2	0

Then, the k_w coefficient was calculated as in [34], as follows:

$$k_w = 1 - \frac{\sum v_{ij} f_{oij}}{\sum v_{ij} f_{cij}}, \quad (1)$$

where f_{oij} is the observed frequency in confusion matrix cell c_{ij} , f_{cij} is the respective chance-expected frequency, and v_{ij} is the disagreement weight in d_{ij} .

In order to calculate the individuals' final fitness score, each one is evaluated, together with normalization and binarization, by using a training set of images. The resulting map is compared to its corresponding ground truth to calculate the k_w classification accuracy. Finally, the fitness value of each individual is the average of all the k_w values obtained from the processing of all the training images and is calculated as follows:

$$F = \frac{1}{n} \sum_{i=1}^n k_{w_i}, \quad (2)$$

where n in the total number of images in the training set.

Once the fitness value of each individual has been calculated, it is necessary to select some for breeding. We use the selection method proposed by Luke and Panait [35] called lexicographic parsimony pressure because it can be helpful to control *bloat*, which is how the phenomenon of uncontrolled growth of the size of individuals is known in the GP community. There are different selection criteria but, usually, they are inspired by Darwin's Theory of Natural Selection, which is summarized in the phrase "the survival of the fittest". Selection allows the fittest individuals to have a greater likelihood of reproducing and passing their genetic material to their descendants; this, in turn, allows the best characteristics to be inherited by the following generations.

After the selection of the parents, they reproduce by means of two genetic operators—crossover (exchange of randomly selected subtrees between two parents) or mutation (substitution of a randomly selected subtree for a new tree created randomly in a parent). Selection and reproduction are repeated until a new complete population is created; each time, a choice is made between crossing and mutation with a probability of 0.7 and 0.3, respectively (see Table 3).

In other words, the following populations will be generated through a stochastic iterative process of evaluation, selection and reproduction steps. As John Koza states,

“Over a period of many generations, populations of computer programs are generated that are increasingly capable of solving the problem in question” [33].

3.5. Comparison Methods

For comparison purposes, two state-of-the-art methods, frequently used in the literature, will be assessed based on our collected imagery. The first one corresponds to the index known as *Structure Intensive Pigment Index 2* (SIPI2). This index has been widely used in the analysis of different types of chlorophyll in plants, estimating differences in the canopy of the trees, added to the identification of plant deficiencies [36,37]. In our research, this index will be employed to differentiate between the pigmentation of *P. velutinum* and the host tree.

In satellite images, the computation of the index is carried out using spectral bands associated to wavelengths 800 nm, 505 nm, and 690 nm, respectively. However, for multispectral imagery collected by UAVs, specific bands can be employed, according to the following expression [38]:

$$\text{SIPI2} = \frac{\text{NIR} - G}{\text{NIR} - R}. \quad (3)$$

The second method corresponds to the well-known Support Vector Machine (SVM) classification technique. The SVM classification method is a supervised learning algorithm frequently used in remote sensing applications, which has demonstrated great advantages in the correct identification of binary classes [39].

In addition, we also compare the accuracy of the GP solutions with two additional indices obtained during the best solution analysis in Section 4.3.

4. Experiments and Results

This section starts with a detailed description of the GP experiments, followed by the analysis of the 50 best solutions obtained, as well as the operation and performance of the best solution. In addition, we compare this solution with other approaches for the detection of mistletoe. At the end, we show an application of the *P. velutinum* detector.

4.1. Experimental Setup

In our experiments, we used a leave-one-out cross-validation methodology with kfold ($k = 5$), dividing the 250 images into 5 folders of 50 images each: 4 folders for training and 1 folder for testing. We executed 50 independent GP runs with five different training sets. At the end of each run, the best individual was evaluated with the corresponding test set to assess the final performance of the best candidates. The implementation of GP to find an optimal solution to *P. velutinum* in aerial images was carried out in the Software MATLAB R2019a implementing GP-Lab [29].

4.2. Structural Analysis of Gp Solutions

We start with an analysis of the structure of the 50 best solutions, with the aim of determining which are the most used terminals and functions are use the most to generate mistletoe detectors. Figures 4 and 5 show the frequencies of the use of terminals and functions employed in the evolutionary process of GP (see Table 3); the bar diagrams show the number of individuals who used each terminal or function, respectively, while the line graph refers to the number of times each terminal or function is used in the set of the best individuals of each generation. Notice that all the spectral bands were employed in the construction of algorithms for the detection of *P. velutinum*, however, both *R* and *B* bands were used constantly in almost all sets of the best individuals. On the other hand, regarding the use of functions, note that the absolute subtraction and division correspond to the most used arithmetic operators, followed by square and round operations. Considering that the GP algorithm is a stochastic method that directs the search for the best solution to subspaces where it is more likely to find better solutions, we could say that the most

common terminals and functions among the 50 best solutions, i.e., *R*, *B*, absolute difference, and division, are useful for the detection of *P. velutinum*.

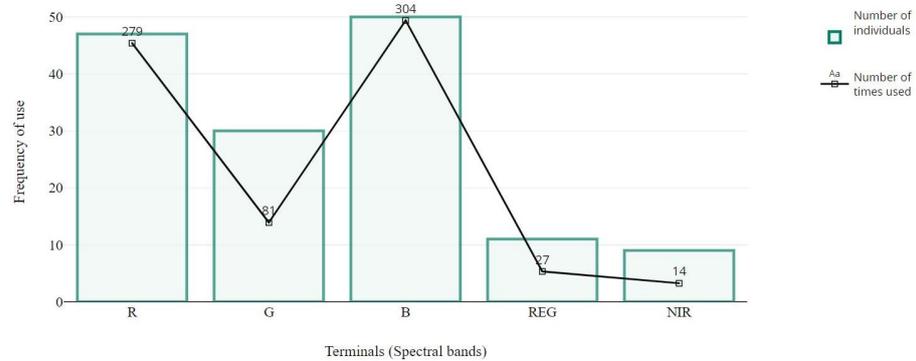


Figure 4. Frequency distribution of the use of terminals in the best individuals evolved by GP. The bar diagram shows how many of the best individuals have used the corresponding terminals; i.e., the *R*, *G*, *B*, *REG*, and *NIR* bands; the line graph shows the number of times that the terminals have been used in total within the best solutions individuals.

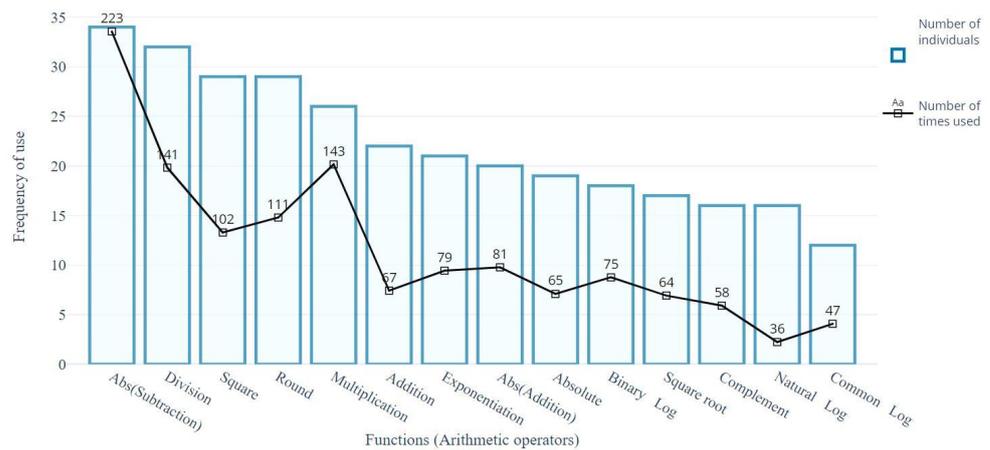


Figure 5. Frequency distribution of the use of functions in the best individuals evolved by GP. The bar diagram shows how many of the best individuals have used the corresponding functions (arithmetic operators), while the line graph shows the number of times that the functions have been used in total within the best-solution individuals.

In Section 3.4.1, we explain the fitness function that we used to guide the evolution of GP and find a solution that not only makes the least number of errors, but also generates the smallest number of false positives, i.e., labels a pixel as mistletoe when it is not. With the aim to assess the effectiveness of our fitness function in this respect, we computed the false positive rate (FPR) and false negative rate (FNR) for the 50 best solutions using their corresponding test image set. As a result, the average FPR was equal to 0.01 and the average FNR was equal to 0.73. In addition, the FPR and FNR for the best individual were 0.007 and 0.68, respectively. Therefore, the FPR of the 50 best solutions is much lower than their FNR.

4.3. Analysis of the Best Solution

Our GP implementation was able to determine several individuals or solutions for pest identification, however, the best individual, performing with the highest accuracy, has the form:

$$\left(\log_2\left(\frac{R}{B}\right) + \left| R - \left(|R - B| + \left| \left| R - \sqrt{\log_{10}\left(\left| \sqrt{\log_{10}(R)} \right| \right)} \right| - \frac{R}{G + (1 - REG^4)} \right) \right| \right) \right) \times \frac{R\sqrt{|\log_{10}(R)|}}{(\log_2(R) - kR)[\log_2(R) - (kR + \log_2(\log_2(G) - k(G + R)))]} \quad (4)$$

where $k = \log_2(e) = 1.4427$ is a constant scalar. The aforementioned expression can be considered as a feature extractor operator designed for mistletoe detection.

To highlight some ideas to help us understand how this solution works, here is a plausible explanation of how the more complicated and informative terms in Equation (4) work.

First, we were struck by the term $R - \sqrt{\log_{10}(|\sqrt{\log_{10}(R)}|)}$ because this results in complex numbers with the form $R - \sqrt{|\log_{10}(|\sqrt{\log_{10}(R)}|)|}i$, where $a = R$ and $c = -\sqrt{|\log_{10}(|\sqrt{\log_{10}(R)}|)|}$ are, respectively, the real and imaginary parts (see the Appendix A for a demonstration). This is a result of all spectral bands having been normalized to the interval $[0, 1]$ with the MATLAB function *rescale*, and the logarithm function for numbers within this interval producing negative values.

Next, the modulus of the resulting complex numbers is calculated, and is equal to the Euclidean distance between $(a, 0)$ and $(0, c)$ points. Note that the values of a and c depend exclusively on the value of R , and that the distance is directly proportional to the value of R . As you can see, the modulus term does not yield any more useful information than the R band alone. There is no other term in Equation (4) that results in complex numbers.

After rounding the modulus to the nearest integer, which means more information loss, the following term is subtracted:

$$\frac{R}{G + 1} \leq \frac{R}{G + (1 - REG^4)} \leq \frac{R}{G};$$

note that the upper limit of the following term is the *Red–Green Ratio Index*, and the lower limit is very close to it too. Hence, in a similar way to the *Red–Green Ratio Index*, this term could be an indicator of the content of anthocyanin, which may be a red, purple, blue, or black pigment, and of the chlorophyll concentration in leaves [40]. The *REG* band is used to accurately estimate the green leaf area index (LAI), as reported in [41]. However, the values of *REG* to the fourth power are very small, so it has very little effect on the result of this ratio, which tends to its lower limit.

The combination (subtraction) of both terms could yield little or no information about the mistletoe location, as shown in the example in Figure 6b. However, more tests are needed to determine this accurately.

On the other hand, we will now focus on analyzing two terms that, as shown in Figure 6a,f, yield a result that facilitates the location of the mistletoe. Considering first the term $|R - B|$, we can say that it represents the distance between the reflectance values of the red and blue bands. The difference $R - B$ is used to estimate the chlorophyll content of leaves, in a similar way to the spectral index analyzed by authors in [42]. The second term is $\log_2\left(\frac{R}{B}\right)$. In the Remote Sensing community, the R/B ratio is known as Iron Oxide Ratio and tends to highlight red or orange-colored materials, which absorb band B and reflect band R [43]. The logarithmic transformation of the input compresses the range of pixel values, which has the effect of improving the contrast and sharpness of regions with low pixel values, and it is frequently used in digital image processing [44,45]. Note that both the distance term and the ratio term between R and B increase as the values of R become larger or those of B become smaller. Based on this information, it stands to reason that these

two terms have high responses to the characteristic pigmentation of mistletoe. Since this assumption does not provide enough evidence, we decided to measure the performance of these two terms when used in isolation for mistletoe detection. The results demonstrate their great contribution to the correct functioning of the solution in Equation (4). These experiments and results are shown in the next Section 4.4.

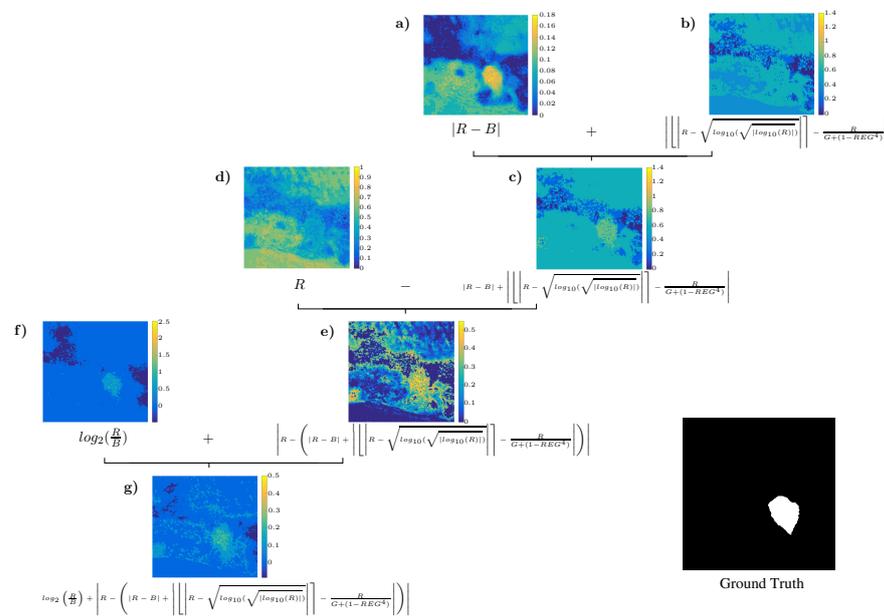


Figure 6. Example of the partial results obtained by applying the first factor of the best solution, shown in Equation (4), on an image. The letters (a–g) show the performance (visual result) of each term in the equation that is expressed below the corresponding image.

Regarding the second factor of the best solution and based on its response to the image shown in Figure 7, we can say that it apparently yields information that allows us to distinguish bare dry soil from the rest. This information could be helpful in ruling out regions of bare dry soil as a location for mistletoe.

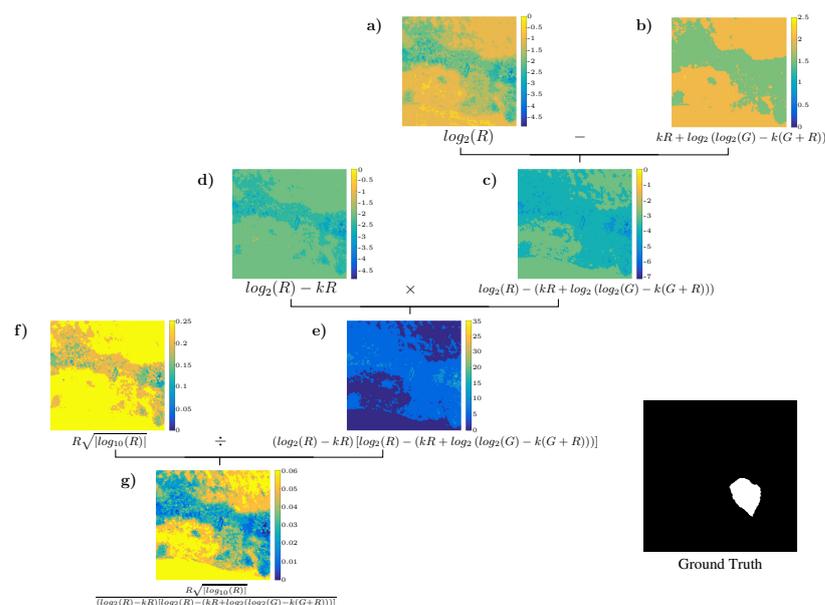


Figure 7. Example of the partial results obtained by applying the second factor of the best solution, shown in Equation (4), on an image. The letters (a–g) show the performance (visual result) of each term in the equation that is expressed below the corresponding image.

The multiplication of the two factors in Equation (4) shows that the greatest values correspond to the location of the mistletoe in the image, although some other smaller peaks can also be observed (see Figure 8). As we mentioned in Section 3.4, after applying the mistletoe detector, an iterative nonlinear normalization operator is applied, which after a few iterations strengthens the initial maximum(s), while at the same time suppressing regions with lower activation [32]. Finally, the map obtained is binarized using a threshold equal to zero.

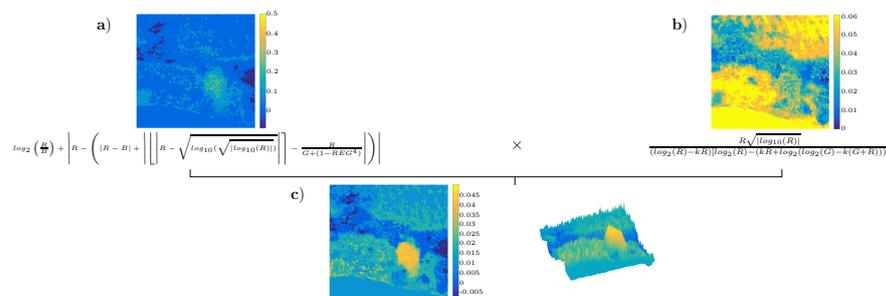


Figure 8. Example of the result of the best solution, shown in Equation (4), when applied on an image. The letters (a–c) show the performance (visual result) of each term in the equation that is expressed below the corresponding image.

4.4. Comparison with Other Methods

The characteristics of the GP metaheuristic make it possible to obtain multiple solutions to the same problem. Therefore, the 50 best algorithms or solutions generated by GP were compared with two traditional methods, frequently used in the literature for the same purpose: the SIPI2 index and the SVM classification technique. These solutions were also compared with two additional indices obtained during the analysis of the best solution of GP, which are $|R - B|$ and $\log_2\left(\frac{R}{B}\right)$, as discussed in Section 4.3.

In the case of SVM, pixel-based classification paradigm [46] was implemented to increase the sample size and facilitate the definition of the support vectors, using the same train and test folders as in GP. The optimal set of parameters for SVM were set using a cross-validation through the Scikit-learn GridSearchCV package, set a polynomial kernel, with a regularization parameter $C = 100$ and a $\gamma = 1$ for this classification. Then, a sample of 1 million pixels, balanced by class, was randomly selected, of which 30% was assigned as validation and 70% was for testing, and from the test set, a random sample of 100 thousand pixels, balanced by class, were selected to perform the model testing. The implementation of this method was done in Google Colab with the *scikit-learn* package.

On the other hand, with the aim of determining the threshold values that maximize the classification of mistletoe with SIPI2, $|R - B|$, and $\log_2\left(\frac{R}{B}\right)$ indices, we performed several experiments with a sample of 100 randomly selected images, varying the threshold values from -1 to 1 ; at the end of the processes, we obtained the corresponding threshold values providing the best classification accuracy score, κ_w , which were the following: 0.9, 0.15, and 0.236, for SIPI2, $|R - B|$, and $\log_2\left(\frac{R}{B}\right)$, respectively.

It is important to note that both GP and SVM were optimized using data from the same set of 200 images. In addition, the selection process of the SVM parameters as well as the threshold values for SIPI2, $|R - B|$, and $\log_2\left(\frac{R}{B}\right)$, ensures that the classification results, shown in Table 5, are the best that can be obtained for the test image set that we used.

The first two rows of Table 5 show the results of the two best individuals determined by the GP method, while the next two rows present the results of the two worst individuals. Notice that the results of the best individual appear in the first line, which obtained an average value of $k_w = 0.45$ at the test stage, an Overall Accuracy (OA) of 96.6%, a recall that indicates that 48% of the total pixels of *P. velutinum* were detected, and a precision indicating that half of the pixels that were labeled as *P. velutinum* were correctly classified.

On the other hand, the SIPI2 method obtained an OA = 54% and 43% of the *P. velutinum* pixels were detected, but only 2.8% of the pixels that were labeled as mistletoe were correctly classified; this can be clearly seen in Table 6 in white and magenta, respectively. The SIPI2 achieved these results by labeling a large number of pixels as mistletoe.

Regarding the SVM method, it achieved an OA of 88%, from which 15% of the mistletoe pixels were detected and 70% of the pixels that SVM classifies as *P. velutinum* were correctly classified. These results indicate that although SVM detects very little mistletoe, this method is highly reliable for doing this task.

Furthermore, Table 6 illustrates the performance of the best solution in Equation (4), obtained for four images. It also presents the performance for SIPI2 and SVM methods. A visual inspection of the corresponding maps confirms what we have stated before.

As we mentioned earlier and supported by the results shown in Figure 6a,f, the information that allows mistletoe to be detected originates from the $|R - B|$ and $\log_2(\frac{R}{B})$ terms. In order to demonstrate this assumption, we decided to measure the performance of these terms while they were used in isolation for the detection of mistletoe. The results in Table 5 show that mistletoe detection using $|R - B|$ reached an OA = 91.3%; 32% of the mistletoe was detected and 37% of what was classified as *P. velutinum* really was. Regarding the results of using $\log_2(\frac{R}{B})$, this term obtained an OA = 73.2%; 65% of the mistletoe was correctly detected, the highest recall of all solutions, but only with 0.25% of precision. Table 7 depicts a few examples of response maps of both $|R - B|$ and $\log_2(\frac{R}{B})$ terms.

Table 5. Classification performance obtained by the two best and the two worst algorithms produced by our GP implementation, in comparison with the Intensive Pigment Index of Structure 2 (SIPI2), and Support Vector Machine (SVM) methods, as well as the $|R - B|$ and $\log_2(\frac{R}{B})$ terms. The first column shows the different methods. The training column shows the results of the training stage of each method employing the Weighted Cohen’s Kappa metric; in the test columns, the results corresponding to the metrics Weighted Cohen’s Kappa (k_w), Cohen’s Kappa (k), Overall Accuracy (OA), Precision (P), and Recall (R) are listed; in the Diff column, the difference between the k_w coefficients in the training and testing stages is shown.

Method	Training	Test		OA	R	P	Diff.
	k_w	k_w	k				
GP ₁	0.465	0.448	0.442	0.966	0.482	0.498	−0.017
GP ₂	0.436	0.416	0.395	0.964	0.389	0.524	−0.021
GP ₄₉	0.298	0.294	0.275	0.963	0.269	0.412	−0.004
GP ₅₀	0.296	0.282	0.284	0.952	0.357	0.319	−0.014
SIPI2	0.0025	−0.005	−0.006	0.543	0.426	0.028	−0.007
$ R - B $	0.267	0.276	0.261	0.913	0.326	0.373	0.009
$\log_2(\frac{R}{B})$	0.236	0.197	0.203	0.732	0.649	0.247	−0.039
SVM	Validation						
	0.198	0.168	0.207	0.881	0.146	0.704	−0.030

Table 6. Classification of *Phoradendron velutinum* (*P. velutinum*) in multispectral images produced by GP, SIPI2 index, and SVM methods. The first column shows the original images evaluated, columns two, three, and four represent, respectively, the classification of *P. velutinum* in the original images using four of the best algorithms found by GP, SIPI2 index, and SVM methods. In these results, the correctly classified mistletoe can be seen in white color, the well-classified background in black color, while shown in green are the pixels that represent errors by omission (FN) and in magenta the errors by commission (FP).

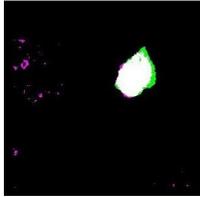
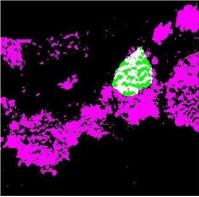
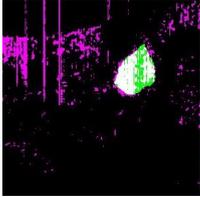
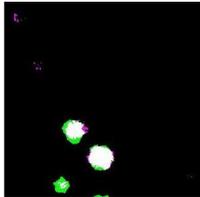
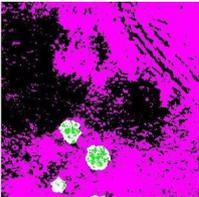
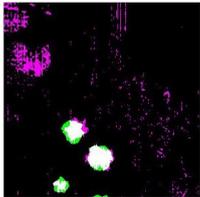
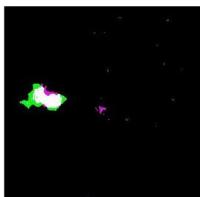
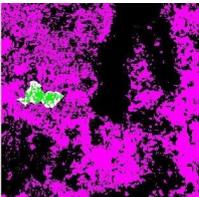
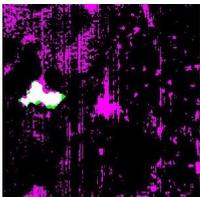
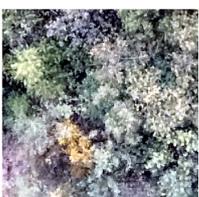
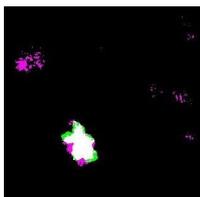
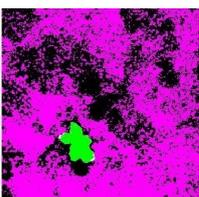
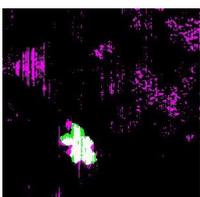
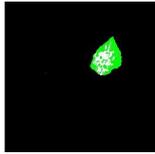
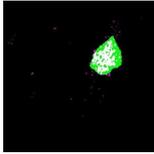
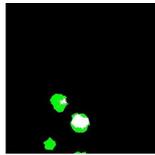
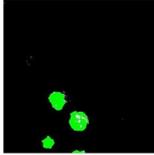
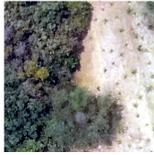
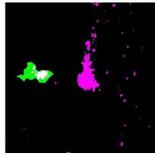
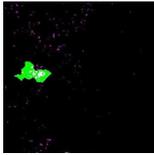
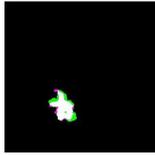
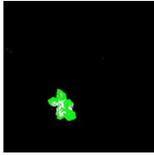
Original Image	Testing Results		
	GP	SIPI2	SVM
	 $k_w = 0.866$ $OA = 0.991$ $P = 0.886$ $R = 0.839$	 $k_w = 0.083$ $OA = 0.796$ $P = 0.090$ $R = 0.566$	 $k_w = 0.328$ $OA = 0.932$ $P = 0.285$ $R = 0.707$
	 $k_w = 0.796$ $OA = 0.988$ $P = 0.894$ $R = 0.663$	 $k_w = 0.011$ $OA = 0.434$ $P = 0.035$ $R = 0.728$	 $k_w = 0.433$ $OA = 0.959$ $P = 0.379$ $R = 0.733$
	 $k_w = 0.712$ $OA = 0.991$ $P = 0.795$ $R = 0.598$	 $k_w = 0.005$ $OA = 0.550$ $P = 0.019$ $R = 0.551$	 $k_w = 0.130$ $OA = 0.885$ $P = 0.106$ $R = 0.838$
	 $k_w = 0.687$ $OA = 0.987$ $P = 0.653$ $R = 0.795$	 $k_w = -0.029$ $OA = 0.366$ $P = 0.002$ $R = 0.046$	 $k_w = 0.325$ $OA = 0.951$ $P = 0.268$ $R = 0.777$



Table 7. Visual term analysis from the best GP algorithm. This comparison is made with the same examples show in Table 6. The first column displays the original images evaluated, the second and third columns show the analysis of $|R - B|$ and $\log_2(\frac{R}{B})$ terms, respectively. Black color is the well-classified background, white is *P. velutinum* correctly classified, green and magenta indicate the omission and commission errors, respectively.

Original Image	Testing Results	
	$ R - B $	$\log_2(\frac{R}{B})$
	 $k_{\text{av}} = 0.71$ $\text{OA} = 0.98$ $\text{P} = 0.91$ $\text{R} = 0.51$	 $k_{\text{av}} = 0.62$ $\text{OA} = 0.98$ $\text{P} = 0.98$ $\text{R} = 0.36$
	 $k_{\text{av}} = 0.16$ $\text{OA} = 0.97$ $\text{P} = 0.81$ $\text{R} = 0.06$	 $k_{\text{av}} = 0.58$ $\text{OA} = 0.98$ $\text{P} = 0.99$ $\text{R} = 0.32$
	 $k_{\text{av}} = 0.27$ $\text{OA} = 0.98$ $\text{P} = 0.39$ $\text{R} = 0.17$	 $k_{\text{av}} = 0.16$ $\text{OA} = 0.96$ $\text{P} = 0.15$ $\text{R} = 0.26$
	 $k_{\text{av}} = 0.49$ $\text{OA} = 0.98$ $\text{P} = 0.92$ $\text{R} = 0.26$	 $k_{\text{av}} = 0.84$ $\text{OA} = 0.99$ $\text{P} = 0.92$ $\text{R} = 0.73$

P. velutinum
 Background
 Omission errors
 Commission errors
 Classified correctly
 Classified correctly
 Omission errors
 Commission errors

4.5. Mistletoe Map Detection

In Figure 9, the result of the implementation of the GP algorithm for the automated detection of *Phoradendron velutinum* in the case study of the Conservation Area of Mexico City is shown in red spots; for comparison purposes, the in situ sampling of *P. velutinum* was overlapped and displayed in blue circles in the same figure. A section of the study area mentioned previously was taken as an experiment, to visualize the areas affected by mistletoe. It is worth mentioning that the implementation refers to the final use of the best algorithm generated with GP (see Equation (4)) directly in the images to be classified, obtaining the location of infestation hot spots by *P. velutinum* related to the flowering stage.

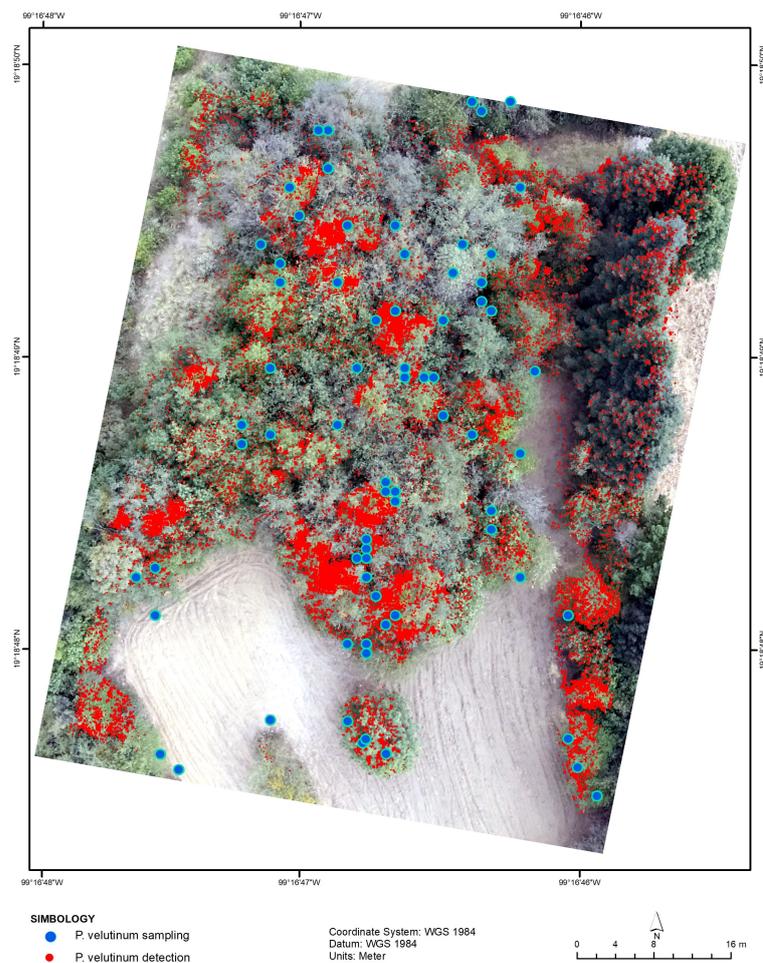


Figure 9. Classification map of *Phoradendron velutinum* (*P. velutinum*) with the best GP algorithm. A section of the Conservation Area of Mexico City is shown, in which the detection of *Phoradendron velutinum* was carried out employing the best individual generated from GP (Equation (4)). The main areas with infestation of *P. velutinum* are shown in red and *P. velutinum* in situ GPS sampling points are represented by blue circles

5. Discussion

The proposed Genetic Programming methodology for the automatic design of an algorithm for detecting *Phoradendron velutinum* in multispectral images of the conservation area yielded a detector that presents better performance and a greater generalization value compared to other methods that have traditionally been used either to detect forestry patterns, such as SIPI2, or to learn the distribution of the data in a defined search space, such as SVM. The SIPI2 was not generated specifically for the detection of this particular mistletoe species, nor for the specific study area, therefore, detection using SIPI2 is very poor, although the threshold that maximizes its accuracy was chosen. Although the SVM is the most accurate detector, it also had the lowest recall, which renders it useless as it misses 85% of the mistletoe. On the other hand, the term $\log_2\left(\frac{R}{B}\right)$ is the most sensitive, it detects 65% of the *P. velutinum* pixels. However, as can be seen in Table 5, the best GP-generated algorithm has the second-best recall and precision, making it the most suitable solution. Furthermore, as indicated by its FPR = 0.007, we obtained a solution that gives very few false alarms, which is what we intended to do by using the disagreement matrix when calculating fitness, see Sections 3.4.1 and 4.2, since it generates high human and financial costs, mobilizing work commissions for mistletoe control, when this parasitic plant is not found in said area. It is also important to mention that the obtained solution has the best OA = 96.6%.

From a visual evaluation of the examples of response maps, see Tables 6 and 7, it is possible to infer that some of the errors by omission near areas correctly classified as *P. velutinum* increase in perimeter areas of the mistletoe, due to the crossing that exists between the final branches of the mistletoe and those of the host tree, limiting the differentiation between species because there is not enough spatial resolution to find these differences.

The solutions obtained by GP are mostly based on the visible spectrum bands, as shown in Figure 4. It is very likely that this is due to the characteristic yellow-orange color of *P. velutinum*, in the flowering stage. However, as shown in Equation (4), the REG band makes specific contributions in the discrimination of *P. velutinum* from other types of vegetation.

The *R* and *B* bands and the absolute difference and division operations are the most used by the best solutions. During the analysis of the best solution of all, we discovered that the terms $|R - B|$ and $\log_2\left(\frac{R}{B}\right)$ could play an important role in its operation, which was confirmed by the results shown in Table 5. Note that both terms are composed precisely by the most used terminals and functions. All these results are consistent. However, neither of these two terms in isolation performs better than the best solution that combines the result of both. This suggests that the two terms could perhaps be used as building blocks for more robust detectors.

6. Conclusions and Future Work

Automation in the detection of the mistletoe species *Phoradendron velutinum* with the use of spectral images is a complex task due to the heterogeneity of vegetation cover found in areas such as the Conservation Area of Mexico City, coupled with the few physical and visual characteristics that are known in the detection in images of this parasitic plant. The detection of pests such as *P. velutinum* is important due to the relevance of host forests in environmental, social, and economic contexts, in addition to the difficult detection of new outbreaks due to the easy spread of the pest.

The algorithm obtained by using a Genetic Programming approach for the detection of *P. velutinum* presented in this manuscript represents a clear and viable solution for the generalization of this task in the proposed study area. This algorithm allows the detection of the pest with 96.6% Overall Accuracy, considering information from both the visible and red edge intervals of the electromagnetic spectrum.

From the findings found with the implementation of GP in the detection of mistletoe, it is possible to deduce that the physical properties of mistletoe and its relationship with the territory allow the early detection of foci of infestation. Additionally, it is possible to appreciate a great competitive advantage over other detection methods, since in addition to solving this problem efficiently, it is able to generalize the solution to the entire study area in different image acquisition conditions, both lighting and exposition.

It is important to emphasize that the development of the proposed algorithm for the detection of *P. velutinum* and the training data are limited to the characteristics of the conservation area of Mexico City, therefore, it will be essential to carry out tests and improvements in the methodological characteristics to study new areas where mistletoe is present. Furthermore, this research reveals that the terms $|R - B|$ and $\log_2\left(\frac{R}{B}\right)$ can be used in the future design of better and more robust *P. velutinum* detectors.

Future research will be oriented towards the application, or even adaptation, of the proposed GP algorithm to other forest areas, including those surrounded by urban environments. In addition, based on the high performance of the best individual determined by GP, it becomes necessary to implement further comparisons with several traditional and novel classification algorithms.

Author Contributions: Conceptualization, P.A.M.-Z., L.D. and J.C.V.-N.; methodology, P.A.M.-Z., L.D. and J.C.V.-N.; software, P.A.M.-Z., L.D. and J.C.V.-N.; validation, P.A.M.-Z., L.D. and J.C.V.-N.; formal analysis, P.A.M.-Z., L.D. and J.C.V.-N.; investigation, P.A.M.-Z., L.D. and J.C.V.-N.; resources, P.A.M.-Z., L.D. and J.C.V.-N.; data curation, P.A.M.-Z., L.D. and J.C.V.-N.; writing—original draft preparation, P.A.M.-Z., L.D. and J.C.V.-N.; writing—review and editing, P.A.M.-Z., L.D. and J.C.V.-N.

N.; visualization, P.A.M.-Z., L.D. and J.C.V.-N.; supervision, P.A.M.-Z., L.D. and J.C.V.-N.; project administration, P.A.M.-Z., L.D. and J.C.V.-N.; funding acquisition, P.A.M.-Z., L.D. and J.C.V.-N. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Centro de Investigación en Ciencias de Información Geoespacial-CentroGeo.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available on request from the authors. The data supporting this study's findings are available from the corresponding author, (Valdiviezo-N., J.C.).

Acknowledgments: The authors would like to thank the members of the San Bartolo Ameyalco Community in Mexico City for allowing us to carry out this research within their community, and the Engineer Israel Jiménez from CORENA, for his support in field work, contribution of technical knowledge of the area study and with the information provided from field sampling of *P. velutinum*. As well as thank the support from "Laboratorio de Supercómputo del Bajío. In addition, Paola Mejia would like to thank the National Council of Science and Technology (CONACYT) for partial support through scholarship grant number 752584.

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

Appendix A

The modulus of a function $y = \log_b(x)$, where $0 \leq x \leq 1$ and $y \leq 0$, is defined as

$$|y| = -y \quad \text{if } y < 0,$$

Therefore,

$$\sqrt{y} = \sqrt{-|y|} = \sqrt{-1} \sqrt{|y|}.$$

By substituting $i = \sqrt{-1}$ and $y = \log_b(x)$, we have that

$$\sqrt{\log_b(x)} = \sqrt{|\log_b(x)|}i$$

Then, by substituting $x = |\sqrt{\log_{10}(R)}|$, which satisfies $0 \leq x \leq 1$, and $b = 10$, we have that

$$\sqrt{\log_{10}(|\sqrt{\log_{10}(R)}|)} = \sqrt{|\log_{10}(|\sqrt{\log_{10}(R)}|)|}i$$

From the above expression, we can say that the term,

$$R - \sqrt{\log_{10}\left(|\sqrt{\log_{10}(R)}|\right)}$$

appearing in the first factor of Equation (4), it can be written as:

$$(R + 0i) - \left(0 + \sqrt{|\log_{10}(|\sqrt{\log_{10}(R)}|)|}i\right) = R - \sqrt{|\log_{10}(|\sqrt{\log_{10}(R)}|)|}i$$

and results in complex numbers with real part $a = R$ and imaginary part $c = -\sqrt{|\log_{10}(|\sqrt{\log_{10}(R)}|)|}$.

References

- Griebel, A.; Watson, D.; Pendall, E. Mistletoe, friend and foe: Synthesizing ecosystem implications of Mistletoe infection. *Environ. Res. Lett.* **2017**, *12*, 115012. [CrossRef]
- Secretaría del Medio Ambiente de la Ciudad de México—SEDEMA. *Suelo de Conservación*, 1st ed.; SEDEMA: Mexico City, Mexico, 2016; pp. 89–95.
- Lee, H.S.; Seo, W.W.; Lee, K.S. Detection of Oak Wilt Disease Using Convolutional Neural Network From Uav Natural Color Imagery. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 6622–6624.
- Dwivedi, R.; Dey, S.; Chakraborty, C.; Tiwari, S. Grape Disease Detection Network based on Multi-task Learning and Attention Features. *IEEE Sens. J.* **2021**, *21*, 17573–17580. [CrossRef]
- Hari-Shankar, R.L.; Veeraraghavan, A. K.; Uvais ; Sivaraman, K.; Shreyas, S. Application of UAV for pest, weeds and disease detection using open computer vision. In Proceedings of the International Conference on Smart Systems and Inventive Technology, Tirunelveli, India, 13–14 December 2018; pp. 287–292.
- Akila, M.; Deepan, P. Detection and classification of plant leaf diseases by using Deep Learning algorithm. *Int. J. Eng. Res. Technol.* **2018**, *6*. Available online: <https://www.ijert.org/detection-and-classification-of-plant-leaf-diseases-by-using-deep-learning-algorithm> (accessed on 30 November 2021).
- Minařík, R.; Langhammer, J.; Lendziach, T. Automatic tree crown extraction from uas multispectral imagery for the detection of bark beetle disturbance in mixed forests. *Remote Sens.* **2020**, *12*, 4081. [CrossRef]
- Nasi, R.; Honkavaara, E.; Blomqvist, M.; Lyytikäinen-Saarenmaa, P.; Hakala, T.; Viljanen, N.; Kantola, T.; Holopainen, M. Remote sensing of bark beetle damage in urban forests at individual tree level using a novel hyperspectral camera from UAV and aircraft. *Urban For. Urban Green.* **2018**, *30*, 72–83. [CrossRef]
- Zhang, N.; Wang, Y.; Zhang, X. Extraction of tree crowns damaged by *Dendrolimus tabulaeformis* Tsai et Liu via spectral-spatial classification using UAV-based hyperspectral images. *Plant Methods* **2020**, *16*, 135. [CrossRef] [PubMed]
- Pernar, R.; Bajic, M.; Ancic, M.; Seletkovic, A.; Idzotic, M. Detection of mistletoe in digital colour infrared images of infested fir trees. *Period. Biol.* **2007**, *109*, 67–75.
- Ančić, M.; Pernar, R.; Bajić, M.; Seletković, A.; Kolić, J. Detecting mistletoe infestation on Silver fir using hyperspectral images. *iForest-Biogeosciences For.* **2013**, *7*, 85–91. [CrossRef]
- Maes, W.H.; Huete, A.R.; Avino, M.; Boer, M.M.; Dehaan, R.; Pendall, E.; Griebel, A.; Steppe, K. Can UAV-based infrared thermography be used to study plant-parasite interactions between mistletoe and eucalypt trees? *Remote Sens.* **2018**, *10*, 2062. [CrossRef]
- Sabrina, F.; Sohail, S.; Thakur, S.; Azad, S.; Wasimi, S. Use of Deep Learning Approach on UAV imagery to Detect Mistletoe Infestation. In Proceedings of the IEEE Region 10 Symposium, TENSYP, Dhaka, Bangladesh, 5–7 June 2020; pp. 556–559.
- León-Bañuelos, L.A.; Endara-Agramont, A.R.; Gómez-Demetrio, W.; Martínez-García, C.G.; Gabino, Nava-Bernal, E. Identification of *Arceuthobium globosum* using unmanned aerial vehicle images in a high mountain forest of central Mexico. *J. For. Res.* **2020**, *31*, 1759–1771. [CrossRef]
- Miraki, M.; Sohrabi, H.; Fatehi, P.; Kneubuehler, M. Detection of mistletoe infected trees using UAV high spatial resolution images. *J. Plant Dis. Prot.* **2021**, *128*, 1679–1689. [CrossRef]
- Mejia-Zuluaga, P.; Dosal, L.; Valdiviezo-N, J.C. Detection of *Phoradendron velutinum*. In Proceedings of the iGISc2021: International Conference on Geospatial Information Sciences, Mexico City, Mexico, 3–5 November 2021.
- Chen, J.; Li, Z.; Liao, Z.; Wang, Y. Distributed service performance management based on linear regression and genetic programming. In Proceedings of the 2005 International Conference on Machine Learning and Cybernetics, Guangzhou, China, 18–21 August 2005; pp. 560–563.
- Zhu, H.-R.; Li, Y.M.; Meng, L.-M. Based on meteorological factors and short-term load forecasting genetic programming. In Proceedings of the 2010 International Conference on Computer Design and Applications, Qinhuangdao, China, 25–27 June 2010; pp. 465–467.
- Huo, L.; Fan, X.; Xie, Y.; Yin, J. Short-term load forecasting based on the method of genetic programming. In Proceedings of the 2007 IEEE International Conference on Mechatronics and Automation, Harbin, China, 5–8 August 2007; pp. 839–843.
- Brameier, M.; Banzhaf, W. A comparison of linear genetic programming and neural networks in medical data mining. *IEEE Trans. Evol. Comput.* **2001**, *5*, 17–26. [CrossRef]
- Icke, I.; Bongard, J.C. Improving genetic programming based symbolic regression using deterministic machine learning. In Proceedings of the 2013 IEEE Congress on Evolutionary Computation, Cancun, Mexico, 20–23 June 2013; pp. 1763–1770.
- Yuan, Y.; Fan, W.; Wang, W.; Liu, H. Robust collaborative optimization of a multi-finger micro-accelerometer based on genetic algorithm. In Proceedings of the 2nd International Conference on Genetic and Evolutionary Computing, Jinzhou, China, 3 October 2008; pp. 105–108.
- Phoradendron velutinum* (DC.) Oliv. 2014. Available online: https://repositorio.unam.mx/contenidos/34phoradendron-velutinum34-dc-oliv-2590151?c=4EQE34&d=false&q=*&i=1&v=1&t=search_0&as=0 (accessed on 30 November 2021).
- González-Gaona, E.; Serrano, Gómez, C.; De Lira, Ramos, K.V.; Quiñonez, Barraza, S.; Sánchez, Martínez, G.; López, Pérez, I.; Sánchez, Lucio, R. *Identificación, Distribución y Control de Muerdago Enano (Arceuthobium spp.) en Bosques de Coníferas*; Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias: Aguascalientes, Mexico, 2017.

25. Gutierrez-Vilchis, L.H.; Resendiz-Martinez, J.F. Fenologia del muerdago enano en el desierto de los leones, D.F. *Rev. Cienc. For. Mex.* **1994**, *19*, 41–62.
26. Rzedowski, J.; Calderon-de-Rzedowski, G. *Flora del Bajío y de regiones adyacentes*; Insituto de Ecologia a.c.: Michoacan, Mexico, 2011.
27. Mathiasen, R.L.; Nickrent, D.L.; Shaw, D.C.; Watson, D.M. Mistletoes: Pathology, systematics, ecology, and management. *Plant Dis.* **2008**, *92*, 988–1006. [[CrossRef](#)]
28. Alvarado-Rosales, D.; Equihua-Martínez, A.; Tagle, E.L.-G.; Rodríguez-Ortega, A. Situacion actual de la declinacion del oyamel en el parque Desierto de los Leones, D.F. In Proceedings of the XIV Simposio Nacional de Parasitología Forestal, Aguascalientes, Mexico, 7–9 November 2007; pp. 24–28.
29. Silva, S.; Almeida, J. GPLAB A Genetic Programming Toolbox for MATLAB. In In Proceedings of the Nordic MATLAB Conference, Stockholm, Sweden, April 2007.
30. Bay, H.; Tuytelaars, T.; Van Gool, L. SURF: Speeded Up Robust Features. In Proceedings of the Computer Vision—ECCV, Graz, Austria, 7–13 May 2006; pp. 404–417. [[CrossRef](#)]
31. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
32. Itti, L.; Koch, C. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis. Res.* **2000**, *40*, 1489–1506. [[CrossRef](#)]
33. Koza, J.R. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, 1st ed.; A Bradford Book: Cambridge, MA, USA, 1992; p. 840.
34. Cohen, J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol. Bull.* **1968**, *70*, 213–220. [[CrossRef](#)] [[PubMed](#)]
35. Luke, S.; Panait, L. Lexicographic parsimony pressure. In Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation, New York, NY, USA, 9–13 July 2002; pp. 829–836.
36. Blackburn, G. A. Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *Int. J. Remote Sens.* **1998**, *19*, 657–675. [[CrossRef](#)]
37. Pu, R.; Gong, P.; Yu, Q. Comparative Analysis of EO-1 ALI and Hyperion, and Landsat ETM+ Data for Mapping Forest Crown Closure and Leaf Area Index. *Sensors* **2008**, *8*, 3744–3766. [[CrossRef](#)]
38. Qi, H.; Wu, Z.; Zhang, L.; Li, J.; Zhou, J.; Jun, Z.; Zhu, B. Monitoring of peanut leaves chlorophyll content based on drone-based multispectral image feature extraction. *Comput. Electron. Agric.* **2021**, *187*, 106292. [[CrossRef](#)]
39. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [[CrossRef](#)]
40. Gamon, J.; and Surfus, J. Assessing Leaf Pigment Content and Activity With a Reflectometer. *New Phytol.* **1999**, *143*, 105–117. [[CrossRef](#)]
41. Delegido, J.; Verrelst, J.; Meza, C.M.; Rivera, J.P.; Alonso, L.; Moreno, J. A red-edge spectral index for remote sensing estimation of green LAI over agroecosystems. *Eur. J. Agron.* **2013**, *46*, 42–52. [[CrossRef](#)]
42. Kawashima, S.; Nakatani, M. An algorithm for estimating chlorophyll content in leaves using a video camera. *Ann. Bot.* **1998**, *81*, 49–54. [[CrossRef](#)]
43. Alasta, A.F. Using Remote Sensing data to identify iron deposits in central western Libya. In Proceedings of the International Conference on Emerging Trends in Computer and Image Processing (ICETCIP'2011), Bangkok, Thailand, 2011; pp. 56–61. Available online: <http://psrcentre.org/images/extraimages/122.%201211924.pdf> (accessed on 30 November 2021).
44. Deng, G.; Cahill, L.W.; Tobin, G.R. The study of logarithmic image processing model and its application to image enhancement. *IEEE Trans. Image Process.* **1995**, *4*, 506–512. [[CrossRef](#)]
45. Negru, M.; Nedeveschi, S.; Peter, R.I. Exponential image enhancement in daytime fog conditions. In Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), Qingdao, China, 8–11 October 2014; pp. 1675–1681.
46. Varma, M.K.S.; Rao, N.K.K.; Raju, K.K.; Varma, G.P.S. Pixel-Based Classification Using Support Vector Machine Classifier. In Proceedings of the 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, India, 27–28 February 2016; pp. 51–55. [[CrossRef](#)]