Automatic estimation of severity level for late blight in images of celery leaves through the firefly algorithm

Niriaska Perozo

niriaska.perozo@ucn.cl

Escuela de Ciencias Empresariales, Universidad Católica del Norte, Coquimbo, Chile

Wilfredo Angulo

antoniowilfredosanchezangulo@gamail.com

PUCE and YACHAY https://orcid.org/0000-0003-1908-9554

Research Article

Keywords: Late blight detection, Celery leaf segmentation, Firefly Algorithm optimization, Disease severity estimation.

Posted Date: October 3rd, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7768279/v1

Additional Declarations: The authors declare no competing interests.

Automatic estimation of severity level for late blight in images of celery leaves through the firefly algorithm

Niriaska Perozo[a,*], Paulimar Angulo[b], Jesús E. Linares[c], Wilfredo Angulo[d]

[a]*Escuela de Ciencias Empresariales, Universidad Católica del Norte, Coquimbo, Chile.*
[b]*Decanato de Ciencias y Tecnología. Universidad Centroccidental Lisandro Alvarado, Barquisimeto, Venezuela.*
[c]*Department of Computer Science, Auburn University at Montgomery, AL 36117, USA.*
[d]*Facultad de Ciencias Exactas, Naturales y Ambientales, Pontificia Universidad Católica del Ecuador, Laboratorio FRACTAL (Fractional Research Analysis, Convexity and Their Applications Laboratory), Quito, Ecuador.*

Abstract

Celery (*Apium graveolens* L. var. Dulce) is a vegetable valued for its flavor and nutritional properties, used both in culinary applications and traditional medicine for its health benefits. Its growing global demand has led to an increase in its cultivation, which is threatened by diseases such as late blight in leaves and stems. The detection and quantification of the severity of this disease are crucial to mitigate crop losses; however, visual observation is imprecise over large cultivation areas. In this work, we present an innovative method for the segmentation and classification of images of celery leaves affected by late blight, aiming to automate the estimation of severity levels. Using thresholded and adaptive segmentation based on the Firefly Algorithm, healthy and diseased areas in the images are delineated. The threshold is optimized through objective functions based on Tsallis, Kapur, and Otsu functions, resulting in an accuracy of 98.8% with Tsallis entropy. Additionally, a scale is proposed to clearly identify color variations associated with late blight, facilitating an accurate estimation of disease severity level.

Keywords: Late blight detection, Celery leaf segmentation, Firefly Algorithm optimization, Disease severity estimation.

1. Introduction

Celery (*Apium graveolens* L. var. Dulce) is a vegetable cultivated for its fleshy leaflets and petioles. It belongs to the Umbelliferae family and is a biennial herbaceous plant that can grow up to 50 cm in height. Its leaves are lobed, smooth, shiny, and range in color from deep green to yellowish green, with long,

*Corresponding author
Email address: `niriaska.perozo@ucn.cl` (Niriaska Perozo)

somewhat triangular petioles [31]. Late blight in celery is the main disease affecting this crop, caused by the fungus *Septoria apiicola* Speng, which attacks the foliage of the plant. Symptoms include chlorotic spots that later become necrotic (dark spots ranging from 3 mm to 10 mm in diameter) surrounded by a yellow chlorotic halo. Studies of the disease have shown that a combination of temperatures between 20 and 22.5°C and periods of relative humidity above 90% favor the development of the pathogen [30].

In intensive celery cultivation, the severity of this disease is a significant concern, leading farmers to establish management strategies, with the percentage of affected area per plant being their main concern. This percentage of affected area is referred to as the severity level of late blight, and its quantification allows farmers to take preventive measures to control the disease in new crops, through the application of controlled doses of fungicides and the implementation of integrated disease management practices. Knowing the severity level of late blight in celery plants also enables farmers to optimize production, avoid significant crop losses, and meet the quality and food safety standards established by regulatory authorities [55]. This aspect is particularly important for farmers selling their products in national and international markets. The method traditionally employed by farmers to detect and quantify the severity of diseases in their crops relies on direct visual observation of the plants or on images captured and digitized using relatively basic techniques. In that approach, the quantification of the severity level of late blight in celery plants necessarily relies on diagrams and severity scales created by experts [7]. These severity scales correlate structural damage with its impact on crop yield by assessing foliar spot development. They also allow for estimating the proportion of the affected area and identifying the crop's growth stage. However, this type of assessment may require the involvement of pathogen experts, leading to additional costs for farmers, and it is prone to inaccuracies due to human error and possible systematic errors in the digitization process [13], [21], [9], [24]. Furthermore, former method may not be effective when monitoring large crop areas, and it increases the sensitivity in the estimation of disease severity levels. For these reasons, visual computing techniques have become essential tools in the analysis and interpretation of images capturing disease severity [28], [48], [42].

Among the various applications of visual computing techniques for assessing disease severity in crops, image segmentation stands out as a prominent approach [1], [39]. Image segmentation generally involves dividing an image into multiple segments also referred to as regions or objects based on criteria such as grayscale intensity, color, shape, or texture [38], [40], [53]. This process enables the separation of plant images into areas affected by disease and those that remain healthy. For this purpose, image segmentation is implemented along with optimization of threshold values through metaheuristic algorithms based on collective intelligence, which provide good approximations, particularly when many images are not available [34]. The ability of collective intelligence algorithms to effectively explore and exploit the solution space enables the achievement of accurate and robust results even with small datasets [18]. Moreover, the adaptability and flexibility of these algorithms make them ideal for solving complex and variable problems [28], [3]. In this context, Viaji [45] proposed a method for segmenting images in the detection of diseases in sunflower plant leaves, based on the Particle Swarm Optimization (PSO) algorithm. The method demonstrated that it does not require prior knowledge of the number of segments, unlike other existing methods based on artificial intelligence heuristics (see for example [28], [2], [35], [52]). Furthermore, optimal results were obtained with low computational effort, and the algorithm proved to be efficient in recognizing and classifying diseases in sunflower leaves.

A similar algorithm to PSO is the Firefly Algorithm (FA) [41], [20]. This algorithm is inspired by the bioluminescence emitted by fireflies and their social behavior. The literature reports that FA has been successfully applied in threshold optimization for digital image segmentation and analysis, showing better results for edge detection compared to methods such as Canny, Sobel, Robert, and Prewitt [16]. Besides, FA has been used for optimization problems associated with medical image processing, including mammography, tomography, and retina images, aimed at noise reduction, contrast enhancement, edge detection, and image segmentation [46], [50], [49], [54], [37], [26], [36]. However, during this research, no works were found applying FA for threshold optimization in agricultural image segmentation or for the automatic detection of severity levels of foliar diseases such as late blight in celery leaves. Therefore, this work presents an innovative method for the segmentation and classification of images of celery leaves (*Apium graveolens* L. var. Dulce) affected by late blight, with the goal of automating the estimation of their severity levels, using FA to optimize the threshold value for the segmentation process adaptively for each image provided from a dataset.

The threshold-based segmentation method proposed in this study enables the automatic selection of a color intensity threshold to distinguish between healthy and diseased tissue in celery leaves. This threshold value is the optimal value determined by FA, considering the images as the search space of the associated optimization problem. Since the brightest regions of the image correspond to healthy tissue, pixel intensity is interpreted as the brightness of fireflies in the FA. This brightness attracts other fireflies and guides their movement toward these areas, thereby allowing the quantification of healthy tissue in the leaf. For this reason, appropriate objective functions for thresholding in the segmentation process using FA were selected, namely Tsallis, Kapur, and Otsu entropies [12], [27], [43]. Additionally, a neighborhood strategy is employed for edge detection and thresholding, also as part of the method's adaptability property.

The structure of this article is organized as follows: Section 2 provides background information on the detection of late blight as a foliar disease and the application of the FA in image processing. Section 3 presents a detailed description of the theoretical foundations of image segmentation by thresholding and optimization, the FA algorithm, the objective functions, and the Horsfall-Barratt scale. In Section 4, we introduce the proposed method for image segmentation by thresholding and optimization using the FA. Section 5 discusses the experimental results, where threshold values for the segmentation procedure were obtained using Tsallis, Kapur, and Otsu functions as objective functions in the FA. Finally, Section 6 presents the conclusions and outlines directions for future research.

2. Literature review

First, some *studies related to image processing using the FA* are reviewed:

Research on the use of meta-heuristic algorithms such as FA for digital image analysis tends to focus on improving efficiency and performance, focusing on lowering processing times and computational costs. In this sense, Vishwakarna and Yerpude [54] propose an image segmentation method based on FA, capable of segmenting gray-scale images with different types of noise. In this work, it was shown that FA has better performance in both execution time and recognition of noise signals in the image, giving an empirical basis for applying FA in the analysis of other types of images.

Das, Quarishi and Barman [11], applied FA to improve the quality of mammographic images, highlighting edge details and contrast by employing a comparative strategy between pixels, where each firefly is located with its adjacent neighbors. They then used other segmentation and clustering techniques to obtain a better mammographic image. Sunny and Pratheba [49], pplied FA for the classification of features found during the segmentation process of mammogram images, this, in order to support breast cancer detection. Gupta et al. [16], proposed a modification of FA with PSO (MFA), using a set of reference images and evaluating its performance in comparison with other techniques widely used to identify edges, such as Canny, Robert, Sobel and Prewitt. The results obtained indicate the superiority of MFA. Vishwakarmay Yerpude [54], proposed a new method using clustering techniques to improve the performance of FA based on the K-means Firefly Algorithm (KFA) technique. In the proposed procedure, the RGB image is converted to an L*A*B color space and then, to a grayscale image to perform histogram equalization and frequency calculation which served as input data for the KFA to cluster the points with similar frequencies. This hybridized method proves to be more optimal than K-means segmentation, requiring less execution time.

Rahebi and Hardalaç [36], propose a method for optic disc detection in human retinal images using FA. Because the optic disc is the brightest area of retinal images, they take the pixel value as the value of firefly light intensity to determine the most attractive individuals and thus, locate the optic disc. Based on the results obtained, they were able to determine that the firefly algorithm is an effective method to locate bright areas of images, considering the brightness of the pixel where each firefly is located as the light intensity of the insect. Moreover, other authors have highlighted in [12] that one of the main challenges in using the FA is the limited number of studies focused on color image processing.

Second, some *studies on automatic estimation of late blight in celery* is reviewed:

At the time of this review, no papers were found in which an automatic estimation of late blight is performed, nor the application of FA for the automatic detection of this disease. In fact, the use of computational tools for the study of this type of disease in tubers and other vegetables is limited only to the use of statistical tools, such as SAS [17], R language [14] to describe and infer aspects related to the general incidence of late blight in celery leaves. Image editing programs such as ImageJ [47] and ImageTool [6], are also used to facilitate the analysis of data obtained during an investigation of the level of severity caused by the disease. In this sense, Ortiz et. al [30] used the ImageJ program to measure the damage intensity of 55 leaves of celery, and then, estimated the number of classes and damage required by the Horsfall-Barratt system [19] used in plant pathology to evaluate the severity level of a disease.

As such, this timely review on the references related to FA for image segmentation and on the deficit of research related to the automated quantification of the severity level of late blight in celery leaves, reaffirms the method proposed in this work as a contribution to modern agriculture, supported by computer visualization and intelligent computing techniques for decision making on the detection and control of crop diseases.

3. Theoretical foundations

3.1. Image segmentation by thresholding and optimization

Image segmentation consists of dividing an image Ω into $s \in \mathbb{N}^*$ regions $\Omega_1, \Omega_2, \ldots,\ \Omega_s$ with homogeneous properties: the union of all regions yields the complete image, and none of them overlap [12]. One of the most commonly used techniques for segmentation is based on *thresholding*, which classifies pixels according to their intensity $f(x, y)$ using a function $g(x, y)$ [33]:

$$g(x,y) = \begin{cases} 255, & \text{if } f(x,y) > U, \\ 0, & \text{if } f(x,y) \leq U. \end{cases}$$

The threshold value U allows distinguishing regions of interest (such as areas affected by late blight on a celery leaf) from the background [15]. This technique can be extended to *multiclass* segmentation by using multiple thresholds for more precise separation. Threshold selection can be done through *parametric* approaches (such as Gaussian models [4]) or *non-parametric* ones [23], such as: maximizing Otsu's between-class variance [56]; maximizing Kapur's entropy [22]; and maximizing Tsallis' generalized entropy [51]. These non-parametric approaches require solving optimization problems, which are addressed in this work through the firefly algorithm (FA) detailed below.

3.2. Firefly algorithm (FA)

The FA, proposed by Yang in 2008 [57], is an optimization algorithm inspired by the bioluminescence behavior of fireflies. Each solution $\boldsymbol{x}$ represents a firefly whose quality is evaluated through an objective function $f_{obj}(\boldsymbol{x})$. Its light intensity I and attractiveness β decrease with the Euclidean distance between two fireflies located at $\boldsymbol{x}_i = (x_{i,1}, \ldots, x_{i,d})$ and $\boldsymbol{x}_j = (x_{j,1}, \ldots, x_{j,d})$, with $d \in \{1, 2, 3\}$, according to the following relations [57]:

$$I(r) = I_0 e^{-\gamma r_{ij}^2} \tag{1}$$

$$\beta(r) = \beta_0 e^{-\gamma r_{ij}^2} \tag{2}$$

where $r_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2$ is the Euclidean distance. Furthermore, the movement of a firefly toward a brighter one is modeled as:

$$\boldsymbol{x}_i^{n+1} = \boldsymbol{x}_i^n + \beta_0 e^{-\gamma r_{ij}^2} (\boldsymbol{x}_j^n - \boldsymbol{x}_i^n) + \alpha \boldsymbol{R}, \quad \forall\, n \in \mathbb{N}^*, \tag{3}$$

where $\boldsymbol{R}$ is a random vector with components $R_k \in [0, 1]$ for each $k = 1, 2, \ldots, d$, and α regulates the exploration capability of the algorithm [58].

From a computational perspective, the steps to implement the FA are outlined in Algorithm 1, where the symbol "$\triangleright$" can be either "$<$" or "$>$".

Algorithm 1 Firefly Algorithm (FA).

Set parameters $MaxGeneration$, N, α, β_0, and γ
Define the objective function $f_{obj}(\boldsymbol{x})$ for $\boldsymbol{x} \in X$
Randomly generate the initial population of fireflies $\mathcal{P}^n = \{\boldsymbol{x}_i^n \in X, i = 1, 2, \ldots, N\}$
Compute the light intensity I_i of each $\boldsymbol{x}_i^n \in \mathcal{P}^n$ using $f_{obj}(\boldsymbol{x}_i)$
while $n < MaxGeneration$ **do**
    for $i = 1$ to N **do**
        for $j = 1$ to N **do**
            if $I_j \triangleright I_i$ **then**
                Move firefly i towards firefly j in all d dimensions using (3)
            end if
            Vary the attractiveness with distance r_{ij} using (2)
            Evaluate new firefly populations $\mathcal{P}^{n+1}$ and update light intensities
        end for
        Rank fireflies and find the current best global solution
    end for
    $n = n + 1$
end while

3.3. Objective functions used with FA

The following objective functions are the three criteria evaluated and compared in this work using the firefly algorithm (FA) as a robust optimization strategy in the multiclass segmentation of images affected by late blight.

3.3.1. Multiclass Tsallis entropy function

Tsallis entropy is expressed as:

$$S_q(U) = \frac{1 - \sum_{k=1}^{U} [p_k(U)]^q}{q - 1} \tag{4}$$

where q is the entropic index. When $q \to 1$, the above expression satisfies Shannon entropy [51].

For i regions, a threshold vector $\boldsymbol{U} = (U_1, U_2, \ldots, U_{i-1})$ is determined with a pseudo-additive model, expressed by the following objective function:

$$J_{T,Mul}(\boldsymbol{U}) = \sum_k S_q^k(\boldsymbol{U}) + (1 - r) \prod_k S_q^k(\boldsymbol{U}). \tag{5}$$

The FA explores solutions $\boldsymbol{U}$ until finding the optimal threshold vector $\boldsymbol{U}^* = (U_1^*, U_2^*, \ldots, U_{i-1}^*)$, which is an element of the argument maximum of (5) [10, 37].

3.3.2. Otsu's between-class variance function

The objective function seeks to maximize the between-class variance given by [8]:

$$J_{O,Mul}(\boldsymbol{U}) = \sum_{k=1}^{i} P_k(\boldsymbol{U})[\bar{\mu}_k(\boldsymbol{U}) - \bar{\mu}_t]^2, \tag{6}$$

using the grayscale histogram, with P_k and $\bar{\mu}_k$ representing the probabilities and means of the classes, respectively. Similarly to the multiclass Tsallis entropy case, the FA finds the threshold vector $\boldsymbol{U}^*$ that maximizes this function.

3.3.3. Multiclass Kapur entropy function

Kapur's function maximizes the sum of entropies of each region, given by [25]:

$$J_{K,Mult}(\boldsymbol{U}) = \sum_{k=1}^{i} S_k^C(\boldsymbol{U}), \tag{7}$$

calculated from the normalized histogram. The FA is applied to optimize this function and determine, as in the two previous functions, the optimal threshold vector $\boldsymbol{U}^* = (U_1^*, U_2^*, \ldots, U_{i-1}^*)$ that is an element of the argument maximum of (7) [43].

3.4. Horsfall-Barratt scale

3.5. 3.4

The Horsfall-Barratt severity scale is a nonlinear visual system used in plant pathology to estimate the percentage of plant tissue affected by diseases. Developed in 1945, it divides severity into 12 classes based on percentage intervals of damaged leaf area, ranging from 0% to 100% (More Details about Horsfall-Barratt scale in [7]). This Horsfall-Barratt scale is still currently used, especially in studies requiring a rapid and standardized visual assessment of disease severity in plants. Although more precise methods using digital tools and image analysis software have emerged, this scale remains valid and practical, particularly under field conditions or when advanced technology is not available. In addition, it is useful in historical or comparative studies where methodological consistency with previous research is needed, thus facilitating the comparison of results over time. Since many classic works in plant pathology used this scale, its continued use ensures that disease severity evaluations remain compatible and comparable across different periods, trials, crops, or regions, which is fundamental for analyzing trends, evaluating the long-term effectiveness of treatments, or validating epidemiological models [32]; these are the reasons why it has been employed in this work.

4. Proposed method

Our method consists of four stages that are implemented using `MATLAB R2024b` and its associated libraries. The stages of the proposed method in this work are described in detail, along with a computational tool based on the FA, emphasizing the key aspects necessary for a clear understanding of the approach. In this regard, we have:

4.1. Stages of the proposed FA-based method with thresholding

Stage 1. **Initialization**: In this stage, the initial values of all the parameters of the FA are provided. The initialization strategy for these parameters consists of a calibration using images of healthy and diseased leaves of celery, the procedure of which is detailed in Appendix A. At this stage, images of celery from an image database reported by Ortiz et. al in [32] are used.

Stage 2. **Image pre-processing**: The pre-processing of the images consists of: adjusting the images to have the same size; background noise reduction and contrast enhancement. For background noise reduction, the image is "cleaned" by eliminating gray, violet, and blue tones using a neighborhood-based strategy. A pixel is considered background if at least six of its eight immediate neighbors are classified as such. Background pixels are then masked according to specific RGB rules: (a) equal R, G, and B values (gray tones); (b) red-dominant pixels ($R > G > B$); and (c) blue-dominant pixels ($B \gg R$ and G). Finally, a second pass masks any pixel surrounded by at least six already-masked neighbors, assigning an RGB value of 255 to all masked pixels.

Stage 3. **Image processing**: In this stage, the FA with Tsallis entropy is applied as the objective function for the segmentation by thresholding of the celery leaves. Due to the characteristics of the disease: yellow tones where the blue channel is minimal and dark spots where there is more blue tone, the segmentation is performed taking these characteristics as thresholds of the blue color channel (RGB). This allows better identification of all healthy parts of the leaf, i.e., where the area of the celery leaf is clearly green as the expected result of processing. Each of the operations proposed at this stage is described below:

- ***Identification of stems***: due to the particular characteristics of the crop and the disease evaluated, it is necessary to identify the stems with the purpose of eliminating them from the images and leaving only the leaves of the celery. The stems are identified considering that the shades of green are regularly lighter than those of the leaves [30]. This operation allows only the leaves where the level of severity of late blight is evident to be processed. However, for other crops where it is necessary to determine the level of severity in both stems and leaves, this step could be omitted.
- ***Obtaining ideal thresholds***: in this operation, the FA equipped with the Tsallis entropy function is executed to adaptively obtain the value of the optimal thresholds in the segmentation of each image. With the information obtained in this operation, the healthy and diseased areas of each celery leaf in the image are identified.
- ***Color image segmentation***: in this step the thresholds obtained in the previous step are used to mask the B component of the pixels of the whole leaf, the stems could be included when they are part of the study, using as search space the image with contrast enhancement obtained in the pre-processing. In this step only the B component of the pixels is masked for the following reasons:

  (a) Plant leaves contain chlorophyll, a pigment that absorbs blue and red light and reflects green, making us able to see this color. The disease to be measured is identified by yellow spots and wilting (brown shades or dark spots), which indicate a lack of chlorophyll in the diseased areas.

  (b) In general, healthy leaves have a higher contrast between green and blue pixels, meaning that there is a more significant difference between the green and blue values reflected by the leaves. In addition to this, focusing on the colors produced by the disease, yellow is characterized by higher values of red and green compared to blue; on the other hand, the brown color is predominantly red and green with a notable presence of blue. Therefore, it can be determined that the B component is the one that provides us with the most information for the extraction of the characteristics of interest in the image.

  The criteria for establishing the parameters for image segmentation are:

  (a) If the blue color value of the pixel is between 0 and the first threshold value obtained in the previous step, then the B component will have the value 0.

(b) If the blue color value of the pixel is between the value of the first and the second threshold, then the B component of the pixel takes the value of the first threshold.

(c) If the value of the blue color of the pixel is between the second and third threshold, then it will take the value of the second threshold and so on, until all the obtained thresholds are evaluated, the maximum range to be evaluated being between the last ideal threshold and 255.

Stage 4. **Post-processing and proposed severity scales**: This stage proposes the procedure for calculating the percentage of healthy and diseased areas, to quantify the level of severity of late blight and the automated establishment of the scales used to specify this level of severity. Specifically, the following operations are performed:

- ***Transformation of the segmented image to the HSV model***: this transformation allows the use of the Hue (H) of the model to identify more easily the different shades of green and the saturation (S) becomes a key factor for the identification of possibly healthy areas. This HSV color model is much more accurate than other existing models [29]. This accuracy makes it possible to distinguish between the characteristic green tones of young healthy leaves and those of mature, diseased leaves, which may appear similar. More details about HSV model can be found in [44].
- ***Identification of healthy and diseased areas***: this identification is performed considering the following two criteria:

  (a) Pixels whose R-channel value are greater than their G-channel values are identified, which indicates a clear indication of infection.

  (b) The degrees of "greenness" and "yellowness" are determined using Figure 1. These color intensities serve as classification rules for the health or disease level of the leaf. Based on this approach, the proposed severity levels for the disease are established.
- ***Quantification of the severity or health level in each leaf***: The number of pixels in each leaf region is calculated to estimate the percentage of highly healthy, healthy, possibly healthy, and diseased areas as shown in table 1. This classification is based on the scale proposed in this study, which is based on the Horsfall-Barrat system with the HSV model.

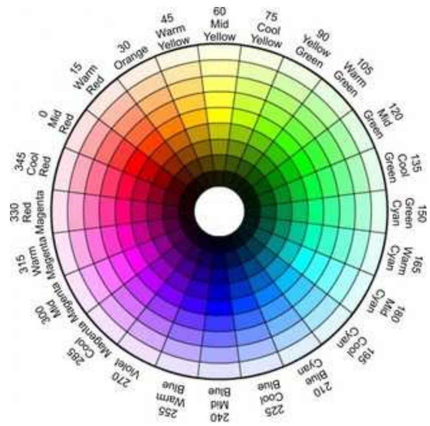

Figure 1: HSV (Hue, Saturation, Value) color model. Reproduced from [44].

Table 1: Proposed scale for leaf health and disease Severity: color markers for green and yellow intensity levels.

Level	H range	Classification
Yellow 1	From 30° to 45°	Diseased area
Yellow 2	From 45° to 61°	Diseased area
Yellow 3	From 61° to 65°	Possibly healthy area if saturation is less than 30
Yellow 4	From 65° to 70°	Possibly healthy area
Green 1	From 70° to 75°	Possibly healthy area
Green 2	From 75° to 90°	Possibly healthy area
Green 3	From 90° to 105°	Healthy area
Green 4	From 105° to 120°	Healthy area
Green 5	From 120° to 130°	Highly healthy area
Green 6	From 135° to 200°	Highly healthy area

5. Results

The values of all the parameters of the FA with Tsallis thresholding obtained through the implemented calibration process are shown in Table 2. Additional details are provided in Appendix A).

To evaluate the performance of the proposed method, the 55 celery leaf images reported by Ortiz et al. [32] were used. These images were obtained from a selection of young cv. Utah 52/70 celery leaves grown in a shade house under

Table 2: Parameters obtained after the calibration process.

Parameter	**Value**
Number of fireflies	150
Total iterations	1000
α	0.2
γ	1
Number of thresholds	5

controlled nutritional, irrigation, and shading conditions, allowing uniform plant development and free progression of the pathogen responsible for late blight.

To illustrate the results of the pre-processing stage, five randomly selected images are presented in Figure 2. The raw images already show signs of late blight in some leaves (dark spots) and healthy areas (green regions without discoloration). Notably, samples in images #16 and #39 exhibit a higher degree of severity. The enhancement of red and green channels further improves the visibility of affected areas, especially in samples where symptoms are subtle or barely perceptible such as images #26 and #55, with greater emphasis on the latter.

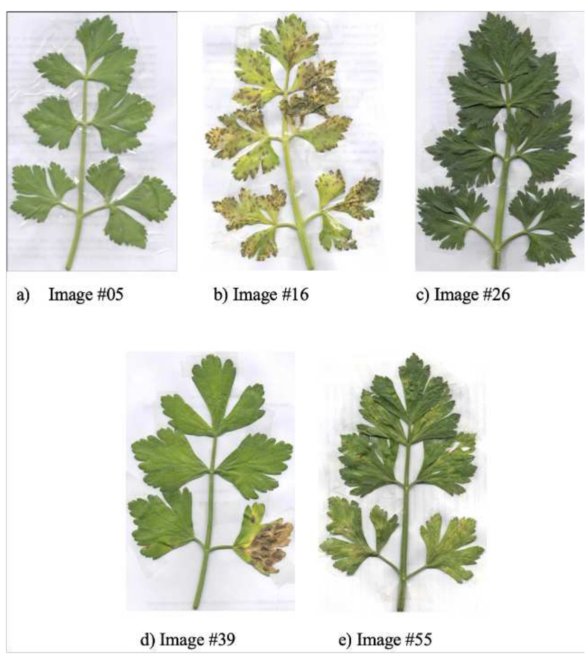

Figure 2: Images #05, #16, #26, #39 and #55 chosen from Ortiz et al. [32].

Figure 3 presents the processing results of the 5 images, selected to exemplify the outcomes obtained from the full set of 55 images after preprocessing. On the left side of the figure, the original images are shown; on the right, the segmentation results using the FA with Tsallis thresholding are displayed. As observed, masking the blue channel with the computed thresholds enables segmentation that highlights high-intensity regions in areas of the celery leaves where severe late blight symptoms are expected, corresponding to higher percentages on the proposed scale. Lower intensities are observed in regions with milder symptoms or in healthy leaf areas.

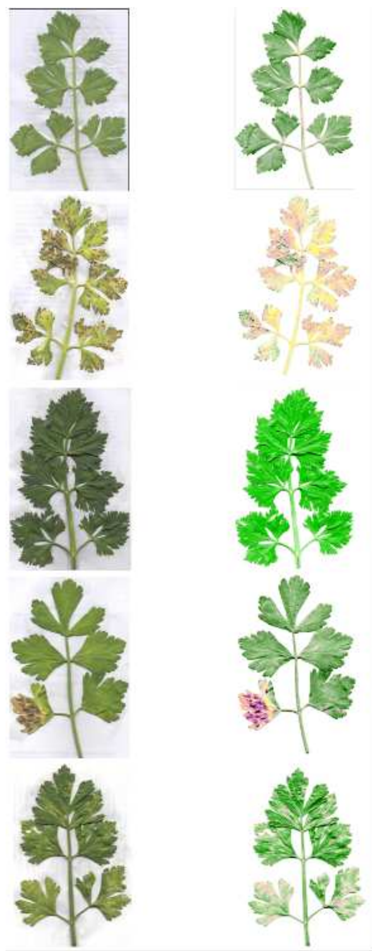

Figure 3: Original leaf images (left) and their output segmented images (right).

Once the algorithm parameters and classification rules are defined, the resulting percentages are compared with those proposed by Ortiz et al. [32] and the classifications presented in table 2. Additionally, each sample is assigned a Horsfall-Barratt class based on the detected percentage of disease.

In an initial evaluation, some leaves show notable differences between the disease percentages detected by the proposed method and those reported by Ortiz et al.[32], with an average discrepancy of 5.97%. A detailed analysis of each classification category reveals that, in some cases, Ortiz et al. classify certain regions as possibly healthy, whereas the HSV color model identifies them as yellow indicating early disease. In other cases, regions considered possibly healthy are instead classified as diseased by this proposal. These discrepancies arise in leaves where healthy and diseased areas are in very close proximity. Figures4 and 5 show the results without and with the corresponding adjustment for this situation, respectively.

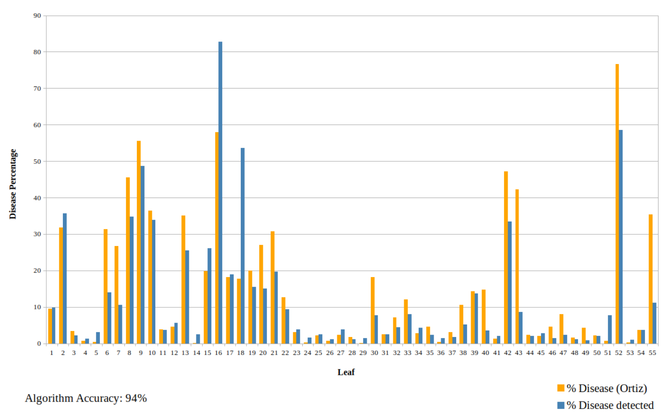

Figure 4: Comparison of leaf disease percentages without adjustment for close proximity between healthy and diseased areas using Tsallis thresholding.

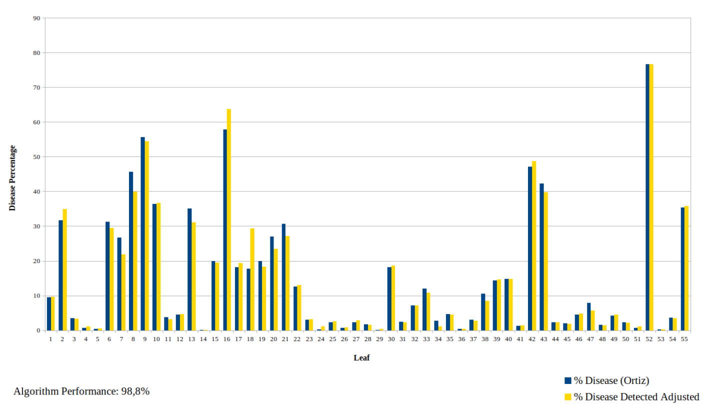

Figure 5: Comparison of leaf disease percentages with adjustment for close proximity between healthy and diseased areas using Tsallis thresholding.

Thus, the difference between the percentages reported by Ortiz et al. [32] and those detected by the proposed method is reduced to just 1.19%. Additionally, 48 leaves are classified in the same Horsfall-Barratt class, with only 7 leaves differing by a single class (Figure 6).

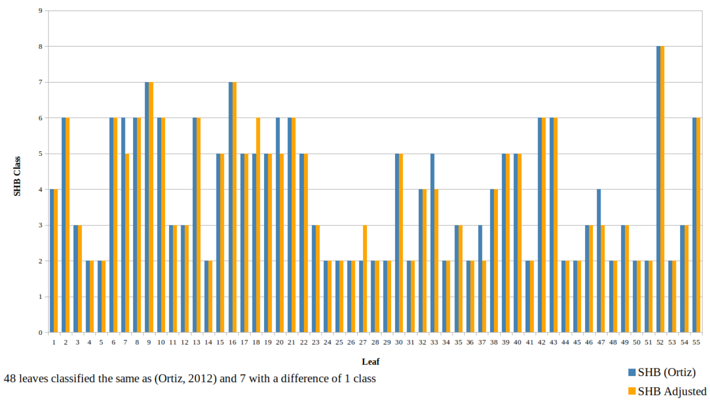

Figure 6: Class-Level Comparison Based on the Horsfall-Barratt System Using Tsallis Thresholding.

Through this evaluation, the method presented in this study is shown to automatically segment celery leaf images and accurately identify healthy, highly healthy, possibly healthy, possibly diseased, and diseased areas according to the proposed scale (Table 1). It also calculates the percentage of disease in celery leaves affected by late blight (Figure 7).

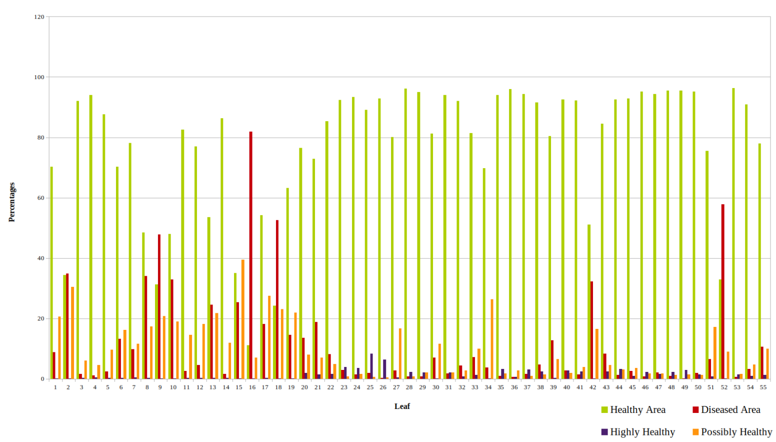

Figure 7: Percentages of healthy and diseased areas according to the proposed classification scale.

Validation of the results using Otsu's and Kapur's entropy as objective function

The results obtained using the Tsallis entropy as the objective function in the FA were validated by repeating the entire process using the Otsu entropy and then the Kapur entropy as alternative objective functions. This comparison was conducted based on reports that Kapur entropy offers faster convergence in image segmentation, while Otsu entropy, although more computationally intensive, yields efficient results when combined with the FA [12].

Table 3 presents the initial parameters for the FA with thresholding under the three objective functions, retaining the values calibrated with Tsallis entropy. It also includes the detection performance percentages, using the results reported by Ortiz et al. [32].

Table 3: Initial parameters for the FA applying the Tsallis, Kapur and Otsu objective functions.

Parameter	**Tsallis**	**Kapur**	**Otsu**
Number of thresholds	5	7	4
Iterations	1000	750	1000
Number of firefly	150	200	100
α	0,2	0,2	0,2
γ	1	1	1
Performance (Detection percentage compared to Ortiz et.al. 2012)	98,8%	98,6%	98,5%

Figure 8 shows the comparison of the leaf disease percentages obtained using Kapur, Otsu and Tsallis entropies as objective function, contrasted with the results reported by Ortiz et al. [32]. The results demonstrate that Tsallis entropy is particularly effective for adaptive segmentation of real and color images, achieving thresholds in a fast and accurate manner. With an accuracy of 98.8%, it significantly outperforms both Kapur and Otsu entropies when used as objective functions in the FA.

These findings align with previous research [27], which also highlights the superior performance of Tsallis entropy in experimental and grayscale image segmentation. Although Kapur entropy offers faster convergence, it does not match the accuracy of Tsallis. Otsu entropy, while requiring fewer thresholds to approximate the results reported by Ortiz et al. [32], yields the lowest accuracy among the three. These results underscore the effectiveness of Tsallis entropy as the most suitable choice for accurate and adaptive segmentation in this type of color imagery.

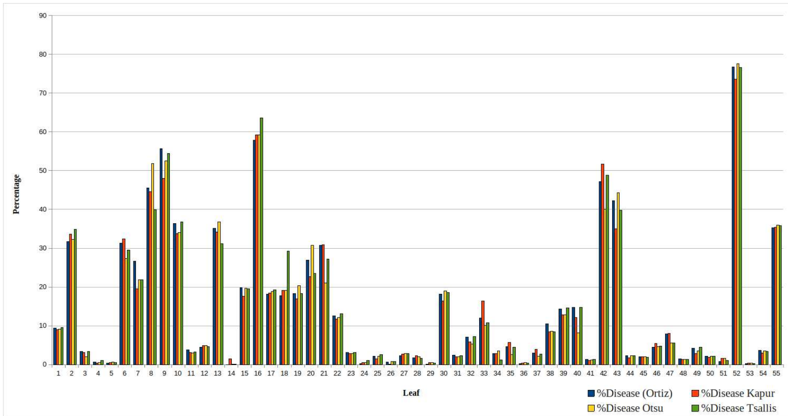

Figure 8: Comparison of leaf disease percentages with Kapur's, Otsu's and Tsallis' entropies.

6. Conclusions and future work

This work presents an innovative method based on the FA using Tsallis, Kapur and Otsu entropies as objective functions for the segmentation of color images of celery leaves affected by late blight. The severity level of the disease (i.e. the percentage of infection in each leaf), along with its classification in the Horsfall-Barratt System and the proposed scale, is accurately estimated. The resulting segmented images clearly highlight color variations associated with the presence of late blight in celery leaves.

On the other hand, the segmentation of color images achieved by the proposed method represents a significant contribution, as most published works in the field rely on grayscale image segmentation to reduce processing complexity. To address this, the method's implementation as a MATLAB-based computational tool allows users to extract additional information based on differences in green and yellow shades, which serve as indicators of varying degrees of disease presence in the leaf. This goes beyond binary segmentation of healthy versus diseased areas, leveraging color information to capture more nuanced gradations.

The results demonstrate that Tsallis entropy enables fast and efficient threshold selection by adaptively segmenting each image. Moreover, it achieves higher accuracy (98.8%) compared to Kapur and Otsu entropies when used as objective functions in the FA. These findings are consistent with those reported in related studies [27] on experimental and grayscale images, where Tsallis entropy outperforms the others, while Kapur entropy is noted for faster convergence. In contrast, although Otsu entropy requires fewer thresholds to produce results that are relatively close to those reported by Ortiz et al. [32], it yields the lowest accuracy among the three.

Finally, as future work, it is proposed to apply and evaluate the FA with Tsallis entropy for the segmentation of images from other crops and foliar diseases with different color characteristics. Additionally, the development of a generic mobile application is proposed, capable of quantifying structural damage across various crop types in real time.

References

[1] Abbasi, R., Martinez, P., Ahmad, R., 2023. Crop diagnostic system: A robust disease detection and management system for leafy green crops grown in an aquaponics facility. Artificial Intelligence in Agriculture 10, 1–12.

[2] Abd Algani, Y.M., Caro, O.J.M., Bravo, L.M.R., Kaur, C., Al Ansari, M.S., Bala, B.K., 2023. Leaf disease identification and classification using optimized deep learning. Measurement: Sensors 25, 100643.

[3] Al-Qudah, R., Almuhajri, M., Suen, C.Y., 2025. Unveiling the potential of sustainable agriculture: A comprehensive survey on the advancement of ai and sensory data for smart greenhouses. Computers and Electronics in Agriculture 229, 109721.

[4] Arora, S., Acharya, J., Verma, A., Panigrahi, P.K., 2008. Multilevel thresholding for image segmentation through a fast statistical recursive algorithm. Pattern Recognition Letters 29, 119–125.

[5] Arora, S., Singh, S., 2013. The firefly optimization algorithm: convergence analysis and parameter selection. International Journal of Computer Applications 69.

[6] Baker, S.A., Gardner, S.D., Rogers, M.L., Sanders, F., Tunnell, T.W., 2001. Imagetool, in: High-Speed Imaging and Sequence Analysis III, SPIE. pp. 1–9.

[7] Campbell, C.L., Madden, L.V., 1990. Introduction to Plant Disease Epidemiology. John Wiley & Sons. URL: `https://ppc.nwafu.edu.cn/docs/2016-06/20160628142315445651.pdf`.

[8] Chakraborty, J., Mukhopadhyay, S., Singla, V., Khandelwal, N., Rangayyan, R.M., 2012. Detection of masses in mammograms using region growing controlled by multilevel thresholding, in: 2012 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS), IEEE. pp. 1–6.

[9] Chaudhary, A., Thakur, R., Kolhe, S., Kamal, R., 2020. A particle swarm optimization based ensemble for vegetable crop disease recognition. Computers and Electronics in Agriculture 178, 105747.

[10] Cuevas, A., Diaz-Ramirez, V.H., Kober, V., Trujillo, L., 2014. Facial recognition using composite correlation filters designed by multiobjective combinatorial optimization, in: Applications of Digital Image Processing XXXVII, SPIE. pp. 268–275.

[11] Das, G., Quraishi, M.I., Barman, M., 2013. Firefly algorithm based mammographic image analysis. Asian Journal of Computer Science and Information Technology 3, 56–59.

[12] Dey, N., et al., 2020. Applications of firefly algorithm and its variants. Springer, USA.

[13] Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. Computers and electronics in agriculture 145, 311–318.

[14] Giorgi, F.M., Ceraolo, C., Mercatelli, D., 2022. The r language: an engine for bioinformatics and data science. Life 12, 648.

[15] Gonzalez, R.C., E, W.R., 2008. Digital image processing. Prentice Hall, USA.

[16] Gupta, M., Tazi, S.N., Jain, A., Deepika, 2014. Edge detection using modified firefly algorithm, in: Sixth International Conference on Computational Intelligence and Communication Networks, pp. 1–6. URL: https://www.researchgate.net/publication/282740295.

[17] Hanrahan, L.P., Mirkin, I., Olson, H.A., Fiore, B.J., 1990. Smrfit: a statistical analysis system (sas) program for standardized mortality ratio analyses and poisson regression model fits in community disease cluster investigations. American journal of epidemiology 132, 116–122.

[18] Hassanien, A.E., Gaber, T., Mokhtar, U., Hefny, H., 2017. An improved moth flame optimization algorithm based on rough sets for tomato diseases detection. Computers and electronics in agriculture 136, 86–96.

[19] Horsfall, J.G., Barratt, R.W., 1945. An improved grading system for measuring plant diseases. Phytopathology 35, 655. URL: https://www.cabidigitallibrary.org/doi/abs/10.5555/19461100104.

[20] Hsieh, F.S., 2024. Comparison of a hybrid firefly–particle swarm optimization algorithm with six hybrid firefly–differential evolution algorithms and an effective cost-saving allocation method for ridesharing recommendation systems. Electronics 13, 324.

[21] Hu, G., Yang, X., Zhang, Y., Wan, M., 2019. Identification of tea leaf diseases by using an improved deep convolutional neural network. Sustainable Computing: Informatics and Systems 24, 100353.

[22] Kapur, J.N., Sahoo, P.K., Wong, A.K., 1985. A new method for gray-level picture thresholding using the entropy of the histogram. Computer vision, graphics, and image processing 29, 273–285.

[23] Kumar, A., Kumar, A., Vishwakarma, A., Singh, G.K., 2022. Multilevel thresholding for crop image segmentation based on recursive minimum cross entropy using a swarm-based technique. Computers and Electronics in Agriculture 203, 107488.

[24] Kundu, N., Rani, G., Dhaka, V.S., Gupta, K., Nayaka, S.C., Vocaturo, E., Zumpano, E., 2022. Disease detection, severity prediction, and crop loss estimation in maizecrop using deep learning. Artificial intelligence in agriculture 6, 276–291.

[25] Mahendran, N., Vincent, P.D.R., 2024. Alzheimer's disease classification via pre-trained convolutional neural network variants utilizing multi-level thresholding based on kapur's entropy, in: 2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON), IEEE. pp. 1–6.

[26] Malathi, K., Nedunchelian, R., 2016. Detecting and classifying diabetic retinopathy in fundus retina images using artificial neural networks-based firefly clustering algorithm. ARPN J Eng Appl Sci 11, 3419–3426.

[27] Manic, K.S., Priya, R.K., Rajinikanth, V., 2016. Image multithresholding based on kapur/tsallis entropy and firefly algorithm. Indian Journal of Science and Technology 9, 89949.

[28] Metre, V.A., Sawarkar, S.D., 2021. Research review on plant leaf disease detection utilizing swarm intelligence. Turkish Journal of Computer and Mathematics Education 12, 177–185.

[29] Noor, A.I., Mokhtar, M.H., Rafiqul, Z.K., Pramod, K.M., 2012. Understanding color models: a review. ARPN Journal of science and technology 2, 265–275.

[30] Ortiz, F., Rodríguez, D., Sanabria, M., 2013. Etiología de tizón tardío del apio (apium graveolens l. var. dulce) en las zonas altas del estado táchira, venezuela. Fitopatología Colombiana 37, 1–6. URL: `https://www.researchgate.net/publication/318469382`.

[31] Ortiz, F., Rodríguez, D., Sanabria, M., Pineda, J., 2016. Manejo del tizón tardío del célery (apium graveolens l. var. dulce) usando extractos vegetales bajo condiciones controladas. Bioagro 28, 87–94. URL: `https://ve.scielo.org/scielo.php?pid=S1316-33612016000200003&script=sci_arttext`.

[32] Ortiz, F., Ulacio, D., Rodríguez, D., Urdaneta, L., 2012. Validación y comparación de dos escalas diagramáticas para la medición del tizón tardío del apio (apium graveolens l. var. dulce). Fitopatología Colombiana 36, 37–40.

[33] Palomino, N.L.S., Concha, U.N.R., 2009. Técnicas de segmentación en procesamiento digital de imágenes. Revista de investigación de Sistemas e Informática 6, 9–16.

[34] Prakash, N., Balaji, V., 2021. Detection of plant disease using swarm intelligence optimization algorithm, in: 2021 International conference on

advancements in electrical, electronics, communication, computing and automation (ICAECA), IEEE. pp. 1–5.

[35] Qadri, S.A.A., Huang, N.F., Wani, T.M., Bhat, S.A., 2024. Advances and challenges in computer vision for image-based plant disease detection: a comprehensive survey of machine and deep learning approaches. IEEE Transactions on Automation Science and Engineering .

[36] Rahebi, J., Hardalaç, F., 2016. A new approach to optic disc detection in human retinal images using the firefly algorithm. Medical & biological engineering & computing 54, 453–461.

[37] Raja, N.S.M., Lakshmi, P.V., Gunasekaran, K.P., 2018. Firefly algorithm-assisted segmentation of brain regions using tsallis entropy and markov random field, in: Innovations in Electronics and Communication Engineering: Proceedings of the Fifth ICIECE 2016, Springer. pp. 229–237.

[38] Resma, K.B., Nair, M.S., 2018. Multilevel thresholding for image segmentation using krill herd optimization algorithm. Journal of king saud university-computer and information sciences .

[39] Ritharson, P.I., Raimond, K., Mary, X.A., Robert, J.E., et al., 2024. Deeprice: A deep learning and deep feature based classification of rice leaf disease subtypes. Artificial Intelligence in Agriculture 11, 34–49.

[40] Rodríguez-Esparza, E., Zanella-Calzada, L.A., Oliva, D., Heidari, A.A., Zaldivar, D., Pérez-Cisneros, M., Foong, L.K., 2020. An efficient harris hawks-inspired image segmentation method. Expert Systems with Applications 155, 113428.

[41] Sanmorino, A., 2018. A brief comparison of particle swarm optimization algorithm and firefly algorithm. Jurnal Informatika Upgris 4.

[42] Sengupta, S., Das, A.K., 2017. Particle swarm optimization based incremental classifier design for rice disease prediction. Computers and Electronics in Agriculture 140, 443–451.

[43] Sharma, A., Chaturvedi, R., Kumar, S., Dwivedi, U.K., 2020. Multi-level image thresholding based on kapur and tsallis entropy using firefly algorithm. Journal of Interdisciplinary Mathematics 23, 563–571.

[44] Singh, A.K., Swarup, A., Agarwal, A., Singh, D., 2019. Vision based rail track extraction and monitoring through drone imagery. Ict Express 5, 250–255.

[45] Singh, V., 2019. Sunflower leaf diseases detection using image segmentation based on particle swarm optimization. Artificial Intelligence in Agriculture 3, 62–68.

[46] Sivaramakrishnan, A., Karman, M., 2014. Medical image segmentation using firefly algorithm and enhanced bee colony optimization, in: International Conference on Information and Image Processing, pp. 316–321. URL: https://www.researchgate.net/publication/326439428.

[47] Sparavigna, A.C., et al., 2020. On the use of imagej segmentation. Zenodo Repos 1, e11.

[48] Su, W.H., 2020. Crop plant signaling for real-time plant identification in smart farm: A systematic review and new concept in artificial intelligence for automated weed control. Artificial Intelligence in Agriculture 4, 262–271.

[49] Sunny, S., Pratheba, M., 2014. Detection of breast cancer using firefly algorithm. International Journal of Emerging Technologies and Engineering (IJETE) Vol 1.

[50] Synthuja, M., Padmasuresh, L., Bosco, J., 2016. Firefly based region growing and region merging for image segmentation, in: International Conference on Emerging Technological Trends, pp. 1–9. URL: https://www.researchgate.net/publication/287469226.

[51] Tsallis, C., 1988. Possible generalization of boltzmann-gibbs statistics. Journal of statistical physics 52, 479–487.

[52] Upadhyay, A., Chandel, N.S., Singh, K.P., Chakraborty, S.K., Nandede, B.M., Kumar, M., Subeesh, A., Upendar, K., Salem, A., Elbeltagi, A., 2025. Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture. Artificial Intelligence Review 58, 1–64.

[53] VijayaLakshmi, B., Mohan, V., 2016. Kernel-based pso and frvm: An automatic plant leaf type detection using texture, shape, and color features. Computers and Electronics in Agriculture 125, 99–112.

[54] Vishwakarma, B., Yerpude, A., 2014. A new method for noisy image segmentation using firefly algorithm. Int. J. Sci. Res.(IJSR) 3, 1721–1725.

[55] Wang, X.J., Luo, Q., Li, T., Meng, P.H., Pu, Y.T., Liu, J.X., Zhang, J., Liu, H., Tan, G.F., Xiong, A.S., 2022. Origin, evolution, breeding, and omics of apiaceae: a family of vegetables and medicinal plants. Horticulture Research 9, uhac076.

[56] Xu, X., Xu, S., Jin, L., Song, E., 2011. Characteristic analysis of otsu threshold and its applications. Pattern recognition letters 32, 956–961.

[57] Yang, X.S., 2010. Nature-inspired metaheuristic algorithms. Luniver press, UK.

[58] Yang, X.S., He, X., 2013. Firefly algorithm: recent advances and applications. International journal of swarm intelligence 1, 36–50.

Appendix A. Calibration of initial parameters using healthy leaves.

The parameters of the Firefly Algorithm (FA) with Tsallis entropy include the number of fireflies, the total number of iterations, the randomization factor α, the light absorption coefficient γ, and the number of thresholds within a specified range. The initial values of these parameters are critical, as the algorithm must converge rapidly while avoiding premature convergence to infeasible local optima. Therefore, proper parameter initialization is essential and requires a thorough testing and calibration process.

In this work, suitable initial values for the FA with Tsallis entropy were determined through an empirical calibration procedure, using the baseline values suggested by Yang [5] and employing 55 celery leaf images. The algorithm was run 25 times for each parameter combination listed in Table A.4, until an acceptable convergence rate was achieved. Once the optimal combination was identified, these parameters were used in tests on leaves with unknown levels of late blight severity.

Table A.4: Parameters used in calibration with healthy leaves.

Parameter	**Range of values to be tested**
Number of fireflies	[100, 150, 200]
Total iterations	[750, 1000, 1500]
α	[0.1, 0.2]
γ	[0.5, 1, 2]
Tsallis thresholds	[5, 6, 7]

Threshold value variation

The number of thresholds to be tested-5, 6, and 7-was determined based on the visual differentiation of healthy and diseased areas in celery leaves, as well as the characteristics of late blight. After testing, five thresholds were found to be ideal, with values distributed across the range (0, 255). Regarding the behavior of the disease in the images, the following criteria were established: (a) In dark spots, where late blight symptoms are most severe, the affected pixels appear very dark or nearly black and are assigned the lowest threshold values; (b) In yellowing areas of the leaf, pixels are brighter due to higher red and green components and lower blue values, thus receiving the highest threshold values; and (c) In healthy green regions, there is no fixed rule since the green tone varies depending on each leaf. Typically, healthy areas exhibit a dominant green component over red and blue, but the final classification relies on the shades of green detected via the HSV model and the classification system proposed in this work. As such, threshold assignment in green areas may vary across the entire range of threshold values.

Variation of the total number of iterations

The total number of iterations selected is closely related to the number of fireflies used in the FA; the greater the number of fireflies, the more iterations are required to achieve an adequate convergence rate. During testing, it was observed that for the image sizes used in this study, fewer than 750 iterations resulted in highly variable threshold outputs across tests, indicating insufficient convergence. Conversely, when more than 3000 iterations were tested, results tended to repeat over many iterations, increasing processing time without improving convergence. Based on these observations, the optimal range for the total number of iterations was determined to be between 750 and 1500, particularly when using between 100 and 200 fireflies.

Variation of the number of fireflies

The values considered for the number of fireflies were: 100, 150 and 200. After the 25 runs performed, the best value obtained is 150 fireflies.

Variation of α and γ parameters

The parameter α is dynamically adjusted during iterations. Considering the image size, the previously selected number of thresholds, and the balance between the number of fireflies and total iterations, the most suitable range for α was found to be between 0.1 and 0.2 when using images of healthy leaves, with optimal convergence achieved at 0.2. Similarly, under the same calibration conditions, the parameter γ was tested and consistently maintained a value of 1 across all scenarios involving different numbers of fireflies and iterations. Therefore, a value of 1 was selected as the initial setting for γ. Table 2 summarizes the final values obtained for each of the initial parameters through this calibration process.

Validation of initial parameters obtained using diseased leaves

To validate the initial parameter values obtained through calibration, 55 images of celery leaves with varying degrees of late blight damage taken from reference [32] were used. Each image was successfully segmented using the five ideal thresholds, clearly distinguishing green regions (indicative of healthy tissue), dark spots (associated with disease), and yellow areas (suggesting potential infection). The disease percentages reported by Ortiz et al. [32], along with the expected values for each leaf based on the Horsfall-Barratt system and the HSV color model, were used to define the classification rules for healthy and diseased regions in the segmented images processed with the FA using Tsallis entropy, as summarized in Table 2.