



Original papers

Method for maize plants counting and crop evaluation based on multispectral images analysis

Wilbur N. Chiuyari Veramendi*, Paulo E. Cruvinel

*Embrapa Instrumentation, São Carlos, SP, Brazil**Federal University of Sao Carlos – Post Graduation Program in Computer Science São Carlos, SP, Brazil*

ARTICLE INFO

Keywords:

Digital image processing
Plant counting
Distance transform
Template matching
Multispectral images
SVM

ABSTRACT

The processing of multispectral images acquired with embedded cameras in unmanned aerial vehicles (drones) has brought new opportunities for precision agriculture. In this study a method for evaluating the number of corn plants (*Zea mays L*) in a crop area is presented. Plant density is one of the most important yield factors, yet its precise measurement after the emergence of plants is impractical in large and medium-scale production, since significant amount of labor is required. For validation, a dataset of spectral images was gathered from flights over an agricultural area, and digital image processing techniques were applied, taking into account the concept of intelligent processing. Therefore, pattern recognition and models to aid decision-making through machine learning were also used. After image acquisition, the processing of orthomosaics in the spectral channels, i.e., red (R), green (G), and blue (B), was performed, making it possible to register and organize all the images. Likewise, techniques for geometric transformation, brightness, and contrast adjustments were evaluated globally, whereas local adjustments were evaluated based on the use of adaptive equalization techniques, which were explored based in the choice of the HSV color space. For the post-processing step, segmentation based on the best observed color threshold technique, in conjunction with Gaussian filtering and morphological operations, were considered. To enable pattern recognition, techniques that use distance maps were evaluated, considering the use of Euclidean distance. Thus, the locations of canopy patterns in maize plants were studied using a template matching algorithm and Chamfer pattern mask. For feature extraction, chain code and circular pattern map techniques were considered. The analyses made it possible to establish vectors of features based on patterns related to the number of maize plants occurrences. Finally, three calibration steps were considered, one related to the plant height versus the canopy opening radius, other related to the number of maize plants for each position in the crop area versus the radii identified by the developed model, and the third related to the cross-correlation between the plant counting by human vision and the new method. In addition, the classification step was established using a set of classifiers based on support vector machine (SVM). Results have shown an accurate and timely counting methodology for maize plants, which can guide cultivation to ensure high yield. The results showed that as a new method it can effectively count the number of maize plants with an average accuracy rate equal to 88.47%. Besides, both selected SVM classifiers have presented accuracy higher than 84% and precision higher than 83%. Furthermore, the cross-correlation between the plant counting by human vision and the new method has presented a linear correlation coefficient equal to 0.98. Thus, the developed method proved to be adequate for counting the maize plants in the post-emergence stage.

1. Introduction

Increases in the world population have magnified the challenges to food security. Prospective studies have indicated that there will be a need to increase agricultural production by approximately 70% to feed the inhabitants of the planet in 2050 (Popp, 2013). The United Nations

Food and Agriculture Organization (FAO) reports that in recent years, food security has been facing challenges due to many factors, such as climate change and economic recessions, among others (FAO et al., 2020). The production of maize, rice and wheat have been observed with special attention, because of their important role in food security

* Corresponding author at: Federal University of Sao Carlos – Post Graduation Program in Computer Science São Carlos, SP, Brazil
E-mail addresses: wilbur.chiuyari@estudante.ufscar.br (W.N.C. Veramendi), paulo.cruvinel@embrapa.br (P.E. Cruvinel).

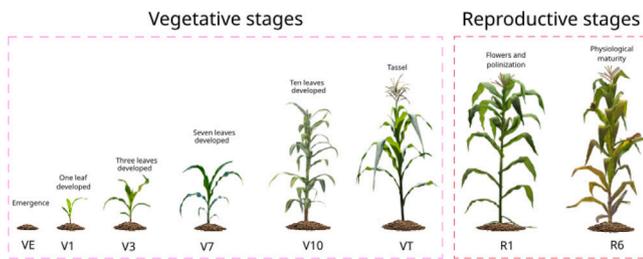


Fig. 1. Maize growth stages or phenological phases.

compared to those played by all other grains, according to the Brazilian National Supply Company (CONAB) report (Laborde et al., 2020).

One of the most important pieces of information needed to improve production in crop fields of maize (*Zea mays* L.) is related to the quantity and size of the maize plants. This information, known as plant density (Liu et al., 2017), is used for planning and estimating growth and for the maintenance of crop fields (Koc-San et al., 2018). The first traditional methods of counting plants (Pound and Clements, 1898) systematically use an area measurable according to their population, which in the state of the art have different counting variations (Oliver and Tansley, 1904; Weaver, 1918). Muller has shown that maize culture presents an exponential behavior in terms of growing for the different phenological stages (Müller et al., 2005). Likewise, systematic studies of maize phenology (Ritchie and Hanway, 1989) define the stages of both growth and reproduction (Fig. 1).

In recent years, there has been a greater use of advanced technology in agriculture, thus resulting in the emergence of applied technologies such as precision agriculture (PA), which minimizes losses and increases the productivity of crops; and remote sensing, which enables the application of non-invasive methods in agriculture through the use of optical instruments, radar (Hasituya et al., 2020), and satellites (Ahmad et al., 2018), among others. However, these solutions may be limited by the presence of clouds (Su et al., 2020), and thus, low-altitude remote sensing methods, including through the use of drones, have become increasingly employed non-invasive PA, albeit at smaller scales. On other hand, low-altitude remote sensing may encounter challenges in the evaluation of plant density, such as dealing with occlusions, appearances, lighting variations, and different scales.

Many researchers have been using a variety of models to estimate plant density, and these studies can be categorized into three approaches: identification, counting, and estimation of phenological phases. These models, in turn, are based on different digital image processing (DIP) techniques, including: correlation analysis (Cruvinel and Minatel, 2002), color spaces (Gnädinger and Schmidhalter, 2017), regions of interest (ROI) (Koc-San et al., 2018), spectral signatures (Hall et al., 2018), number of leaves (Praveen Kumar and Domnic, 2019), conic sections for canopy estimation (Veramendi and Cruvinel, 2021), and vegetation indices (Bah et al., 2020), among others. On the other hand, estimating the heights of plants or trees on farms using drones is based on techniques such as three-dimensional point clouds using LIDAR sensors (Malambo et al., 2018), point clouds using RGB-D camera (Qiu et al., 2022), digital elevation models (DEM) using multispectral sensors (Johansen et al., 2018), sun elevation angles with azimuth values (Koc-San et al., 2018), among others.

Recent advances in drone technology have exhibited the properties of cameras at different bands, such as in multispectral (Marinello, 2017) and hyperspectral (Lu et al., 2020) images. In one study, a method for estimating the densities of wheat plants in the emergence stage using images obtained by drone at altitudes of 3 m, 5 m and 7 m. Their experiments showed that the ground resolution of the images should be greater than 0.40 mm to facilitate the classification of green pixels. In Pallottino et al. (2021), the use of low-cost drones in the phenotyping of cereals enabled information about the actual color (with a lower

accuracy of 12/256) to be obtained at a height of 15 m. For the above mentioned methods the hyperspectral sensors normally generate large amount of data, and should be used only when necessary. Such a condition may increase computational cost. Additionally, by using low-altitude drones flights may involve increasing in battery consumption, as well as limitation for data acquisition in large crop area, which can be disadvantages.

A mobile application for the classification of plants and soils (Hernández-Hernández et al., 2017) integrated multiple computer vision techniques to be able to segment of plants in crop images. The segmentation algorithm was based on rectangular markers, with color information and probabilistic histograms. The counting objects were based on connected components, and the accuracy of the rectangle detection algorithm was evaluated to be below 0.7 cm, with a maximum error of 1.64 cm. Moreover, and the cut areas had an average of 98.3% intersection with the correct areas. In Wang et al. (2018), a segmentation method was developed based on the Chan–Vese model and the Sobel operator, using a feature that identifies tones with relatively high levels of green to extract the leaf region, and implementing the Chan–Vese model and the improved Sobel operator to extract the contours of the leaves. With regard to its efficiency, the algorithm exhibited an average error rate of 0.0428, which corresponds to a decrease of 6.54% in relation to the average error rate of the level method established by the authors. In Khan et al. (2019), the use of Clifford’s geometric algebra was proposed to improve segmented images acquired by drones for different agricultural fields, overcoming limitations related to the RGB space. Specifically, the authors used the subalgebra called Clifford quaternions for foreground and background segmentation, in the form of a 3×3 quaternion mask. In Riehle et al. (2020), an algorithm that performed robust automatic background and plant segmentation under various imaging conditions was developed. The algorithm was based on a vegetation index method for approximate pre-segmentation. This first approximation was used to calculate the threshold value for segmentation in conjunction with the CIE Lab color space. The accuracy of the algorithm reached 97.4%. According to the presented segmentation methods, challenges still persist for plants segmentation. In fact, there are overlaps among the leaves within a single color space. Consequently, it is relevant to explore other color spaces that can be suitable for analysis related to captured reflectance.

An algorithm for counting oil palms using template matching (TM) and cross-correlation using treetop templates of drone images was proposed in Kalantar et al. (2017), were a reduction in the estimation error of approximately 27% of the trees counted was obtained. In Xie et al. (2019), a new extended phase correlation algorithm based on Log-Gabor filtering for imaging with nonlinear radiometric differences and large-scale geometric differences between pairs of images was proposed. The authors found certain limitations when the models used complex transformations, but otherwise demonstrated the rationality and effectiveness of the algorithm. In Zhang et al. (2022), a dynamic programming strategy was used with TM based on normalized cross-correlation to overcome the computational complexity in matching algorithms. For the detection and counting of citrus fruits from trees to estimate the citrus productivity of a single tree before harvest, the use of markers for the object on the Watershed algorithm was proposed. The results indicated errors less than 8.31% and a coefficient of regression of $R^2 = 0.98$. In Fernandez et al. (2018), a method for the automatic detection of cucumbers grown in agricultural fields was developed using distance transformation and morphological operations to separate overlaps, in conjunction with object clipping based on convex hull and classifier based on bag of visual words. The algorithm was validated a detection rate of 91.79%, a false positive rate of 2.56%, and an accuracy that reached 85.65%. As observed in the literature, the TM algorithms with sub-images patterns have been widespread usage. Nevertheless, it is worthwhile to consider the potential for investigating the implementation of distance-based patterns or even the use of optimized masks.

For the classification problem of counting ears of wheat in a field using images obtained by a manned ground vehicle, a method was developed based on a twin-SVM model, resulting in a precision of 0.79–0.82 and average calculation time of 0.1 s (Zhou et al., 2018). In Chen et al. (2019), a segmentation method based on monocular vision and SVM algorithm was proposed for the segmentation of citrus trees. The results for the regions were calculated using SVM. The proposed method was verified on datasets of different brightnesses and weeds, and achieved accuracies of $85.27\% \pm 9.43\%$. In Pereira et al. (2019), an algorithm for recognize peaches was introduced; dimensions such as volume and weights in the natural conditions of trees in the field were calculated. Rounding metrics were used as characteristics to train the SVM, where the contours were found with the aid of the Sobel operator. Using second-degree polynomial functions for prediction, the algorithm achieved an accuracy of 72% for orchards with 808 trees/ha. In Wang et al. (2019), a method for the detection and enumeration of individual oil palm trees using a hierarchical approach was proposed. The proposed method classified images as vegetation and non-vegetation using an SVM classifier and features extractor HOG with samples blocks of 16×16 pixels and offsets of 8 pixels. An evaluation was conducted based on manual results, revealing that the proposed method detected 2590 oil palm trees with an overall accuracy of 99.21%. The literature review indicated the use of the SVM classifier for a broad range of agricultural applications. Nevertheless, it would be of interest to assess its a viability for pattern classifications in agricultural problems relate to vegetative stages, which would be relevant for AP techniques.

In the context of this study, we present a methodology for evaluation the count maize in the post-emergence stage (V2), as well as their heights to be sure about such a phenological stage, based on the use of conic sections, DIP techniques, pattern recognition (PR), and SVM on multispectral images of agricultural fields obtained by Drone. The rest of the paper is organized as follows: Section 2 introduces the materials and methods; Section 3 presents the results and discussions; and Section 4 provides the conclusion.

2. Materials and methods

The organization of the methodology, from drone imagery for counting plants to classification of the post-emergence phenological stage (V2), and the main techniques used for its development, are illustrated in Fig. 2.

The experimental area where the present research was conducted following the study standards of Embrapa Instrumentação, located 860 m from the geographic coordinates: $21^\circ 57'13.9''$ S and $47^\circ 51'10.9''$ W at the National Laboratory of Reference in Precision Agriculture (LANAPRE) in São Carlos, SP/Brazil.

The experiment for the evaluation of the method was organized in an agricultural area with maize (*Zea mays* L.), having 4000 m², and sampling grid equal to 10 m \times 10 m. Fig. 3 shows each of these site-specific management, i.e., divided into 40 blocks. In fact, this study was conducted taking into account not only 37 blocks of maize plants belonging to the post-emergence stage V2 for computational analysis but also 3 random blocks for manual counting by an agronomic expert.

The experimental area for validation was conducted with a batch of maize seeds. Pioneer P4285 hybrid VYHR, characterized by its flat shape. Sowing was carried out using 5 seeds per linear meter, resulting in estimated density of 600 plants per blocks. Additionally, the experimental arrangement was carefully designed, considering a spacing of 90 cm between rows and the presence of 12 rows per blocks, which led to the formation of a total experimental area of 60 rows. With all this information available, the calculation of the number of seeds used in the experimental area was equal to 24 000.

The experimental area for validation was conducted with a batch of maize seeds. Pioneer P4285 hybrid VYHR, characterized by its flat shape. Sowing was carried out using 5 seeds per linear meter, resulting in estimated density of 600 plants per blocks. Additionally,

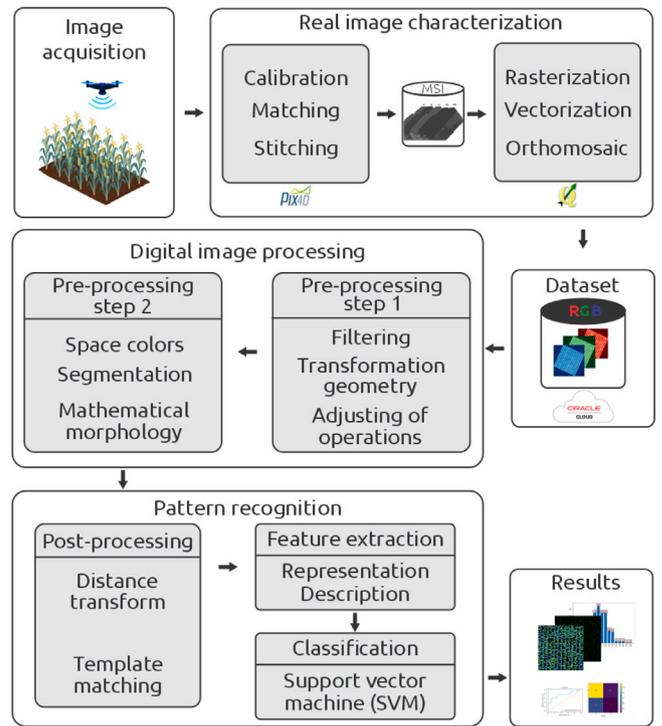


Fig. 2. Overall method workflow.

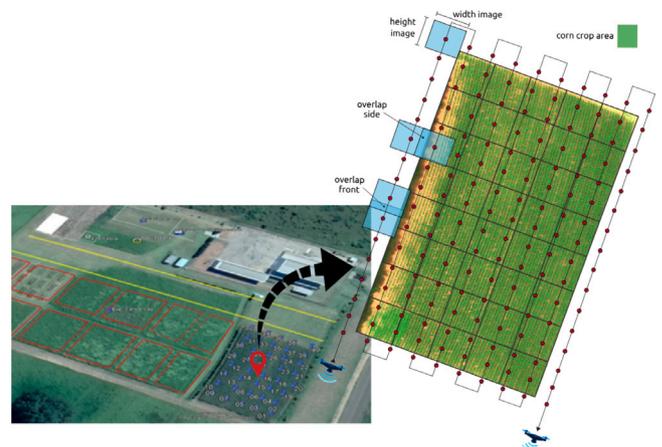


Fig. 3. Experimental detail with 40 blocks (from b-01 to b-40), where 37 blocks were evaluated computationally and 3 from them were examined manually for control (b-08, b-16, and b-33).

the experimental arrangement was carefully designed, considering a spacing of 90 cm between rows and the presence of 12 rows per blocks, which led to the formation of a total experimental area of 60 rows. With all this information available, the calculation of the number of seeds used in the experimental area was equal to 24 000.

2.1. Image acquisition

For this module, ground control points (GCP) in the experimental area were collected by high-precision GPS in conjunction with an RTK receiver that recorded their geographic coordinates to an accuracy of ± 1 cm. In this way, the GCPs were used as input to control the flight missions of a multicopter Drone DJI Matrice 100. For imaging a sensor multispectral Micasense RedEdge-M was embedded, and a multispectral sensor was provided onboard. Additionally, the imaging equipment

Table 1
Micasense RedEdge-M specifications.

Description	Values	Units
Raw image	12	bit
Sensor size	4.8 × 3.6	mm
Focal length	5.4	mm
Frame capture	1	frame/seg
Wavelength red	668	nm
Wavelength green	560	nm
Wavelength blue	475	nm
Spatial resolution	1280 × 960	pixels

Table 2
Dataset RGB.

Description	Values
File type	GeoTIFF
Bits per pixel	32 Bit
Resolution	96 ppi
Spatial resolution	546 × 546 pixels
Images per band: R, G e B	37
Total images per block	111

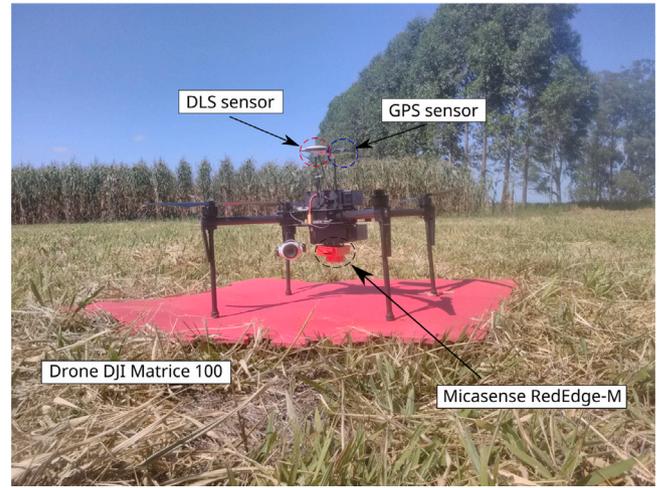


Fig. 4. Drone and hardware setup for image acquisition.

includes to a DLS sensor for measuring the influences of the sun's brightness, or changes in contrast due to superimposition of clouds in the sky, thus providing the capacity to correct global changes in light, the same ones that fluctuate during flight. The specifications of the MicanSense sensor are detailed in [Table 1](#). [Fig. 4](#) shows details of the Drone and the multispectral camera for the images acquisition used in this study .

Aerial mapping reproduces the phenomenon of stereoscopic vision, through the superimposition of multiple vertical and horizontal images, which can be acquired in blocks and with multiple lines during the flight stage. Overlapping aerial images are strongly linked with the number of flight lines, and routes used during flight planning, and are expressed in percentages frontally and laterally. Altogether, for the present study, the lateral and frontal overlaps were 80% each, following a flight pattern of parallel lines, along specified routes with waypoints. Besides, in accordance with the drone settings and the onboard sensor, the flights were conducted within a time interval from 11 A.M. to 12 A.M., i.e., during the morning periods.

2.2. Characterization of images

In this stage, orthoimages (digital model of the experimental area) were produced in the bands blue (B), red (R), and green (G). In this regard, the Pix4D software ([Deng et al., 2018](#)) was used for filtering, matching, and stitching the images in a cloud of points with geometric correction in the geocentric reference system for the Americas 2000 (SIRGAS2000). Finally, the orthomosaic was processed from the orthoimages of the experimental area with the assistance of QGIS software, separating them into blocks, thus composing the data set in the RGB channels, as detailed in [Table 2](#).

2.3. Digital image processing

It should be considered that the orthomosaic blocks have much black information because of the vectorization process using the QGIS software. Furthermore, because the images, at their real intensities, tend to have a lack of brightness and contrast, pre-processing operations are necessary in this module. In this sense, an extraction of the region of interest (ROI) was computed mainly by Gaussian filtering (Eq. (1)) and a minimum rectangular area that calculates the angle approximation for a rotation operation around the object center (R - Eq. (2)).

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

where the Gaussian function G at coordinates (x, y) is controlled by the variance σ^2 , and mean (μ) of zero.

$$R = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -t_x \\ 0 & 1 & -t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

where $-t_x, -t_y$ correspond to the translation of the object to the origin, whereas t_x, t_y shift the object to its original position.

Adjustment operations for brightness and contrast change pixel values without changing the size, geometry, or local structure of the image. Thus, Eq. (3) is used to adjust brightness (α) and contrast (β), and is defined as

$$G(x, y) = \alpha F(x, y) + \beta \quad (3)$$

where $F(x, y)$ and $G(x, y)$ are input and output images, respectively.

Histogram adjustments are subject to a first-order probability function $p_r(r_k; x, y) = n_k$ indicating the probability of pixels (x, y) with r_k occurrences at the k th intensity level, in a number of pixels n_k . In this regard, contrast limited adaptive histogram equalization (CLAHE) was used, distributing the pixel values in the histogram accumulation phases. Its equation is defined as

$$h_{b_*, l_*}(F(x, y)) = \sum S'_{x,y}(x, y, b_*, l_*) \quad (4)$$

where b_* is the block size, l_* is the threshold applied to contrast in a given block, and $S'_{x,y}$ is a bilinear equation of neighborhood eight.

Starting from the RGB color space, other spaces were explored, as follows: XYZ, YCbCr, HSV, and CIELab ([Gnädinger and Schmidhalter, 2017](#)), with the purpose of looking for a relationship between the intensities of the maize plants and the reflectances collected from the pre-processed images. Of these, the HSV color space (H: hue, S: saturation and V: value) is highlighted, where hue determines tonality, saturation determines purity of color (i.e., from faint to intense), and value determines its intensity. Eq. (5) defines the HSV color space from RGB color space as follows

$$H = \begin{cases} 60^\circ \left(\frac{G-B}{M'-m'} \right), & \text{if } M' = R \\ 60^\circ \left(\frac{B-R}{M'-m'} \right) + 120^\circ, & \text{if } M' = G \\ 60^\circ \left(\frac{R-G}{M'-m'} \right) + 240^\circ, & \text{if } M' = B \end{cases} \quad (5)$$

$$S = \begin{cases} \frac{M'-m'}{M'}, & \text{if } M' \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$V = M'$$

where M' and m' represent the maximum and minimum, respectively, of the RGB tristimulus.

Images can contain several regions, which are connected and have different features or patterns in relation to some measure, such as spatial proximity or similarity of their attributes. Therefore, for this work, segmentation by binarization was used with intervals of intensities $[L_1, L_2]$ (Eq. (6)), subject to intervals in a color space and the reflectance of the maize plant.

$$G(x, y) = \begin{cases} 1, & \text{if } F(x, y) \in [L_1, L_2] \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

That being said, to ensure a global segmentation with good results, it was decided to use a morphological closing operation (Eq. (7)) with an isotropic elliptical structuring element (B) to provide translation invariance in addition to noise blurring.

$$A \bullet B = (A \oplus B) \ominus B \quad (7)$$

where \oplus and \ominus represent dilation and erosion, respectively. That is, object A is being translated by the structuring element B .

2.4. Pattern recognition

The results of the segmentation step do not necessarily represent the topology of the canopy shape of the maize plant, being separated or joined by regions of segmented pixels in raw form. In this regard, it is necessary to have methods that can represent the segmented regions in a topological or geometric way, as in the case of the distance transform (DT) technique, defined as

$$D(x, y) = \min_{s,t:k(s,t)} d(x-s, y-t) \quad (8)$$

where k is a suitable mask to be used in DT, and $d\{(x, y), (s, t)\}$ is the Euclidean distance.

Fundamental tasks, such as shape or pattern detection, are performed using template matching (TM), which is based on the correlation between a mask face (k) and the global image. In this work, TM employed a correlation coefficient (\check{c}) was used, which is defined as

$$\check{c}(x, y) = \frac{\sum_{s,t} (F(x+s, y+t) - \bar{F})(k(s, t) - \bar{k})}{\sqrt{\sum_{s,t} (F(x+s, y+t) - \bar{F})^2} \sqrt{\sum_{s,t} (k(s, t) - \bar{k})^2}} \quad (9)$$

where $\check{c} \in [-1, 1]$, if $\check{c} \approx 1$ indicates a high correlation, otherwise no similarity. Furthermore, \check{c} describes a piecewise local correlation between k and the current subpicture.

The mask used for the search for maize patterns was used with a Chamfer algorithm (Thiel and Montanvert, 1992) that calculates a weighted local distance, defined as:

$$d(x, y; B') = \max \{d_1(x, y), d_2(x, y)\} \quad (10)$$

where $B' = \{b(i) : 1, 2, \dots, m\}$, and d_1, d_2 are defined as:

$$d_1(x, y) = p_* + \sum_{j=1}^{p_*} \left[\frac{x-y-f'(y-1)-1}{f'(p_*)} \right] \quad (11)$$

$$d_2(x, y) = \max(x, y)$$

where p_* is a periodic sequence of length B' , $e f'(i) = \sum_{j=1}^i b(j) \forall i 1 \leq i \leq p_*$.

The sequence of steps of Pseudocode 1 compute the map of distances in relation to the locations of possible canopies of the maize plant, in addition to the TM with the Chamfer algorithm.

Locations of local maxima obtained in the TM phase take information from features that can be evaluated through contour descriptors, such as the chain code defined by Eq. (12):

$$P'_i = H(\Delta s_i, \Delta t_i) \quad (12)$$

$$(\Delta s_i, \Delta t_i) = \begin{cases} (s_{i+1} - t_i, s_{i+1} - t_i) & 0 \leq i < M - 1 \\ (s_0 - t_i, s_0 - t_i) & i = M - 1 \end{cases}$$

Pseudocode 1: Digital post-processing

Input: Segmented matrix: I_b
Output: Maize plant locations matrix: M_{loc}

```

1 begin
2    $M, N = \text{size}(I_b)$ 
3    $D' = \text{zeros}(M, N)$ 
4    $k \leftarrow \text{create Chamfer mask with size } b_*$ 
5   for  $i, j$  to  $M, N$  do
6      $d_1, d_2 = k_{l,r} + D'_{(i,j)}, k_{l,r} + D'_{(i,j)}$ 
7      $d_3, d_4 = k_{l,r} + D'_{(i,j)}, k_{l,r} + D'_{(i,j)}$ 
8      $D'(i, j) = \min(d_1, d_2, d_3, d_4)$ 
9   end
10   $k \leftarrow \text{create Chamfer mask with size } p_*$ 
11  for  $i, j$  to  $M, N$  do
12     $V_{match} = \check{c}_N(D'(i, j), k)$ 
13     $V_{loc} = V_{match}(i, j)$ 
14  end
15   $M_{match}, M_{loc} = [V_{match}, V_{loc}], \max(M_{match} \times I_*)$ 
16 end
```

where M is the total number of pixels in a contour.

The ROIs can be represented by the chain code, and for this it was necessary to extract geometric characteristics, which are related to the radii of the canopy openings of the maize plant. Therefore, we define x, y points for each $x_{min} < x < x_{max}$ e $y_{min} < y < y_{max}$, so a bounding box is defined as

$$C'_d = \langle x_{min}, x_{max}, y_{min}, y_{max} \rangle \quad (13)$$

where x_{min}, y_{min} and x_{max}, y_{max} , are minimum and maximum pixels, respectively, of the ROI (A').

The radius is calculated using diameter information, defined as

$$\hat{d} = \max_{i,j} [d_e(x_i, y_j)] \quad (14)$$

where d_e is the Euclidean distance, and x_i, y_i are the minimum and maximum points, respectively, of the contour of the region.

A vector of (V_r) characteristics was also defined for the opening of post-emergence maize plants, such as

$$V_r = 1/2 \times [\max(d_e(C'_d))] \quad (15)$$

Thereby, Pseudocode 2 describes the vector feature extraction process and the generation of maps with circles for the purpose of visualizing and detecting maize plants (Veramendi and Cruvinel, 2021).

Pseudocode 2: Feature extraction

Input: Map of distances and local max: D', M_{loc}
Output: Feature matrix: M_c

```

1 begin
2    $M, N = \text{size}(D')$ 
3    $P_k = \text{zeros}(M, N)$ 
4   for each point  $(x, y) \in M_{loc,i}$  do
5      $P_k(x, y) = \text{binarize}(D'(x, y))$ 
6      $A'(x, y) = H(P_k(x, y))$ 
7   end
8   for each point  $(x, y) \in A'$  do
9      $x_{min}, x_{max}, y_{min}, y_{max} = C'_d(A'(x, y))$ 
10     $V_r = 1/2 [\max(d_e(x_{min}, x_{max}, y_{min}, y_{max}))]$ 
11     $M_c(x, y) = \text{createMap}(V_r)$ 
12  end
13 end
```

For the counting of maize plants, a histogram analysis of the plants occurrences can be performed. Therefore, the final quantity of plants

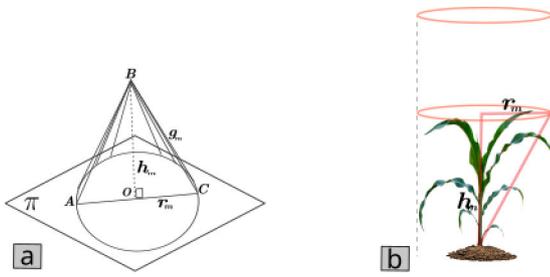


Fig. 5. In (a) the conic model, and in (b) the radius, and the height of the maize plant.

for each block is calculated based on the model:

$$N_p = f_r \sum_{i=1}^N \sqrt{1.04(r_i)} \quad (16)$$

where f_r is the number of occurrences obtained from the histograms, i.e., for each block. The variable r_i represents the radius of the canopy aperture.

Indeed, the relationship between the radius and the canopy aperture of the plant was experimentally obtained, considering the different phenological stages of the maize plant. In this way, a calibration curve was established, making it possible to use the conic model to estimate the plant heights as an indicator of their phenological stage. In this work, the information is applied to the post-emergence stage (V2).

In this way, it is possible to establish a plane π that incorporates a circular base, where a central point (vertex B) located outside the plane corresponds orthogonally to the center of a such base, thus allowing to fine the height distance h_m . It is noteworthy that a triangle BOC (Fig. 5) generates a conic section from a rotation axis (vertex O), determining the meridian section. By integrating these components, the plant's height is defined by the properties of the generatrix, given by $h_m = r\sqrt{3}$

SVM classifiers provide good generalization even with a small number of training samples. To use these classifiers, one should find a hyperplane of parameters (w, b) such that the distance between the hyperplane and training samples is maximized. The training samples are represented by a set of radii vectors (x_i), derived from the previously mentioned histograms, and are organized through a threshold with the computed standard deviation. However, the data may have some sparsity, which could cause some support vectors to be in very small marginal hyperplanes. Because of this, some atypical features may be misclassified. In such cases, relaxing constraints on primal optimization may provide separability (Eq. (17)).

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i \quad \forall i = 1, \dots, d \quad (17)$$

where ξ_i are slack or relaxation variables, such that $\xi_i \geq 0$.

The slack variables correspond to selecting a hyperplane that minimizes empirical error. Because of this, possible minimum errors can maximize the margin of separation between classes. Thus, a linear classifier with maximum margin in optimization is defined as:

$$\rho = \min_{w, b, \xi} \frac{1}{2} \|w\|^2 + C^* \sum_{i=1}^d \xi_i \quad (18)$$

where C^* is a weight constant for training error minimization.

The distribution of the feature space is not necessarily linearly separable, and a solution may require the application of a data mapping for larger spaces, i.e., leading to space vectors of higher dimensions $\Phi \subseteq \mathbb{R}^d$. Thus, it is necessary to apply the kernel defined by Eq. (19):

$$k^*(m_i, m_j) = \Phi(m_i)^T \Phi(m_j) \quad (19)$$

where Φ is a representation of non-linear functions of the original mapped features.

Table 3

Kernels evaluated.

Kernel	Function	Parameters
Polynomial	$(x_i^T \cdot x_j + 1)^p$	p
Gaussian	$\exp(-\gamma \ x_i - x_j\ ^2)$	γ
Sigmoid	$\tanh(k_1 x_i^T \cdot x_j + k_2)$	k_1, k_2

p is the degree of the polynomial, γ is equivalent to $1/2\sigma^2$ which controls flexibility of kernel function, k_1 is a positive weight and k_2 is an offset value.

In this work, three different types of kernel functions were evaluated (Table 3), aiming to find the best results in terms of accuracy, precision, recall, and F1-score.

For validation, the data vectors (x_i) were split into training and test datasets in proportions of 80%–20%, 70%–30% and 50%–50%. Thereby, Pseudocode 3 shows the training and testing processes for classifying the canopy openings of maize plants in post-emergence V2.

Pseudocode 3: Feature classification

Input: Feature matrix: M_c

Output: Accuracy, precision, recall, F1-score

```

1 begin
2    $V_{tr} = \text{trainingVector}(M_c)$ 
3    $V_{ts} = \text{testVector}(M_c)$ 
4   for each classifier  $\mathcal{K}$  do
5     if  $\mathcal{K} \in \text{Polynomial}$  then
6        $\mathcal{K} = (x_i^T \cdot x_j + 1)^p$ 
7     end
8     if  $\mathcal{K} \in \text{Sigmoid}$  then
9        $\mathcal{K} = \tanh(k_1 x_i^T \cdot x_j + k_2)$ 
10    end
11    if  $\mathcal{K} \in \text{Gaussian}$  then
12       $\mathcal{K} = \exp(-\gamma \|x_i - x_j\|^2)$ 
13    end
14     $params = \text{fit SVM}(V_{tr}, C, \mathcal{K})$ 
15     $acc, prec, rec, F-1 = \text{test SVM}(V_{ts}, params)$ 
16  end
17 end
```

3. Results and discussions

For the process of image acquisition from the V2 stage, it was necessary to perform radiometric calibration to convert the metadata of the digital image to a physical scale. On the other hand, the geometry of the aerial image was established by the size of the camera sensor, the focal length, and height of the drone flight, which together determine its ground sample distance (GSD) (Table 4). The GSD provides the corresponding measure for the pixels of the surface of the experimental area or the area covered by the image.

Because of this, it was necessary to establish the percentages of lateral and frontal overlapping of the aerial images to guarantee good densities in the image characterization module. The number of registered images was 300 images for each spectral band, or in a total of 1500 images for the 5 spectral bands. The total required storage capacity was 3.69 GB (*gigabyte*), because the surface width and height were equal to 27 m \times 20 m respectively, and the distances between each front and side capture were 4 m and 5 m respectively.

For phenological growth analysis, and to estimate the heights of the maize plants (conic section model), five measurements relevant to the canopy opening and maize height were manually collected at random. Thus, to carry out such an analysis, a calibration curve was obtained, taking in account measurements for different phenological stages of the maize plants, as the examples shown in Table 5,

Table 4
Survey specifications.

Description	Values	Units
Flying altitude	138	m
Mission flying time	12	min
Max. speed of flying	11	m/s
Front and side overlap	80	%
Ground sample distance	5.95	cm/pixel

Table 5
Examples of data manually collected in the experimental area.

Stages	Variables	Measures (cm)
V2	Lengths	35.0; 33.0; 29.0; 32.0; 31.0
	Radii	17.5; 16.5; 14.5; 16.0; 15.5
	Heights	27.0; 32.0; 29.0; 33.0; 28.0
V5	Lengths	88.0; 90.0; 86.0; 91.0; 86.0
	Radii	44.0; 45.0; 43.0; 45.5; 43.0
	Heights	48.0; 50.0; 45.0; 53.0; 47.0
V7	Lengths	104.0; 84.0; 99.0; 83.0; 104.0
	Radii	52.0; 52.0; 49.5; 41.5; 52.0
	Heights	165.0; 177.0; 168.0; 184.0; 198.0

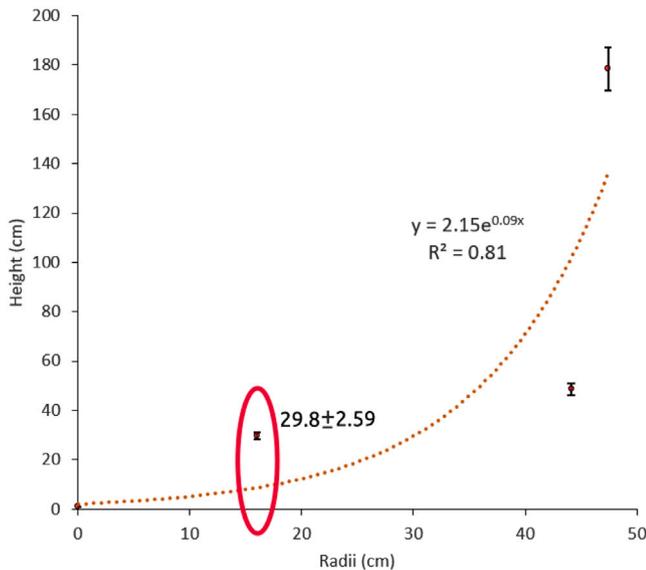


Fig. 6. Estimation model for the growth of maize plants, obtained from real measurements in an experimental field conducted for method validation.

Based on these data collected in the field for a random sampling of maize plants, the mean radius ($\mu_r = 16$ cm), mean height ($\mu_h = 29.08$ cm), and the standard deviation of the height ($\sigma_h = 2.59$ cm) were calculated. The resulting behavior of the calibration curve agree with the model presented in Müller et al. (2005). Besides, the correlation coefficient obtained was $R^2 = 0.81$, and for the post-emergence stage (V2) the mean value was found equal to 29.8 ± 2.59 cm (Fig. 6).

Collected images were sampled, filtered, and stitched with the help of Pix4D software, with which orthoimages were generated for each spectral band of interest (RGB). To generate of orthomosaics from the experimental field, the experimental area (area where the maize crop is located), which had a spatial resolution of 3935×5019 pixels, Cuts each channel R, G, and B from the experimental area in blocks of $10 \text{ m} \times 10 \text{ m}$, using the QGIS software. The image bank was evaluated based on specific site units, also called image blocks, where the red ($\lambda = 668 \text{ nm}$), green ($\lambda = 560 \text{ nm}$), and blue ($\lambda = 475 \text{ nm}$) spectral bands corresponding to block 30 are depicted in pseudo colors for visualization purposes. Each mosaic has a width and height of 546×546 pixels, radiometric resolution of 32 bits and digital resolution of 96 dpi.

Table 6
Parameters for geometric transformation.

Description	Parameters	Values
Size of kernel	b_*	3×3
Opening for filtering	σ	0
Threshold for binarization	l_*	10
Actual angle for rotation	θ	$\approx 25^\circ$

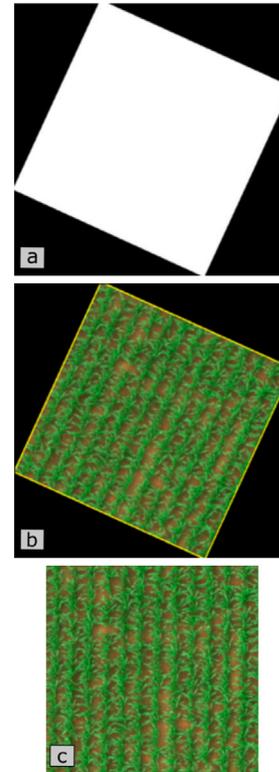


Fig. 7. Geometric transformation of image: (a) filtering and thresholding; (b) contour detection considering the region of interest (ROI); (c) rotation around center of object and final image matrix with 410×410 pixels (area equal to 100 m^2).

To closely study only the maize crop, additional information on the intensities of other elements, such as soil, stones, and other surrounding elements, should be excluded. The area to be cut was identified, and an angle of inclination was calculated, subject to three sub-processes: filtering and thresholding, contour detection, and rotation around the center of the object. Fig. 7 shows the process of geometric transformation.

Gaussian filtering blurring the edges of the area of interest. Therefore, a kernel size of $kernel \ b_* = 3 \times 3$ was selected, in conjunction with the opening standard deviation of $\sigma = 0$. The thresholding step used a degree of intensity $l_* = 10$, and the actual angle used was $\theta \approx 25^\circ$ rotating the ROI around the center of the object (Table 6). Thus, the new dimension of the ROI is defined as 410×410 pixels in relation to the sample grid.

An examination of the ROI histogram showed the need to preprocess the image with adjustments for brightness and contrast, in addition to local operations due to the variability of luminance in small regions. Thus, CLAHE was used in a set of color spaces: HSV, YCbCr, and CIELab. Results for the brightness and contrast parameters were $\alpha = 1.25$ and $\beta = 5$ in the RGB space. For CLAHE, a window size $b_* = 16$ was used, together with the local threshold $l_* = 2$; the difference is that these parameters are applied locally in the mentioned color spaces. This approach was applied an iterative number of times $k = 2$. Table 7 details these parameters.

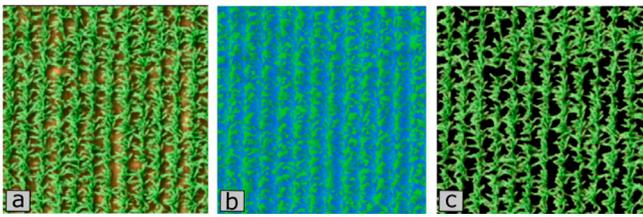


Fig. 8. Results of segmentation in HSV color space, (a) enhanced image, (b) image in HSV space, and (c) segmentation results for maize plants; (area equal to 100 m²).



Fig. 9. Results of distance transform: (a) segmented image, (b) binary mask, (c) Euclidean distance map; (area equal to 100 m²).

Table 7
Parameters used for brightness, contrast, and CLAHE adjustments.

Description	Parameters	Values
Contrast	α	1.25
Brightness	β	5
Window size of CLAHE	b_*	16
Local threshold CLAHE	l_*	2

Table 8
Minimum and maximum parameters in HSV.

Parameters	Channel H	Channel S	Channel V
Minimum	45	100	50
Maximum	75	250	250

Experiments were conducted to improve brightness, contrast, and a Region of Interest (ROI) on a selected image. In addition, an analysis of histograms on the H, S, and V channels provided a set of minimum and maximum intensities in reference to the color spectrum reflected by the maize plants, which are detailed in Table 8. These data make it possible for the segmentation phase to be used in other applications involving the natural environment.

In the implementation of the segmentation technique based on color space, a set of techniques were applied together in the segmentation step, such as Gaussian filtering with size 5 × 5. Ranges identified for HSV images were operated using a morphological closure operator following an isotropic structuring element with size 5 × 5 and of circular shape. Following this method, Fig. 8 illustrates the discrimination of the maize plant in relation to soil, straw, and other elements possibly existing in the considered region.

Based on the spatial and geometric variation, in addition to the heterogeneous patterns in the regions of the maize plants, a Euclidean distance transform was applied to the segmented image, which originated a distance map with local intensities representing the object of interest in relation to its closest distances from the central opening of the canopy. Fig. 9 illustrates an example of results obtained with the application of the distance transform technique, where it is possible to observe that the distance transform technique satisfactorily localized the presence of maize plants in relation to their geometrically shaped central distance.

To spatially locate the canopies of maize plants, TM with normalized correlation was employed. In this context, Chamfer masks were established considering the p_* parameters ($p_1 = 8, p_2 = 13$ and

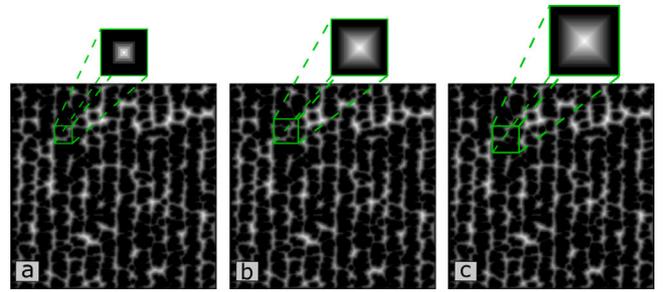


Fig. 10. Distance map for Chamfer mask, with (a) 8 × 8, (b) 13 × 13, and (c) 18 × 18 pixels; (area equal to 100 m²).

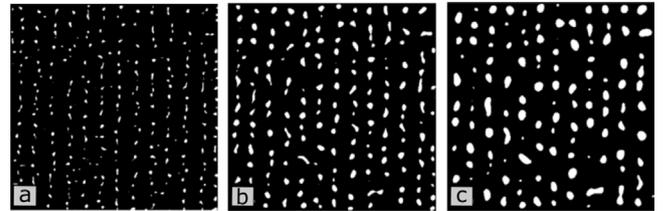


Fig. 11. Peaks located after using TM with correlation coefficient and Chamfer mask, for (a) 8 × 8, (b) 13 × 13, and (c) 18 × 18 pixels; (area equal to 100 m²).

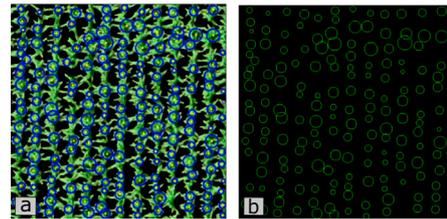


Fig. 12. Post-emergence V2 locations with geometric descriptors, (a) circular map locations, (b) circular feature map; (area equal to 100 m²).

$p_3 = 18$). The choice of these values for the parameter p_* was made experimentally, considering adjustments to avoid both overlap and distancing of neighboring circles. Furthermore, given the different canopy sizes, three different Chamfer masks were used. Fig. 10 illustrates the standards applied for the use of these masks in conjunction with the TM technique. Specifically, the three masks were applied to the image that presents the Euclidean distance map, considering the Chamfer masks from smallest to largest.

The approach resulted in a range of low and high correlation values, saving spatial locations to guarantee the presence of a plant according to the correlation. To evaluate these values in the correlation range, a threshold of at least 50% of these values was used. Fig. 11 illustrates the locations derives using Chamfer masks as subimages via the TM algorithm. Based on the results obtained, the Chamfer pattern for parameter $p_* = 13$ made it possible to obtain local information more uniformly compared to those obtained using the other patterns.

Using the descriptor chain code that defines the contours of these regions, to be binarized in relation to the minimum area of the canopy opening, made it possible to extract the contour characteristics of the peaks found in the TM stage. Thus, the presence of maize plants in the post-emergence stage V2 was visually determined (Fig. 12).

Features of vectors of circular radii, corresponding to the blocks of the maize crop, were analyzed to validate the proposed method. Specifically, all the blocks were analyzed to validate the method, whereas three blocks were analyzed also considering the manual counting of plants from emerged seeds. In this process, a vector with a smaller dimension of 116 radii (block 7) and a vector with a larger dimension

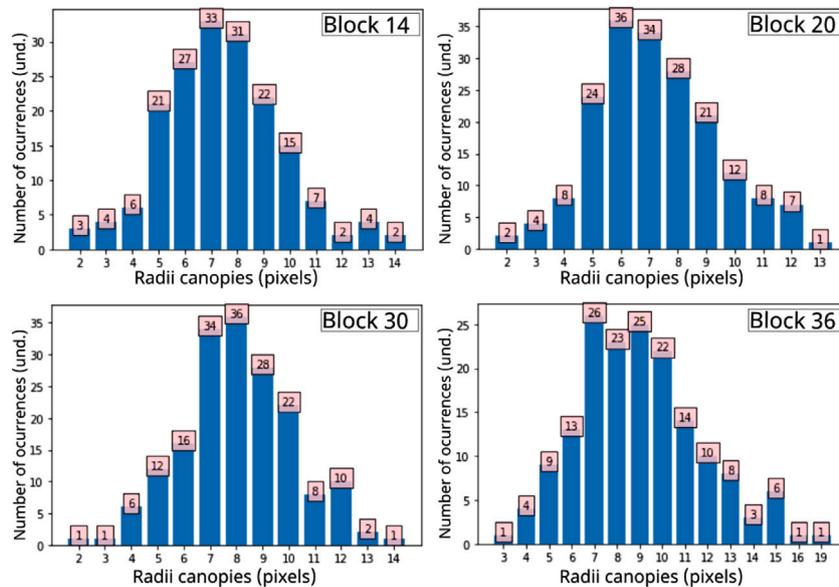


Fig. 13. Examples of histograms for number of maize plants occurrences and the observed radii in the blocks.

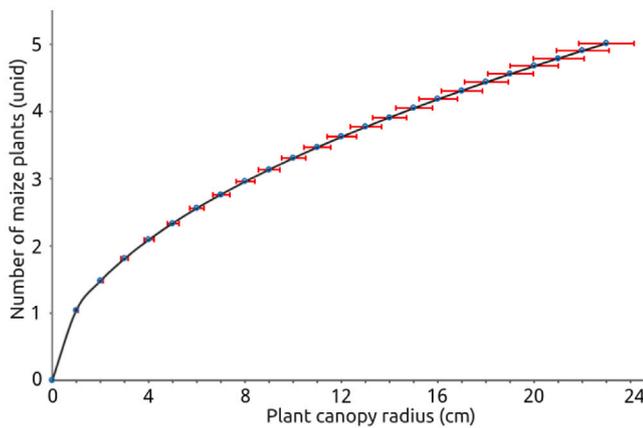


Fig. 14. Estimation model for counting maize plants, obtained from computational method.

of 186 radii (block 39) were found. Another characteristic evaluated was the elaboration of the number of occurrences of circular radii in relation to the analyzed blocks, to know the canopies of the detected plants (radii of openings) and, in turn, to infer the number of the maize plants in each block. Fig. 13 illustrates examples, randomly choice, of the feature histograms of the blocks analyzed in this work.

To count the maize plants present in the area, it was necessary to perform a normalization with the minimum ($r_{min} = 1$) and maximum ($r_{max} = 23$) radii. The smallest radius, corresponded to 1 plant of maize whereas the largest radius 23 corresponded to 5 plants of maize. Meanwhile, the blocks that had the lowest and highest number of maize plants were block 7, which had 247, and block 39, which had 527, respectively.

However, results of counting the maize plants are based on the calibration of a projected model, which is shown in Fig. 14. This model included standard deviations for the radii of openings of the canopies.

Therefore, based on the average value of the radii obtained computationally, which was equal to 12 units of measure, corresponding to a height of 20.78 units of measure (model of conic sections), a coefficient $\alpha = 0.69$ between these units was found. Based on the calibration of this model, Fig. 15 illustrates the number of plants emerged per block in the experimental field.

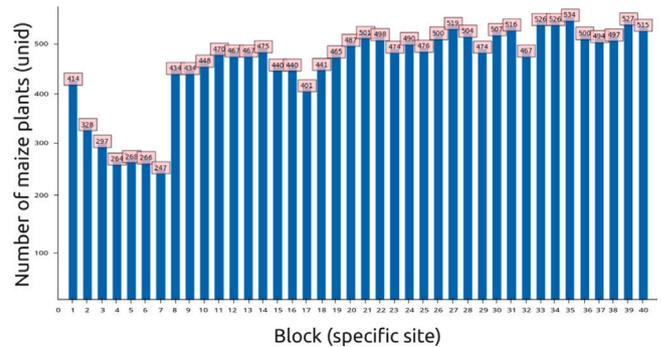


Fig. 15. Number of maize plants counted by the computational model for each specific site.

The analyzed histograms describe a pattern that is close to a Gaussian distributions, which led to the organization of 503 vectors, each consisting of a radius and number of occurrences. These values were categorized according to the ranges of standard deviation values ($\sigma_{min} = 1.86$ and $\sigma_{max} = 3.42$) and means ($\mu_{min} = 5.06$ and $\mu_{max} = 9.09$), with which a threshold ($\sigma_t = 2.5$) was then established to determine which observations had dispersions that were greater or lesser than the occurrences of the distributions. Thus, histograms with $\sigma \geq 2.5$ formed class 1, and histograms with $\sigma < 2.5$ formed class 2. Where, class 1 was composed of 230 vectors, and class 2 of 273 vectors.

To continue the trials for training and testing the classifiers, a proportion of vectors from each class were assigned to the classifiers. In this case, the proportions were 80% and 20% from class 1 and class 2, respectively. Thus, Classifier 1 was composed of 184 and 55 vectors from class 1 and class 2, respectively. On the other hand, Classifier 2 assigned the complementary set of vectors, i.e., 20% and 80% of the vectors from class 1 and class 2, respectively. Thus, Classifier 2 was composed of 46 and 218 vectors from class 1 and 2, respectively. To validate the SVM classifiers, they were trained and tested on data divided according to proportions of 50% for training and 50% for testing, 70% for training and 30% for testing, and 80% for training and 20% for testing. Table 9 presents a comparison of the results obtained by the applications of different kernels in the SVM classifiers for the post-emergence stage V2 patterns.

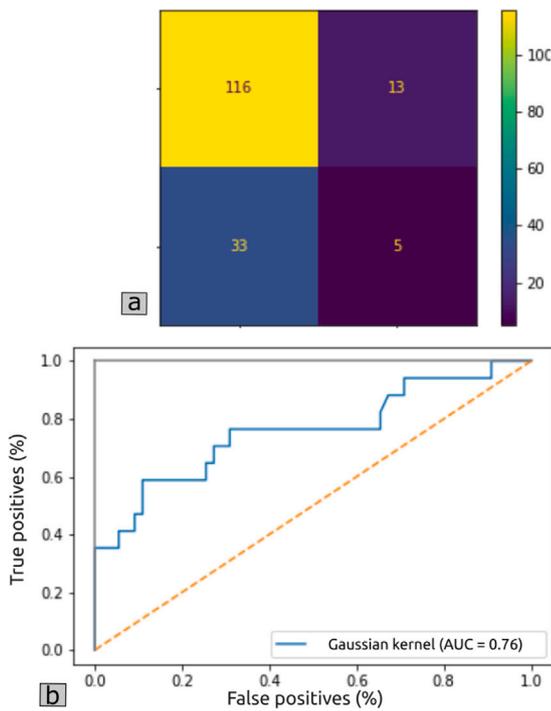


Fig. 16. (a) Confusion matrix, and (b) ROC Curve of Classifier 1.

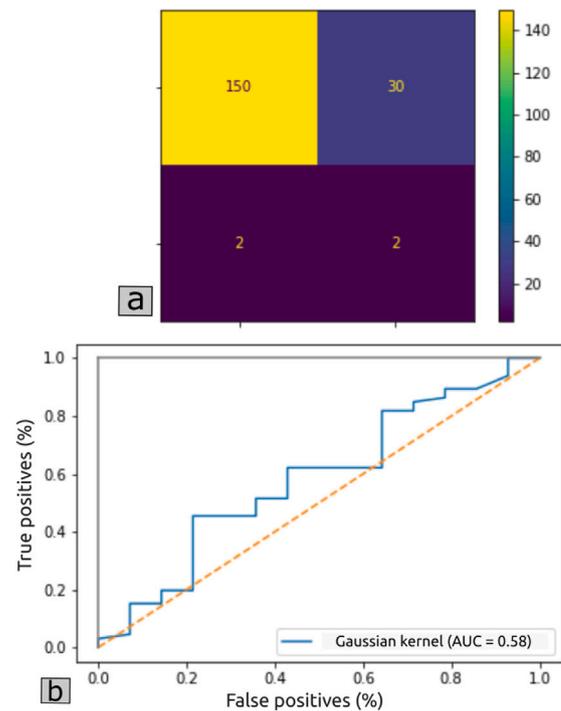


Fig. 17. (a) Confusion matrix, and (b) ROC Curve of Classifier 2.

Table 9
Obtained results for the SVM selection.

Kernel	Proportions	Metrics	Classifier 1	Classifier 2
Sigmoid	50%:50%	Precision	0.81	0.82
		Recall	0.92	0.94
		Accuracy	0.78	0.79
		F1-Score	0.86	0.88
		Support vectors	92	109
	70%:30%	Precision	0.74	0.78
		Recall	0.87	0.76
		Accuracy	0.67	0.62
		F1-Score	0.80	0.77
		Support vectors	55	66
	80%:20%	Precision	0.84	0.83
		Recall	0.97	0.98
Accuracy		0.83	0.81	
F1-Score		0.90	0.89	
Support vectors		37	44	
Gaussian	50%:50%	Precision	0.82	0.84
		Recall	0.97	0.94
		Accuracy	0.82	0.80
		F1-Score	0.89	0.88
		Support vectors	92	198
	70%:30%	Precision	0.83	0.84
		Recall	1	1
		Accuracy	0.85	0.84
		F1-Score	0.91	0.91
		Support vectors	55	66
	80%:20%	Precision	0.82	0.84
		Recall	1	0.95
Accuracy		0.83	0.81	
F1-Score		0.90	0.89	
Support vectors		37	44	

From Table 9 it was possible to observe the best results for selection of the SVM classifiers which are highlighted in bold. The Figs. 16 and 17 shows the confusion matrices and ROC curves for the selected Classifier 1 and Classifier 2.

Through analysis of the results of the experiments conducted to select the SVM Classifiers, considering the training and testing stages, it

was possible to verify that the precision remained at a satisfactory value for both, that is, above 74%. This occurred for the classification of the post-emergence patterns V2, i.e., considering the opening canopies of maize plants (*Zea mays L.*).

Furthermore, when a Sigmoid Kernel was used for both classifiers, as well as when based on percentages for training and testing equal to 70%:30%, low accuracies were observed, i.e., 67% and 62% for Classifier 1 and Classifier 2 respectively.

Likewise, when a Gaussian kernel was used for both classifiers, that means, considering the percentages for training and testing equal to 70%:30%, high accuracies were observed, i.e., 85% and 84% for Classifier 1 and Classifier 2 respectively. Besides, we also observed better precision values for such configurations that mean, 83% and 84% for Classifier 1 and Classifier 2 respectively.

3.1. Discussion of the experimental results

The method developed relates the count of maize plants (*Zea Mays L.*) for the V2 phenological stage and then with the percentage of emergence in the crop area. In fact, emergence occurs when the first leaves, called coleoptiles, appear above the soil surface due to the germinating seeds having absorbed water (approximately 30% to 35% of their weight) and oxygen. Then, for those that germinate, the rootlets emerge quickly near the tips of the seeds, depending on the temperature and humidity conditions of the soil.

Results have shown the usefulness by using drone flights and data acquisition considering 138 m of altitude with a GSD equal to 5.95 cm/pixel for the maize plant counting directly in a crop area. Additionally, the use of the MicaSense camera equipped with a DLS sensor to correct potential shadows was also adequate for such application. However, it was also observed that when mapping the acquired images, the minimum and maximum reflectance values should be normalized before being stored as a geotiff images. The use of RGB spectral bands have confirmed the feasibility of the developed method in the visible spectrum. Furthermore, to ensure the assessment of maize plant density, subdividing the experimental area into blocks units allowed

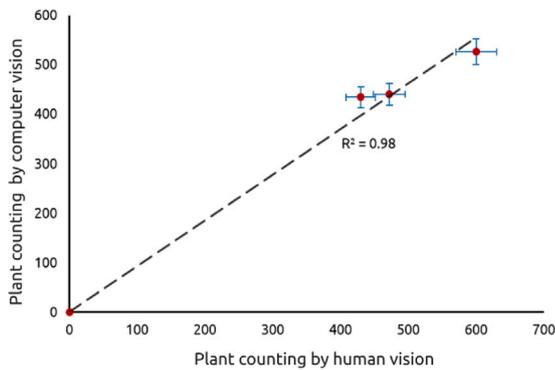


Fig. 18. Comparing plant counting models: Human Vision vs. Computer Vision.

the optimization of the computational process for the sequence of algorithms employed in the developed method.

In addition, as observed in the results, to enable the counting of maize plants, it was necessary to calibrate the field measurements. For this study, such a validation was carried out considering simultaneously for the blocks 8, 16 and 33 not only the computational but also human measurements doing the post-emergence stage V2 (Fig. 18).

To obtain the plant heights, it was necessary to take in account information from the ground truth and a calibration curve. Additionally, for the analyzed blocks, 17 998 maize plants were identified, according for 75% of the estimated total. Thus, there was a 25% of loss in emergency.

4. Conclusions

The maize plant population is of great importance to the agriculture cycle. In this study, a method for evaluate maize plants using images of maize crop acquired by drone flights was developed and validated. The proposed method is based on structuring orthomosaics of RGB channels obtained by embedded multispectral sensors, which enabled verification of the quality of images collected by drone with a GSD of 5.95 cm/pixel.

The geometric transform of rotation around the object in relation to the orthomosaics enabled the initial stages of image processing, for which an inclination angle $\theta \approx 25^\circ$ was determined to be suitable. Improving the brightness, contrast, and CLAHE led to improvements in intensities, and it was concluded that the HSV space provides better viability for the subsequent steps of the image processing. In the image segmentation step, techniques based on the use of HSV space established validated color ranges, which, together with morphological operators, enabled separation between maize plants. Locating of the maize plants in the crops was made possible by the integrated use of DT, and the normalized correlation coefficient of TM. With the inclusion of templates based on Chamfer masks, it became possible to determine candidates for the location of local maxima in relation to the DT.

From feature extraction, it was possible concluded that the chain codes described contours that represented circular features. Likewise, spectral histogram patterns that describe more open or closed lengths and widths of leaves, both related to the V2 stage, were observed.

For the classification of the post-emergence stage V2 of maize plants, the best results were obtained using a training to testing ratio of 70%–30% for SVM with Gaussian kernel, which demonstrated accuracies of 85% and 84% and a precision of 83% and 84%, for Classifier 1 and Classifier 2, respectively.

Finally, can be concluded that the organization of the evaluated techniques made it possible to validate the count of maize plants in the post-emergence stage or phenological stage V2, which is of interest to agriculture 4.0, research and service to the productive sector.

For future work, we are considering the possibility to evaluate the use of sensors that can provide depth information from the aerial images (Z-axis information) for automatically height maize plants calculation. Also, the evaluation of the use of semantic segmentation methods in conjunction with artificial intelligence for the evaluation of the spatial variability in the maize plants emergency and their productivity.

CRedit authorship contribution statement

Wilbur N. Chiuyari Veramendi: Conceptualization, Methodology/Study design, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Paulo E. Cruvinel:** Conceptualization, Methodology/Study design, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research was partially supported by the São Paulo Research Foundation, Brazil (FAPESP, Process No. 17/19350-2), and the National Council for Scientific and Technological Development, Brazil (CNPq - Process No. 133517/2019-5). We thank the institutional support received from the Brazilian Corporation for Agricultural Research (Embrapa) and the Post Graduation Program in Computer Science of the Federal University of Sao Carlos (UFSCar). The authors would also like to thank the agricultural engineer and senior researcher Luis Henrique Basso from Embrapa for the useful scientific discussions on maize crop management.

References

- Ahmad, I., Saeed, U., Fahad, M., Ullah, A., ur Rahman, M.H., Ahmad, A., Judge, J., 2018. Yield forecasting of spring maize using remote sensing and crop modeling in Faisalabad-Punjab Pakistan. *J. Indian Soc. Rem. Sens.* 46, 1701–1711.
- Bah, M.D., Hafiane, A., Canals, R., 2020. CRoWNet: Deep network for crop row detection in UAV images. *IEEE Access* 8, 5189–5200.
- Chen, Y., Hou, C., Tang, Y., Zhuang, J., Lin, J., He, Y., Guo, Q., Zhong, Z., Lei, H., Luo, S., 2019. Citrus tree segmentation from UAV images based on monocular machine vision in a natural orchard environment. *Sensors* 19 (24), 5558.
- Cruvinel, P., Minatel, E., 2002. Image processing in automated pattern classification of oranges. In: *World Congress of Computers in Agriculture and Natural Resources, Proceedings of the 2002 Conference*. American Society of Agricultural and Biological Engineers, p. 56, URL <https://elibrary.asabe.org/abstract.asp?aid=8312&t=1>.
- Deng, L., Mao, Z., Li, X., Hu, Z., Duan, F., Yan, Y., 2018. UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras. *ISPRS J. Photogramm. Remote Sens.* 146, 124–136, URL <https://www.sciencedirect.com/science/article/pii/S0924271618302533>.
- FAO, IFAD, UNICEF, WFP, WHO, 2020. *The State of Food Security and Nutrition in the World 2020*. The Publishing Group - FAO, pp. 1–320.
- Fernandez, R., Montes, H., Surdilovic, J., Surdilovic, D., Gonzalez-De-Santos, P., Armada, M., 2018. Automatic detection of field-grown cucumbers for robotic harvesting. *IEEE Access* 6, 35512–35527.
- Gnäding, F., Schmidhalter, U., 2017. Digital Counts of Maize Plants by Unmanned Aerial Vehicles (UAVs). *Remote Sens.* 9, 544.
- Hall, O., Dahlin, S., Marstorp, H., Archila Bustos, M.F., Öborn, I., Jirstrom, M., 2018. Classification of maize in complex smallholder farming systems using UAV imagery. *Drones* 2 (3), 22.

- Hasituya, Zhongxin, C., Fei, L., Yuncai, H., 2020. Mapping plastic-mulched farmland by coupling optical and synthetic aperture radar remote sensing. *Int. J. Remote Sens.* 41, 7757–7778.
- Hernández-Hernández, J., Ruiz-Hernández, J., García-Mateos, G., González-Esquiva, J., Ruiz-Canales, A., Molina-Martínez, J., 2017. A new portable application for automatic segmentation of plants in agriculture. *Agricult. Water Manag.* 183, 146–157.
- Johansen, K., Raharjo, T., McCabe, M.F., 2018. Using multi-spectral UAV imagery to extract tree crop structural properties and assess pruning effects. *Remote Sens.* 10 (6), 854.
- Kalantar, B., Idrees, M., Mansor, S., Abdul Halin, A., 2017. Smart counting – oil palm tree inventory with UAV. In: *Coordinates e-Navigation Assistance System for Cooperative Collision Avoidance at Sea*. Vol. XIII. pp. 17–22.
- Khan, P.W., Xu, G., Latif, M.A., Abbas, K., Yasin, A., 2019. UAV's agricultural image segmentation predicated by clifford geometric algebra. *Ieee Access* 7, 38442–38450.
- Koc-San, D., Selim, S., Aslan, N., San, B.T., 2018. Automatic citrus tree extraction from UAV images and digital surface models using circular hough transform. *Comput. Electron. Agric.* 150, 289–301.
- Laborde, D., Martin, W., Swinnen, J., Vos, R., 2020. COVID-19 risks to global food security. *Science* 369 (6503), 500–502.
- Liu, S., Baret, F., Andrieu, B., Burger, P., Hemmerlé, M., 2017. Estimation of wheat plant density at early stages using high resolution imagery. *Front. Plant Sci.* 8, 739.
- Lu, B., Dao, P.D., Liu, J., He, Y., Shang, J., 2020. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sens.* 12, 2659.
- Malambo, L., Popescu, S.C., Murray, S.C., Putman, E., Pugh, N.A., Horne, D.W., Richardson, G., Sheridan, R., Rooney, W.L., Avant, R., et al., 2018. Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *Int. J. Appl. Earth Obs. Geoinf.* 64, 31–42.
- Marinello, F., 2017. Last generation instrument for agriculture multispectral data collection. *Agric. Eng. Int.: CIGR J.* 19 (1), 87–93.
- Müller, A.G., Bergamaschi, H., Bergonci, J.I., Radin, B., França, S., Silva, M.d., 2005. Estimativa do índice de área foliar do milho a partir da soma de graus-dia. *Revista Brasileira de Agrometeorol.* 13 (1), 65–71.
- Oliver, F.W., Tansley, A.G., 1904. Methods of surveying vegetation on a large scale. *New Phytol.* 3 (9), 228–237.
- Pallottino, F., Figorilli, S., Cecchini, C., Costa, C., 2021. Light drones for basic in-field phenotyping and precision farming applications: RGB tools based on image analysis. In: *Crop Breeding*. Springer, New York, NY, pp. 269–278.
- Pereira, T.M., Gaspar, P.D., Simões, M.P., 2019. Fruit recognition and classification based on SVM method for production prediction of peaches-preliminary study. In: *IV Balkan Symposium on Fruit Growing 1289*. Leuven, Belgium, pp. 141–150.
- Popp, J., 2013. Global socio-economic and environmental dimensions of GM maize cultivation. *Food Nutr. Sci.* 4, 8–20.
- Pound, R., Clements, F.E., 1898. *A Method of Determining the Abundance of Secondary Species*. Minnesota, USA.
- Praveen Kumar, J., Domnic, S., 2019. Image based leaf segmentation and counting in rosette plants. *Inform. Process. Agric.* 6 (2), 233–246.
- Qiu, R., Zhang, M., He, Y., 2022. Field estimation of maize plant height at jointing stage using an RGB-D camera. *Crop J.* 10 (5), 1274–1283.
- Riehle, D., Reiser, D., Griepentrog, H.W., 2020. Robust index-based semantic plant/background segmentation for RGB-images. *Comput. Electron. Agric.* 169, 105201.
- Ritchie, S.W., Hanway, J.J., 1989. *How a Corn Plant Develops*. Tech. Rep., USA, URL <https://dr.lib.iastate.edu/handle/20.500.12876/90185>.
- Su, J., Coombes, M., Liu, C., Zhu, Y., Song, X., Fang, S., Guo, L., Chen, W.-H., 2020. Machine learning-based crop drought mapping system by UAV remote sensing RGB imagery. *Unmanned Syst.* 08 (01), 71–83.
- Thiel, E., Montanvert, A., 1992. Chamfer masks: Discrete distance functions, geometrical properties and optimization. In: *Proceedings., 11th IAPR International Conference on Pattern Recognition*. Vol. III. Conference C: Image, Speech and Signal Analysis. pp. 244–247.
- Veramendi, W.N.C., Cruvinel, P.E., 2021. Algorithm for the countering maize plants based on UAV, digital image processing and semantic modeling. In: *2021 IEEE 15th International Conference on Semantic Computing*. ICSC, IEEE, Laguna Hills, CA, USA, pp. 393–397.
- Wang, Z., Wang, K., Yang, F., Pan, S., Han, Y., 2018. Image segmentation of overlapping leaves based on Chan–Vese model and Sobel operator. *Inf. Process. Agric.* 5 (1), 1–10.
- Wang, Y., Zhu, X., Wu, B., 2019. Automatic detection of individual oil palm trees from UAV images using HOG features and an SVM classifier. *Int. J. Remote Sens.* 40 (19), 7356–7370.
- Weaver, J.E., 1918. The Quadrat method in teaching ecology. *Plant World* 21 (11), 267–283.
- Xie, X., Zhang, Y., Ling, X., Wang, X., 2019. A novel extended phase correlation algorithm based on Log-Gabor filtering for multimodal remote sensing image registration. *Int. J. Remote Sens.* 40 (14), 5429–5453.
- Zhang, X., Toudeshki, A., Ehsani, R., Li, H., Zhang, W., Ma, R., 2022. Yield estimation of citrus fruit using rapid image processing in natural background. *Smart Agric. Technol.* 2, 100027, URL <https://www.sciencedirect.com/science/article/pii/S2772375521000277>.
- Zhou, C., Liang, D., Yang, X., Yang, H., Yue, J., Yang, G., 2018. Wheat ears counting in field conditions based on multi-feature optimization and TWSVM. *Front. Plant Sci.* 9, 1024.