Definition of flight height for image monitoring of pupunha palm cultivation

Daniel Alexandre Botta de Siqueira[1], José Paulo Molin[2], Bianca Batista Barreto[3], Mateus Lima Silva[1], Peterson Ricardo Fiorio[2]

[1] Master's student at Department of Biosystems Engineering, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba 13418-900, SP, Brazil.

[2] Professor doctor at Department of Biosystems Engineering, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba 13418-900, SP, Brazil.

[3] Postdoctoral at Department of Biosystems Engineering, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba 13418-900, SP, Brazil.

A paper from the Proceedings of the
17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture
13-15 July 2026
Porto Alegre, Brazil

Abstract.

The use of Unmanned Aerial Vehicles (UAVs) associated with computer vision has broadened the applications of precision agriculture, especially in perennial crops that require detailed spatial monitoring. In the peach palm (Bactris gasipaes Kunth), the automated detection of clusters from aerial images represents an efficient alternative to conventional field surveys, supporting crop management and production estimates. This study evaluated the influence of flight height on the performance of the computer vision model YOLO11L (You Only Look Once) for automatic cluster detection, aiming to identify the most appropriate altitude for this application. It was hypothesized that increasing the flight height would improve operational efficiency while reducing detection performance due to the loss of spatial resolution. The experiment was carried out in a commercial field of 0.4 ha containing approximately 1,955 clumps spaced 2 × 1 m apart for the production of peach palm heart. The images were taken at flight heights of 20, 30, 40 and 60 m using a Mavic 2 Enterprise Advanced UAV equipped with a 48 MP RGB camera. Increasing flight height reduced acquisition time and increased ground sampling distance (GSD), ranging from 5.6 mm at 20 m to 16.7 mm at 60 m. Image selection was performed through a clustering-based sampling strategy, inspired by active learning, and approximately 50 representative images per altitude were annotated using the Smart Polygon tool in Roboflow. The datasets were divided into training (70%), validation (20%), and testing (10%), with data augmentation applied during pre-processing. YOLO11L models were trained using an input size of 1280 pixels, batch size of 2, and 100 epochs. Model performance was evaluated using Accuracy, Recall, mAP@0.5, and mAP@0.5:0.95. A progressive reduction in detection performance was observed as the flight height increased. The definition of the best altitude was based on mAP metrics, which jointly evaluate classification performance and location quality at different Intersection over Union (IoU) thresholds. The best results were obtained at 20 m (mAP@0.5 = 62.3% and mAP@0.5:0.95 = 32.3%), while the lowest performance occurred at 60 m (mAP@0.5 = 44.5% and mAP@0.5:0.95 = 19.7%). Although higher flight altitudes improved mapping efficiency, the reduction of spatial *detail negatively affected detection performance. The results demonstrate that flight height directly influences the accuracy of deep learning-based detection, with 20 m providing the best balance between spatial resolution and detection performance under the conditions evaluated.*

Keywords. *Bactris gasipaes, Deep learning, Unmanned Aerial Vehicles, Precision agriculture, Remote sensing*

1. INTRODUCTION

Modern agriculture faces the challenge of increasing the productivity and efficiency of production systems, while reducing environmental impacts and optimizing the use of available resources. In this context, Precision Agriculture (PA) has consolidated itself as a management strategy based on the collection, processing, and analysis of spatial and temporal data to support decision-making in the field (Molin et al. 2015). According to the International Society of Precision Agriculture (ISPA 2024), PA consists of an approach that integrates georeferenced data, agronomic information, and digital technologies to improve the production efficiency, profitability, and sustainability of agricultural systems.

The advancement of digital technologies has expanded the monitoring capacity of agricultural crops. In this context, Unmanned Aerial Vehicles (UAVs) stand out for the acquisition of high spatial resolution images and operational flexibility, allowing applications in Precision Agriculture, crop monitoring and productivity estimates (Jorge and Inamasu 2014; Nhamo et al. 2020; Rejeb et al. 2022). Compared to orbital sensors, UAVs enable greater spatial detail, favoring the individualized identification of plants and other applications that demand high spatial resolution (Bollas et al. 2021; Bukowiecki et al. 2021).

At the same time, deep learning techniques, especially those aimed at detecting objects, have been widely used in the automated analysis of agricultural images. Among the models available, the YOLO (You Only Look Once) family stands out for its balance between speed and accuracy, being used in different applications related to plant detection and monitoring (Redmon et al. 2016; Bochkovskiy et al. 2020; Khan et al. 2023). Recent studies demonstrate the potential of integrating UAVs and artificial intelligence for the automatic identification of agricultural crops, including plant counting and detection of individuals in aerial imagery (Bazamé et al. 2023; Putra and Wijayanto 2023; Kipli et al. 2023; Lee et al. 2024).

The pupunha palm (*Bactris gasipaes* Kunth) stands out among the perennial species cultivated in Brazil due to its importance for the sustainable production of hearts of palm. The crop has rapid growth, high tillering capacity and the possibility of successive harvests, constituting an alternative to the extraction of native palm trees (Bovi, 1998; Clement et al., 2010; Embrapa, 2019). Because it has a habit of continuous growth, the harvest of the heart of palm is carried out selectively and asynchronously throughout the year, based on the individualized cutting of the stems that have reached the ideal point of industrial maturation. This dynamic generates a constant variation in the architecture and leaf density of the clumps, making stand monitoring and yield estimates complex tasks. In addition to its economic relevance, the species plays an important role in the diversification of production systems and in the generation of income in agricultural properties. Monitoring the spatial distribution of clumps and plant stand dynamics is essential for production planning and productivity estimates.

In perennial systems, such as peach palm, the ability to identify and locate clumps individually over time is a fundamental requirement for evolutionary monitoring applications, allowing the monitoring of changes in the stand, expansion of tillers and possible development failures. Thus, the definition of acquisition protocols that preserve the spatial detail of the images becomes a determining factor for the practical application of these technologies in agricultural monitoring programs.

Despite the advances in computer vision applied to agriculture, the quality of the information extracted by the models depends directly on the characteristics of the images acquired. Among the main factors, the flight height stands out, which simultaneously influences spatial resolution, the area covered by the image, flight time and operational efficiency. In surveys carried out with UAVs, the increase in altitude provides greater coverage of the area and reduction in the time required for image acquisition, but results in a decrease in spatial resolution and detail of the targets of interest (Dash et al. 2017; Nhamo et al. 2020; Bollas et al. 2021; Bukowiecki et al. 2021). In applications that involve individual plant identification, this relationship can directly influence the ability of detection models to correctly recognize objects present in the images.

There are still few studies that specifically evaluate the influence of flight height on the performance of detection models applied to crops such as peach palm, as well as its implications for temporal crop monitoring applications. Considering that the spatial resolution decreases as the flight altitude increases, it is expected that changes in the level of detail of the images can influence the ability of the models to correctly identify the clumps present in the area.

Thus, the hypothesis of this study was that the increase in flight height provides greater operational efficiency in image acquisition, but reduces the ability to individualize and identify clumps due to the decrease in spatial resolution. Thus, the objective of this work was to evaluate the influence of flight height on the quality of the information obtained for the monitoring of the peach palm crop from images acquired by UAVs and analyzed with the YOLO11L model, identifying the most appropriate altitude for agricultural monitoring applications based on computer vision.

2. MATERIAL AND METHODS

The experiment was carried out in a commercial area for the production of peach palm (*Bactris gasipaes* Kunth) located in the municipality of Limeira, SP, Brazil (**Fig. 1**), at 585 m altitude and central coordinates 22°33'50.9" S and 47°18'15.8" W. The experimental area has approximately 0.4 ha, containing 1,955 clumps for the production of heart of palm. established at a spacing of 2 × 1 m. According to the Köppen climate classification, the region has a Cwa climate, characterized as subtropical with dry winter and hot summer. The average annual rainfall is approximately 1,300 to 1,400 mm and the average annual temperature is around 21 to 22 °C.

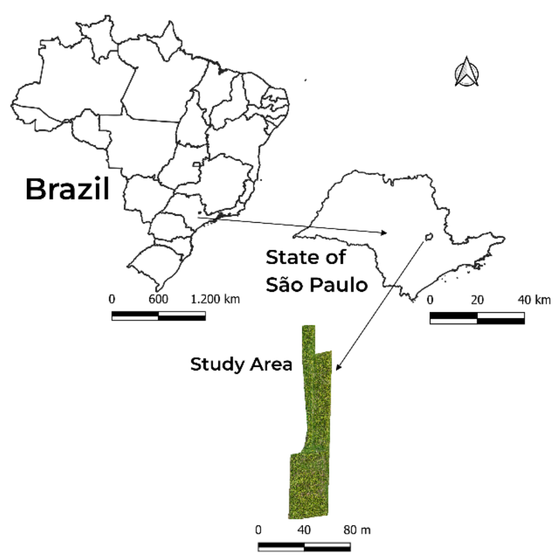

Fig. 1 Study area

The images were acquired on September 03, 2025 at around 10:00 h, using a Mavic 2 Enterprise Advanced unmanned aerial vehicle (UAV) (DJI Technology Co., Shenzhen, China), equipped with a 48 MP RGB camera stabilized by triaxial gimbal, producing images with a resolution of 8000 × 6000 pixels. The flights were carried out at heights of 20, 30, 40 and 60 m, using 80% frontal overlap and 70% lateral overlap. The heights resulted in different Ground Sample Distances (GSDs) and flight time (**Table 1**).

Table 1 Flight time and spatial resolution (GSD) for the different flight heights evaluated

Flight height (m)	Flight time	GSD (mm pixel^{-1})
20	17 min 05 s	5.6
30	13 min 17 s	8.3
40	4 min 00 s	11.1
60	3 min 12 s	16.7

For the georeferencing of the area, seven control points were strategically distributed at the site. The coordinates were obtained by a pair of RTK SP85 GNSS receivers (Spectra Geospatial, Westminster, Colorado, United States), with the base remaining in tracking for two hours over a point installed in the center of the experimental area, performing PPP processing and post-processing of the other coordinates obtained with the Rover. The equipment has multiconstellation capability (GPS, GLONASS, Galileo, BeiDou, QZSS, IRNSS and SBAS) and nominal accuracy of 8 mm + 1 ppm in the horizontal component and 15 mm + 1 ppm in the vertical component in RTK mode, ensuring high accuracy in the positioning of the control points used in photogrammetric processing. Additionally, all pupunha palm collected during the week prior to image acquisition were georeferenced for subsequent spatial validation of the detections. After the acquisitions, the images of each height were processed in the Metashape Professional software (Agisoft LLC, Russia) to generate the orthomosaics.

After the initial processing, a sample selection stage based on clustering was carried out using the K-Means algorithm, following a strategy inspired by *active learning* to reduce visual redundancy and increase the representativeness of the dataset. For this, normalized RGB histograms and Principal Component Analysis (PCA) were used, and approximately 50 representative images were selected for each flight height.

The selected images were inserted into the Roboflow platform (**Fig. 2**) (Roboflow Inc., USA), where the clumps were annotated using the Smart Polygon tool. For each height, an independent dataset was created, divided into training (70%), validation (20%) and testing (10%). During data preparation, data *augmentation techniques were applied*, including autoorientation, horizontal inversion, 90° rotations, and saturation adjustments between −20% and +20%. In addition, a three-fold magnification factor was applied to the training set, resulting in the correct division of the data, aiming to increase the variability of the samples and improve the generalization capacity of the models.

Fig. 2 Image being labeled on the Roboflow platform using the Smart Polygon tool (yellow labels)

The models were trained using the YOLO11L architecture, with an input resolution of 1280 pixels, *batch size* equal to 2 and 100 epochs. The evaluation was performed exclusively on the test set, using the metrics *Precision*, *Recall*, *F1-score*, *mAP@0.5* and *mAP@0.5:0.95*. Considering the objective of evaluating the influence of flight height on detection performance, metrics *mAP@0.5* and *mAP@0.5:0.95* were adopted as the main indicators of comparison between the heights evaluated, while *Precision*, *Recall* and *F1-score* were used as complementary metrics.

As a complementary analysis, inference was performed in complete orthomosaics using the YOLO11L model associated with the SAHI (*Slicing Aided Hyper Inference*) library, in order to evaluate the behavior of detections at field scale. Different slicing parameters, confidence threshold, overlap rate and post-processing methods were evaluated, and the configuration that presented the best performance for application in the orthomosaics of the four heights evaluated was selected.

Finally, the centroids of the bounding boxes generated by the model were converted to UTM coordinates via the Trace library and exported in CSV format. The spatial analyses and the validation of the detections in relation to the real booth were carried out in the QGIS software (QGIS Development Team, Switzerland).

3. RESULTS AND DISCUSSION

The acquisition of images at different flight heights resulted in notable changes in spatial resolution and operational efficiency of the survey. As the altitude increased, it was observed a reduction in the time required to cover the experimental area and an increase in the GSD, directly reflecting on the level of detail of the images.

The height of 20 m had the lowest GSD (5.6 mm) and the longest flight time (17 min and 05 s), while the height of 60 m had the longest GSD (16.7 mm) and the shortest acquisition time (3 min and 12 s). The intermediate heights of 30 m and 40 m showed intermediate values of spatial resolution and operational efficiency.

This behavior evidences the compromise between the quality of spatial information and the operational productivity of surveys carried out with UAVs. Although greater heights allow the mapping of more extensive areas in less time, the reduction of spatial resolution can compromise the ability to identify individual plants, especially in perennial crops conducted at high population density, such as pupunha palm.

According to Nhamo et al. (2020), Bollas et al. (2021) and Bukowiecki et al. (2021), the definition of flight height represents one of the most important factors in the planning of agricultural surveys, since it simultaneously influences the quality of the data obtained and the operational efficiency of the acquisition.

Table 2 presents the values of *Precision*, *Recall*, *F1-score*, *mAP@0.5* and *mAP@0.5:0.95* obtained for each flight height evaluated in the test set. In general, a progressive reduction in the performance of the models was observed with the increase in the altitude of image acquisition, evidencing the direct influence of spatial resolution on the detection capacity of peach palm clumps.

Table 2 Detection performance metrics across the different evaluated flight heights

Flight height (m)	Precision (%)	Recall (%)	F1-score (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)
20	**58.7**	**61.0**	**59.8**	**62.3**	**32.3**
30	**53.6**	**59.5**	**56.4**	**60.6**	**27.4**
40	**48.3**	**55.3**	**51.6**	**49.9**	**22.4**
60	**50.0**	**44.4**	**47.0**	**44.5**	**19.7**

The best results were obtained at the height of 20 m, which presented the highest values for all the metrics evaluated, with *mAP@0.5* of 62.3% standing out and *mAP@0.5:0.95* of 32.3%. On the other hand, the height of 60 m presented the lowest performances, with *mAP@0.5*% of 44.5% and *mAP@0.5:0.95* of 19.7%. These results indicate that the reduction in spatial resolution resulting from the increase in flight height compromises both the ability to identify the clusters and the accuracy of the positioning of the bounding boxes, as well as the difficulty in labeling for the generation of the dataset for training.

The results obtained corroborate observations described by Kipli et al. (2023) and Lee et al. (2024), which report a direct influence of spatial resolution on the performance of deep learning models employed in the detection of perennial crops. In higher resolution images, the morphological characteristics of plants are represented by a higher number of pixels, favoring the extraction of relevant visual patterns during training and inference.

In addition to the main detection metrics, the behavior of the *F1-score* showed a similar trend of reduction with the increase in flight height, as shown in **Fig. 3**. The indicator decreased from 59.8% at the height of 20 m to 47.0% at the height of 60 m, demonstrating a gradual loss in the balance between *Precision* and *Recall* as the spatial resolution became more limited.

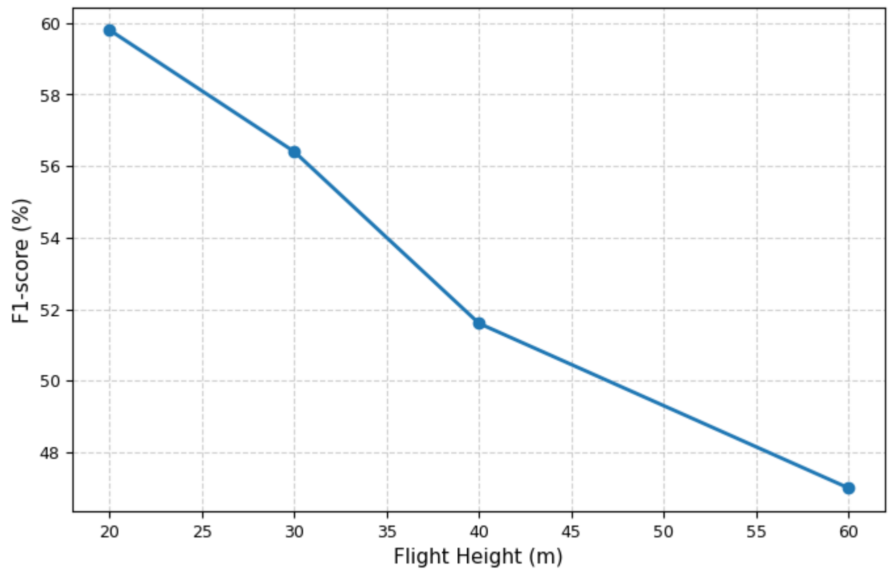

Fig. 2 Comparison chart as a function of *F1-score* and flight height

The reduction observed in the *F1-score* between 20 m and 60 m was approximately 21%, indicating a significant loss of the model's ability to correctly identify the clumps when images with lower spatial resolution were used. This reduction is associated with the loss of detail resulting from the increase in the GSD. There is also a consistent trend of decrease in the *F1-score* with the increase in flight height, evidencing the impact of spatial resolution on the predictive capacity of the model.

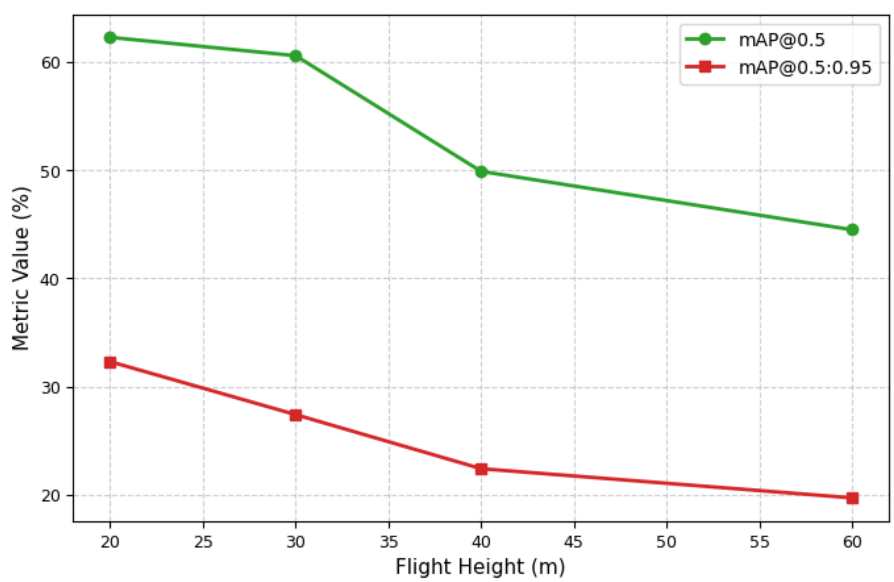

Fig. 3 Graph of the performance of *mAP* values as a function of detections at different flight heights

Analysis of location metrics revealed similarly consistent behavior. The indicator *mAP@0.5* decreased from 62.3% at the height of 20 m to 44.5% at the height of 60 m. Similarly, the *mAP@0.5:0.95* decreased from 32.3% to 19.7% in the same period. These results demonstrate that, in addition to reducing the ability to detect targets, the increase in flight height also compromises the spatial accuracy of the bounding boxes generated by the model, reflecting the loss of visual detail caused by the increase in GSD.

Considering that *mAP* is widely used as a reference metric for the evaluation of object detection models, the results obtained indicate that the height of 20 m provided the best overall performance of the YOLO11L in the conditions evaluated. Thus, the reduction in spatial resolution resulting from the increase in flight height simultaneously compromises the identification of clumps and the quality of their location in the images.

The observed reduction in *the F1-score* between the heights of 20 m and 60 m was approximately 21%, evidencing the decrease in the balance between *Precision* and *Recall* as the spatial resolution of the images was reduced by the increase in GSD. The dynamics of this behavior can be observed in **Fig. 2**, which shows the variation of the *F1-score* as a function of flight height. There is a general trend of reduced performance with increasing altitude, indicating that the loss of spatial detail progressively affects the model's ability to correctly identify pupunha palm clumps.

The visual analysis of the inferences allowed us to verify that the spatial resolution exerted a direct influence on the individualization capacity of the peach palm clumps. In the surveys carried out at a height of 20 m, it was observed that the crowns were better defined and that the adjacent individuals were better distinguished, allowing the model to identify the plants with greater precision. The vegetative structures presented more defined contours, favoring the extraction of relevant traits during training and inference.

On the other hand, in the images obtained at 40 m and 60 m, there was a progressive loss of visual details. The reduction in the number of pixels representing each clump made it difficult to differentiate between nearby plants and increased the occurrence of incorrect detections or omissions. The direct impact of this visual degradation on the divergent behavior between the accuracy of the predictions (*Precision*) and the recallability of the model is illustrated in **Fig. 4**.

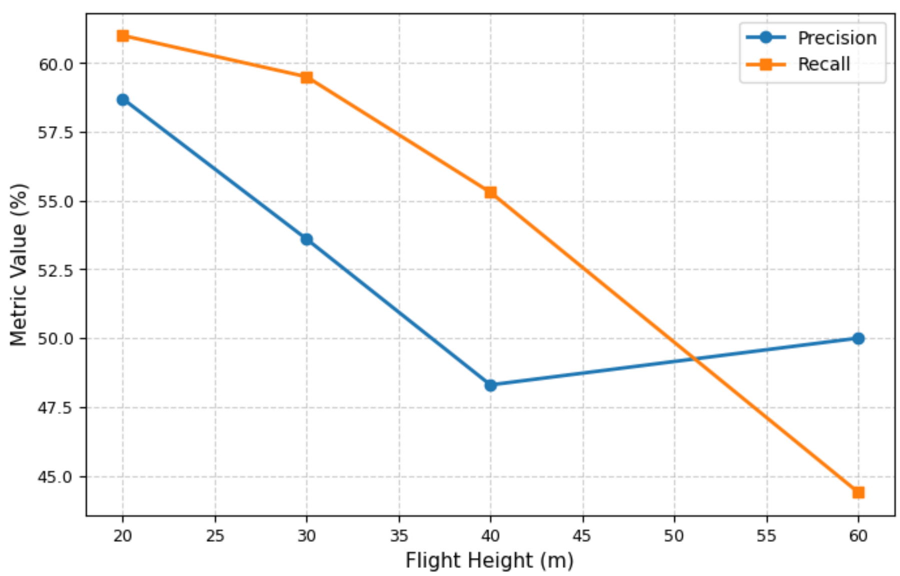

Fig. 4 Graph *of Precision* and *Recall* as a function of flight height

It is observed that both metrics showed a general trend of reduction with the increase in flight height. The *Recall* showed a behavior more sensitive to the degradation of spatial resolution, reducing from 61.0% at 20 m to 44.4% at 60 m. This result indicates that the model started to omit an increasing number of clumps as the GSD increased. *Precision* also showed a reduction between 20 m and 40 m, although there was a slight recovery at a height of 60 m. This behavior suggests that the model became more conservative at high altitudes, reducing the number of detections performed, but failing to identify part of the plants present in the area.

The decrease observed in performance metrics is not due to an isolated factor, but to the interaction between the morphological characteristics of the crop and the limitations imposed by the spatial resolution of the images. The cespitosa architecture of the peach palm, characterized by the continuous emission of tillers, favors the partial overlapping of the crowns and reduces the visual separation between adjacent individuals. At flight heights of 40 m and 60 m, corresponding to GSDs of 11.1 mm and 16.7 mm, respectively, the decrease in spatial detail reduces the ability to discriminate the boundaries between neighboring plants, resulting in an increase in the occurrence of undue groupings of clumps in a single detection or the omission of smaller and partially shaded individuals.

This result has great relevance for temporal crop monitoring applications. The correct individualization of the clumps is a fundamental requirement for the comparison of surveys carried out at different times and for the monitoring of the dynamics of the stand throughout the production cycle. Ferreira et al. (2016) highlight that spatial resolution represents one of the main factors associated with the quality of analyses performed with high-resolution images, especially when the objective involves the individual identification of biological targets.

The results showed the existence of a compromise between the quality of the information extracted by the model and the operational efficiency of the image acquisition. The lower flight heights provided better detection results due to the greater spatial detail of the images, but required more time to cover the experimental area. On the other hand, the increase in altitude significantly reduced the time needed to acquire the data, but resulted in a decrease in the model's performance.

The height of 20 m presented the best detection results, reaching *an F1-score* of 59.8%, while the height of 60 m presented the lowest performance, with *an F1-score* of 47.0%. Regarding operational efficiency, the opposite behavior was observed, with a reduction in flight time from 17 min and 05 s to only 3 min and 12 s between heights of 20 m and 60 m. This behavior shows the compromise between accuracy and operational productivity, as illustrated in **Fig. 5**.

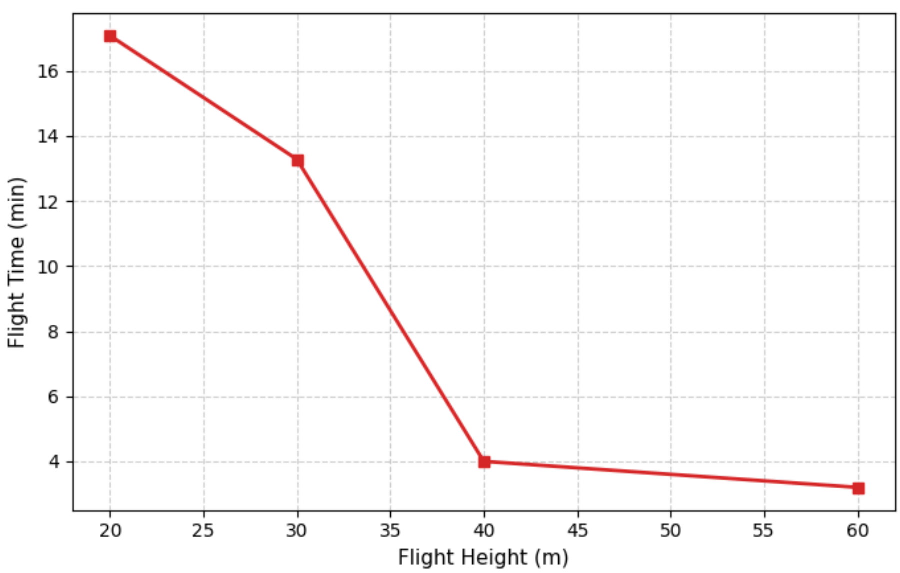

Fig. 5 Operational efficiency chart for image acquisition

Thus, the choice of flight height must consider the purpose of the application. For activities aimed at the evolutionary monitoring of the crop, in which the correct identification of clumps is essential, smaller heights tend to be more appropriate. On the other hand, applications focused solely on quickly obtaining general information about large areas can benefit from greater heights, as long as the performance reduction is acceptable.

The results obtained showed that the flight height directly influences the quality of the information extracted by computer vision models. Under the conditions evaluated, the height of 20 m provided the best overall performance of the model, presenting the highest values of *Precision, Recall, F1-score, mAP@0.5 and mAP@0.5:0.95*. These results indicate that the higher spatial resolution favored the individualization of the clumps and the quality of the detections, making this height the most promising for future crop monitoring applications.

Although the present study was conducted using images acquired in a single period, the results demonstrated the methodological potential for crop monitoring applications. The ability of the YOLO11L model to automatically identify the clumps present in the experimental area provides a basis for periodic surveys carried out throughout the production cycle, allowing to monitor changes in the spatial distribution of the plants and in the dynamics of the stand. Among the possible applications are the identification of planting failures, the monitoring of plant mortality and the monitoring of the evolution of the clumps over the years of exploration of the area. In addition, the use of UAVs significantly reduces the need for manual surveys, making it possible to obtain information at the field scale with greater speed and standardization.

In order to evaluate the applicability of the field-scale methodology, inference was made on complete orthomosaics using the YOLO11L model associated with the SAHI (Slicing Aided Hyper Inference) library. Initially, the orthomosaic obtained at a height of 20 m was selected, corresponding to the condition that presented the highest *mAP values* during the evaluation with isolated images. The direct application of the model on the complete orthomosaic resulted in only 563 clumps detected, compared to a real stand of approximately 1,955 clumps, evidencing limitations related to the scale of the images and the structural complexity of the culture.

For this, a progressive calibration of the slicing and post-processing parameters of the SAHI library was performed. The best results were obtained using slices of 4096 × 4096 pixels, match threshold of 0.45, confidence threshold of 0.15, overlap ratio of 0.27 and the GREEDYNMM algorithm for consolidation of detections. This configuration provided greater spatial stability of the bounding boxes and reduced the occurrence of redundant boxes on the same clump, improving the spatial consistency of the results obtained in the reference orthomosaic.

When the optimized configuration was applied in an exploratory way to the other orthomosaics at different flight heights, a marked reduction in detection capacity was observed, with 44 clumps detected at 30 m, 32 clumps at 40 m and no detection at 60 m. This behavior highlights the strong dependence of the slicing parameters in relation to the spatial resolution of the images and demonstrates that the degradation of visual detail resulting from the increase in GSD significantly compromises the performance of inference at a tile scale. It is noteworthy that this application had an exploratory character, being used mainly to investigate the influence of spatial resolution and the transferability of inference parameters between different flight heights, and not to establish a definitive operational methodology for counting in orthomosaics.

In addition to the geometric factors associated with the processing of the orthomosaics, the architecture of the culture itself contributed to limiting the performance of the model. The pupunha palm presents cespitous growth, continuous emission of tillers and high leaf density, favoring overlap between adjacent individuals. In more developed areas, the interlacing of the canopies reduces the visual separation between the clumps and hinders their individualization in images obtained in zenith perspective, leading to the grouping of multiple individuals in a single detection or the omission of partially covered plants. This behavior becomes even more evident as the flight height increases and the spatial resolution decreases, reducing the amount of information available for the discrimination of each individual.

These results demonstrate that, although the methodology has potential for agricultural monitoring applications based on computer vision, its efficiency depends directly on the spatial resolution of the images and the structural characteristics of the analyzed crop. The flight height of 20 m presented the best results both in the evaluation of the models and in the application on orthomosaics, preserving the level of detail necessary for the individualization of the clumps. According to Massruhá et al. (2020), the integration between digital agriculture, artificial intelligence, and remote sensing represents one of the main trends for the modernization of agricultural production systems. In this context, the results obtained indicate that the proposed methodology can contribute to future applications of monitoring this crop, as long as the limitations imposed by the spatial resolution and the high leaf density characteristic of the species are considered. As a future perspective, it is recommended the integration of three-dimensional information obtained by photogrammetry or LiDAR sensors, in order to complement optical data and expand the ability to identify plants in areas with high leaf density.

4. CONCLUSIONS

The flight height directly influenced the performance of the model YOLO11L in the automatic detection of peach palm clumps. The increase in altitude reduced the spatial resolution of the images and promoted a progressive decrease in the *mAP* values, evidencing the importance of spatial detailing for the individual identification of plants. Under the conditions evaluated, the height of 20 m presented the best detection results, representing the most suitable alternative for crop monitoring applications based on computer vision. Although greater heights increase the operational efficiency of image acquisition, the reduction in spatial quality compromises the performance of the model, demonstrating the need to balance spatial resolution and operational productivity in the definition of flight protocols.

Acknowledgments

The authors would like to thank Fapesp (Proc. 2022/09319-9) for the funding and CAPES for the first author scholarship (Proc. 88887.952591/2024-00).

References

Bazamé, H. C., Althoff, D., & Filgueiras, R. (2023). Deep learning and unmanned aerial vehicles for autonomous plant detection and counting in perennial agriculture. *Precision Agriculture*, 24(4), 1320–1342.

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint*, doi:10.48550/arXiv.2004.10934

Bollas, N., Koknou, E., & Tsouros, D. (2021). High-spatial-resolution imagery from UAVs for individual plant identification in dense agricultural systems. *Remote Sensing*, 13(18), 3642.

Bovi, M. L. A. (1998). *Cultivo da pupunheira para produção de palmito*. Campinas, SP, Brazil: Instituto Agronômico (IAC).

Bukowiecki, M., Studer, A., & Walter, A. (2021). The impact of flight altitude and spatial resolution on individual tree crown detection using UAV imagery. *Drones*, 5(3), 89.

Clement, C. R., Santos, L. A., & Flores, W. B. (2010). Domesticação e dispersão da pupunha (*Bactris gasipaes* Kunth) na Amazônia. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, 5(1), 67–83.

Dash, J. P., Watt, M. S., & Pearse, G. D. (2017). Influence of flight altitude and camera parameters on spatial resolution and object detection accuracy in UAV remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 62, 144–153.

Embrapa. (2019). *A cultura da pupunheira: sistemas de produção para o mercado de palmito*. Colombo, PR, Brazil: Embrapa Florestas.

Ferreira, M. P., Zortea, M., & Amarante, G. (2016). Resolução espacial e processamento de imagens de alta resolução aplicados à identificação individual de alvos biológicos. *Revista Brasileira de Cartografia*, 68(5), 987–1002.

ISPA. (2024). *Official definition of Precision Agriculture*. International Society of Precision Agriculture. https://www.ispag.org/about/definition. Accessed 12 January 2026.

Jorge, L. A. C., & Inamasu, J. Y. (2014). *Uso de veículos aéreos não tripulados em agricultura de precisão e monitoramento remoto*. São Carlos, SP, Brazil: Embrapa Instrumentação.

Khan, A., Khan, S., & Ahmed, M. (2023). Object detection models in smart agriculture: A review of the YOLO family from v1 to v8. *Journal of Agricultural Engineering Research*, 71(2), 115–138.

Kipli, K., Hoque, M. A., & Mahajon, S. (2023). Influence of flight altitude on the performance of object detection algorithms for deep-learning-based perennial crop monitoring. *Computers and Electronics in Agriculture*, 212, 108115.

Lee, S. H., Park, J. H., & Kim, D. H. (2024). Balancing spatial resolution and operational efficiency in drone-based remote sensing using YOLOv11 architectures. *IEEE Access*, 12, 45120–45135.

Massruhá, S. M. F. S., Leite, M. A. A., Oliveira, S. R. M., Meira, C. A. A., Luchiari Junior, A., & Bolfe, E. L. (Eds.). (2020). *Agricultura digital: pesquisa, desenvolvimento e inovação nas cadeias produtivas*. Brasília, DF, Brazil: Embrapa.

Molin, J. P., Amaral, L. R., & Colaço, A. F. (2015). *Agricultura de Precisão*. São Paulo, SP, Brazil: Oficina de Textos.

Nhamo, L., Magidi, J., Nyamugama, A., Clulow, A. D., Sibanda, M., Chimonyo, V., et al. (2020). Prospects of improving agricultural and water productivity through unmanned aerial vehicles. *Agriculture*, 10(7), 256.

Putra, Y. C., & Wijayanto, A. W. (2023). Automatic detection and counting of oil palm trees using remote sensing and object-based deep learning. *Remote Sensing Applications: Society and Environment*, 29, 100914.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779–788). Las Vegas, NV, USA.

Rejeb, A., Abdollahi, A., & Rejeb, K. (2022). Unmanned aerial vehicles (UAVs) in agriculture: A development and application trend analysis. *Smart Agricultural Technology*, 2, 100038.