World Journal of Information Systems 1(3) 71-82
DOI: https://doi.org/10.17013/wjis.v1i3.20

Article

Urban tree surveying using aerial UAV images and machine learning algorithms

Juan P. D'Amato [1,3], Pablo Rinaldi [1,2] and Gustavo Boroni [1,3]

[1] PLADEMA Institute - Universidad Nacional del Centro Prov. Bs. As

[2] CNEA

[3] CONICET

***** Correspondence: juan.damato@gmail.com

(Article history: Received June 07, 2024; Received in revised form July 15, 2024; Accepted August 02, 2024; Available online September 15, 2024)

Abstract: In this work, a novel approach to surveying urban trees based on automatic detecting from images captured using unmanned aerial vehicles (UAVs) is presented. Such a method is a cost-effective alternative to traditional measurement techniques. Through autonomous flights, UAVs capture detailed aerial imagery of urban areas, which is then processed to generate high-resolution raster images and elevation models. Machine learning algorithms are then applied to these images to identify trees, refining the detection process by eliminating false positives and estimating tree heights. Dealing with challenges such as flight time limitations and the irregularity of urban trees, the method achieves great accuracy in tree identification, not only covering the tree detection but also the separation between sidewalk and block interior trees, the estimation of height among other important data. Although the methodology does not cover all aspects of tree surveying, such as trunk health or diameter, it serves as a complementary tool to ground survey systems.

Keywords: machine learning, UAV, urban studies

1. Introduction

Urban trees are essential for enhancing and sustainably managing urban environments. For effective monitoring of such trees, consistent geo-location is necessary. Traditional techniques for measuring tree canopy characteristics, including height, diameter, canopy area, and volume, can be costly in terms of time and labor, particularly when precise measurements are needed across a city jurisdiction. In this context, aerial imagery presents a quicker and more cost-effective solution. Several works such as [3,4,6] remarks the important role of trees in planning urban spaces.

© 2025 ITMA. All rights reserved.

Unmanned aerial vehicles (UAVs), also known as drones, have significantly transformed various sectors by offering innovative solutions that were either impossible or very costly with traditional methods. Their ability to fly autonomously or be remotely controlled has made them incredibly versatile tools across numerous fields. Their use has several impacts in wildlife monitoring and forest management, they offer a non-intrusive method. In cities, they are also useful, as they can fly around places without the interruption of traffic, cars or other common elements. UAVs are generally equipped with high-resolution cameras and sophisticated sensors. At the same time, using drones has several challenges, as they have limitations specially related to flight time. If the areas are large, the UAV could not cover all in one pass, and some image combination and flight strategies are required.

In this work, a workflow for the surveying and detection of urban trees using autonomous flights with UAVs is presented. The work consists of several stages. The first involves autonomous flights with a commercial UAV that captures a sequence of photographs as it flies over the study area. The second is the generation of different raster-type images by combining the aerial photographs and the generation of the corresponding elevation model. In a third step, machine learning algorithms are run to identify trees in the mosaic-type images. Lastly, the detection of urban trees is refined by eliminating false positives, which include trees within block interiors. At the same time, the main features such as height are estimated with the help of the elevation model.

Despite each step in the process introducing some level of error due to its characteristics, and the existing trees in the study area being highly irregular, a precision level in identification close to 80% was achieved. The result of the process is the detection and characterization of the trees existing on the sidewalks of the surveyed streets, but a series of intermediate products are also obtained that are extremely useful for the municipality, such as updated aerial images with a high level of detail, the elevation model, and the detection of trees within block interiors. While there are some aspects of interest in the survey of specimens, mainly related to the trunk, such as health status or diameter, that are not covered by this methodology, it can be complementary to some other ground survey system. This work is part of a pilot test carried out in an urban area in the city of Tandil within a program managed by the municipality and financed by the IDB, aimed at demonstrating the feasibility of using these technologies for the efficient and rapid survey of urban trees.

2. State of art

Over the past decade, remote sensing techniques have become increasingly popular for studying forests for various purposes, offering a more efficient and less expensive alternative to traditional ground-based studies. As presented in [1] satellite and aerial imagery has enabled large-scale forest investigations, allowing for assessments of forest biomass, monitoring of changes in forest cover, and classification of different biomes. The advent of airborne multispectral and hyperspectral sensors has enhanced the level of detail in forest

© 2025 ITMA. All rights reserved.

observation, facilitating the classification of tree species, monitoring of forest health, and estimation of chlorophyll content. The introduction of Unmanned Aerial Vehicles (UAVs) like is discussed in [2,5] has further enriched remote sensing capabilities, providing higher-resolution data, versatile application, and the ability to carry diverse sensors like multispectral cameras for detailed vegetation health monitoring. Many studies have leveraged the synergy of UAVs equipped with multispectral sensors to examine agricultural crops for precision farming practices. These studies commonly utilize Vegetation Indices (VIs) and its variants. Such technology could effectively combine with existent satellital imaginary, as is discussed in [7].

In assessing the foliage characteristics of trees, it's critical to consider factors like tree size, branch structure, and density, as well as how leaves cluster together. The complexity increases with certain classes of trees that affect the sensor's received signal. There are only a handful of studies that have explored forest health assessment using UAVs and multispectral cameras. For example, [13] demonstrated the effectiveness in detecting stress in mature plantation trees using a qualitative classification approach. In the realm of quantitative, spectral-based approaches like Authors of [14] utilize very high-spatial resolution data, accurately defining the boundaries of individual trees is essential. The process of delineating tree crown boundaries has been the subject of various studies, leading to the development of multiple methods for detecting individual trees. In [10] demonstrated the use of a segmentation method on RGB orthoimages and a Height Model (CHM) derived from UAV data. In [8,12] expanded on this by applying a similar technique to multispectral data.

More recent advancements include the segmentation of LiDAR point clouds, as explained in [1,15], which has been shown to produce highly accurate results in identifying individual tree structures. Additionally, emerging techniques leveraging deep learning have been introduced for tree detection, illustrating the field's rapid evolution towards more sophisticated and precise methods [2]. These developments mark significant progress in the ability to quantitatively analyze forested areas using high-resolution spectral data

3. Styles and Formats

For carrying out such survey in cities, we have implemented a capturing workflow consisting in four streps:

Flight Planning: The process begins with planning the drone flight path to ensure comprehensive coverage of the targeted area. This step involves defining study area, setting waypoints and flight parameters based on the terrain and the specific requirements of the survey.

Data Acquisition: During the flight, drones capture a vast amount of data through their onboard sensors. High-resolution imagery data are collected as the drone autonomously follows its pre-defined route.

© 2025 ITMA. All rights reserved.

Image Stitching: The collected images and data are then processed using specialized software. This includes stitching together aerial photographs to create a continuous, high-resolution map of the area and analyzing temporal overlapped images to generate 3D representations of the terrain and tree structures. We have taken much attention to weather and the battery capacities, as drones could not flight under certain conditions and the limited flight capacities. Part of this methodology was already applied for discovering areas that could be affected by floods [11].

Data processing workflow: once the images are combined, a Convolutional Neural network (CNN) model is run for detecting trees. Once the potential trees are obtained, a post-processing is carried out, for improving the confidence on results, using additional geographical information, like streets or green maps.

The overall solution is presented in Figure 1.

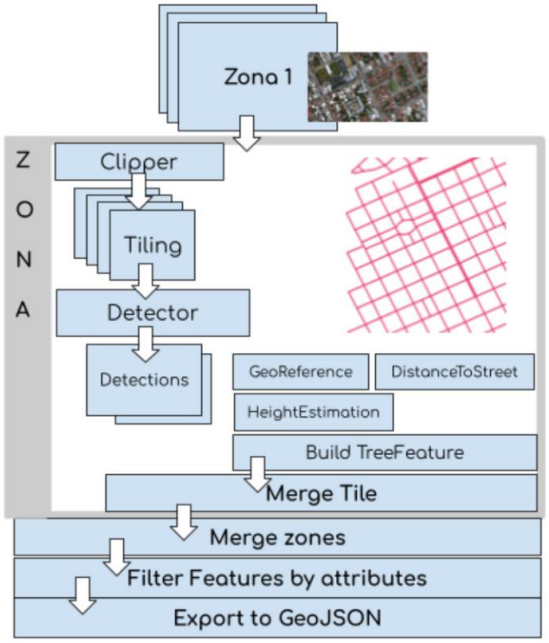

Figure 1. Architecture of the solution.

3.1. Subsection

The surface defined for the survey includes an area of 2.95 km2 (approximately 300 hectares) delimited by four (4) main avenues, as can be seen in figure 2. The Coordinates of the rectangle are stored in UTM (WGS84) Zone 21S.

© 2025 ITMA. All rights reserved.



Figure 2. Selected study area.

The equipment used for the survey is a DJI Phantom 4 Advanced multirotor drone with an approximate maximum autonomy of 25 minutes under ideal conditions. This translates into flights of a maximum of 20 minutes to maneuver with the necessary safety margin. It is equipped with a digital camera featuring a 1-inch, 20 Megapixel CMOS sensor, stabilized on a three-axis gimbal, and with integrated GPS to automatically geolocate the photos during capture. The equipment is shown in Figure 3.



Figure 3. DJI Phantom 4 Advanced employed for capturing images

For flight planning, a mobile app management software named Pix4DCapture, in conjunction with a Control+DJI interface software, was used. The app allows off-line flight planning (prior to the activity) by taking into account the specific equipment limitations (flight speed, autonomy, camera) and simulating possible routes. This flight plan, along with the configuration of its parameters, is transferred to the UAV for the equipment to perform the flight autonomously, always accompanied by an operator.

One of the most important parameters is the flight altitude at which the mission is programmed; higher altitudes cover a larger area in a single flight, but as a downside, the ground sampling distance or GSD is also greater since each image pixel will represent a larger surface area. The flight altitude is limited by current regulations to 122 meters (400 feet) above ground, and at this flight altitude, the Phantom 4 Advanced camera achieves a ground

© 2025 ITMA. All rights reserved.

sampling distance (GSD) close to 5 cm per pixel (5.34). This spatial resolution is more than sufficient for the task of recognizing tree vegetation and can even be reduced to work more efficiently with the images.

Since the UAV establishes the flight altitude at the takeoff point, it is necessary to analyze the altitude differences in the flight area to avoid violating regulations (flying above 120 meters) or flying too low. Based on our experience [11] and for areas like the one in this study where the altitude variation is minimal, it is most advisable to set a maximum flight altitude of 90 meters above the highest point of the area to be surveyed. Once the takeoff and landing point is defined, the flight altitude is corrected based on the altitude difference between this point and the highest point. The remaining parameters are set optimally: maximum flight speed, vertical camera alignment, and 70% image overlap.

3.2. Data collection

To carry out the survey of the proposed study area, initially, 6 flight plans or missions with similar characteristics were proposed, based mainly on the UAV's autonomy for the defined flight parameters but also on the characteristics of the area. In this division, it was also considered necessary to have a sufficient time margin between the estimated travel time and the UAV's autonomy. Then, once in the field, flight conditions and the distance to the chosen takeoff points necessitated some plans to be reconfigured, so the final missions amounted to 7.

Figure 4 shows the 7 final flight plans, where the simulated route for the UAV's flight in each of them can also be seen. This simulation was performed with the Pix4DCapture App. The areas are planned with a slight overlap to include a common strip in both flights.



© 2025 ITMA. All rights reserved.

Figure 4. Selected areas as part of the flight planning process

For safety reasons, it is necessary for the takeoff and landing area, where the remote pilot will be located, to have a line of sight with the UAV throughout the flight. Therefore, in urban areas with some tall buildings, this location must be in a clear and far area or on the terrace of one of the buildings. The chosen spot must be a traffic-free zone, without nearby trees, and away from antennas and overhead wires. It is also important that there be no nearby power transformers or other sources that could generate magnetic fields capable of altering the normal operation of the UAV's internal compass. The time of year the flights were conducted was specifically planned so that deciduous trees would have their green leaves, which would allow for the use of the images with automatic classification algorithms. On average, each flight covered an area of approximately 43 hectares, taking 230 photographs with a resulting Ground Sampling Distance (GSD) of 5.4 cm.

As the surveys include the total area surveyed, it is necessary to identify the streets of the city to somehow differentiate the urban road trees from the internal trees of the lots. To identify the city streets of the city, a file is included in geojson format with a list of linear “figures”, where each figure has different attributes, such as: name, type, height, among others. The streets sector corresponding to the UAV path is shown below.

3.2.1 Ground Control Points:

Ground Control Points (GCPs) are used to correct distortions in data and reference them in known coordinate systems. A Ground Control Point is a measurement of a mark made on the ground, typically using high-precision GPS, but they can also be obtained from previous measurements or other georeferencing systems.

If the data is available before the flights are conducted, these can be visibly marked to facilitate their identification in the images. When integration occurs after the flight, it's important to identify points that are easily locatable in the images. For this project, control points extracted from Google Earth were used since the height values correspond to the SRTM surface models of recognized accuracy.

The importance of these points for the final objective of this project is relative since it's only required that all products generated in each mission are aligned with each other and with respect to the street system. The terrain height values of the DSMs are only relatively important when estimating tree heights. For each of the missions, between 5 and 6 control points were chosen, distributed homogeneously and ensuring that there are more than three images per flight where they appear.

3.3 Image Stitching

The final output of image process are two big images:

© 2025 ITMA. All rights reserved.

- Orthomosaic color
- Digital Surface Model (DSM)

An orthomosaic is a georeferenced image created by combining and processing multiple aerial images, taken in this case from the UAV. This georeferenced image of the flown-over area is the primary product for the automated detection of trees. The orthomosaic itself also represents a highly valid tool for manual surveys from a GIS tool or for any other type of applications that require an updated aerial view with precision, such as the detection of undeclared buildings.

The DSM or elevation model is also a georeferenced grid of the flight area but represents the height value of each point in each pixel. The DSM obtained from UAV images always shows the highest element in the image at each point, which could be the roof of a house, the terrace of a building, or the canopy of a tree. This image is also useful for other applications in topography, urban planning, and different studies related to terrain geometry. Additionally, once a tree is detected, the canopy height and surrounding terrain data can be used to estimate the approximate height of the plant. It is normal for some noise to occur along the boundaries of the zones due to the lower number of available images, which is why an overlap area was left between each mission. A typical DSM looks like Figure 5.

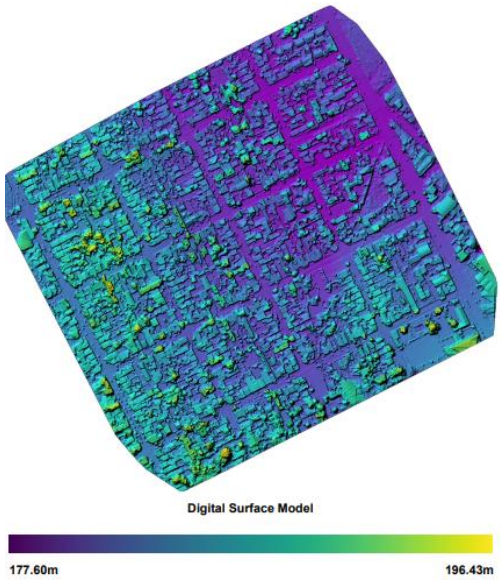

Figure 5. Preview del modelo de superficie generado de la zona 8.

3.4 Data Processing workflow

The final step consists in processing the images for getting almost the exact position of trees, while they can also be classified by different criteria's. For classification, it was used a Convolutional Neural Network architecture based on Yolo [16]. Such architecture has shown a great performance and detection relation. As images are very big. more than

© 2025 ITMA. All rights reserved.

20,000x20,000 pixels, we have split them into smaller areas, run the detector and finally mix the results.

The images from each survey flight are a map created from multiple photographs, which forms an orthomosaic. The study area was divided into several zones (one for each flight) generating a set of products from each one, among which is the orthomosaic of each zone. As the ideal image size for processing is between approximately 10,000 and 20,000 pixels per side to be able to be processed on a standard computer, a larger size may require a server with greater performance.

3.4.1 Training the CNN model

As it was mentioned, to perform tree detection, a Convolutional Neural Network (CNN) model based on the YOLO architecture in its version 8 was chosen. This is relatively easy to train and achieves very good results in image classification to detect specific objects, with high performance, capable of working on both CPU and GPU. This network was trained using RGB images dataset (orthophoto) available on the web. To achieve more efficient training, additional tasks such as collecting source images and labeling them are necessary, manually selecting where the objects to be detected (trees in this case) are located. The labeling process can be slow but is indispensable. The generated annotations are exported as a JSON file for each image.



Figure 6. Streets layer.

3.4.2 Refining detections

After applying the CNN model on all UAV images, a list of potential detections is obtained with the corresponding confidence score. Such detections could contain false positives (non-real trees), false negatives (non-detected) and overfits, product of multiple passes on the same image. For removing such cases, there are computed additional attributes,

© 2025 ITMA. All rights reserved.

that are finally evaluated and accepted or discarded using different thresholds. Such computed attributes are:

- Amount of green of each detection
- Distance to closest street
- Height below a threshold

Additionally, an Intersection-over-union strategy (IoU) [8] is applied for discarding overlapping detections, keeping the ones with higher scores.

For computing the distance to the street, metric an extra geo layer is added, that defines each street of the area. In figure 6 an overlay of the selected layer is shown.

4. Results

To conduct a comprehensive analysis of all areas, a total of four days were required for flying and obtaining all necessary images using the UAV, particularly due to some weather conditions, such as wind, and in search of optimal lighting conditions. As final a result, 12 images were generated, 6 orthomosaics and 6 DSMs, which required approximately 30GB of storage. Such computation took about 1 day.

To carry out the complete analysis process, it was used Python based open source libraries such as OpenCV, Yolo, and Shapely for geographic data processing. We use the well-known metrics as Precision and Accuracy in **equation 1** for measuring results, where:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \qquad Preccision = \frac{TP}{TP+FP} \tag{1}$$

where TP = True positive; FP = False positive, TN = True negative and FN = false negative

Each pair of images (RGB+DSM) were processed independently, applying the whole proposed pipeline and finally merged into a single GeoJson type data layer, which contains all the detected trees, along with information such as estimated height, estimated width, and distance to the nearest street. Processing these images took about 30 minutes on a laptop with an Intel i7-13620h processor, 32GB of RAM, and an Nvidia RTX 4060 GPU.

© 2025 ITMA. All rights reserved.



Figure 7. Showing tree detections (red) potential (green) selected ones.

In total, 2,400 trees were detected, of which a total of 74% were considered to belong to trees on public streets and a 26% belonged to private ones. Part of the detections were manually validated, and precision and accuracy was computed, obtaining a value of precision = 92% and accuracy of 80%. This relatively lower accuracy was due to the trained model could not detect all trees in quite large tree areas. A partial detection is shown in Figure 7.

5. Conclusions

Urban tree coverage is an important metric that all governments try to improve. For obtaining such metric, we have described a method based on computer vision algorithms applied on UAVs images. Manipulating such UAVs could be challenging, so the plan for covering large areas was also included. It also is a cost-effective method, that could be applied in any city of any country.

The results seem promising, as the accuracy was higher enough and detail description corresponds with manual verification. Even though, there are still plenty of things to work on. The height estimation stills should be validated and measured. Also, temporal analysis could be valuable, for verifying tree health.

Acknowledgements

The current work is part of an international project funded by the Inter-American Development Bank, Argentina 2023.

References

[1] Li, J.; Yang, B.; Cong, Y.; Cao, L.; Fu, X.; Dong, Z. 3D Forest Mapping Using A Low-Cost UAV Laser Scanning System: Investigation and Comparison. *Remote Sens.* 2019, *11*, 717.

© 2025 ITMA. All rights reserved.

[2] Neuville, R.; Bates, J.S.; Jonard, F. Estimating Forest Structure from UAV-Mounted LiDAR Point Cloud Using Machine Learning. *Remote Sens.* 2021, *13*, 352. https://doi.org/10.3390/rs13030352

[3] Berland, A.; Shiflett, S.A.; Shuster, W.D.; Garmestani, A.S.; Goddard, H.C.; Herrmann, D.L.; Hopton, M.E. The role of trees in urban stormwater management. *Landsc. Urban Plan.* 2017, *162*, 167–177

[4] Ritter, B. Use of Unmanned Aerial Vehicles (UAV) for Urban Tree Inventories. Ph.D. Thesis, Clemson University, Clemson, SC, USA, 2014. Available online: https://tigerprints.clemson.edu/all_theses/1890

[5] Tomaštík, J.; Mokroš, M.; Saloň, Š.; Chudý, F.; Tunák, D. Accuracy of photogrammetric UAV-based point clouds under conditions of partially-open forest canopy. *Forests* 2017, *8*, 151.

[6] Osborn, J.; Dell, M.; Stone, C.; Iqbal, I.; Lacey, M.; Lucieer, A.; McCoull, C. Photogrammetry for Forest Inventory: Planning Guidelines. Deployment and Integration of Cost-Effective, High. Spatial Resolution, Remotely Sensed Data for the Australian Forestry Industry; Australian Government Department of Agriculture, Fisheries and Forestry (DAFF): Canberra, Australia, 2017.

[7] Dash, J.; Pearse, G.; Watt, M. UAV Multispectral Imagery Can Complement Satellite Data for Monitoring Forest Health. *Remote Sens.* 2018, *10*, 1216

[8] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, S. Savarese : Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression, Computer Vision and Pattern Recognition, 2019. https://doi.org/10.48550/arXiv.1902.09630

[9] Gallardo-Salazar, J.L.; Pompa-García, M. Detecting Individual Tree Attributes and Multispectral Indices Using Unmanned Aerial Vehicles: Applications in a Pine Clonal Orchard. *Remote Sens.* 2020, *12*, 4144.

[10] Lim, Y.S.; La, P.H.; Park, J.S.; Lee, M.H.; Pyeon, M.W.; Kim, J.-I. Calculation of Tree Height and Canopy Crown from Drone Images Using Segmentation. *J. Kor. Soc. Survey. Geodesy Photogram. Cartogr.* 2015, *33*, 605–614.

[11] Rinaldi, P. DAmato, J, Larrabide, I. "Drone Based DSM Reconstruction for Flood Simulations in Small Areas: A Pilot Study", New Knowledge in Information Systems and Technologies, 2019, DOI: 10.1007/978-3-030-16181-1_71

[12] Díaz-Varela, R.; de la Rosa, R.; León, L.; Zarco-Tejada, P. High-Resolution Airborne UAV Imagery to Assess Olive Tree Crown Parameters Using 3D Photo Reconstruction: Application in Breeding Trials. Remote Sens. 2015, 7, 4213–4232.

[13] Dash, J.P.; Watt, M.S.; Pearse, G.D.; Heaphy, M.; Dungey, H.S. Assessing Very High Resolution UAV Imagery for Monitoring Forest Health during a Simulated Disease Outbreak. ISPRS J. Photogramm. Remote Sens. 2017, 131, 1–14

[14] Franklin, S.E. Pixel- and Object-Based Multispectral Classification of Forest Tree Species from Small Unmanned Aerial Vehicles. J. Unmanned Veh. Syst. 2018, 6, 195–211.

[15] G. Parmehr, E.; Amati, M. Individual Tree Canopy Parameters Estimation Using UAV-Based Photogrammetric and LiDAR Point Clouds in an Urban Park. Remote Sens. 2021, 13, 2062. https://doi.org/10.3390/rs13112062

[16] F. Xiaofeng Fu, W. Guoting, Y. Xia, L. Yongshun, and Yuming Bo. Efficient yolov7-drone: An enhanced object detection approach for drone aerial imagery. Drones, 7(10), 2023. ISSN 2504-446X. doi: 10.3390/drones7100616. URL https://www.mdpi.com/2504-446X/ 7/10/616

© 2025 ITMA. All rights reserved.