

RESEARCH TITLE

**IMPROVE THE DETECTION OF TREE CROWNS ON DRONE
IMAGES BY MACHINE LEARNING ALGORITHM**

Mohammed Issa Mohammed Al-hayani¹

Doç. Dr. HAKAN SEPET²

¹ Ahi Evran University, Türkiye.

Email: Mohammed.pilot10@gmail.com

² Ahi Evran University, Türkiye.

Email: hakansepet@gmail.com

HNSJ, 2023, 4(5); <https://doi.org/10.53796/hnsj4529>

Published at 01/05/2023

Accepted at 25/04/2023

Abstract

Huge advances in technology have allowed developers to modernize the way they handle data. Data management is one of the most important areas of AI today. AI systems rely on creating a way to analyze and manage data. The data we consider in our research is data from areas with trees such as forests. We collect and process data containing images taken by aircraft from above. The reason for the increase in land pressure, splitting of tree crowns under the high canopy, and deep forest areas remain unresolved tasks. High-density short-range unmanned aerial vehicles can help solve this problem to some extent. We develop a smart model based on machine learning in which image processing tools are used to analyze forest aerial images. A comprehensive description of the image can be provided and the trees in it can be identified. Captured trees are also classified and statistics are provided for each classification.

Key Words: Drones, Object detection, Mask R-CNN, Prediction

Introduction

Forests are a vital component of world ecology and constitute one of the most significant land areas on Earth. Quantitative study of the forest's structure, environmental modeling, biomass assessment, and deforestation assessment all require timely and precise readings of forest characteristics at the level of individual trees, such as the number of trees, tree height, and canopy size. It can be very helpful for determining tree age to manage the future to be able to identify trees from high-resolution overhead photos by looking at the shape and structure of tree tops. Locating trees that could fall on a house or other structure, or calculating the percentage of a property's land that is covered by trees.

Correct woodland tracking and research help to minimize the dangers we may encounter [1]. Furthermore, the availability of systems for studying and editing forest pictures provide prior knowledge about the condition of forests, their abundance, and the types of plants found within them. This data can be used as a management tool to help save woodlands. Because forests are the primary source of pure oxygen and a clean atmosphere in our lives, preserving them and giving them the tools needed to lower their risk is one of the most essential mechanisms in assisting all humankind.

A system is described that examines the state of plants in woods using a picture processing method based on machine learning and artificial intelligence techniques. The position of treetops is found by removing their location from pictures taken by the drone. It is an essential instrument to acquire a thorough study of the character of the forest and knowledge about the abundance and existence of trees by using our suggested method, thus preserving them and providing an important service to the people.

To know the importance of forest data and image management, we must know the importance of forests in our daily life. People benefit greatly from the existence of forests [2]. Forests produce oxygen and reduce carbon dioxide and are a great source of food, water, paper, and wood. It is also important for animals, they live in the shade of the forest during summer days which helps protect them from the winter wind. Despite the importance of forests, they are currently being destroyed and withdrawn at an alarming rate; Where forests have been destroyed by fires and in many ways for the timber industry and agriculture. 18.7 million acres of forest are disappearing annually One of the largest rainforests in the world, the Amazon has lost at least 17% of its forests in the last half century [3].

There are three stages in our proposed algorithm: Crowns and nested areas can be highlighted as part of preprocessing. The input to the algorithm can be an image in RGB format, in natural or artificial color, with visible channels. Panchromatic images can also be used. We will develop the luminance histogram for simpler crown segmentation as part of preprocessing. In the second step, the treetops can be segmented using a multiple-threshold approach. According to the results, we are classifying trees after using morphological techniques to smooth out the segment boundaries

Related researches

Many academics focus on tree detection because of image processing and machine learning. Ke et al. [4] addressed crown-based tree genus detection. In this paper, the experts stressed the importance of thorough and prompt information for forest management. High-resolution spatial photos made forest survey and analysis accurate and cost-effective. Several studies on forest survey standards and data have yielded tree crown spotting algorithms. This essay discusses negative remote sensing and its methods. Automatic spotting methods were categorized by tree top. These methods are applied to examine photos and their kinds, along with the features of the studied areas. This study examines measurement accuracy to spot trees and its impacts on methods.

Lindberg et al. [5] linked the fast development of remote surveillance methods to the availability of a lot of three-dimensional information through digital point clouds and aerial laser scanning. The paper discussed 3D tree discovery and its use in forestry and environmental studies. Recent research has focused on tree-level three-dimensional data analysis. Previous methods used two-dimensional tree crown data. These tree analysis and information methods were powerful and efficient. Three-dimensional data methods demand technical property creation. This method is affected by tree length

because tall trees give exact results using surface models, but short trees lying under tall trees require more precise and effective ways to study and examine them. Many devices can detect treetops, which can be used to create plant maps. but small trees under tall trees need more precise and effective analysis and discovery.

Larsen et al. [6] test six treetop detection/demarcation algorithms on an image sample of six forest kinds in three European nations. Local maximum detection, valley tracking, area growth, template matching, scale-space theory, and random frame methods are used in the algorithms. The drone photos show a uniform crop, lonely trees, and a dense deciduous forest. No program can assess every scenario. Breaking images into regular forest samples is necessary before using tree spotting algorithms, according to the study. It also shows the need for a uniform, openly available test photo collection and test methods to assess tree detection/limit detection algorithms. Finally, even human translators have trouble distinguishing tree tops in complex forest kinds using binocular photos.

Terrestrial surveys often fail to properly record tropical forest canopies, which have a variety of treetops of different heights and forms, according to Aubry-Kientz et al.[7]. Airborne laser scanning data can identify these tops, but analysis methods for arctic or temperate forests may need to be adapted for tropical forests. The researchers compared six segmentation methods on six 39-hectare areas in French Guiana. Researchers disrupted crown projections in high-resolution photos.

Ecosystem benefit dispersal requires tree crown tracing. Others based their work on it [8]. The time-consuming nature of area sampling and regional variety make exact, up-to-date urban crown mapping tough. For large-scale, low-cost forest map analysis, data cost is another issue. The mask region-based convolutional neural network model (Mask R-CNN) and Google Earth data were used to build a new framework for spotting treetop cover in New York Central Park, a tested model for a wooded environment. Mostly urban with uneven tree top cover. For the total study area, the crown region had an 81.8% detection rate and the Mask R-CNN crown detection model projected 82.8% of trees. The program identified lone trees and closed forest tree areas after removing 87.5% and 81.6% of tree numbers. The treetop recognition program can map urban treetop edges and identify trees in highly complex scenarios, according to the study.

Chang et al. [9] used LiDAR data to locate tree tops by geometrically linking forest local maximums and minimums. LiDAR data was subtracted for tree tops and crown limits. Treetops were put on the tree rings nearest to the four area minimums. We identified 77% of reference tree tops using LiDAR data from thick and mixed woods in Korea at 4.3 points/m². The regression line between field data and results underestimated tree height and crown girth. Forest conditions and high-point-density data should be reviewed.

Terms definition

1 - GeoTIFF

By associating a raster picture with a specified model space or map projection, GeoTIFF allows georeferencing and geocoding information to be stored in raster files that are TIFF 6.0 compatible. Inheriting the file format specified in the relevant section of the TIFF standard, GeoTIFF files are TIFF 6.0 [TIFF_6] files [10]. Images acquired via satellite imaging systems, digitized aerial photos, scanned maps, digital elevation models, or the results of geographic studies can all have their related cartography data described in the GeoTIFF format utilizing a predetermined set of TIFF tags. Figure 1 is an example of a tree area picture in GeoTIFF format.



Figure1: GeoTIFF drone image [16]

It is possible to keep a wide variety of georeferencing information in GeoTIFF, including both real and future coordinate systems. Projection classes, coordinate systems, datums, ellipsoids, etc. in GeoTIFF are all described via integer identifiers. Refer to Section 2.4. GeoTIFF File and Key Structure and the Appendices of the Format Specification, Revision 1.0 for information on Tag ID, Key ID, and number codes. Similar to the TIFF format, GeoTIFF can only be up to 4 GB in size because of the 32-bit coordinates required [11]. BigTIFF is a variation of TIFF that was created to meet the requirements of geographic information systems (GIS), large format scanning (LFS), medical imagery, and other areas by allowing for images larger than 4 GB through the use of 64-bit offsets.

2 - Mask R-CNN

Segmenting detected objects at the pixel level is a common task for deep learning instance segmentation techniques, and Mask R-CNN is a prominent method [12]. There is no problem with overlapping objects or numerous classes for the Mask R-CNN algorithm. When it comes to segmenting images, the state-of-the-art tool is the Convolutional Neural Network (CNN) known as Mask R-CNN. This specific type of Deep Neural Network can recognize items within a picture and produce a reliable segmentation mask for each one. When an item is identified multiple times, an instance segmentation [13] map is generated to better categorize each object. Independent of their class, individual objects are handled separately in instance segmentation. Semantic segmentation, on the other hand, treats all instances of a given class as though they belonged to the same thing.

To train a network to recognize new groups or to fine-tune its current settings, transfer learning can be used. In most cases, the masking object is used to set up an existing Mask R-CNN [14] model for transfer learning. Next, the trainMaskRCNN algorithm is used to train the model based on the provided training data. As Figure 2 and according to the architecture; the Mask R-CNN model will be graded based on how well it performs the assessment Instance Segmentation task.

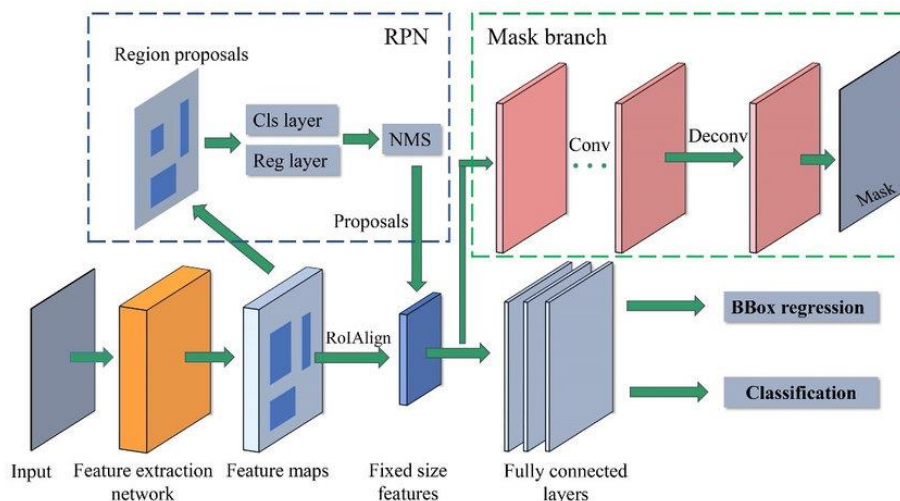


Figure 2: Mask R-CNN model architecture [17]

To set up a Mask R-CNN network for transferring data, you must build a mask R-CNN object and give it the titles of the classes and the anchor boxes to use. Additional network characteristics, such as the network input size and the ROI pooling values, can be specified if desired. In the Mask R-CNN network, there are two phases. In the first step, boundary boxes for proposed objects are predicted using an anchor box [15]. An R-CNN detector is then used in the second step to further refine the suggestions, classify them, and calculate the segmentation at the pixel level. With our proposed model Mask R-CNN is used to identify the areas with trees in the aerial photos. Through the previously fully trained neural network, the images entered into the system are compared. Parts that contain trees can be obtained in the form of distinct and specific sections.

Masking is an image processing method in which we define a small 'image piece' and use it to modify a larger image. Masking is the process that is underneath many types of image processing, including edge detection, motion detection, and noise reduction. Mask R-CNN is a deep neural network designed to address object detection and image segmentation, one of the more difficult computer vision challenges. The Mask R-CNN model generates bounding boxes and segmentation masks for each instance of an object in the image.

Methodology of the Prediction model

Prediction in image processing systems depends on prior recognition of the type of images and the required results. Machine learning algorithms can be used to build systems that predict processing outcomes. A project model based on the integration of machine learning algorithms and tools is used to identify trees. Its main components are reading aerial imagery data, implementation of Mask R-CNN to perform tree crown delineation from RGB images, and identification of tree crowns from used image models and scripts for tree age analysis. Obtaining precise delineations of trees through the application of machine learning frameworks that provide cutting-edge deep learning recognition and segmentation techniques. To effectively distinguish tree crowns, an established deep learning model known as Mask R-CNN was utilized. A top-down RGB picture that was recorded by a drone is used in conjunction with a pre-trained model to determine the position of tree canopies as well as their overall extent. The proposed model's structure method is shown in Figure 3.

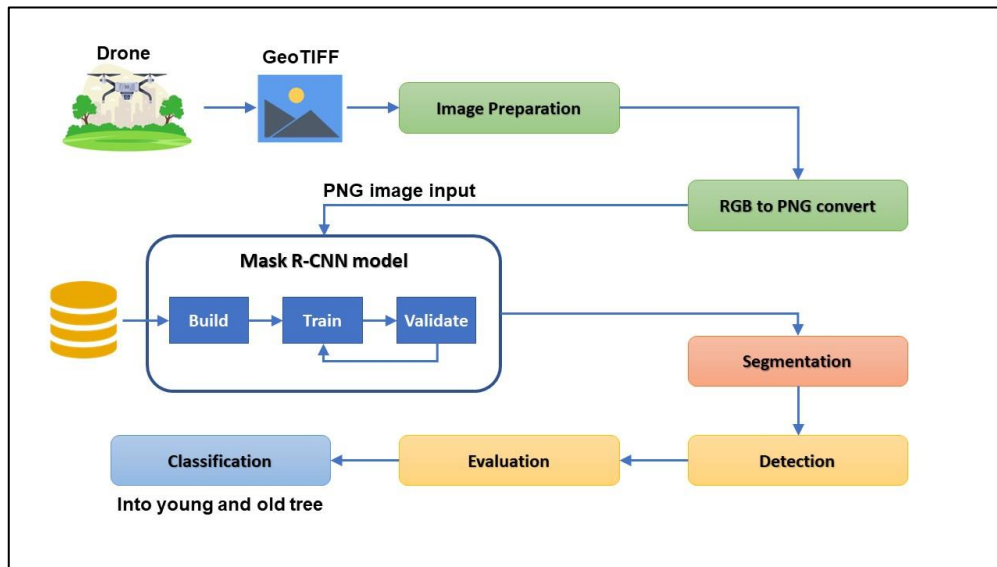


Figure 3: proposed model’s structure

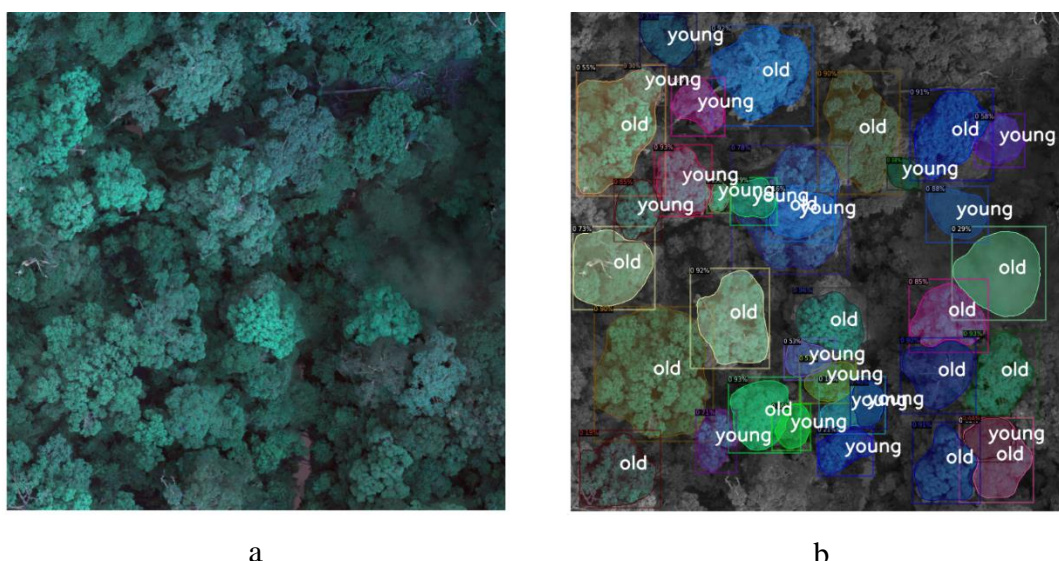
The system works on pictures taken of trees using drones. Images are taken at different heights to establish a very suitable angle of view to increase accuracy and efficiency. The images are extracted in GeoTIFF format to be the first step in the proposed system. After entering the images into the system, preparation and preparation of these images is done. Pre-processing includes image conditioning, sharpening, image scaling, and blemish removal. In the following steps, the images will be entered into the Mask R-CNN model, so the images will be converted from RGB format to PNG. The pre-trained Mask R-CNN works on specific image formats.

The next stage is to segment the picture of the trees so that we can identify specific tree regions. The most fundamental model outcome for finding trees is segmentation. The model's results are used to identify wooded regions, with a percentage precision assigned to each. There is a system in place to categorize the identified plants, separating them into two groups: ancient and new. The tree's girth, which can be used to estimate its age, and its overall form from above are the primary factors used in the categorization.

Experiments and results analysis

We evaluate the accuracy of the system depending on three basic experiments. Each one of these experiments is known as a case. The three cases are visualized by displaying the input images of the system as well as the results as follows

Case 1: Figure 4 represents the images for this case. Image (a) is the main photo taken with the drone. Image (b) is the image generated using our proposed model.



a b

Figure 4: Case1 (a) is the input image and (b) is the output

In this image 36 trees are detected in the system. The images are identified by segmenting the main image. Because it contains large and small trees, the trees were classified into 17 old and 19 young trees.

Case 2: this case is represented by the pictures in Figure 5, which can be found below. The primary photograph obtained from the drone is shown in the image (a). Image (b) is the result of applying our suggested model to the image generation process.

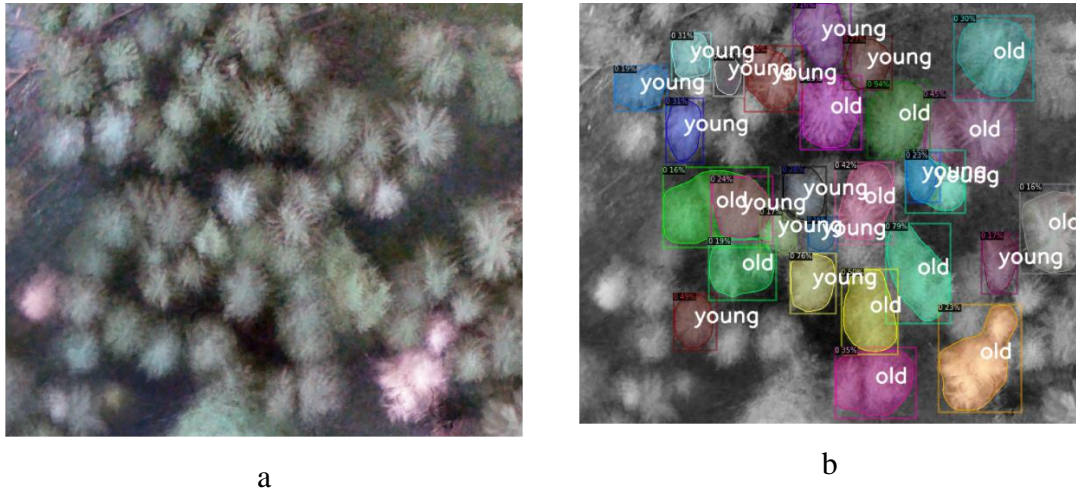


Figure 5: Case2 (a) is the input image and (b) is the output

This picture contains a total of 28 plants that have been identified by the system. The primary picture is segmented, and each of the images can then be recognized. The trees were divided into 12 older trees and 16 young trees since there were both large and small trees present.

Case 3: this case is depicted in Figure 6, which can be found further down this page. The principal image captured by a drone is shown in the image. (a). Image (b) is the consequence of implementing our proposed model in the process of image generation.

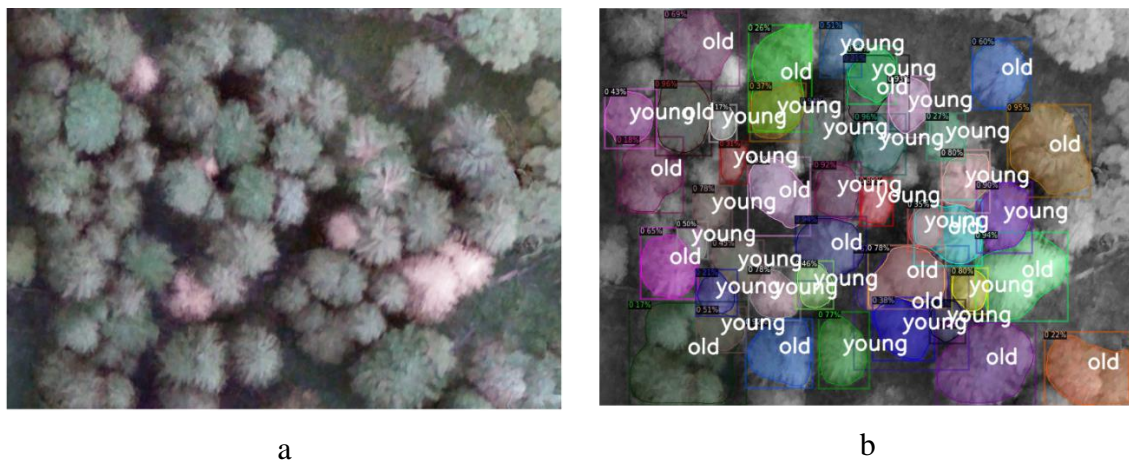


Figure 6: Case3 (a) is the input image and (b) is the output

There are a total of 44 plants identified by the algorithm in this picture. Segmenting the primary image is used to locate the pictures. The trees were divided into two groups, ancient and young because there are both big and little trees there: 18 old trees and 26 young trees. The following table (Table 1) is the outcome of all experiments.

Table 1: three cases outcomes

Case no	Total detections	Old trees	Young trees
Case 1	36	17	19
Case 2	28	12	16
Case 3	44	18	26

Accuracy in the identification process is a component of each step in the process of locating trees in the pictures. The numbers of the system's accuracy can be determined by using these percentages as a basis for comparison. When that number is relatively close to one hundred, it indicates that the system is reliable. In Table 2, we take a look at how accurate the algorithm was for each of the three different experiments.

Table 2: model accuracy

Case no	Highest accuracy	Mean accuracy
Case 1	96%	74.36%
Case 2	94%	53.90%
Case 3	96%	78.95%

Measurement precision of the algorithm during the tree detection step allows us to achieve a best-case scenario accuracy of 96% based on the model values.

Conclusions and future work

Our proposal can be categorized as one of the intelligent tree site image processing systems. This project's primary objective was to develop algorithms for classifying individual treetops by species. To accomplish this, individual tree heads must be chosen or separated. The objective of the segmentation phase is to divide the tree crowns into segments whose shape corresponds to that of the crown in the original image. In the classification phase, the morphology of the tree crowns is used.

Using image processing and machine learning techniques, we propose developing a new mechanism to detect and classify trees by age by scrutinizing and analyzing images captured from above, as part of forest protection. Images captured by drones in actual forests are used to configure the system. Google's "Google Maps" and "Google Earth" services can also take advantage of the cloud technologies it provides. The objective that has been met is the creation of a useful instrument that uses the collected data to determine the types and density of trees, to provide the most appropriate identification and analysis, thereby facilitating their protection in the event of a threat. Experiments indicate that the system functions with exceptional precision. The accuracy rate for the cases examined was 95.33 %. In the future, it may be recommended to use optimization algorithms to determine the most effective methods to develop tree recognition by determining the age-based relationship between trees.

References

- [1] Wessman, C. A., Archer, S., Johnson, L. C., & Asner, G. P. (2004). Woodland expansion in US grasslands: assessing land-cover change and biogeochemical impacts. *Land change science: observing, monitoring and understanding trajectories of change on the Earth's surface*, 185-208.
- [2] Bamwesigye, D., Hlavackova, P., Sujova, A., Fialova, J., & Kupec, P. (2020). Willingness to pay for forest existence value and sustainability. *Sustainability*, 12(3), 891.
- [3] Vancutsem, C., Achard, F., Pekel, J. F., Vieilledent, G., Carboni, S., Simonetti, D., ... & Nasi, R. (2021). Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Science Advances*, 7(10), eabe1603.
- [4] Ke, Y., & Quackenbush, L. J. (2011). A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. *International Journal of Remote Sensing*, 32(17), 4725-4747.
- [5] Lindberg, E., & Holmgren, J. (2017). Individual tree crown methods for 3D data from remote sensing. *Current forestry reports*, 3(1), 19-31.
- [6] Larsen, M., Eriksson, M., Descombes, X., Perrin, G., Brandtberg, T., & Gougeon, F. A. (2011). Comparison of six individual tree crown detection algorithms evaluated under varying forest conditions. *International Journal of Remote Sensing*, 32(20), 5827-5852.
- [7] Aubry-Kientz, M., Dutrieux, R., Ferraz, A., Saatchi, S., Hamraz, H., Williams, J., ... & Vincent, G. (2019). A comparative assessment of the performance of individual tree crowns delineation algorithms from ALS data in tropical forests. *Remote Sensing*, 11(9), 1086.
- [8] Yang, M., Mou, Y., Liu, S., Meng, Y., Liu, Z., Li, P., ... & Peng, C. (2022). Detecting and mapping tree crowns based on convolutional neural network and Google Earth images. *International Journal of Applied Earth Observation and Geoinformation*, 108, 102764.
- [9] Chang, A., Eo, Y., Kim, Y., & Kim, Y. (2013). Identification of individual tree crowns from LiDAR data using a circle fitting algorithm with local maxima and minima filtering. *Remote sensing letters*, 4(1), 29-37.
- [10] Treboux, J., Genoud, D., & Ingold, R. (2018, May). Decision tree ensemble vs. nn deep learning: efficiency comparison for a small image dataset. In *2018 International Workshop on Big Data and Information Security (IWBIS)* (pp. 25-30). IEEE.
- [11] Gumelar, O., Saputra, R. M., Yudha, G. D., Payani, A. S., & Wahyuningsih, S. D. (2020, June). Remote sensing image transformation with cosine and wavelet method for SPACeMAP Visualization. In *IOP Conference Series: Earth and Environmental Science* (Vol. 500, No. 1, p. 012079). IOP Publishing.
- [12] He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "Mask R-CNN." *ArXiv:1703.06870 [Cs]*, January 24, 2018 - arxiv.org/pdf/1703.06870.
- [13] Hafiz, A. M., & Bhat, G. M. (2020). A survey on instance segmentation: state of the art. *International journal of multimedia information retrieval*, 9(3), 171-189.
- [14] Bharati, P., & Pramanik, A. (2020). Deep learning techniques—R-CNN to mask R-CNN: a survey. *Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2019*, 657-668.
- [15] Dollár, K. H. G. G. P., & Girshick, R. (2017, October). Mask r-CNN. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).
- [16] besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.13472
- [17] Yang, Z., Dong, R., Xu, H., & Gu, J. (2020). Instance segmentation method based on improved mask R-CNN for the stacked electronic components. *Electronics*, 9(6), 886.