Advanced Engineering Days

aed.mersin.edu.tr

Deep learning based poplar tree detection and counting using multispectral UAV images

Ismail Colkesen*[1], Taskin Kavzoglu [1], Umut Gunes Sefercik [1], Osman Yavuz Altuntas[1], Mertcan Nazar [1], Muhammed Yusuf Ozturk [1], Mustafacan Saygı [1]

[1] *Gebze Technical University, Engineering Faculty, Department of Geomatics Engineering, Kocaeli, Türkiye, icolkesen@gtu.edu.tr*

Cite this study: Colkesen, I., Kavzoglu, T., Sefercik, U.G., Altuntas, O. Y., Nazar, M., Ozturk, M. Y. & Saygı, M. (2023). Deep learning based poplar tree detection and counting using multispectral UAV images. Advanced Engineering Days, 6, 64-67

Keywords
Poplar trees
Deep learning
Tree detection
Tree counting
YOLOv7

Abstract

Poplars (Populus sp.), a member of fast-growing and short-lived tree species, have been widely planted since ancient times. Identification and mapping of poplar planted areas on global and local scales, as well as the automatic crown detection and counting of individual poplar trees in a given area provide valuable information to decision-makers in developing strategies for planting area estimation, growth monitoring and yield estimation. At this point, offering significant advantages compared to traditional methods, remote sensing technologies, especially unmanned aerial vehicle (UAV) systems, have become a prominent data source in individual tree crown detection. In this study, one of the latest You Only Look Once (YOLO) algorithms, YOLOv7, was applied to high-resolution multispectral UAV captured images to detect and count individual poplar (P. deltoides) trees. For this purpose, a UAV-derived orthomosaic image covering dense hybrid poplar tree plantations in the Akyazı district of Sakarya province was used as the primary data source. Training and validation datasets were created from the orthomosaic with a total of 260 images and 19.989 instances. Results showed that the YOLOv7 model achieved the precision, recall, mean average precision, and F-score values for the bounding boxes of poplar trees as 86.50%, 86.80%, 87.80%, and 88.20%, respectively.

Introduction

Poplar trees, a leading member of the family of fast-growing trees, are extensively used as raw material (e.g., timber and fibre), especially in the forest industry, and are also of great importance in terms of environmental and agricultural benefits [1]. Therefore, identification and mapping of poplar planted areas on global and local scales, as well as the automatic crown detection and counting of individual poplar trees in a given area provide valuable information to decision-makers in developing strategies for planting area estimation, growth monitoring and yield estimation. At this point, offering significant advantages compared to traditional methods, remote sensing technologies, especially unmanned aerial vehicle (UAV) systems, have become a prominent data source in individual tree crown detection studies in the literature [2-5]. The availability and widespread use of very high-resolution imagery has led to a proliferation of studies around the application of machine learning techniques, including deep learning (DL) models, due to their robustness in information extraction from the imagery [6]. Various DL models have been successfully applied in the literature, reporting the effectiveness in tree crown detection and counting. For example, Yu et al. [7] evaluated the performance of the Mask Region-based Convolutional Neural Networks (Mask R-CNN) model in broad-leaved tree crown extraction from multispectral-UAV imagery. Zhu et al. [8] applied the YOLO (You Only Look Once) v4 algorithm for fruit tree detection from very high spatial resolution UAV imagery, while Yildirim et al. [9] utilized the YOLOv4 to detect stone pine crowns from the RGB-UAV imagery. Gan et al. [3] utilized DeepForest and Detectree2 DL-based algorithms to delineate individual tree crowns from the UAV-captured RGB imagery. In this study, YOLOv7, the relatively new state-of-the-art DL-based object detection model, was applied to detect and count individual poplar trees using multispectral UAV-derived images.

Study area and dataset

The study area covering dense hybrid poplar tree plantation sites is located in Çıldırlar neighbourhood of Akyazı district, Sakarya province, Türkiye (Figure 1). Although agriculture is the main economic activity, poplar cultivation is an essential forestry activity for the region's economy. In order to achieve the ultimate goal of this study, five bands multispectral (MS) (red, green, blue, red edge and near-infrared) images of the study area were taken by DJI Phantom 4 MS UAV on 05 November 2021. The UAV has a RTK GNSS receiver and own positioning capability that means, no need ground control points for absolute orientation of the aerial photos. The spatial resolution of the UAV camera is 2.08 MP however, by means of the short focal length, the ground sampling distance (GSD) reaches half of the 20 MP RGB UAV cams. The aerial photos were collected from 90 m flight altitude with 80% forward and %60 side overlap, resulting in GSD of 4.86 cm. Structure from motion (Sfm)-based image matching software Agisoft Metashape Professional was used for geometric correction, camera reflectance calibration and whole photogrammetric processing. Finally, a 16-bit resolution orthomosaic was produced in original grid spacing.

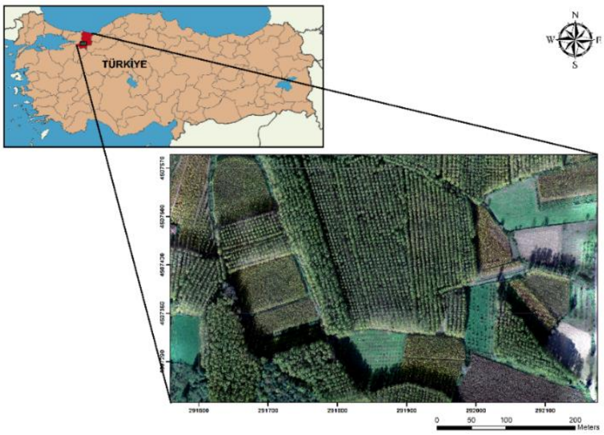

Figure 1. The location of the study area and the UAV ortho-mosaic imagery

Methodology

In this study, individual poplar trees were detected and counted from the multispectral orthomosaic using the YOLOv7 deep learning model. The YOLO model structure usually consists of backbone, neck, and head parts. In YOLO models, image frames are highlighted thanks to the backbone [10]. Another purpose of the backbone part is to extract the basic features of the image. Therefore, the selection of backbone architecture is essential to increase object detection accuracy [11]. The features extracted from the backbone are fused and mixed in the neck part and transmitted to the head of the network. The final estimate was utilized using the non-maximum suppression technique. The technique mainly removes the boxes with low overlap by selecting the most appropriately represented among the duplicate and overlapping suggestion boxes of the objects [12]. Following these steps, YOLO models estimate the probabilities of the location and class membership for each object detected with bounding boxes. Among the real-time object detection models, the recently released YOLOv7 has been introduced as an improved model concerning processing speed and prediction accuracy compared to its previous versions [13]. Unlike its previous versions, YOLOv7 uses enhanced efficient layer aggregation networking (E-ELAN) architecture in the backbone part, based on expand, shuffle, and merge cardinality for continuously increasing the learning ability of the model.

Results

To construct a YOLOv7 model, 50 images of 640x640 and 1280x1280 pixel sizes were selected from the orthomosaic and labelled as poplar polygons in Roboflow. Data augmentation methods (i.e., Gaussian blur, vertical and horizontal cropping, 90 degrees clockwise, and counter-rotation) were also applied to increase the labelled training data size artificially. As a result, 260 images containing 19.989 poplar samples were created, and the dataset was randomly divided into 80% training and 20% validation. PyTorch deep learning framework was used to construct the YOLOv7 model and implemented on a high-capacity workstation with NVIDIA GeForce RTX 3090 graphics card, Intel® Core™ i9-12900K 3.2GHz 24-Core ~3.2GHz processor, and 128GB Ram available at GTU Geomatics Engineering Department's Advanced Remote Sensing Technology Laboratory (ARTLAB). The model parameters to be set by the user side in the training phase were selected as 640x640 image size, 1000 epoch, 4 batch size, 0.01 learning rate, 0.937 momentum, and 0.005 weight decay.

Two loss measures, including the bounding box and objectiveness, were calculated for the YOLOv7 model in the training and validation processes. No overfitting was observed in the constructed model, and loss values showed a decreasing trend during the training stage (Figure 2). Standard accuracy measures, namely precision, recall, mean average precision (mAP), and F-score values, were also calculated for the validation dataset, and the accuracy values for the bounding boxes of poplar trees were found to be 0.8650, 0.8680, 0.8780, and 0.8820, respectively. In order to further evaluate the performance of the model, precision, recall, and F-Score values were also estimated for the two test datasets covering different areas with varying image sizes selected apart from the training and validation dataset. First test image contained 1024x1024 pixels and 115 samples labelled as poplar trees. When the constructed YOLOv7 model was applied to the test image, 111 True Positive (TP), 4 False Positive (FP) and 4 False Negative (FN) trees were predicted. As a result, precision, recall and F-score values for poplar trees were 0.9652, 0.9652 and 0.9652, respectively. The second test image comprised 512x512 pixel sizes and 54 samples labelled as poplar trees. The trained YOLOv7 model correctly predicted 48 test samples as poplar trees (TP), whereas the remaining 6 were miss-detected (FP). For this test dataset, estimated precision, recall and F-score values were 0.8888, 1.0000 and 0.9411, respectively.

Figure 2. Loss graphs of the trained YOLOv7 model.

To visually analyze the performances of the trained YOLOv7 model, the resulting predictions representing bounding boxes for detected poplar trees for the first test image were shown in Figure 3 where the red boxes represent the accurately detected poplar trees, whereas the blue and purple boxes indicate FP and FN predictions, respectively. If the bounding boxes predicted by the model overlapped more than 50% with ground truth, the detected bounding boxes were marked as TP, otherwise as FP. On the other hand, other land cover types or trees miss-detected as poplar by the model were labelled as FN. As can be seen from the figure, poplar trees that are very close to each other were labelled as false positives, while some other broadleaf trees in the area were detected as false negatives. Finally, 119 objects were detected in the first test image, 111 of which were correctly labelled as poplar trees.

Figure 3. The results of poplar detection and counting on the UAV image

Conclusion

The main goal of this study is to automatically detect and count individual poplar trees from the UAV-derived multispectral imagery using the state-of-art YOLOv7 deep learning-based object detection model. The qualitative and quantitative evaluations based on accuracy measures and visual interpretation showed that the YOLOv7 model with mAP of about 88% showed robust performance in detecting individual poplar trees from the UAV-derived imagery. Accuracy results calculated for the two test images indicated that the YOLOv7 model correctly detected the individual poplar trees over the %94 in terms of the F-score measure. While the results verified the usefulness and effectiveness of the DL-based model in automatic individual tree detection and counting problems, comparative studies are required to assess the performances of different DL models on poplar tree detection and segmentation of tree crowns using very high-resolution remotely sensed imagery.

References

1. Isebrands, J. G. & Richardson, J. (2014). Poplars and Willows: Trees for Society and the Environment CAB International and Food and Agriculture Organization of the United Nations (FAO), CABI, p. 634.
2. 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.
3. Gan, Y., Wang, Q., & Iio, A. (2023). Tree Crown Detection and Delineation in a Temperate Deciduous Forest from UAV RGB Imagery Using Deep Learning Approaches: Effects of Spatial Resolution and Species Characteristics. Remote Sensing, 15(3), 778.
4. Ateşoğlu, A., Kavzoğlu, T., Çölkesen, İ., Özlüsoylu, Ş., Tonbul, H., Yılmaz, E. Ö. & Öztürk, M. Y. (2022). Türkiye'de Hızlı Büyüyen Türlere Ait Spektral Kütüphane Kurulması: Kavak Türleri Çalışması. Bartın Orman Fakültesi Dergisi, 24 (2), 324-338.
5. Colkesen, I., Kavzoglu, T., Atesoglu, A., Tonbul, H., & Ozturk, M. Y. (2022). Multi-seasonal evaluation of hybrid poplar (P. Deltoides) plantations using Worldview-3 imagery and state-of-the-art ensemble learning algorithms. Advances in Space Research. https://doi.org/10.1016/j.asr.2022.10.044
6. Diez, Y., Kentsch, S., Fukuda, M., Caceres, M. L. L., Moritake, K., & Cabezas, M. (2021). Deep learning in forestry using UAV-acquired RGB data: A practical review. Remote Sensing, 13(14), 2837.
7. Yu, K., Hao, Z., Post, C. J., Mikhailova, E. A., Lin, L., Zhao, G., ... & Liu, J. (2022). Comparison of classical methods and mask R-CNN for automatic tree detection and mapping using UAV imagery. Remote Sensing, 14(2), 295.
8. Zhu, Y., Zhou, J., Yang, Y., Liu, L., Liu, F., & Kong, W. (2022). Rapid Target Detection of Fruit Trees Using UAV Imaging and Improved Light YOLOv4 Algorithm. Remote Sensing, 14(17), 4324.
9. Yildirim, E., Nazar, M., Sefercik, U. G., & Kavzoglu, T. (2022). Stone Pine (Pinus Pinea L.) Detection from High-Resolution UAV Imagery Using Deep Learning Model. In IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium (pp. 441-444). IEEE.
10. Vimal, S., & Subbulakshmi, P. (2016). Secure data packet transmission in MANET using 18 enhanced identity-based cryptography. International Journal of New Technologies in Science and Engineering, 3(12), 35-42.
11. Chen, Y., Yang, T., Zhang, X., Meng, G., Xiao, X., & Sun, J. (2019). Detnas: Backbone search for object detection. Advances in Neural Information Processing Systems, 32.
12. Rothe, R., Guillaumin, M., & Van Gool, L. (2015). Non-maximum suppression for object detection by passing messages between windows. In Computer Vision–ACCV 2014: 12th Asian Conference on Computer Vision, Singapore, Singapore, November 1-5, 2014, Revised Selected Papers, Part I 12 (pp. 290-306). Springer International Publishing.
13. Chen, J., Liu, H., Zhang, Y., Zhang, D., Ouyang, H., & Chen, X. (2022). A Multiscale Lightweight and Efficient Model Based on YOLOv7: Applied to Citrus Orchard. Plants, 11(23), 3260.