|  |
| --- |
| **Title:** Artificial Intelligence on Aerial Imagery (AI2) for individual tree inventory in mixed boreal forests: is it better to work on the individual drone photos or the orthomosaic? |
| **Introduction/Aim:** Drones and artificial intelligence (AI) are increasingly used for enhanced forest inventories that provide information on every tree in a stand required for precision forestry. Here we demonstrate and compare two possible workflows that start with either individual drone photos or an orthomosaic derived from them, and end in a geospatial census of trees, namely a GIS layer consisting of points representing individual trees with attributes such as species, height, crown diameter and others such stem volume that could be derived from published allometric equations. We also introduce a new, open projection code for transferring annotations to a GIS, and discuss pros and cons of each workflow and the factors affecting accuracy.  **Methods:** We initially used a YOLOv8 (a popular AI model) trained, validated and tested (60/20/20 % split) on ~10k bounding boxes corresponding to tree crowns of five classes (pine, spruce, poplar, larch and dead trees) from ~150 manually annotated drone true color (RGB) images of ~1cm GSD from an 8 ha mixedwood stand in the boreal plains of Alberta, Canada. The trained YOLOv8 was applied to images acquired by the same system (a DJI M300RTK with a Zenmuse P1 camera) over the same stand in a different date. The detected boxes were projected onto an orthomosaic derived from those same photos based on the position, aim, internal parameters (focal length and radial and tangential distortion) of the camera, plus its distance to the treetop of the corresponding box (obtained from ray tracing). Duplicate detections of the same tree coming from different photos were removed using an ad hoc non-maxima suppression (NMS). The result was manually revised as to create a new reference dataset with ~ 16k trees that was used to train another YOLOv8 model, this time based on the orthomosaic rather than the original drone photos. Two versions of the orthomosaic-based model were tested, the conventional RGB, and another where the blue channel was replaced by a Canopy Height Model (CHM) derived from the photogrammetric point cloud.  **Results:** The model trained on the original drone images achieved precision, recall and mean average precision at an IoU threshold of 50% (MAP50) of 0.60, 0.67 and 0.65, respectively. The same accuracy metrics for the model trained on GIS-projected, refined boxes were 0.63, 0.64, and 0.64, respectively. At the class level, the F1 score ranged from 0.84 for pine, the most common tree in the stand, to 0.44 for larch, the least common tree which is easily confused with spruce. The modified orthomosaic that included a CHM did not lead to better results.    **Conclusion:** Both workflows (applying the AI to the original drone images vs. the orthomosaic) seem to lead to similar results. However, the analysis of reference ground data, which will be completed before the summer, may provide additional insights. Working on the original drone images has the advantage that they are free from artifacts common in orthomosaics derived from Structure from Motion (SfM). However, it requires precisely aligned images and camera calibration parameters. |