Seedling Detection on Seismic Lines using Convolutional Neural Networks: Case Study of Kirby, Alberta

Terentieva I[1], Shellian C[2], Linke J[3], McDermid G[4]

[1]Department of Geography, University of Calgary, irina.terenteva@ucalgary.ca
[2]Department of Geography, University of Calgary, colette.shellian@ucalgary.ca
[3]Department of Geography, University of Calgary, julia.linke@ucalgary.ca
[4]Department of Geography, University of Calgary, mcdermid@ucalgary.ca

ABSTRACT:

Assessing tree-regeneration status on industrial disturbances is a key element of forest-restoration planning and assessment. Current operational approaches use a manual interpretation of stereo imagery in a softcopy environment, but this is a labor-intensive process that is difficult to scale. To address the problem, we applied a convolutional neural network to automatically detect coniferous seedlings in a study area near Lac La Biche, Alberta (Canada). Using drone and ground-truth data, we were able to identify 82% of coniferous seedlings during the leaf-off season. Different learning rates, tile sizes, training sample sizes, and distributions were tested to reveal the best combinations of model parameters for further studies.

1. Introduction:

Alberta contains 1.8 million kilometers of seismic lines – petroleum-exploration corridors – that segment the forests and influence a host of ecosystem dynamics (Dabros et al., 2022). The problem is particularly concerning sites preventing seedling survival and growth. Unfortunately, many of these lines are in a state of *arrested succession* and are no longer making progress toward recovery. The problem is particularly common in peatlands, where the equipment used to cut seismic lines has compressed the soil and produced conditions unfavorable to 'normal' patterns of recovery (Dabros et al., 2022).

Restoring seismic lines to pre-disturbance conditions is of high interest to researchers and managers. A conventional silvicultural approach – mounding – provides drier microsite conditions for tree planting, but is invasive and expensive. To prioritize lines in need of direct intervention and to assess the following restoration success, we require effective ways of identifying seedlings and characterizing their attributes.

Drone data provides an easy and affordable way to accurately assess vegetation on a line-by-line basis. In Alberta, this is conventionally done using manual softcopy (3D) interpretation of imagery, which leads us to the important question – can we do that faster and more efficiently using machine learning?

Following the rapid development of convolutional neural networks (CNN) to numerous architectures and applications in 2022, it seems obvious that a synergy between drones and CNN provides promising future perspectives for vegetation remote sensing (Kattenborn et al., 2021). Castilla et al. (2020) advocate that a trained artificial intelligence (AI) could detect the location of the seedlings, and other modules could make automated calls on species, health status, and

 Volume DOI: 10.17605/OSF.IO/N6DU5

height. In the long run, CNN may provide end-to-end-learning approaches with no need in sophisticated preprocessing.

To achieve the best CNN model performance, we require a proper choice of input data and hyperparameters. In this study, we experimented with fine-tuning of RetinaNet – a conventional architecture for object detection – to detect coniferous seedlings on seismic lines near Lac La Biche, Alberta (Canada). We experimented with different learning rates, tile sizes, training sample sizes, and distributions to reveal the best combination of model parameters for further studies.

2. Methods

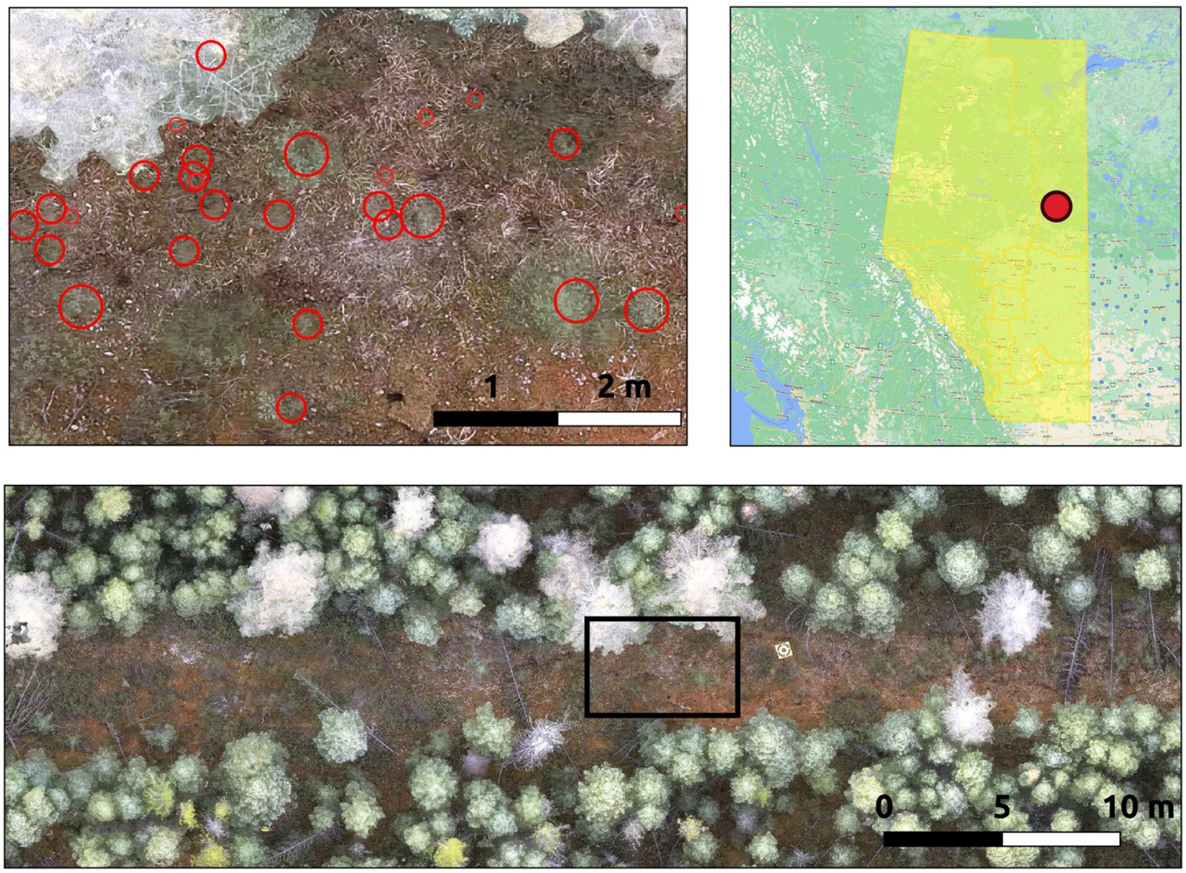

Figure 1: Input data for the seedling detection using CNN: drone orthomosaic with field-measured locations and size of seedlings (a, red circles) within the Seismic Line №464 (b) in a study area near Lac La Biche, Alberta (c)

 Volume DOI: 10.17605/OSF.IO/N6DU5

2.1 Study Area

2.1.1 Study Site

The study area is located in a boreal zone of Alberta, Canada (Fig. 1) with prevailing transitional and lowland ecosystems. The study was conducted in the east-west oriented seismic line №464. The line is 4-6 meters wide through transitional forest characterized by mesic moisture regime, 30% shrub cover, and dominance of bryophytes and herbs in dominant understory vegetation.

2.1.2 Reference Data

Reference data were acquired in the summer season of 2021. Diameters and heights of a total of 418 individual regenerating seedlings were measured across two subsites (located 70 meters within the line). Each subsite covered 100 m^2 and comprised ten equal-area subplots. The precise location of seedlings and the subplot vertices was recorded with a Hemisphere s631 real-time kinematic Global Navigation Satellite System (RTK GNSS).

Among seedlings, only 17 young trees (> 1 m height) were deciduous, the rest of the seedlings were coniferous with an abundance of germinants (37%) and small seedlings from 30 to 60 cm height (44%). Fewer seedlings were from 60 cm to 1 meter (9%) or higher than 1 meter (5% of seedlings). The median seedlings' height was 25 cm for the first and 47 cm for the second subsite.

2.1.3 Drone Data

The drone photographs were acquired on 17^{th} May 2021 with a DJI Matrice 300 with a P1 (RGB) sensor. The photographs were captured at a mean altitude of 40 m. The camera focal length was 35 mm, yielding mean ground-sampling distances (GSD) of 0.48–0.60 cm across the sites. All the flights took place under diffuse light conditions (designed to minimize shadows) and with minimal winds.

2.2 RetinaNet

To detect seedlings, we used the DeepForest python package developed for delineating and predicting individual tree crowns using RGB drone imagery (Weinstein et al., 2019). The package comes with a prebuilt model trained on data from the National Ecological Observatory Network. The model is based on RetinaNet – a powerful one-stage detector using a resnet-50 classification backbone and focal loss – to address the dataset imbalance during training.

To build a training dataset, the visual interpretation of RGB drone imagery was used. Visible seedlings were hand-annotated by drawing bounding boxes in QuantumGIS. Then, the dataset was randomly split into training (80%) and validation (20%) data. Orthomosaics were tiled with 20% overlap to capture all seedlings; since the model is sensitive to tile size, we experimented with window sizes using tiles of 500, 1000, and 1700 pixels.

The RetinaNet model (with a batch size of 8) was trained for up to 40 epochs; to reduce overfitting, early stopping was used by monitoring a validation loss. Predicted bounding boxes were filtered when they strongly overlapped (> 90%) or had a low prediction score (less than 0.3). Also, we removed bounding boxes from the tile edge of the final predictions to overcome the edge effect.

To assess model accuracy, we calculated precision and recall using field-collected data of seedling locations, diameters, and heights. We used a 20% intersection-over-union threshold to address the natural discrepancy between stem and crown locations (seedlings are not strictly vertical; Fig. 2). Since target tree species must be clearly visible in the image (Kattenborn et al., 2021), we manually filtered the ground-truth dataset to remove location errors and obscured seedlings. Field-based measurements of seedling heights from the reference dataset were used to assess the height-specific accuracy of individual models within the test set.

2.3 Experiments

To simulate how the sample distributions affect the model accuracy and convergence, we trained 5 models for 30 epochs with a learning rate of 0.005 and tiles of 1000-px size using random splits to train (80%) and validation sets (20%).

To test how a tile size affects model accuracy and training time, we experimented with tiles of 500, 1000, and 1700-pixel sizes. For each tile size, we trained 3 models with the same hyperparameters (30 epochs with a learning rate of 0.005) using different random splits to train (80%) and validation sets (20%). The best models were applied to the field-based dataset to reveal height-specific accuracies.

The training set was decreased by 15, 30, 50, and 70% for tiles of 500 and 1000 pixels to reveal how training set size affects the model performance. To test model robustness when unlabeled seedlings are present in the training set, also we applied two methods to reduce the training set: i) random seedlings were deleted from random tiles (leaving unlabeled seedlings), ii) deleting all seedlings from randomly chosen tiles.

For revealing the best learning rate, we experimented with 0.005, 0.001, 0.0005, and 0.0001 learning rates using different sample distributions and the number of training data. To assess the accuracy of the final model, we trained it for 15 epochs using a full training set, a learning rate of 0.005, and a tile size of 500 pixels.

3. Results

To assess our ability to detect coniferous seedlings during the leaf-off season, we trained several RetinaNet models using different learning rates, tile sizes, training sample sizes, and distributions. In most cases, models converged during the first 10-20 epochs; the validation recall increased rapidly during the first epochs and then stabilized.

Model performance was assessed using ground-truth data with precise locations of seedlings and field-measured heights and crown diameters. Overall, we were able to correctly recall 82% of visible coniferous seedlings. Among them, the lowest seedlings (15-30 cm) were the most challenging class with a recall reaching 58%. Best models could reliably detect up to 90% of coniferous seedlings of more than 60 cm in height. The overall precision was high for most algorithms, reaching 90% on average.

To simulate how the sample distributions affect the performance of individual models, we trained five models with the same set of hyperparameters but random training/validation splits of the same size. It affected individual models during the first training epochs, but then the averaged performance rapidly stabilized with a standard deviation of 4% for validation recall within models.

Learning rate affected model performance and speed. High learning rates (0.005 and 0.001) performed similarly well in most cases. Training with them resulted in a slightly higher recall of 81±2% for the last 10 epochs compared to 76±5% for lower learning rates (0.0001).

RetinaNet is sensitive to tile size; but unlike the previous research in tree delineation (Weinstein et al., 2019), we found that smaller patch size worked better for seedling detection. The overall recall increased from 79% to 92% due to decreasing the tile size from 1700 to 1000 pixels. A further decrease to 500 pixels resulted in a slightly higher recall and precision (from 76% to 81%). Using a field-based dataset, we revealed that smaller tile size especially boosted the detection accuracy of the smallest seedlings: after decreasing tile size from 1000 to 500 pixels, recall grew

 Volume DOI: 10.17605/OSF.IO/N6DU5

from 23% to 55% for the lowest coniferous seedlings (15-30 cm), and from 67% to 88% - for seedlings of 30-45 cm.

Decreasing the training set size affected the validation accuracy not as we expected. Even a 50% decrease in the training sample did not worsen the performance of most models. We were able to train a model reaching 64% recall in detecting the smallest seedlings (15-30 cm) using only 30% of the initial training data (Fig. 3). Two approaches for decreasing the dataset – removing random tiles with all seedlings and random seedlings from tiles (leaving them unlabeled) – also resulted in the similarly good accuracy. The presence of unlabeled seedlings didn't slow down the model convergence or make it less stable.

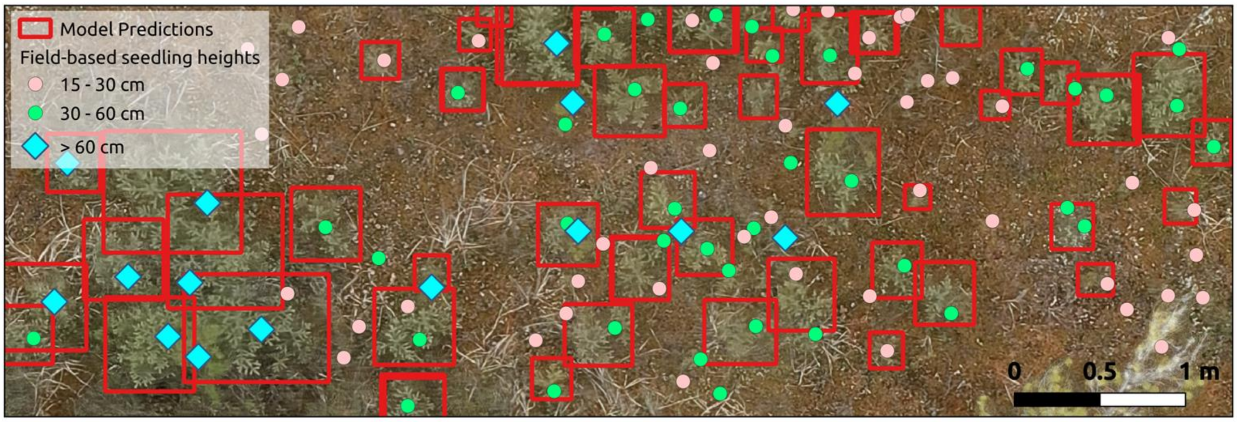

Fig. 2. Accuracy assessment of conifer detection within the test area: bounding boxes of CNN model predictions (red rectangular) and the reference dataset of coniferous seedling locations and heights

4. Discussion & Conclusion:

Being under intense oil exploration for many decades, Alberta will benefit from novel approaches to restoration assessment of its industrial disturbances. In this study, we reached state-of-art accuracies in the detection of coniferous seedlings within a seismic line in a case study near Lac La Biche, Alberta. To reveal how tile size, learning rate, and training set distributions affect the model performance, several DeepForest models were trained using hand-annotated RGB drone data. We assessed the height-specific accuracies of the best models using field-measured data on seedling heights and locations. Further large-scale studies are needed to develop a workflow for operational assessment of the restoration success on seismic lines.

Acknowledgments:

This research is part of the Boreal Ecosystem Recovery and Assessment (BERA) project (www.bera-project.org), and was supported by a Natural Sciences and Engineering Research Council of Canada Alliance Grant (ALLRP 548285 - 19) in partnership with Alberta-Pacific Forest Industries Inc., Canadian Natural Resources Ltd., Cenovus Energy, ConocoPhillips Canada Resources Corp., Imperial Oil Resources Ltd., Canadian Forest Service's Northern Forestry Centre, and the Alberta Biodiversity Monitoring Institute.

References:

Dabros, A., Higgins, K. L., & Pinzon, J. (2022). Seismic line edge effects on plants, lichens and their environmental conditions in boreal peatlands of Northwest Alberta (Canada). Restoration Ecology, 30(4), e13468.

Castilla, G., Filiatrault, M., McDermid, G. J., & Gartrell, M. (2020). Estimating individual conifer seedling height using drone-based image point clouds. Forests, 11(9), 924.

Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 173, 24-49.

Weinstein, B. G., Marconi, S., Bohlman, S., Zare, A., & White, E. (2019). Individual tree-crown detection in RGB imagery using semi-supervised deep learning neural networks. Remote Sensing, 11(11), 1309.