1   Geographic Generalization in Airborne RGB Deep Learning Tree Detection

2   Ben. G. Weinstein[1], Sergio Marconi[1], Stephanie A. Bohlman[2], Alina Zare[3], Ethan P.

3   White[1]

4   [1] Department of Wildlife Ecology and Conservation, University of Florida, Gainesville,

5   Florida, USA

6   [2] School of Forest Resources and Conservation, University of Florida, Gainesville,

7   Florida, USA.

8   [3] Department of Electrical and Computer Engineering, University of Florida, Gainesville,

9   Florida, USA

10   Keywords: Tree Crown Detection, RGB Deep learning; object detection; Airborne LiDAR

11   Abstract

12   Tree detection is a fundamental task in remote sensing for forestry and ecosystem

13   ecology applications. While many individual tree segmentation algorithms have been

14   proposed, the development and testing of these algorithms is typically site specific, with

15   few methods evaluated against data from multiple forest types simultaneously. This

16   makes it difficult to determine the generalization of proposed approaches, and limits tree

17   detection at broad scales. Using data from the National Ecological Observatory Network

18   we extend a recently developed semi-supervised deep learning algorithm to include

19   data from a range of forest types, determine whether information from one forest can be

20   used for tree detection in other forests, and explore the potential for building a universal

21   tree detection algorithm. We find that the deep learning approach works well for

22   overstory tree detection across forest conditions, outperforming conventional LIDAR-

23   only methods in all forest types. Performance was best in open oak woodlands and

24   worst in alpine forests. When models were fit to one forest type and used to predict

25   another, performance generally decreased, with better performance when forests were

26   more similar in structure. However, when models were pretrained on data from other

27   sites and then fine-tuned using a small amount of hand-labeled data from the evaluation

28   site, they performed similarly to local site models. Most importantly, a universal model fit

29   to data from all sites simultaneously performed as well or better than individual models

30   trained for each local site. This result suggests that RGB tree detection models that can

31   be applied to a wide array of forest types at broad scales should be possible.

32   1. Introduction

33   Tree detection is a critical step in remote sensing of forested landscapes. Identifying

34   individual crowns in airborne imagery allows ecologists, foresters, and land managers to

35   increase the extent of sampling compared to terrestrial surveys. While many LIDAR-

36   based tree segmentation algorithms have been proposed (Aubry-Kientz et al., 2019),

37   the field has been slow to adopt automated methods due to concerns over accuracy,

38   transferability and transparency (Vaglio Laurin et al., 2019). As a result, existing

39   methods are rarely evaluated on multiple forests simultaneously, making it unclear how

40   they will perform in the novel contexts required for large scale application. The

41   availability of LIDAR data can also be limiting for large scale application. In contrast,

42   RGB imagery is more widely available, but relatively few RGB algorithms have been

43   proposed (González-Jaramillo et al., 2019) due, in part, to challenges with closed

44   canopies and the diverse appearance of trees across forest types.

45   Current tree segmentation approaches are primarily based on user-defined

46   algorithms that describe the appearance of trees in a hierarchical sequence of rules.

47   These rule-based approaches rely on combinations of shape features (Gomes et al.,

48   2018), template matching (Dai et al., 2018), network analysis (Williams et al., 2019),

49   and watershed routines (Silva et al., 2016) that are applied to either LIDAR point clouds

50   or RGB photogrammetric imagery (Brieger et al., 2019). By describing the parameters

51   that define an individual tree, unsupervised algorithms attempt to match these rules

52   when predicting unlabeled data. This can make applying these algorithms across

53   different forest types challenging because the rules describing a tree vary depending on

54   the type of forest, leading to overfitting for a particular geographic area. For example,

55   some methods use allometric relationships between crown area and tree height to

56   improve algorithm performance (Coomes et al., 2017; Williams et al., 2019), but these

57   relationships vary with forest type and species. Recent attempts to mitigate this

58   variation have used approaches that choose from a pool of potential tree shapes

59   (Gomes et al., 2018). However, the need to define the full pool of possible tree shapes

60   before analyzing each new site will be prohibitive over large geographic areas that

61   incorporate diverse assemblages. As a result of these limitations, most tree detection

62   algorithms have been applied and tested on similar forest types with little exploration of

63   how the algorithms generalize to other natural settings. Therefore, despite the intense

64   work in airborne tree detection over the last decade (Coomes et al., 2017; Heinzel and

65   Huber, 2018; Jakubowski et al., 2013; Li et al., 2012; Williams et al., 2019), there

66   remains no clear consensus on best practices (Aubry-Kientz et al., 2019).

67   Within the field of computer vision, there has been a broad shift away from rule-

68 based approaches (i.e., user-designed features) towards approaches that learn features

69 from data using deep learning neural networks (Agarwal et al., 2018). There have been

70 few attempts to use learned features in tree detection (Dai et al., 2018) due to the need

71 for a large amounts of labeled training data, which is often difficult or impossible to

72 collect in ecological contexts. Overall, generalization of deep learning algorithms across

73 applications in airborne remote sensing remains a challenging task (Zhu et al., 2017). A

74 typical neural network has millions of parameters and is therefore at risk of overfitting

75 when using small datasets. Given the diversity of trees, finding general features will

76 require a combination of large training datasets and algorithmic approaches that allow

77 the neural networks to learn the combination of features that characterize trees across

78 forest types.

79   Weinstein (et al. 2019) recently developed a deep learning approach for tree

80 detection using RGB data that has the potential to address these requirements for

81 identifying trees across forest types. The semi-supervised method first uses

82 unsupervised LIDAR-based tree detection (e.g., Silva et al. 2016) to generate millions of

83 labeled trees that are used to pretraining of the neural network. This pretraining stage is

84 followed by retraining the network based on a small number of high-quality hand-

85 annotations. This addresses the need for large training data by generating millions of

86 annotations of moderate quality for model pretraining and the method has been shown

87 to perform well on a single oak woodland site. Due to its deep learning architecture, this

88 method has the potential to learn general features of trees across forest types, but this

89 remains an untested possibility.

90   Here we build on Weinstein et al. (2019) explore the potential of this tree

91   detection method to generalize across sites by evaluating its performance on a range of

92   forest types, assessing the transferability of tree features across forest types, and

93   exploring the possibility of building a single unified tree detection model. Our aim is to

94   test a deep learning approach 1) for identifying trees in four different forest types when

95   trained on that forest type ('within-site'); 2) for identifying trees when trained on data

96   from other forest types ('cross-site'); 3) for combining pretraining data from other sites

97   with hand-annotated data from a new site ('transfer learning'); and 4) for comparing the

98   performance of a within-site model to a universal model fit to data on all forest types

99   simultaneously ('universal'). We also explore the sensitivity of the self-supervised

100   method to the number of hand annotations, to determine the amount of time-intensive

101   work needed to produce accurate results. By answering these questions, we will

102   improve our understanding of the potential for universal tree detection methods and

103   potentially advance RGB-based tree detection from algorithm development to large

104   scale application for better understanding forests at scale.
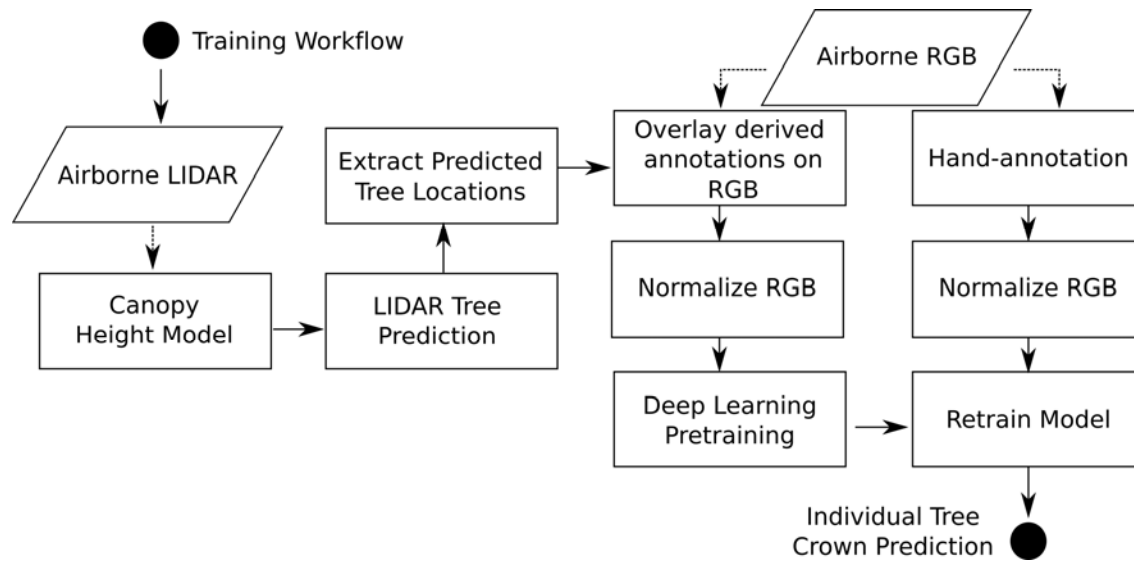
105   METHODS

106

107 Figure 1. Conceptual workflow of proposed approach for airborne detection of individual

108 tree crowns. Pretraining data is generated by overlaying predicted trees from a LIDAR-

109 based unsupervised algorithm on to RGB imagery. These RGB images are used to

110 pretrain a deep learning neural network. The resulting model is retrained based on RGB

111 hand-annotations.

112 DATA COLLECTION AND SITE DESCRIPTIONS

113 The aerial remote sensing data products were provided by the National Ecology

114 Observation Network (NEON) Airborne Observation Platform. We used the NEON 2018

115 "classified LiDAR point cloud" data product (NEON ID: DP1.30003.001) and the

116 "orthorectified camera mosaic" (NEON ID: DP1.30010.001). The LiDAR data consist of

117 3D spatial point coordinates with an average of 4-6 points/m^2. These data provide high

118 resolution information about crown shape and height. The RGB data are a 1km x 1km

119 mosaic of individual images with a cell size of 0.1 meters. All data are publicly available

120 on the NEON Data Portal (http://data.neonscience.org/). For hand-annotations, we

121 selected two 1km x 1km RGB tiles and used the program RectLabel

122   (https://rectlabel.com/) to draw bounding boxes around each visible tree. For a count of

123   tree annotations per site, see Table 1. All code for this project is available on GitHub

124   (https://github.com/weecology/DeepLiDAR) and archived on Zenodo, and all

125   annotations are available as part of the forthcoming NEON Tree Benchmark

126   (https://github.com/weecology/NeonTreeEvaluation).

127   We selected four sites from the NEON network to capture a range of canopy

128   complexity and forest types. The 'Oak Woodland' is the San Joaquin Experimental

129   Range, California. The site contains live oak (*Quercus agrifolia*), blue oak (*Quercus*

130   *douglasii*) and foothill pine (*Pinus sabiniana*) forest. The majority of the site has a single-

131   story canopy with mixed understory of herbaceous vegetation. The "Mixed Pine" site is

132   Teakettle Canyon, California (37.00583, -119.00602) which contains red fir (*Abies*

133   *magnifica*) and white fir (*Abies concolor*), jeffrey pine (*Pinus jeffreyi*) and Lodgepole

134   Pine (*Pinus contorta*). The "Alpine" site is Niwot Ridge Mountain Research Station,

135   Colorado (40.05425, -105.58237). This high elevation site (3000m) is near treeline with

136   clusters of subalpine fir (*Abies lasciocarpa*) and englemann spruce (*Picea engelmanii*).

137   Finally, the "Eastern Deciduous" site is the Mountain Lake Biological Station, Virginia

138   (37.37828, -80.52484). Here the dense canopy is dominated by red maple (*Acer*

139   *rubrum*) and white oak (*Quercus alba*). Each site presents its own challenges, with

140   broad flat-topped trees in the Oak Woodland, tight clusters of trees in the Mixed Pine

141   forest, thin conifers in the Alpine forest, and completely connected crowns in the

142   Eastern Deciduous forest.

143   For each site, we manually annotated training tiles using the program RectLabel

144   (Table 1). Training tiles were selected at random from the NEON data portal. At higher

145   tree density sites, we cropped the 1km^2 tiles to create more tractable sizes for hand-

146   annotation. To enforce a minimum size threshold for tree annotations, we compared the

147   hand-annotations to a LiDAR canopy height model and removed any trees less than 3m

148   in height. The resulting annotations were compared to the LiDAR point cloud for further

149   assessment. No attempt was made to delineate understory trees that were not visible in

150   the RGB imagery.

151   For model evaluation, we used the NEON "tower" plots, which are a set of

152   40x40m plots placed throughout each site. For the Eastern Deciduous site, it was

153   difficult to determine tree boundaries in both the RGB and LiDAR images. For this site,

154   we overlaid a 1m resolution three-band hyperspectral composite image to highlight

155   differences among co-occurring tree species in the area. The composite image came

156   from NEON's orthorectified surface reflectance (ID: DP1.30006.001) and contained

157   bands in the infrared (940nm), red (650nm), and blue (430nm) spectrum. This allowed

158   us to more accurately annotate the training and evaluation data in the closed canopy

159   conditions.

160   LiDAR TREE DETECTION

161   We tested three existing unsupervised LiDAR algorithms (Dalponte and Coomes, 2016;

162   Li et al., 2012; Silva et al., 2016), as implemented in the lidR R package (Roussel and

163   David Auty, 2019), as both a comparison to the semi-supervised approach, and as

164   potential algorithms to generate tree labels for the self-supervised portion of the

165   workflow. We selected the best performing method (Silva et al., 2016) to create initial

166   tree predictions in the LiDAR point cloud. This approach uses a canopy height model

167   and an allometry of tree height to crown width to cluster the LiDAR cloud into individual

168   trees. We used a canopy height model of 0.5m horizontal resolution to generate local

169   treetops and an allometry of 90% of crown diameter to height for deciduous forests

170   (Oak Woodland and Eastern Deciduous) and 20% of crown diameter to height for the

171   coniferous forests (Mixed Pine and Alpine). LiDAR algorithms perform segmentation on

172   a per-point basis, so we converted the output to a bounding box that covered the entire

173   set of LiDAR points assigned to each tree to create training data equivalent to the hand

174   annotated bounding boxes.

175 Table 1. The number of tree annotations used for semi-supervised pretraining,
176 retraining and evaluation. Pretraining annotations are generated automatically using a
177 LiDAR-based unsupervised algorithm. Training and evaluation annotations were hand-
178 drawn.

Forest Type	Pretraining Annotations	Training Annotations	Evaluation Annotations
Oak Woodland	550,905	2,533	293
Mixed Pine	2,522,855	3,405	747
Alpine	3,121,036	9,730	1,699
Eastern Deciduous	3,131,283	1,231	489

179 SEMI-SUPERVISED DEEP LEARNING

180 We used our previously developed self-supervised algorithm for RGB-based tree

181 identification (Weinstein et al. 2019). This method uses the Retinanet one-stage object

182 detector (Gaiser et al 2017) with a Resnet-50 classification backbone, which allows

183 pixel information to be shared at multiple scales, from individual pixels to groups of

184 connected objects. We used a Resnet-50 classification backbone pretrained on the

185 ImageNet dataset (He et al., 2016). Since the entire 1km RGB tile cannot fit into GPU

186 memory, we cut each tile into 40m by 40m windows with an overlap of 5% (n=729). The

187 order of tiles and windows were randomized before training to minimize overfitting

188 among epochs. To reduce potential spatial autocorrelation in tree appearance between

189 evaluation plots and pretraining data, we removed any training tiles within 1km of an

190 evaluation tile. Using the pool of unsupervised LiDAR-based tree predictions, we

191 pretrained the network with a batch size of 20 on 2 Tesla K80 GPU for 5 epochs. To

192 align these unsupervised classifications with the ImageNet pretraining weights, we

193 normalized the RGB channels by subtracted the ImageNet mean from each channel.

194 We then retrained the network using the hand-annotated data for 40 epochs. For more

195 details of this semi-supervised approach see Weinstein et al. (2019).

196 MODEL EVALUATION

197 Using the evaluation plots, we chose two metrics to assess model performance. For

198 comparison with the existing LiDAR-only implementations, we used precision and recall

199 statistics with a bounding box marked as true positive if it had an intersection-over-union

200 (IoU) of greater than 0.5. Intersection-over-union is the ratio of the area of bounding box

201 overlap to the area of bounding box union between the predicted tree crown and the

202 visually annotated crowns in the evaluation data. For each bounding box prediction, the

203 deep learning model reports a confidence score between 0 and 1. To transform these

204 scores into precision and recall statistics, we need to define a threshold of box scores to

205 accept. As we lower the threshold for acceptance, a greater number of trees will be

206 captured, but at the expense of decreased precision. To highlight this relationship, we

207 showed the performance of the deep learning approach across all bounding box

208 probability thresholds between 0 and 1 with an interval of 0.1.

209 While IoU precision and recall are intuitive statistics, they are reported separately

210 and do not capture differences in bounding box confidence scores. When comparing the

211 different generalization approaches, it is useful to have a single metric to compare. We

212 used the Average Precision (AP) metric commonly used for object detection tasks in

213 computer vision, which is the area under the precision-recall curve computed at the 11

214 fixed 0.1 intervals between 0 and 1 (Lin et al., 2017).

215 ASSESSING GENERALIZATION, TRANSFERABILITY, AND UNIVERSAL MODEL FIT
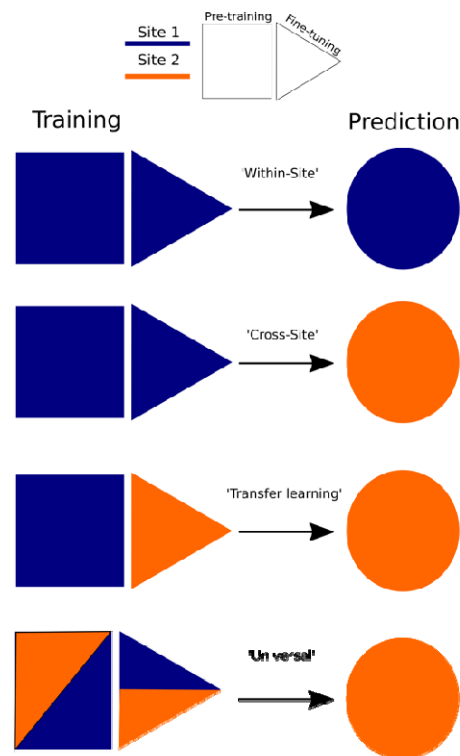
216

217 Figure 2. Approaches to geographic generalization in semi-supervised model training:

218 1) 'Within-site' training in which training data from site 1 is used to predict site 1; 2)

219 'Cross-site' training in which training data from site 1 is used to predict site 2; 2)

220 'Transfer learning' in which a model is first trained on site 1 data, followed by finetuning

221 on site 2 training data, and 3) 'Universal' model in which training data from both site 1

222 and site 2 are used to predict evaluation data from site 2.

223   To assess generalization among sites, we performed three types of experiments that

224   used different combinations for hand-annotations and pretraining data (Figure 2). The

225   first experiment is to use pretraining and hand-annotated data to predict the evaluation

226   data from the same site ('within-site'). The next setup is to use the pretraining data and

227   hand-annotated from the same site to predict the evaluation data from a different site

228   ('cross-site'). For example, using each of the within-site models, we can test the ability

229   for a model to predict tree conditions in each of the other geographic sites, creating a

230   matrix of cross-site predictions. To assess generalization without local pretraining data,

231   we tested a model training using pretraining data from all other sites, but hand

232   annotations from the same site as the evaluation data ('transfer-learning'). For example,

233   the transfer learning model for Oak Woodland used the hand-annotations from Oak

234   Woodland, but the pretraining data for Alpine, Mixed Pine, and Eastern Deciduous.

235   Finally, to test the potential for a universal model, we tested a model pretrained on all

236   sites, followed by retraining on all hand-annotations. We then compared this model with

237   each of the within-site model to test whether the addition of data from other sites

238   improved predictions of trees from the same site.

239   SENSITIVITY TO THE NUMBER OF HAND-ANNOTATIONS

240   Collecting a sufficient number of training samples will often be a bottleneck in

241   developing supervised methods in airborne imagery. It is therefore useful to test the

242   number of local training samples needed to achieve peak performance. We performed a

243   sensitivity study by training models using different proportions of training data. We

244   selected 5%, 25%, 50% and 75% of the total hand-annotations to compare to the full

245   dataset for the within-site results for each site. We reran this experiment five times to

246   account for the random subsampling of annotations. In addition, we ran the evaluation

247   plots for the pretraining model only (i.e. 0% hand-annotated data) to assess whether the

248   addition of hand-annotated data improved the within-site pretraining.

249   RESULTS

250   Based on the highest performing probability cutoff (Figure 4), within-site predictions

251   ranged from 0.60 recall and 0.75 precision in Mixed Pine to 0.34 recall and 0.55

252   precision in Alpine. The Oak Woodland and Mixed Pine sites consistently performed

253   better than the Eastern Deciduous and Alpine sites. Visual inspection of the results

254   showed that the vast majority of false positives were positively identified trees, but

255   whose crown boundaries were either too large or too small for the intersection-over-

256   union score of 0.5. Repeated training runs for each model showed relatively little

257   variance, despite heterogeneity in tree types at all sites (Figure 4).
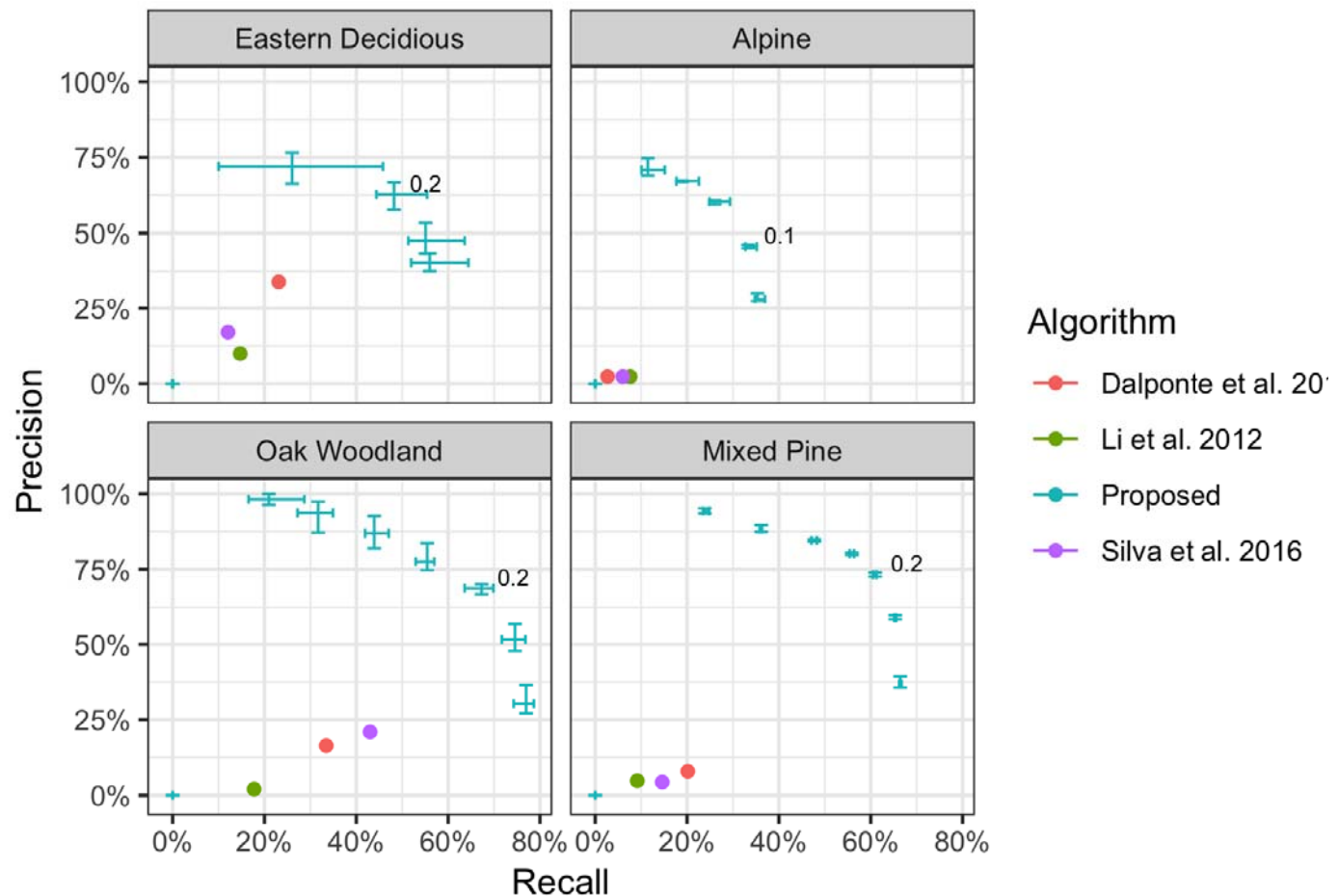
258

Figure 4. For each site, results of our proposed workflow compared to three existing LiDAR-only implementations from the commonly used lidR package. The proposed approach was evaluated at each of the 0.1 probability score intervals between 0-1. The probability threshold of the best performing model in our approach, calculated by f-score, is shown in black. Error bars show the variance in recall and precision based on five runs of hand-annotation training for each site.

265   The semi-supervised deep learning approach significantly outperformed the

266   available LiDAR-only implementations from the lidR R package at all sites (Roussel and

267   David Auty, 2019) (Figure 4). When comparing model performance with the Silva et al.

268   (2016) algorithm used to generate the pretraining data, the deep learning model was

269   more successful at delineating complex crown boundaries and avoiding clumping

270   together small trees with narrow gaps (Figure 5). For the Oak Woodland site, the deep

271   learning model was better able to capture crown area for the flat-topped canopy and

272   avoided erroneously labeling bushes as trees. For the Eastern Deciduous site, the deep

273   learning model more accurately found trees in the closed canopies, despite strong

274   overlap in bounding box predictions and similarity in neighboring tree appearance. The

275   Alpine site was the worst performing model, and many small trees were missed. This is

276   likely due to the minimum anchor box size in the object detector and the arbitrary cutoff

277   at 3m height for defining a 'Tree' class in the Alpine training data. To view predictions

278   overlaid on each of the plots for the within-site models, see supplemental dataset S1.
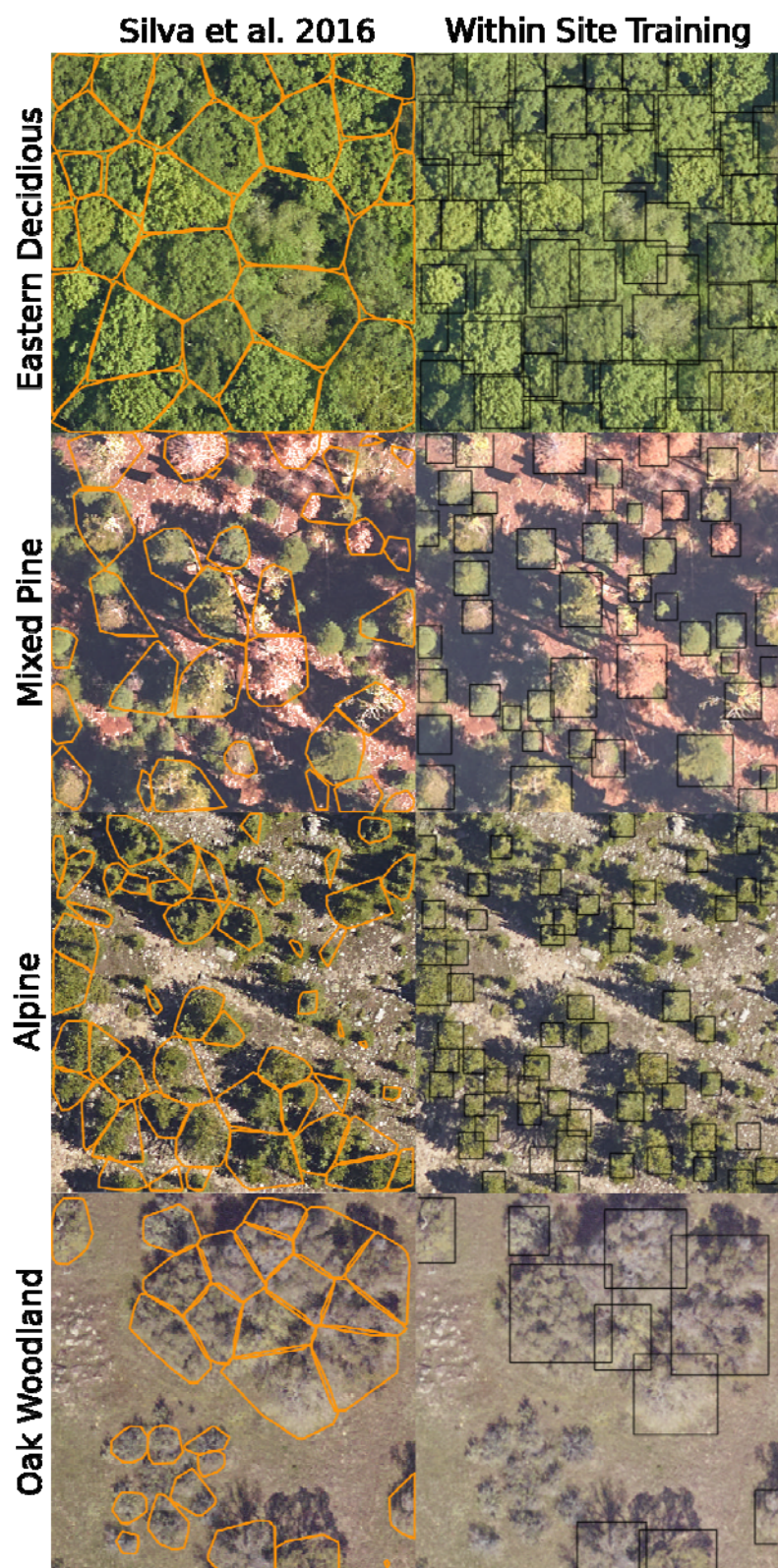
279

280   Figure 5. Example predictions for the LiDAR-only pretraining algorithm, and the deep
281   learning detection network trained within-site.

282   When applying a model fit at one site to make predictions at other sites, we found

283   generalization of the single-site models to be weak (Figure 7 A-E). Tree stems were

284   often correctly identified among sites with similar forest conditions (Coniferous versus

285   Deciduous), but the resulting crown boundaries were rarely accurate (Figure 6 – "Cross-

286   Site"). The one exception was the prediction of Alpine evaluation plots using a model

287   built from the Mixed Pine site. This model outperformed all other cross-site experiments

288   and was superior even to the Alpine within-site model.

289   Combining local hand-annotated data with unsupervised pretraining data from

290   the other three sites demonstrated good transferability, with performance almost as

291   good as using local pretraining data. The transfer learning experiments performed better

292   than cross-site predictions for every site. This suggests that the pretraining model

293   allows for generalized features that can be fine-tuned to local conditions.

294   Fitting a single universal model using data from all sites resulted in the best

295   predictions for every individual site. The Average Precision of the Eastern Deciduous

296   site improved from 0.44 to 0.54 (22%), Mixed Pine from 0.56 to 0.59 (5.4%), Alpine from

297   0.24 to 0.26 (8.3%) and Oak Woodland from 0.6 to 0.61 (1.6%). Visual inspection of

298   predictions for the Mixed Pine site illustrates why the universal model improves

299   performance. In Figure 7B, the within-site model (Mixed Pine) erroneously labels a large

300   boulder in the bottom right hand corner of the image as a tree. This error was made in

301   all other cross-site models, except for Oak Woodland. In the Universal model, this error

302   was not made, suggesting that the universal model learned information about the

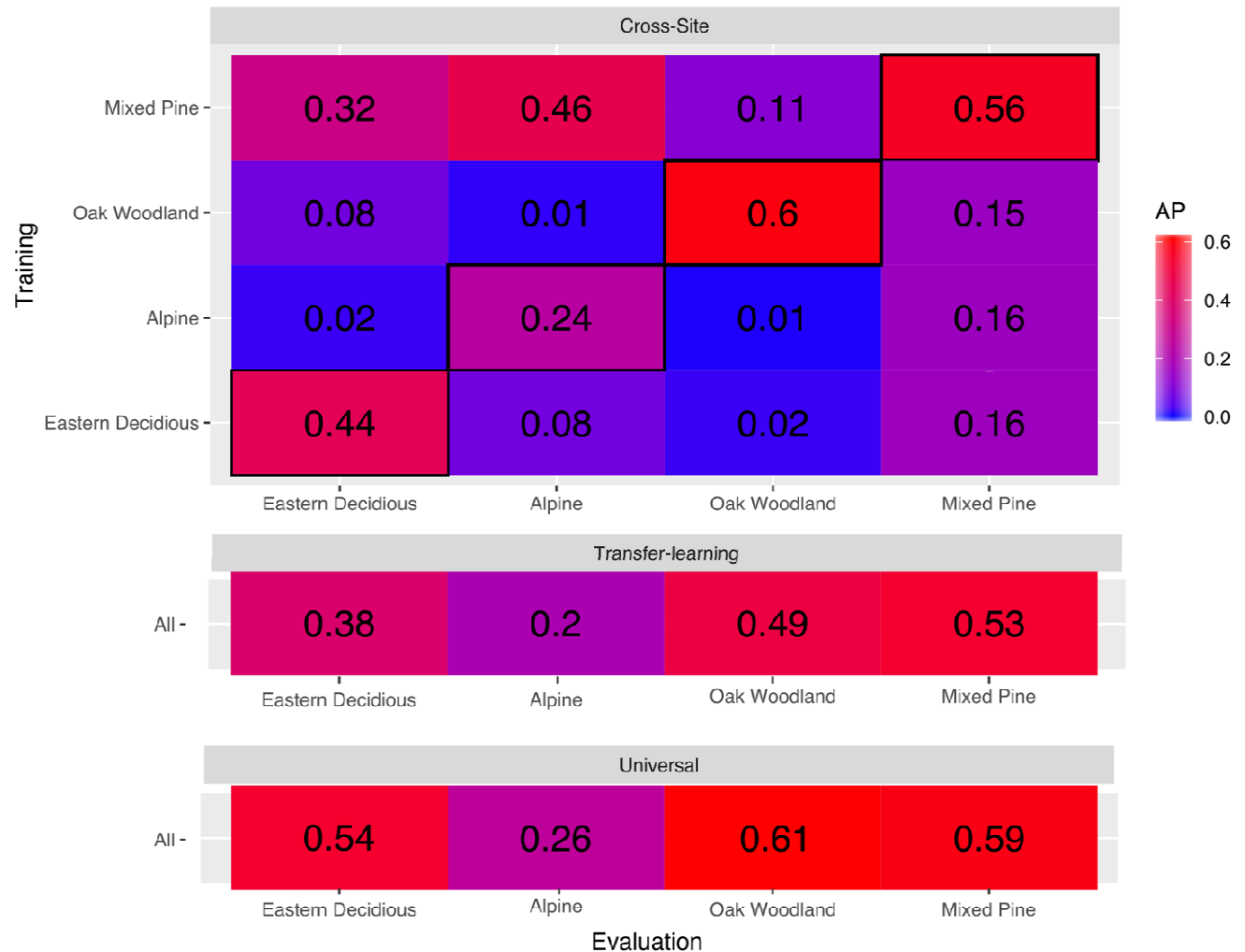303   background from other sites to improve predictions.

304

Figure 6. Comparison single-site, cross-site, transfer, and universal model performance based on Average Precision (AP). Single site predictions are on the bolded diagonal of the cross-site section and represent fitting and predicting on the same site. Cross-site predictions are for models trained on the one site (listed on the left side of the results matrix) and evaluated on a second site (listed across the bottom of the results matrix). Transfer learning takes a model pretrained on all sites except the focal site and retrained using the hand-annotations of the evaluation site. The universal model uses pretraining and hand-annotation data from all sites.
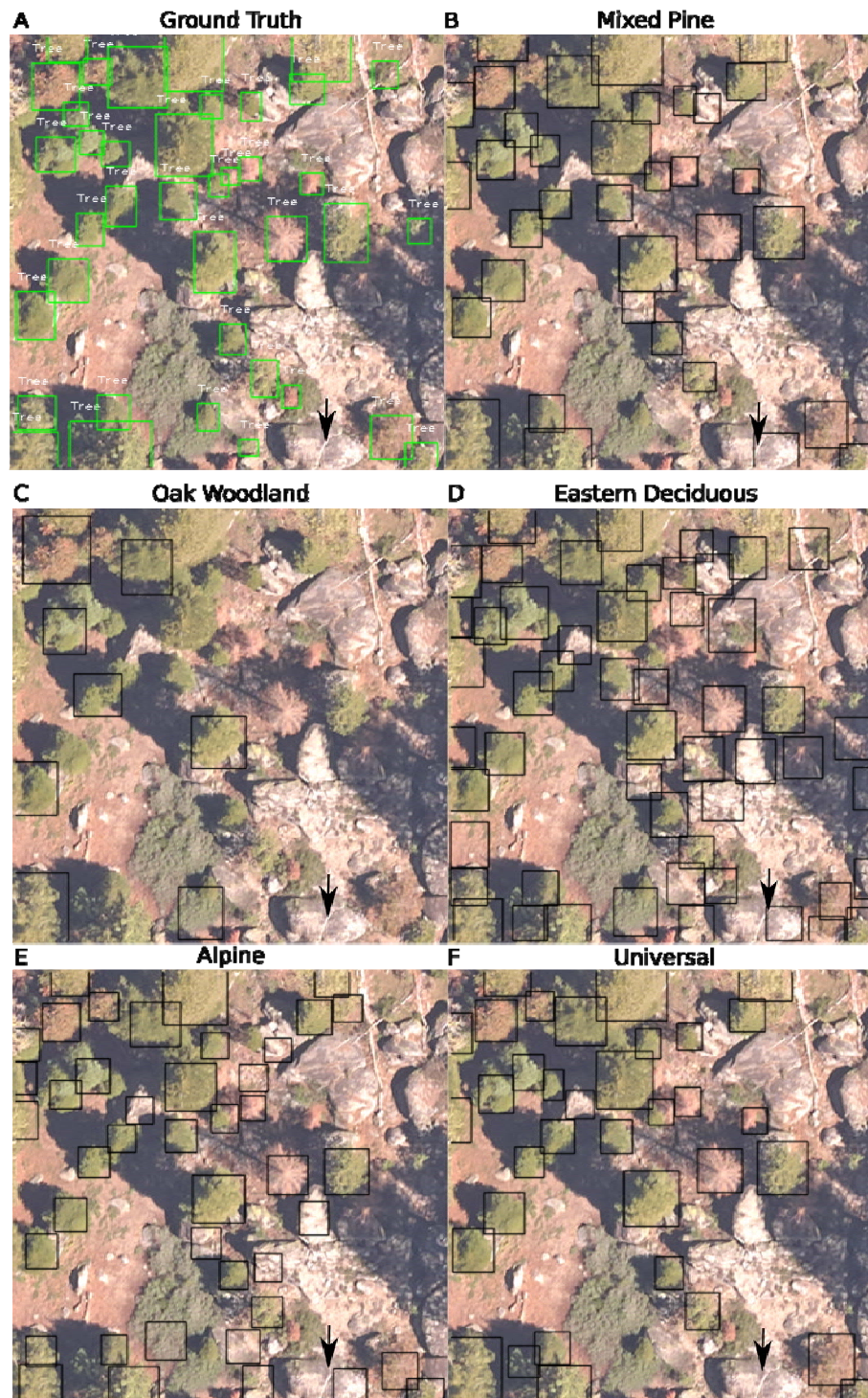
313

Figure 7. A sample evaluation plots from the Mixed Pine site predicted by a model built from training data from the same site, from each other site, and a universal model. Ground truth boxes are shown in green (A). Individual trees with a predicted probability greater than 15% are shown in black (B-E). The universal model (F) built from all annotations slightly outperformed all other models, including the model trained only from

319 the Mixed Pine site. For example, the boulder in the bottom right corner is incorrectly
320 classified as a tree by the models trained from Mixed Pine, Alpine, and Eastern
321 Deciduous sites, but is correctly ignored in the Oak Woodland and Universal models.

322   Assessment of the number of hand-annotations needed to improve model

323 performance, indicated that while some hand-annotated data was important at all sites,

324 the number of hand-annotated trees needed to improve model performance was

325 typically relatively small. For example, the recall in the Mixed Pine site was < 0.2 with no

326 hand-annotated data and was over 0.6 using approximately 2000 hand labeled crowns.

327 Only minimal gains in performance occurred using up to an additional 2000 hand-

328 annotated crowns. Overall, the shape of the ablation curves suggest that the model is

329 fairly robust and needs only approximately 1000 crowns in most cases to create a

330 model close to full performance. The exception is the Alpine model, which improved by

331 more than 30% after 3,000 crowns. In general, the precision was more robust than

332 recall, suggesting that the hand annotations mostly improve the predictions of crown

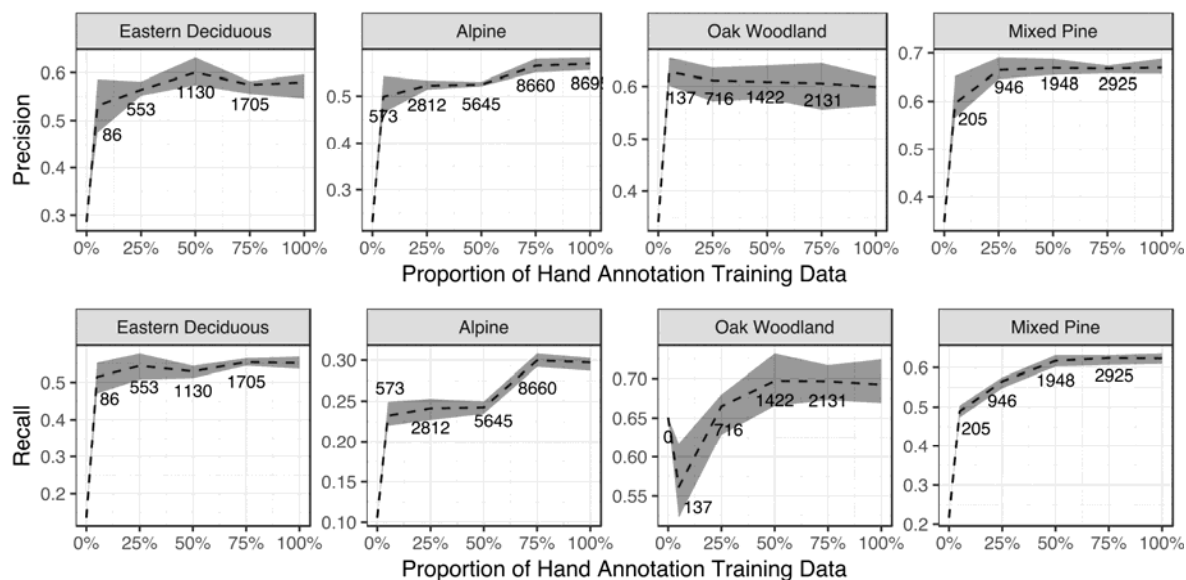333 boundaries rather than additional tree locations.

334

335   Figure 8. Sensitivity curves of the proportion of hand-annotation training data for each
336   site. Values indicate the number of trees in the training dataset for each cutoff. Shaded
337   area is the range of results from rerunning the analysis five times for each site. Note
338   that due to the random sampling among runs, the exact number of trees will vary
339   slightly. For simplicity, we show the mean number of training trees for each threshold.

340   # DISCUSSION

341   Airborne tree detection promises to unlock ecological and forestry data at

342   unprecedented spatial extents compared to traditional ground surveys. To turn remote

343   sensing data into ecological information, there is a need for a unified tree detection

344   model that can be applied to a broad array of forest conditions. Using a semi-supervised

345   deep learning approach, we trained individual tree detection models for four geographic

346   sites and studied the transferability of tree features among forest types. Our results

347   show significant improvements over commonly used LiDAR-only implementations.

348   Despite challenging conditions including overlapping canopies and a range of

349   acquisition environments, the proposed approach holds promise for automated tree

350   location and size detection at scale. On average across sites, the universal model

351   correctly identified crown extent with approximately 65% recall and 70% precision. The

352   remaining false positives were almost always correctly detected individual trees, but

353   whose crown boundaries did not meet the intersection over union score of greater than

354   0.5.

355   One goal was to assess the proposed crown detection approach in variety of

356   canopy conditions to better understand which factors limit performance. We find

357   performance is best in open environments with large, well-spaced, trees as in the Oak

358   Woodland site. We had anticipated the performance of the algorithm would be worst at

359   the closed canopy Eastern Deciduous site. However, it was at the Alpine site that the

360   algorithm had the poorest performance, suggesting that short dense trees, rather than

361   complex, interconnected tree boundaries are the biggest challenge. One possible

362   explanation is that the trees in the Alpine site are more sensitive to the resolution of the

363   RGB image due to their small size. Since we use an evaluation metric of intersection-

364   over-union of 0.5, a difference of one pixel is inconsequential for large trees but may

365   push small trees under the threshold for predicted positive.

366   One of the advantages of deep learning approaches to tree detection is the potential

367   to learn cross-site features. We conducted three types of generalization experiments to

368   assess the transferability among forest types. The first was to use models trained from

369   one site to predict an unseen site. Prediction to unseen conditions is a challenging task

370   in computer vision, especially when the sites were specifically chosen to represent

371   distinct forest types. Overall, we saw a significant decrease in performance between

372   cross-site and within-site models. This means that generalization between two forest

373   types without local training data remains unlikely to provide acceptable results. The one

374   exception was the prediction of the Alpine site, which had superior performance when

375   predicted by the Mixed Pine site, rather than using the Alpine hand annotations. This

376   may stem from the difficulty of hand annotating the small trees that are common in the

377   Alpine site. It is possible that the model was better at transferring the features from the

378   large conifers in Mixed Pine to the smaller conifers in Alpine than a human was in

379   annotating the crown boundaries in Alpine. A second possibility is that the significant

380   heterogeneity in the pretraining data for the Alpine site led to poor results. The LiDAR-

381   based pretraining algorithm did not perform well at this site, with consistent under-

382   segmentation among small trees. It is possible that the superior quality of the pretraining

383   data at the Mixed Pine site allowed for better predictions in the Alpine site, compared to

384   using lower quality data from the same site.

385   To provide the model with more information on local tree conditions, we conducted

386   transfer learning experiments to assess whether models pretrained at other sites could

387   be used in conjunction with training data from a local to site to fine tune the model to

388   that site. We find that building from existing models of tree detection is a promising

389   avenue towards cross-site generalization. Adding only a small amount of local training

390   data greatly increased performance and nearly recovered performance of the within-site

391   model. The results were mostly logical; combining pretraining from a deciduous site

392   (e.g. Oak Woodland) to predict another deciduous site (e.g. Eastern Deciduous) is

393   better than using pretraining from a coniferous site (e.g. Alpine). This opens up the

394   possibility of regional tree detection models that connect ecotypes based on their

395   dominant canopy structure and species.

396   The ultimate goal of the proposed approach is to move toward a single unified

397   model that can produce individual tree predictions in a variety of ecosystems. Our

398   analysis shows promising results for a universal model trained from all pretraining and

399   hand annotations from every site. In all sites, a universal model provided equivalent or

400   better predictions than a within-site model, with improvements of up to 20% in one site.

401   Given that the sites were selected to be as different as possible, and encompass a

402   range of tree canopy conditions, this result highlights the ability of convolutional neural

403   networks to learn flexible deep features. We expect that as more sites are included, the

404   universal model will continue to improve. This means that a way forward is to combine

405   pretraining from as many sites as possible. Given that each NEON site has millions of

406   trees, and there are dozens of sites with trees collected annually, there is a possibility of

407   pretraining on continental scale. Further work is needed to know the balance between

408   the number of training images per site and the number of sites to most efficiently train

409   generalized features.

410   In addition to universal model development, transferring knowledge beyond the

411   NEON sites may be useful for many applied problems. It is currently unknown to what

412   extent features learned from the 0.1 m resolution data used here can be applied to

413   lower resolution satellite data (Karlson et al., 2014) or higher resolution UAV data

414   (Brieger et al., 2019). Cross resolution training has not been fully explored in

415   environmental remote sensing, but Li et al., (2018) recently showed that deep learning

416   networks can learn scale invariant land classifications that can be matched among data

417   sources. Given the ability to collect virtually unlimited pretraining data using our semi-

418   supervised approach, NEON sites can be seen as an ideal training sources for RGB

419   tree models that could then be applied to other data types.

420   Our semi-supervised deep learning method uses LiDAR-based pretraining and

421   RGB deep learning to perform individual tree segmentation (Weinstein et al. 2019). The

422   NEON Airborne platform also collects hyperspectral information that may improve

423   generalization across sites with similar species composition. Due to foliar and physical

424   properties, tree species often have distinct spectral signatures which may facilitate

425   distinguishing adjacent tree crowns. Hyperspectral features for tree species

426   classification are relatively common (e.g. Maschler et al., 2018), but few papers have

427   focused on integrating hyperspectral data into tree detection alongside other sensors.

428   Hyperspectral data is available for all NEON sites, and we utilized a three-band

429   composite image to assist in annotating the Eastern Deciduous site (Figure 9),

430   illustrating the usefulness of hyperspectral data to distinguish adjacent tree crowns with

431   human vision. Choosing the best way to represent high-dimensional hyperspectral data

432   in conjunction with the LiDAR and RGB data is non-trivial and will be important for

433   improvements in individual tree detection at broad scales.
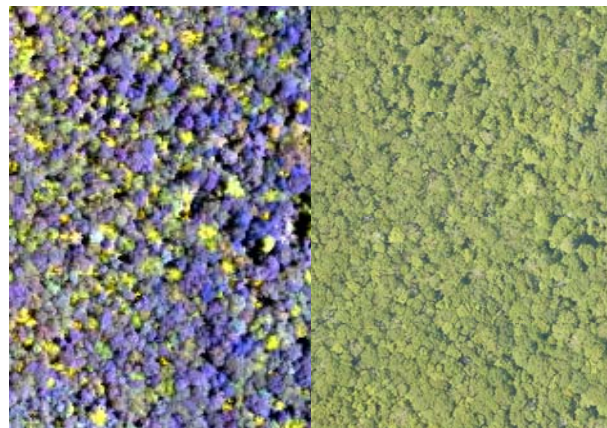
434

435   Figure 9. Composite hyperspectral image and corresponding RGB image for the
436   Eastern Deciduous site. The composite image contained near infrared (940nm), red
437   (650nm), and blue (430nm) bands. Forests that are difficult to segment in RGB imagery
438   may be more separable in hyperspectral imagery due to the differing foliar chemical
439   properties of co-occurring trees.

440   Methods to extract ecological information from airborne sensors are maturing

441   due to advancements in computer vision, data availability and sensor quality. Given our

442   results, what are the strengths and limitations ecologists should consider when adding

443   airborne-derived data? Our results, and prior works (Aubry-Kientz et al., 2019), suggest

444   that small and subcanopy trees will likely be overlooked. We therefore expect that

445   studies in which the results are driven by the upper canopy, sun-exposed trees will

446   benefit the most from remote sensing at broad scales. For example, the total amount of

447   biomass in most forests depends strongly on the largest trees and will be less sensitive

448   to potential non-detections of smaller subcanopy trees (Asner et al., 2012, Stegen et al.

449   2011, Bastin et al. 2018). The inclusion of RGB data may benefit existing large-scale

450   LiDAR-based studies of tree growth (Caughlin et al., 2016), taxonomy (Féret and Asner,

451   2012) and disturbance (Garcia Millan and Sanchez-Azofeifa, 2018), since improved

452   individual segmentation will lead to a more accurate matching of individual trees to

453   metadata on taxonomy and health status. Studies of post-landscape disturbance, such

454   as post-fire, will be aided by the broader perspective of airborne data, as well having

455   significantly reduced risk compared to field surveys in recently burned forests. Most

456   disturbances, such as fire and windstorms, alter the size distribution of forests, including

457   large trees, and thus our approach can provide valuable, detailed landscape scale

458   information about disturbance intensity and impacts (Kamoske et al., 2019). To address

459   these questions, we envision a future in which airborne data on tree locations and sizes

460   are a complement to local field surveys in broadening the scale of sampling in complex

461   landscapes.

462 ACKNOWLEDGEMENTS

468 LITERATURE CITED

469   Agarwal, S., Terrail, J.O. Du, Jurie, F., 2018. Recent Advances in Object Detection in

470   the Age of Deep Convolutional Neural Networks 1–106. https://doi.org/10.1016/B0-

471   12-369397-7/00385-X

472   Asner, G.P., Clark, J.K., Mascaro, J., Garc, G.A.G., 2012. High-resolution mapping of

473   forest carbon stocks in the Colombian Amazon 2683–2696.

474   https://doi.org/10.5194/bg-9-2683-2012

475   Aubry-Kientz, M., Dutrieux, R., Ferraz, A., Saatchi, S., Hamraz, H., Williams, J.,

476   Coomes, D., Piboule, A., Vincent, G., 2019. A Comparative Assessment of the

477   Performance of Individual Tree Crowns Delineation Algorithms from ALS Data in

478   Tropical Forests. Remote Sens. 11, 1086. https://doi.org/10.3390/rs11091086

479   Bastin, J.F., Rutishauser, E., Kellner, J.R., Saatchi, S., Pélissier, R., Hérault, B., Slik, F.,

480   Bogaert, J., De Cannière, C., Marshall, A.R., Poulsen, J., Alvarez-Loyayza, P.,

481   Andrade, A., Angbonga-Basia, A., Araujo-Murakami, A., Arroyo, L., Ayyappan, N.,

482   de Azevedo, C.P., Banki, O., Barbier, N., Barroso, J.G., Beeckman, H., Bitariho, R.,

483   Boeckx, P., Boehning-Gaese, K., Brandão, H., Brearley, F.Q., Breuer Ndoundou

484   Hockemba, M., Brienen, R., Camargo, J.L.C., Campos-Arceiz, A., Cassart, B.,

485   Chave, J., Chazdon, R., Chuyong, G., Clark, D.B., Clark, C.J., Condit, R., Honorio

486   Coronado, E.N., Davidar, P., de Haulleville, T., Descroix, L., Doucet, J.L., Dourdain,

487   A., Droissart, V., Duncan, T., Silva Espejo, J., Espinosa, S., Farwig, N., Fayolle, A.,

488   Feldpausch, T.R., Ferraz, A., Fletcher, C., Gajapersad, K., Gillet, J.F., Amaral, I.L.

489   do, Gonmadje, C., Grogan, J., Harris, D., Herzog, S.K., Homeier, J., Hubau, W.,

490   Hubbell, S.P., Hufkens, K., Hurtado, J., Kamdem, N.G., Kearsley, E., Kenfack, D.,

491   Kessler, M., Labrière, N., Laumonier, Y., Laurance, S., Laurance, W.F., Lewis, S.L.,

492   Libalah, M.B., Ligot, G., Lloyd, J., Lovejoy, T.E., Malhi, Y., Marimon, B.S., Marimon

493   Junior, B.H., Martin, E.H., Matius, P., Meyer, V., Mendoza Bautista, C.,

494   Monteagudo-Mendoza, A., Mtui, A., Neill, D., Parada Gutierrez, G.A., Pardo, G.,

495   Parren, M., Parthasarathy, N., Phillips, O.L., Pitman, N.C.A., Ploton, P., Ponette,

496   Q., Ramesh, B.R., Razafimahaimodison, J.C., Réjou-Méchain, M., Rolim, S.G.,

497   Romero-Saltos, H., Rossi, L.M.B., Spironello, W.R., Rovero, F., Saner, P., Sasaki,

498   D., Schulze, M., Silveira, M., Singh, J., Sist, P., Sonke, B., Soto, J.D., de Souza,

499   C.R., Stropp, J., Sullivan, M.J.P., Swanepoel, B., Steege, H. ter, Terborgh, J.,

500   Texier, N., Toma, T., Valencia, R., Valenzuela, L., Ferreira, L.V., Valverde, F.C.,

501   Van Andel, T.R., Vasque, R., Verbeeck, H., Vivek, P., Vleminckx, J., Vos, V.A.,

502   Wagner, F.H., Warsudi, P.P., Wortel, V., Zagt, R.J., Zebaze, D., 2018. Pan-tropical

503   prediction of forest structure from the largest trees. Glob. Ecol. Biogeogr. 27, 1366–

504   1383. https://doi.org/10.1111/geb.12803

505   Brieger, F., Herzschuh, U., Pestryakova, L.A., Bookhagen, B., Zakharov, E.S., Kruse,

506   S., 2019. Advances in the Derivation of Northeast Siberian Forest Metrics Using

507   High-Resolution UAV-Based Photogrammetric Point Clouds. Remote Sens. 11,

508   1447. https://doi.org/10.3390/rs11121447

509   Caughlin, T.T., Graves, S.J., Asner, G.P., Van Breugel, M., Hall, J.S., Martin, R.E.,

510   Ashton, M.S., Bohlman, S.A., 2016. A hyperspectral image can predict tropical tree

511   growth rates in single-species stands. Ecol. Appl. 26, 2367–2373.

512   https://doi.org/10.1002/eap.1436

513   Coomes, D.A., Dalponte, M., Jucker, T., Asner, G.P., Banin, L.F., Burslem, D.F.R.P.,

514   Lewis, S.L., Nilus, R., Phillips, O.L., Phua, M.H., Qie, L., 2017. Area-based vs tree-

515   centric approaches to mapping forest carbon in Southeast Asian forests from

516   airborne laser scanning data. Remote Sens. Environ. 194, 77–88.

517   https://doi.org/10.1016/j.rse.2017.03.017

518   Dai, W., Yang, B., Dong, Z., Shaker, A., 2018. A new method for 3D individual tree

519   extraction using multispectral airborne LiDAR point clouds. ISPRS J. Photogramm.

520   Remote Sens. 144, 400–411. https://doi.org/10.1016/j.isprsjprs.2018.08.010

521   Dalponte, M., Coomes, D.A., 2016. Tree-centric mapping of forest carbon density from

522   airborne laser scanning and hyperspectral data. Methods Ecol. Evol. 7, 1236–1245.

523   https://doi.org/10.1111/2041-210X.12575

524   Féret, J.B., Asner, G.P., 2012. Semi-supervised methods to identify individual crowns of

525   lowland tropical canopy species using imaging spectroscopy and lidar. Remote

526   Sens. 4, 2457–2476. https://doi.org/10.3390/rs4082457

527   Garcia Millan, V., Sanchez-Azofeifa, A., 2018. Quantifying Changes on Forest

528   Succession in a Dry Tropical Forest Using Angular-Hyperspectral Remote Sensing.

529   Remote Sens. 10, 1865. https://doi.org/10.3390/rs10121865

530   Gomes, M.F., Maillard, P., Deng, H., 2018. Individual tree crown detection in sub-meter

531   satellite imagery using Marked Point Processes and a geometrical-optical model.

532   Remote Sens. Environ. 211, 184–195. https://doi.org/10.1016/j.rse.2018.04.002

533   González-Jaramillo, V., Fries, A., Bendix, J., 2019. AGB estimation in a tropical

534   mountain forest (TMF) by means of RGB and multispectral images using an

535   unmanned aerial vehicle (UAV). Remote Sens. 11, 1–22.

536   https://doi.org/10.3390/rs11121413

537   Hans Gaiser, Maarten de Vries, Valeriu Lacatusu, Ashley Williamson, Enrico Liscio,

538   D.D., 2018. fizy-r/Keras-retinanet. https://doi.org/zenodo.1464720

539   He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image

540   Recognition. Comput. Vis. Pattern Recognit. (CVPR), 2016 770–778.

541   https://doi.org/10.3389/fpsyg.2013.00124

542   Heinzel, J., Huber, M.O., 2018. Constrained Spectral Clustering of Individual Trees in

543   Dense Forest Using Terrestrial Laser Scanning Data.

544   https://doi.org/10.3390/rs10071056

545   Jakubowski, M.K., Li, W., Guo, Q., Kelly, M., 2013. Delineating individual trees from

546   lidar data: A comparison of vector- and raster-based segmentation approaches.

547   Remote Sens. 5, 4163–4186. https://doi.org/10.3390/rs5094163

548   Kamoske, A.G., Dahlin, K.M., Stark, S.C., Serbin, S.P., 2019. Leaf area density from

549   airborne LiDAR: Comparing sensors and resolutions in a temperate broadleaf

550   forest ecosystem. For. Ecol. Manage. 433, 364–375.

551   https://doi.org/10.1016/j.foreco.2018.11.017

552   Karlson, M., Reese, H., Ostwald, M., 2014. Tree crown mapping in managed woodlands

553   (Parklands) of semi-arid West Africa using WorldView-2 imagery and geographic

554   object based image analysis. Sensors (Switzerland) 14, 22643–22669.

555   https://doi.org/10.3390/s141222643

556 Li, W., Guo, Q., Jakubowski, M.K., Kelly, M., 2012. A New Method for Segmenting

557   Individual Trees from the Lidar Point Cloud. Photogramm. Eng. Remote Sens. 78,

558   75–84. https://doi.org/10.14358/PERS.78.1.75

559 Li, Y., Zhang, Y., Huang, X., Ma, J., 2018. Learning source-invariant deep hashing

560   convolutional neural networks for cross-source remote sensing image retrieval.

561   IEEE Trans. Geosci. Remote Sens. 56, 6521–6536.

562   https://doi.org/10.1109/TGRS.2018.2839705

563 Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P., 2017. Focal Loss for Dense Object

564   Detection. Proc. IEEE Int. Conf. Comput. Vis. 2017-Octob, 2999–3007.

565   https://doi.org/10.1109/ICCV.2017.324

566 Maschler, J., Atzberger, C., Immitzer, M., 2018. Individual Tree Crown Segmentation

567   and Classification of 13 Tree Species Using Airborne Hyperspectral Data. Remote

568   Sens. 10, 1218. https://doi.org/10.3390/rs10081218

569 Roussel, J.-R., David Auty, 2019. lidR: Airborne LiDAR Data Manipulation and

570   Visualization for Forestry Applications.

571 Silva, C.A., Hudak, A.T., Vierling, L.A., Loudermilk, E.L., O'Brien, J.J., Hiers, J.K., Jack,

572   S.B., Gonzalez-Benecke, C., Lee, H., Falkowski, M.J., Khosravipour, A., 2016.

573   Imputation of Individual Longleaf Pine (Pinus palustris Mill.) Tree Attributes from

574   Field and LiDAR Data. Can. J. Remote Sens. 42, 554–573.

575   https://doi.org/10.1080/07038992.2016.1196582

576   Stegen, J.C., Swenson, N.G., Enquist, B.J., White, E.P., Phillips, O.L., Jørgensen, P.M.,

577   Weiser, M.D., Monteagudo Mendoza, A., Núñez Vargas, P., 2011. Variation in

578   above-ground forest biomass across broad climatic gradients. Glob. Ecol.

579   Biogeogr. 20, 744–754. https://doi.org/10.1111/j.1466-8238.2010.00645.x

580   Vaglio Laurin, G., Ding, J., Disney, M., Bartholomeus, H., Herold, M., Papale, D.,

581   Valentini, R., 2019. Tree height in tropical forest as measured by different ground,

582   proximal, and remote sensing instruments, and impacts on above ground biomass

583   estimates. Int. J. Appl. Earth Obs. Geoinf. 82, 101899.

584   https://doi.org/10.1016/j.jag.2019.101899

585   Williams, J., Schönlieb, C.-B., Swinfield, T., Lee, J., Cai, X., Qie, L., Coomes, D.A.,

586   2019. Three-dimensional Segmentation of Trees Through a Flexible Multi-Class

587   Graph Cut Algorithm (MCGC) 1–33.

588   Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep

589   Learning in Remote Sensing: A Comprehensive Review and List of Resources.

590   IEEE Geosci. Remote Sens. Mag. 5, 8–36.

591   https://doi.org/10.1109/MGRS.2017.2762307

592

593   Asner, G.P., Clark, J.K., Mascaro, J., Garc, G.A.G., 2012. High-resolution mapping of

594   forest carbon stocks in the Colombian Amazon 2683–2696.

595   https://doi.org/10.5194/bg-9-2683-2012

596   Aubry-Kientz, M., Dutrieux, R., Ferraz, A., Saatchi, S., Hamraz, H., Williams, J.,

597   Coomes, D., Piboule, A., Vincent, G., 2019. A Comparative Assessment of the

598   Performance of Individual Tree Crowns Delineation Algorithms from ALS Data in

599   Tropical Forests. Remote Sens. 11, 1086. https://doi.org/10.3390/rs11091086

600   Bastin, J.F., Rutishauser, E., Kellner, J.R., Saatchi, S., Pélissier, R., Hérault, B., Slik, F.,

601   Bogaert, J., De Cannière, C., Marshall, A.R., Poulsen, J., Alvarez-Loyayza, P.,

602   Andrade, A., Angbonga-Basia, A., Araujo-Murakami, A., Arroyo, L., Ayyappan, N.,

603   de Azevedo, C.P., Banki, O., Barbier, N., Barroso, J.G., Beeckman, H., Bitariho, R.,

604   Boeckx, P., Boehning-Gaese, K., Brandão, H., Brearley, F.Q., Breuer Ndoundou

605   Hockemba, M., Brienen, R., Camargo, J.L.C., Campos-Arceiz, A., Cassart, B.,

606   Chave, J., Chazdon, R., Chuyong, G., Clark, D.B., Clark, C.J., Condit, R., Honorio

607   Coronado, E.N., Davidar, P., de Haulleville, T., Descroix, L., Doucet, J.L., Dourdain,

608   A., Droissart, V., Duncan, T., Silva Espejo, J., Espinosa, S., Farwig, N., Fayolle, A.,

609   Feldpausch, T.R., Ferraz, A., Fletcher, C., Gajapersad, K., Gillet, J.F., Amaral, I.L.

610   do, Gonmadje, C., Grogan, J., Harris, D., Herzog, S.K., Homeier, J., Hubau, W.,

611   Hubbell, S.P., Hufkens, K., Hurtado, J., Kamdem, N.G., Kearsley, E., Kenfack, D.,

612   Kessler, M., Labrière, N., Laumonier, Y., Laurance, S., Laurance, W.F., Lewis, S.L.,

613   Libalah, M.B., Ligot, G., Lloyd, J., Lovejoy, T.E., Malhi, Y., Marimon, B.S., Marimon

614   Junior, B.H., Martin, E.H., Matius, P., Meyer, V., Mendoza Bautista, C.,

615   Monteagudo-Mendoza, A., Mtui, A., Neill, D., Parada Gutierrez, G.A., Pardo, G.,

616   Parren, M., Parthasarathy, N., Phillips, O.L., Pitman, N.C.A., Ploton, P., Ponette,

617   Q., Ramesh, B.R., Razafimahaimodison, J.C., Réjou-Méchain, M., Rolim, S.G.,

618   Romero-Saltos, H., Rossi, L.M.B., Spironello, W.R., Rovero, F., Saner, P., Sasaki,

619   D., Schulze, M., Silveira, M., Singh, J., Sist, P., Sonke, B., Soto, J.D., de Souza,

620   C.R., Stropp, J., Sullivan, M.J.P., Swanepoel, B., Steege, H. ter, Terborgh, J.,

621   Texier, N., Toma, T., Valencia, R., Valenzuela, L., Ferreira, L.V., Valverde, F.C.,

622   Van Andel, T.R., Vasque, R., Verbeeck, H., Vivek, P., Vleminckx, J., Vos, V.A.,

623   Wagner, F.H., Warsudi, P.P., Wortel, V., Zagt, R.J., Zebaze, D., 2018. Pan-tropical

624   prediction of forest structure from the largest trees. Glob. Ecol. Biogeogr. 27, 1366–

625   1383. https://doi.org/10.1111/geb.12803

626 Brieger, F., Herzschuh, U., Pestryakova, L.A., Bookhagen, B., Zakharov, E.S., Kruse,

627   S., 2019. Advances in the Derivation of Northeast Siberian Forest Metrics Using

628   High-Resolution UAV-Based Photogrammetric Point Clouds. Remote Sens. 11,

629   1447. https://doi.org/10.3390/rs11121447

630 Caughlin, T.T., Graves, S.J., Asner, G.P., Van Breugel, M., Hall, J.S., Martin, R.E.,

631   Ashton, M.S., Bohlman, S.A., 2016. A hyperspectral image can predict tropical tree

632   growth rates in single-species stands. Ecol. Appl. 26, 2367–2373.

633   https://doi.org/10.1002/eap.1436

634 Coomes, D.A., Dalponte, M., Jucker, T., Asner, G.P., Banin, L.F., Burslem, D.F.R.P.,

635   Lewis, S.L., Nilus, R., Phillips, O.L., Phua, M.H., Qie, L., 2017. Area-based vs tree-

636   centric approaches to mapping forest carbon in Southeast Asian forests from

637   airborne laser scanning data. Remote Sens. Environ. 194, 77–88.

638   https://doi.org/10.1016/j.rse.2017.03.017

639 Dai, W., Yang, B., Dong, Z., Shaker, A., 2018. A new method for 3D individual tree

640   extraction using multispectral airborne LiDAR point clouds. ISPRS J. Photogramm.

641   Remote Sens. 144, 400–411. https://doi.org/10.1016/j.isprsjprs.2018.08.010

642   Dalponte, M., Coomes, D.A., 2016. Tree-centric mapping of forest carbon density from

643   airborne laser scanning and hyperspectral data. Methods Ecol. Evol. 7, 1236–1245.

644   https://doi.org/10.1111/2041-210X.12575

645   Féret, J.B., Asner, G.P., 2012. Semi-supervised methods to identify individual crowns of

646   lowland tropical canopy species using imaging spectroscopy and lidar. Remote

647   Sens. 4, 2457–2476. https://doi.org/10.3390/rs4082457

648   Garcia Millan, V., Sanchez-Azofeifa, A., 2018. Quantifying Changes on Forest

649   Succession in a Dry Tropical Forest Using Angular-Hyperspectral Remote Sensing.

650   Remote Sens. 10, 1865. https://doi.org/10.3390/rs10121865

651   Gomes, M.F., Maillard, P., Deng, H., 2018. Individual tree crown detection in sub-meter

652   satellite imagery using Marked Point Processes and a geometrical-optical model.

653   Remote Sens. Environ. 211, 184–195. https://doi.org/10.1016/j.rse.2018.04.002

654   González-Jaramillo, V., Fries, A., Bendix, J., 2019. AGB estimation in a tropical

655   mountain forest (TMF) by means of RGB and multispectral images using an

656   unmanned aerial vehicle (UAV). Remote Sens. 11, 1–22.

657   https://doi.org/10.3390/rs11121413

658   Hans Gaiser, Maarten de Vries, Valeriu Lacatusu, Ashley Williamson, Enrico Liscio,

659   D.D., 2018. fizy-r/Keras-retinanet. https://doi.org/zenodo.1464720

660   He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image

661   Recognition. Comput. Vis. Pattern Recognit. (CVPR), 2016 770–778.

662   https://doi.org/10.3389/fpsyg.2013.00124

663   Heinzel, J., Huber, M.O., 2018. Constrained Spectral Clustering of Individual Trees in

664   Dense Forest Using Terrestrial Laser Scanning Data.

665   https://doi.org/10.3390/rs10071056

666   Jakubowski, M.K., Li, W., Guo, Q., Kelly, M., 2013. Delineating individual trees from

667   lidar data: A comparison of vector- and raster-based segmentation approaches.

668   Remote Sens. 5, 4163–4186. https://doi.org/10.3390/rs5094163

669   Kamoske, A.G., Dahlin, K.M., Stark, S.C., Serbin, S.P., 2019. Leaf area density from

670   airborne LiDAR: Comparing sensors and resolutions in a temperate broadleaf

671   forest ecosystem. For. Ecol. Manage. 433, 364–375.

672   https://doi.org/10.1016/j.foreco.2018.11.017

673   Karlson, M., Reese, H., Ostwald, M., 2014. Tree crown mapping in managed woodlands

674   (Parklands) of semi-arid West Africa using WorldView-2 imagery and geographic

675   object based image analysis. Sensors (Switzerland) 14, 22643–22669.

676   https://doi.org/10.3390/s141222643

677   Li, W., Guo, Q., Jakubowski, M.K., Kelly, M., 2012. A New Method for Segmenting

678   Individual Trees from the Lidar Point Cloud. Photogramm. Eng. Remote Sens. 78,

679   75–84. https://doi.org/10.14358/PERS.78.1.75

680   Li, Y., Zhang, Y., Huang, X., Ma, J., 2018. Learning source-invariant deep hashing

681   convolutional neural networks for cross-source remote sensing image retrieval.

682   IEEE Trans. Geosci. Remote Sens. 56, 6521–6536.

683   https://doi.org/10.1109/TGRS.2018.2839705

684   Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P., 2017. Focal Loss for Dense Object

685   Detection. Proc. IEEE Int. Conf. Comput. Vis. 2017-Octob, 2999–3007.

686   https://doi.org/10.1109/ICCV.2017.324

687   Maschler, J., Atzberger, C., Immitzer, M., 2018. Individual Tree Crown Segmentation

688   and Classification of 13 Tree Species Using Airborne Hyperspectral Data. Remote

689   Sens. 10, 1218. https://doi.org/10.3390/rs10081218

690   Roussel, J.-R., David Auty, 2019. lidR: Airborne LiDAR Data Manipulation and

691   Visualization for Forestry Applications.

692   Silva, C.A., Hudak, A.T., Vierling, L.A., Loudermilk, E.L., O'Brien, J.J., Hiers, J.K., Jack,

693   S.B., Gonzalez-Benecke, C., Lee, H., Falkowski, M.J., Khosravipour, A., 2016.

694   Imputation of Individual Longleaf Pine (Pinus palustris Mill.) Tree Attributes from

695   Field and LiDAR Data. Can. J. Remote Sens. 42, 554–573.

696   https://doi.org/10.1080/07038992.2016.1196582

697   Stegen, J.C., Swenson, N.G., Enquist, B.J., White, E.P., Phillips, O.L., Jørgensen, P.M.,

698   Weiser, M.D., Monteagudo Mendoza, A., Núñez Vargas, P., 2011. Variation in

699   above-ground forest biomass across broad climatic gradients. Glob. Ecol.

700   Biogeogr. 20, 744–754. https://doi.org/10.1111/j.1466-8238.2010.00645.x

701   Vaglio Laurin, G., Ding, J., Disney, M., Bartholomeus, H., Herold, M., Papale, D.,

702   Valentini, R., 2019. Tree height in tropical forest as measured by different ground,

703   proximal, and remote sensing instruments, and impacts on above ground biomass

704   estimates. Int. J. Appl. Earth Obs. Geoinf. 82, 101899.

705   https://doi.org/10.1016/j.jag.2019.101899

706 Weinstein, B.G., Marconi, S., Bohlman, S., Zare, A., White, E., 2019. Individual Tree-

707   Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural

708   Networks. Remote Sens. 11, 1309. https://doi.org/10.3390/rs11111309

709 Williams, J., Schönlieb, C.-B., Swinfield, T., Lee, J., Cai, X., Qie, L., Coomes, D.A.,

710   2019. Three-dimensional Segmentation of Trees Through a Flexible Multi-Class

711   Graph Cut Algorithm (MCGC) 1–33.

712 Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep

713   Learning in Remote Sensing: A Comprehensive Review and List of Resources.

714   IEEE Geosci. Remote Sens. Mag. 5, 8–36.

715   https://doi.org/10.1109/MGRS.2017.2762307

716

717 Asner, G.P., Clark, J.K., Mascaro, J., Garc, G.A.G., 2012. High-resolution mapping of

718   forest carbon stocks in the Colombian Amazon 2683–2696.

719   https://doi.org/10.5194/bg-9-2683-2012

720 Aubry-Kientz, M., Dutrieux, R., Ferraz, A., Saatchi, S., Hamraz, H., Williams, J.,

721   Coomes, D., Piboule, A., Vincent, G., 2019. A Comparative Assessment of the

722   Performance of Individual Tree Crowns Delineation Algorithms from ALS Data in

723   Tropical Forests. Remote Sens. 11, 1086. https://doi.org/10.3390/rs11091086

724 Bastin, J.F., Rutishauser, E., Kellner, J.R., Saatchi, S., Pélissier, R., Hérault, B., Slik, F.,

725   Bogaert, J., De Cannière, C., Marshall, A.R., Poulsen, J., Alvarez-Loyayza, P.,

726   Andrade, A., Angbonga-Basia, A., Araujo-Murakami, A., Arroyo, L., Ayyappan, N.,

727   de Azevedo, C.P., Banki, O., Barbier, N., Barroso, J.G., Beeckman, H., Bitariho, R.,

728   Boeckx, P., Boehning-Gaese, K., Brandão, H., Brearley, F.Q., Breuer Ndoundou

729   Hockemba, M., Brienen, R., Camargo, J.L.C., Campos-Arceiz, A., Cassart, B.,

730   Chave, J., Chazdon, R., Chuyong, G., Clark, D.B., Clark, C.J., Condit, R., Honorio

731   Coronado, E.N., Davidar, P., de Haulleville, T., Descroix, L., Doucet, J.L., Dourdain,

732   A., Droissart, V., Duncan, T., Silva Espejo, J., Espinosa, S., Farwig, N., Fayolle, A.,

733   Feldpausch, T.R., Ferraz, A., Fletcher, C., Gajapersad, K., Gillet, J.F., Amaral, I.L.

734   do, Gonmadje, C., Grogan, J., Harris, D., Herzog, S.K., Homeier, J., Hubau, W.,

735   Hubbell, S.P., Hufkens, K., Hurtado, J., Kamdem, N.G., Kearsley, E., Kenfack, D.,

736   Kessler, M., Labrière, N., Laumonier, Y., Laurance, S., Laurance, W.F., Lewis, S.L.,

737   Libalah, M.B., Ligot, G., Lloyd, J., Lovejoy, T.E., Malhi, Y., Marimon, B.S., Marimon

738   Junior, B.H., Martin, E.H., Matius, P., Meyer, V., Mendoza Bautista, C.,

739   Monteagudo-Mendoza, A., Mtui, A., Neill, D., Parada Gutierrez, G.A., Pardo, G.,

740   Parren, M., Parthasarathy, N., Phillips, O.L., Pitman, N.C.A., Ploton, P., Ponette,

741   Q., Ramesh, B.R., Razafimahaimodison, J.C., Réjou-Méchain, M., Rolim, S.G.,

742   Romero-Saltos, H., Rossi, L.M.B., Spironello, W.R., Rovero, F., Saner, P., Sasaki,

743   D., Schulze, M., Silveira, M., Singh, J., Sist, P., Sonke, B., Soto, J.D., de Souza,

744   C.R., Stropp, J., Sullivan, M.J.P., Swanepoel, B., Steege, H. ter, Terborgh, J.,

745   Texier, N., Toma, T., Valencia, R., Valenzuela, L., Ferreira, L.V., Valverde, F.C.,

746   Van Andel, T.R., Vasque, R., Verbeeck, H., Vivek, P., Vleminckx, J., Vos, V.A.,

747   Wagner, F.H., Warsudi, P.P., Wortel, V., Zagt, R.J., Zebaze, D., 2018. Pan-tropical

748   prediction of forest structure from the largest trees. Glob. Ecol. Biogeogr. 27, 1366–

749   1383. https://doi.org/10.1111/geb.12803

750   Brieger, F., Herzschuh, U., Pestryakova, L.A., Bookhagen, B., Zakharov, E.S., Kruse,

751   S., 2019. Advances in the Derivation of Northeast Siberian Forest Metrics Using

752   High-Resolution UAV-Based Photogrammetric Point Clouds. Remote Sens. 11,

753   1447. https://doi.org/10.3390/rs11121447

754   Caughlin, T.T., Graves, S.J., Asner, G.P., Van Breugel, M., Hall, J.S., Martin, R.E.,

755   Ashton, M.S., Bohlman, S.A., 2016. A hyperspectral image can predict tropical tree

756   growth rates in single-species stands. Ecol. Appl. 26, 2367–2373.

757   https://doi.org/10.1002/eap.1436

758   Coomes, D.A., Dalponte, M., Jucker, T., Asner, G.P., Banin, L.F., Burslem, D.F.R.P.,

759   Lewis, S.L., Nilus, R., Phillips, O.L., Phua, M.H., Qie, L., 2017. Area-based vs tree-

760   centric approaches to mapping forest carbon in Southeast Asian forests from

761   airborne laser scanning data. Remote Sens. Environ. 194, 77–88.

762   https://doi.org/10.1016/j.rse.2017.03.017

763   Dai, W., Yang, B., Dong, Z., Shaker, A., 2018. A new method for 3D individual tree

764   extraction using multispectral airborne LiDAR point clouds. ISPRS J. Photogramm.

765   Remote Sens. 144, 400–411. https://doi.org/10.1016/j.isprsjprs.2018.08.010

766   Dalponte, M., Coomes, D.A., 2016. Tree-centric mapping of forest carbon density from

767   airborne laser scanning and hyperspectral data. Methods Ecol. Evol. 7, 1236–1245.

768   https://doi.org/10.1111/2041-210X.12575

769   Féret, J.B., Asner, G.P., 2012. Semi-supervised methods to identify individual crowns of

770   lowland tropical canopy species using imaging spectroscopy and lidar. Remote

771   Sens. 4, 2457–2476. https://doi.org/10.3390/rs4082457

772   Garcia Millan, V., Sanchez-Azofeifa, A., 2018. Quantifying Changes on Forest

773   Succession in a Dry Tropical Forest Using Angular-Hyperspectral Remote Sensing.

774   Remote Sens. 10, 1865. https://doi.org/10.3390/rs10121865

775   Gomes, M.F., Maillard, P., Deng, H., 2018. Individual tree crown detection in sub-meter

776   satellite imagery using Marked Point Processes and a geometrical-optical model.

777   Remote Sens. Environ. 211, 184–195. https://doi.org/10.1016/j.rse.2018.04.002

778   González-Jaramillo, V., Fries, A., Bendix, J., 2019. AGB estimation in a tropical

779   mountain forest (TMF) by means of RGB and multispectral images using an

780   unmanned aerial vehicle (UAV). Remote Sens. 11, 1–22.

781   https://doi.org/10.3390/rs11121413

782   Hans Gaiser, Maarten de Vries, Valeriu Lacatusu, Ashley Williamson, Enrico Liscio,

783   D.D., 2018. fizy-r/Keras-retinanet. https://doi.org/zenodo.1464720

784   He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image

785   Recognition. Comput. Vis. Pattern Recognit. (CVPR), 2016 770–778.

786   https://doi.org/10.3389/fpsyg.2013.00124

787   Heinzel, J., Huber, M.O., 2018. Constrained Spectral Clustering of Individual Trees in

788   Dense Forest Using Terrestrial Laser Scanning Data.

789   https://doi.org/10.3390/rs10071056

790   Jakubowski, M.K., Li, W., Guo, Q., Kelly, M., 2013. Delineating individual trees from

791   lidar data: A comparison of vector- and raster-based segmentation approaches.

792   Remote Sens. 5, 4163–4186. https://doi.org/10.3390/rs5094163

793   Karlson, M., Reese, H., Ostwald, M., 2014. Tree crown mapping in managed woodlands

794   (Parklands) of semi-arid West Africa using WorldView-2 imagery and geographic

795   object based image analysis. Sensors (Switzerland) 14, 22643–22669.

796   https://doi.org/10.3390/s141222643

797   Li, W., Guo, Q., Jakubowski, M.K., Kelly, M., 2012. A New Method for Segmenting

798   Individual Trees from the Lidar Point Cloud. Photogramm. Eng. Remote Sens. 78,

799   75–84. https://doi.org/10.14358/PERS.78.1.75

800   Li, Yangyan, Bu, R., Sun, M., Wu, W., Di, X., Chen, B., 2018. PointCNN: Convolution

801   On X Transformed Points.

802   Li, Yansheng, Zhang, Y., Huang, X., Ma, J., 2018. Learning source-invariant deep

803   hashing convolutional neural networks for cross-source remote sensing image

804   retrieval. IEEE Trans. Geosci. Remote Sens. 56, 6521–6536.

805   https://doi.org/10.1109/TGRS.2018.2839705

806   Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P., 2017. Focal Loss for Dense Object

807   Detection. Proc. IEEE Int. Conf. Comput. Vis. 2017-Octob, 2999–3007.

808   https://doi.org/10.1109/ICCV.2017.324

809   Maschler, J., Atzberger, C., Immitzer, M., 2018. Individual Tree Crown Segmentation

810   and Classification of 13 Tree Species Using Airborne Hyperspectral Data. Remote

811   Sens. 10, 1218. https://doi.org/10.3390/rs10081218

812   Qi, C.R., Su, H., Mo, K., Guibas, L.J., 2017. PointNet: Deep learning on point sets for

813   3D classification and segmentation. Proc. - 30th IEEE Conf. Comput. Vis. Pattern

814   Recognition, CVPR 2017 2017-Janua, 77–85.

815   https://doi.org/10.1109/CVPR.2017.16

816   Roussel, J.-R., David Auty, 2019. lidR: Airborne LiDAR Data Manipulation and

817   Visualization for Forestry Applications.

818   Silva, C.A., Hudak, A.T., Vierling, L.A., Loudermilk, E.L., O'Brien, J.J., Hiers, J.K., Jack,

819   S.B., Gonzalez-Benecke, C., Lee, H., Falkowski, M.J., Khosravipour, A., 2016.

820   Imputation of Individual Longleaf Pine (Pinus palustris Mill.) Tree Attributes from

821   Field and LiDAR Data. Can. J. Remote Sens. 42, 554–573.

822   https://doi.org/10.1080/07038992.2016.1196582

823   Stegen, J.C., Swenson, N.G., Enquist, B.J., White, E.P., Phillips, O.L., Jørgensen, P.M.,

824   Weiser, M.D., Monteagudo Mendoza, A., Núñez Vargas, P., 2011. Variation in

825   above-ground forest biomass across broad climatic gradients. Glob. Ecol.

826   Biogeogr. 20, 744–754. https://doi.org/10.1111/j.1466-8238.2010.00645.x

827   Vaglio Laurin, G., Ding, J., Disney, M., Bartholomeus, H., Herold, M., Papale, D.,

828   Valentini, R., 2019. Tree height in tropical forest as measured by different ground,

829   proximal, and remote sensing instruments, and impacts on above ground biomass

830   estimates. Int. J. Appl. Earth Obs. Geoinf. 82, 101899.

831   https://doi.org/10.1016/j.jag.2019.101899

832   Weinstein, B.G., Marconi, S., Bohlman, S., Zare, A., White, E., 2019. Individual Tree-

833   Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural

834   Networks. Remote Sens. 11, 1309. https://doi.org/10.3390/rs11111309

835   Williams, J., Schönlieb, C.-B., Swinfield, T., Lee, J., Cai, X., Qie, L., Coomes, D.A.,

836   2019. Three-dimensional Segmentation of Trees Through a Flexible Multi-Class

837   Graph Cut Algorithm (MCGC) 1–33.

838   Zhou, Y., Tuzel, O., 2017. VoxelNet: End-to-End Learning for Point Cloud Based 3D

839   Object Detection. https://doi.org/1711.06396v1

840   Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep

841   Learning in Remote Sensing: A Comprehensive Review and List of Resources.

842   IEEE Geosci. Remote Sens. Mag. 5, 8–36.

843   https://doi.org/10.1109/MGRS.2017.2762307

844