

NEON NIST data science evaluation challenge: methods and results of team FEM

Michele Dalponte ^{Corresp., 1}, Lorenzo Frizzera ¹, Damiano Gianelle ¹

¹ Dept. of Sustainable Agro-Ecosystems and Bioresources, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Trento, Italia

Corresponding Author: Michele Dalponte
Email address: michele.dalponte@fmach.it

An international data science challenge, called NEON NIST data science evaluation, was set up in autumn 2017 with the goal to improve the use of remote sensing data in ecological applications. The competition was divided into three tasks: 1) segmentation of tree crowns; 2) data alignment; and 3) tree species classification. In this paper the methods and results of team FEM in the NEON NIST data science evaluation challenge are presented. The individual tree crown (ITC) segmentation (Task 1 of the challenge) was done using a region growing method applied to a near-infrared band of the hyperspectral images. The optimization of the parameters of the segmentation algorithm was done in a supervised way on the basis of the Jaccard score using the training set provided by the organizers. The alignment (Task 2) between the segmented ITCs and the ground measured trees was done using an Euclidean distance among the position, the height, and the crown radius of the ITCs and the ground trees. The classification (Task 3) was performed using a Support Vector Machine classifier applied to a selection of the hyperspectral bands. The selection of the bands was done using a Sequential Forward Floating Selection method and the Jeffries Matusita distance. The results in the three tasks were very promising: team FEM ranked first in Task 1 and 2, and second in Task 3. The segmentation results showed that the proposed approach segmented both small and large crowns. The alignment was correctly done for all the test samples. The classification results were good, even if the accuracy was biased towards the most represented species.

1 **NEON NIST data science evaluation challenge: methods and**
2 **results of team FEM**

3 Michele Dalponte, Lorenzo Frizzera, and Damiano Gianelle

4 Dept. of Sustainable Agro-ecosystems and Bioresources, Research and Innovation Centre,
5 Fondazione E. Mach, Via E. Mach 1, 38010 San Michele all'Adige (TN), Italy.

6

7 Corresponding Author:

8 Michele Dalponte

9 Via E. Mach 1, 38010 San Michele all'Adige (TN), Italy

10 Email address: michele.dalponte@fmach.it

11

12

13 Abstract

14 *An international data science challenge, called NEON NIST data science evaluation, was set up*
15 *in autumn 2017 with the goal to improve the use of remote sensing data in ecological*
16 *applications. The competition was divided into three tasks: 1) segmentation of tree crowns; 2)*
17 *data alignment; and 3) tree species classification. In this paper the methods and results of team*
18 *FEM in the NEON NIST data science evaluation challenge are presented. The individual tree*
19 *crown (ITC) segmentation (Task 1 of the challenge) was done using a region growing method*
20 *applied to a near-infrared band of the hyperspectral images. The optimization of the parameters*
21 *of the segmentation algorithm was done in a supervised way on the basis of the Jaccard score*
22 *using the training set provided by the organizers. The alignment (Task 2) between the segmented*
23 *ITCs and the ground measured trees was done using an Euclidean distance among the position,*
24 *the height, and the crown radius of the ITCs and the ground trees. The classification (Task 3)*
25 *was performed using a Support Vector Machine classifier applied to a selection of the*
26 *hyperspectral bands. The selection of the bands was done using a Sequential Forward Floating*
27 *Selection method and the Jeffries Matusita distance. The results in the three tasks were very*
28 *promising: team FEM ranked first in Task 1 and 2, and second in Task 3. The segmentation*
29 *results showed that the proposed approach segmented both small and large crowns. The*
30 *alignment was correctly done for all the test samples. The classification results were good, even*
31 *if the accuracy was biased towards the most represented species.*

32

33 1 Introduction

34 The NEON NIST data science evaluation challenge (Marconi et al., 2018) was an
35 international competition with the goal to challenge international scientists on three tasks that are

36 central in converting remote sensing images into vegetation diversity and structure information
37 traditionally collected by ecologists: 1) individual tree crown (ITC) segmentation, for identifying
38 the location and size of individual trees; 2) alignment to match ground truth data on trees with
39 remote sensing; and 3) species classification to identify trees to species.

40 There is a large amount of literature about crown segmentation (e.g. Popescu, Wynne &
41 Nelson, 2003; Lee & Lucas, 2007; Ene, Næsset & Gobakken, 2012; Hung, Bryson & Sukkarieh,
42 2012; Ferraz et al., 2012; Duncanson et al., 2015), and there have been many studies comparing
43 segmentation methods on different data types (Ke & Quackenbush, 2011; Vauhkonen et al.,
44 2012; Eysn et al., 2015; Dalponte et al., 2015b). Many papers focus on light detection and
45 ranging (LiDAR) data as these remote sensing data are very common in the forestry and ecology
46 domains. Some studies exist on methods for crown segmentation of camera images, while fewer
47 studies exist on segmentation of hyperspectral data (Dalponte et al., 2014).

48 The alignment to match ground truth data of trees with remote sensing was never explored
49 in specific papers and usually only briefly mentioned on papers devoted to crown segmentation.
50 This fact makes alignment very subjective because different approaches are used in every crown
51 segmentation paper, and the alignment is adapted to the data used in the specific work.

52 Tree species classification with remote sensing data is a widely covered topic by the
53 scientific literature (Fassnacht et al., 2016). The first studies on this topic were focusing on large
54 categories of species as they were done using satellite multispectral data, but since the 2000s
55 with the availability of airborne hyperspectral data many studies focused on the separation of tree
56 species (Dalponte, Bruzzone & Gianelle, 2012; Dalponte et al., 2013; Budei et al., 2017). Indeed,
57 hyperspectral data due to their dense sampling of the spectral signatures can separate many

58 different species with high level of accuracy. Moreover, the advances in the remote sensing
59 community on development of hyperspectral image classifiers, and on band selection and
60 reduction have significantly improved the possibility to detect tree species.

61 The objective of this paper is to present the methods and results of team FEM in the NEON
62 NIST data science evaluation challenge. The FEM team belongs to the Forest Ecology and Bio-
63 geochemical cycles unit of the Research and Innovation Centre of the Edmund Mach Foundation
64 in Italy. The research activities of the Forest ecology and Biogeochemical Cycles unit are
65 focused on the interactions between the vegetation canopy and the atmosphere's chemical-
66 physical layer in addition to the soil structure and functionality. In particular, energy and matter
67 (carbon, water, nitrogen) fluxes between the atmosphere and the biosphere are analysed and
68 models simulating vegetation systems and turbulent and radiative transfer are used. These data
69 are up-scaled at a regional level to obtain a carbon balance integrating ground and remote
70 sensing data bases. The remote sensing team of the unit is specialized in LIDAR and
71 hyperspectral image processing both from airborne and satellite sensors, on the forest domain.

72 **2 Materials**

73 For a detailed description of the data used we recall to (Marconi et al., 2018). The data
74 from NEON included the following data products: 1) Woody plant vegetation structure
75 (NEON.DP1.10098); 2) Spectrometer orthorectified surface directional reflectance - flightline
76 (NEON.DP1.30008); 3) Ecosystem structure (NEON.DP3.30015); and High-resolution
77 orthorectified camera imagery (NEON.DP1.30010).

78 **3 Methods**79 **3.1 Task 1: segmentation**

80 The ITCs segmentation was performed on the hyperspectral data using the algorithm
81 presented in (Dalponte et al., 2015b). In greater detail the steps of the segmentation method
82 were:

- 83 1. the hyperspectral band closest to 810 nm was selected for the segmentation;
- 84 2. the normalized difference vegetation index (NDVI) was computed for each pixel, and all the
85 pixels in the band selected at step 1 having NDVI below 0.6 were masked;
- 86 3. seeds points $S = \{s_1, \dots, s_N\}$ was defined using a moving window. An image pixel $H(x,y)$ was
87 a seed point if:

$$H(x,y) \in S \text{ if } H(x,y) = \max(\text{moving window}) \quad (1)$$

- 88 4. initial regions were defined starting from the seed points. A label map L was defined:

$$\begin{cases} L_{i,j} = k & \text{if } H(i,j) \in S \\ L_{i,j} = 0 & \text{if } H(i,j) \notin S \end{cases} \quad (2)$$

- 89 5. starting from L , regions grew according to the following procedure:
 - 90 a. a label map point $L_{i,j} \neq 0$ was considered and its neighbor pixels (NP) in the image
91 were taken:

$$NP = \{H(i,j-1); H(i-1,j); H(i,j+1); H(i+1,j)\} \quad (3)$$

- 92 b. a neighbor pixel $NP(i',j')$ was added to the region n if:

$$NP(i',j') \in \begin{cases} \text{dist}(NP(i',j'), s_n) < \text{DistMax} \\ NP(i',j') > (s_n * \text{PercThresh}) \\ L_{i',j'} \neq 0 \end{cases} \quad (4)$$

93 where $\text{PercThresh} \in (0;1)$, and $\text{DistMax} > 0$;

- 94 c. this procedure was iterated over all pixels that have $L_{i,j} \neq 0$, and was repeated until
95 no pixels were added to any region;
- 96 6. from each region in L the central coordinates of each pixel were extracted, and a 2D convex
97 hull was applied to these points;
- 98 7. the resulting polygons were the final ITCs.

99 The raster image used in this paper was the hyperspectral band at 810 nm, already used in
100 previous studies for this purpose (Clark, Roberts & Clark, 2005; Dalponte et al., 2014). The
101 parameters of the segmentation (i.e. the size of the moving window, *PercThresh*, and *DistMax*)
102 were optimized in a supervised way using a training set made available by the organizers of the
103 challenge: the set of parameters that provided the highest Jaccard score (Real & Vargas, 1996)
104 on the training set was chosen. The parameters used for the delineation on the test set were: a
105 moving window a size of 3x3 pixels, a *PercThresh* of 0.4, and a *DistMax* of 4. The
106 implementation used is the one in *itcIMG* of the *itcSegment* R package (Dalponte, 2016).

107 3.2 Task 2: alignment

108 The alignment between ground measured trees and the delineated ITCs was done using a four
109 step procedure: 1) prediction of missing ground measured crown radius; 2) prediction of missing
110 ITC heights; 3) linking ITCs and ground measured trees using an Euclidean distance based on X
111 and Y coordinates, and height and crown radius; and 4) visual inspection of the results.

112 The crown radius of ground measured trees, for which this attribute was not measured on
113 the ground, was predicted using a relationship linking the field measured crown radius (R_{FIELD})
114 with the tree height (H_{FIELD}) and the stem diameter (D_{FIELD}):

$$R_{FIELD} = a \times (H_{FIELD} \times D_{FIELD})^b \quad (5)$$

115 Eqn. 5 was fitted using the function *nls* of the package *stats* of the R software (R Development
116 Core Team, 2008).

117 The height of the ITCs, for which this attribute was missing, was predicted using a
118 relationship linking the ITCs height (H_{ITC}) and the ITCs crown radius (R_{ITC}):

$$H_{ITC} = a \times R_{ITC}^b \quad (6)$$

119 Eqn. 6 was fitted using the function *nls* of the package *stats* of the R software (R Development
120 Core Team, 2008).

121 Each ITC was linked to the closest ground measured tree according to the Euclidean
122 distance between their position and their attributes (height, and crown radius):

$$D = \sqrt{(X_{ITC} - X_{FIELD})^2 + (Y_{ITC} - Y_{FIELD})^2} + \sqrt{(H_{ITC} - H_{FIELD})^2 + (R_{ITC} - R_{FIELD})^2} \quad (7)$$

123 After the linking, a visual inspection of the results on a GIS software was done and some
124 trees were manually realigned.

125 **Task 3: classification**

126 The classification of the tree species was done with a four step procedure: 1) data normalization;
127 2) feature selection; 3) classification; and 4) aggregation.

128 Data normalization was done to ensure that the pixel values were uniformly distributed
129 across all the crowns. Each pixel value was divided by the sum of the values of that pixel in all
130 the bands (Yu et al., 1999). In this way, we reduced the difference in radiance due to the fact that
131 the samples are distributed on multiple images.

132 The feature selection is necessary in order to select only the bands that are useful to
133 separate the analysed species. A feature selection method is made up of a searching strategy and
134 a separability criterion. In this study, the search strategy we used was the Sequential Forward
135 Floating Selection (SFFS) (Pudil, Novovičová & Kittler, 1994), and the separability criterion
136 was the Jeffries Matusita distance (Bruzzone, Roli & Serpico, 1995). These methods were used
137 successfully in many previous studies (Dalponte, Bruzzone & Gianelle, 2008, 2012, Dalponte et
138 al., 2009, 2013, 2014). The feature selection was applied on the training data, and we used the
139 function *varSelSFFS* in the R package *varSel*.

140 The classification was performed using a Support Vector Machine (SVM) classifier,
141 having as input the features selected at step 2 and the value of the CHM corresponding to each
142 ITC. We used the SVM implemented in the R package *kernelab*.

143 The predicted species labels of each pixel were aggregated at crown level with a majority
144 rule.

145 **4 Results**

146 **4.1 Task 1: segmentation**

147 The Jaccard score for the delineated ITCs over all the plots was 0.3402. The overall
148 confusion matrix (OCM) is showed in Table 1. To analyze the performance over each plot, the
149 confusion matrix for each plot was visualized as a bar chart (see Figure 1). The Jaccard score by
150 crown area is shown in Figure 2. Variability in the crown size did not change the Jaccard score,
151 showing that the method used is behaving in the same way for all crown sizes. The top-6 best
152 and worst delineations of the system are shown in Figure 3 and 4, respectively.

153 4.2 Task 2: alignment

154 All the test ITCs were aligned with the respective ground measured tree.

155 4.3 Task 3: classification

156 In Table 2 a summary of the overall performances is provided. Performance metrics at the
157 class level are also shown: accuracy and specificity (Figure 5), F1 score (Figure 6), precision
158 (Figure 7), and recall (Figure 8). From the overall performances, it is clear that the classification
159 method used was effective, as all the performance metrics are quite good. Looking at the results
160 per class (Table 3) it can be seen that for some classes the performance metrics are really low
161 (e.g. ACRU), while others are really good (e.g. PIPA).

162 5 Discussion

163 Team FEM ranked first for Task 1. As explained in the methods, we chose to segment a
164 hyperspectral band instead of the LiDAR point cloud. This choice was motivated by the fact that
165 looking at the training ITCs provided by the organizers, the hyperspectral data seemed more
166 suitable for this task. The comparison of results across teams showed that the FEM approach
167 outperforms the other approaches in the delineation of the small trees, while it was less efficient
168 for the large trees. This is due to the fact that we decided to use a small moving window (3x3).
169 The use of a variable size moving window, like the one that is implemented for LiDAR data in
170 the *itcSegment* library and used in (Dalponte et al., 2018), would have probably improved the
171 final results. The segmentation method used was compared in a previous study with three
172 segmentation methods based on LiDAR data (Dalponte et al., 2015b) and it was shown that this
173 method outperformed the LiDAR based methods on the delineation of broadleaf trees. This fact
174 can also explain the very good performances of team FEM delineations in the NEON NIST data

175 science evaluation challenge because in the study area species were mainly broadleaf or pine
176 trees. The crown shape of pine trees is quite close to the ones of many broadleaf trees.

177 In Task 2 FEM team ranked again in the first place with all the trees correctly aligned.
178 Surely the choice to consider not only the position, but also the tree characteristics (i.e. height,
179 and crown radius) was the winning choice. Moreover, after the automatic matching a visual
180 inspection of the results helped make the final improvements, as two trees were reassigned after
181 this inspection. A visual inspection of the alignment is not doable over large datasets, even if, in
182 our experience, it is always suggested as it helps in finding macroscopic errors. As mentioned in
183 the introduction, the choice of alignment strategy can depend also on the type of data that can be
184 used for this purpose. The fact that each crown delineation paper uses a different alignment
185 method specific to the dataset is not a good approach. Indeed, there is the need to have a
186 reference alignment method that could be used in every crown segmentation paper that allows a
187 fair comparison among delineation results.

188 The classification task (Task 3) had the most participants and team FEM ranked at the
189 second place. In this case the architecture that we used was effective, even if the results showed a
190 serious problem in distinguishing minority species. This is a limitation of many other works
191 proposed in the literature as many classifiers tend to give priority to highly represented species.
192 A better balance in the training set could have achieved higher classification accuracies. As an
193 example, the use of a semi-supervised classification approach (Dalponte et al., 2015a) could have
194 improved the classification of minority species. Moreover, a feature selection specifically
195 devoted on the identification of the best features to separate minority species cloud have helped.

196 6 Conclusions

197 In this paper the results of team FEM of the NEON NIST data science evaluation challenge were
198 presented. The methods applied were effective as team FEM ranked first in Task 1 and 2, and
199 second in Task 3. The delineation method proposed was based on hyperspectral images, showing
200 that LiDAR data are not always the best data source for ITC delineation. The alignment strategy
201 was based on both location and tree characteristics, and this combination of different information
202 provided the added value to the perfect alignment of the crowns. The classification architecture
203 adopted was quite standard, and it failed to classify more rare species. As a future development,
204 it may be interesting to combine both hyperspectral and LiDAR information in the crown
205 segmentation, and to consider classifiers, like for example semi-supervised ones, that can
206 improve the classification of more rare species.

207 7 Acknowledgements

208 The National Ecological Observatory Network is a program sponsored by the National Science
209 Foundation and operated under cooperative agreement by Battelle Memorial Institute. This
210 material is based in part upon work supported by the National Science Foundation through the
211 NEON Program. The ECODSE competition was supported, in part, by a research grant from
212 NIST IAD Data Science Research Program to D.Z. Wang, E.P. White, and S. Bohlman, by the
213 Gordon and Betty Moore Foundation's Data-Driven Discovery Initiative through grant
214 GBMF4563 to E.P. White, and by an NSF Dimension of Biodiversity program grant (DEB-
215 1442280) to S. Bohlman. This work was partially supported by the HyperBio project (project
216 #244599), which was financed by the BIONÆR program of the Research Council of Norway
217 and by TerraTec AS, Norway.

218 **8 References**

- 219 Bruzzone L., Roli F., Serpico SB. 1995. An extension to multiclass cases of the Jeffreys-
220 Matusita distance. *IEEE Transactions on Geoscience and Remote Sensing* 33:1318–1321.
- 221 Budei BC., St-Onge B., Hopkinson C., Audet F-A. 2017. Identifying the genus or species of
222 individual trees using a three-wavelength airborne lidar system. *Remote Sensing of*
223 *Environment*:0–1. DOI: 10.1016/j.rse.2017.09.037.
- 224 Clark ML., Roberts DA., Clark DB. 2005. Hyperspectral discrimination of tropical rain forest
225 tree species at leaf to crown scales. *Remote Sensing of Environment* 96:375–398. DOI:
226 10.1016/j.rse.2005.03.009.
- 227 Dalponte M. 2016. Package ‘ itcSegment .’
- 228 Dalponte M., Bruzzone L., Gianelle D. 2008. Fusion of Hyperspectral and LIDAR Remote
229 Sensing Data for Classification of Complex Forest Areas. *IEEE Transactions on*
230 *Geoscience and Remote Sensing* 46:1416–1427. DOI: 10.1109/TGRS.2008.916480.
- 231 Dalponte M., Bruzzone L., Gianelle D. 2012. Tree species classification in the Southern Alps
232 based on the fusion of very high geometrical resolution multispectral/hyperspectral images
233 and LiDAR data. *Remote Sensing of Environment* 123:258–270. DOI:
234 10.1016/j.rse.2012.03.013.
- 235 Dalponte M., Bruzzone L., Vescovo L., Gianelle D. 2009. The role of spectral resolution and
236 classifier complexity in the analysis of hyperspectral images of forest areas. *Remote Sensing*
237 *of Environment* 113:2345–2355. DOI: 10.1016/j.rse.2009.06.013.
- 238 Dalponte M., Ene LT., Marconcini M., Gobakken T., Næsset E. 2015a. Semi-supervised SVM

- 239 for individual tree crown species classification. *ISPRS Journal of Photogrammetry and*
240 *Remote Sensing* 110:77–87. DOI: 10.1016/j.isprsjprs.2015.10.010.
- 241 Dalponte M., Frizzera L., Ørka HO., Gobakken T., Næsset E., Gianelle D. 2018. Predicting stem
242 diameters and aboveground biomass of individual trees using remote sensing data.
243 *Ecological Indicators* 85:367–376. DOI: 10.1016/j.ecolind.2017.10.066.
- 244 Dalponte M., Ørka HO., Ene LT., Gobakken T., Næsset E. 2014. Tree crown delineation and tree
245 species classification in boreal forests using hyperspectral and ALS data. *Remote Sensing of*
246 *Environment* 140:306–317. DOI: 10.1016/j.rse.2013.09.006.
- 247 Dalponte M., Orka HO., Gobakken T., Gianelle D., Naeset E. 2013. Tree Species Classification
248 in Boreal Forests With Hyperspectral Data. *IEEE Transactions on Geoscience and Remote*
249 *Sensing* 51:2632–2645. DOI: 10.1109/TGRS.2012.2216272.
- 250 Dalponte M., Reyes F., Kandare K., Gianelle D. 2015b. Delineation of Individual Tree Crowns
251 from ALS and Hyperspectral data: a comparison among four methods. *European Journal of*
252 *Remote Sensing* 48:365–382. DOI: 10.5721/EuJRS20154821.
- 253 Duncanson LI., Dubayah RO., Cook BD., Rosette J., Parker G. 2015. The importance of spatial
254 detail: Assessing the utility of individual crown information and scaling approaches for
255 lidar-based biomass density estimation. *Remote Sensing of Environment* 168:102–112. DOI:
256 10.1016/j.rse.2015.06.021.
- 257 Ene L., Næsset E., Gobakken T. 2012. Single tree detection in heterogeneous boreal forests using
258 airborne laser scanning and area-based stem number estimates. *International Journal of*
259 *Remote Sensing* 33:5171–5193. DOI: 10.1080/01431161.2012.657363.

- 260 Eysn L., Hollaus M., Lindberg E., Berger F., Monnet J-M., Dalponte M., Kobal M., Pellegrini
261 M., Lingua E., Mongus D., Pfeifer N. 2015. A Benchmark of Lidar-Based Single Tree
262 Detection Methods Using Heterogeneous Forest Data from the Alpine Space. *Forests*
263 6:1721–1747. DOI: 10.3390/f6051721.
- 264 Fassnacht FE., Latifi H., Stere??czak K., Modzelewska A., Lefsky M., Waser LT., Straub C.,
265 Ghosh A. 2016. Review of studies on tree species classification from remotely sensed data.
266 *Remote Sensing of Environment* 186:64–87. DOI: 10.1016/j.rse.2016.08.013.
- 267 Ferraz A., Bretar F., Jacquemoud S., Gonçalves G., Pereira L., Tomé M., Soares P. 2012. 3-D
268 mapping of a multi-layered Mediterranean forest using ALS data. *Remote Sensing of*
269 *Environment* 121:210–223. DOI: 10.1016/j.rse.2012.01.020.
- 270 Hung C., Bryson M., Sukkarieh S. 2012. Multi-class predictive template for tree crown
271 detection. *ISPRS Journal of Photogrammetry and Remote Sensing* 68:170–183. DOI:
272 10.1016/j.isprsjprs.2012.01.009.
- 273 Ke Y., Quackenbush LJ. 2011. A review of methods for automatic individual tree-crown
274 detection and delineation from passive remote sensing. *International Journal of Remote*
275 *Sensing* 32:4725–4747. DOI: 10.1080/01431161.2010.494184.
- 276 Lee AC., Lucas RM. 2007. A LiDAR-derived canopy density model for tree stem and crown
277 mapping in Australian forests. *Remote Sensing of Environment* 111:493–518. DOI:
278 10.1016/j.rse.2007.04.018.
- 279 Marconi S., Graves SJ., Gong D., Nia MS., Bras M Le., Dorr J., Fontana P., Gearhart J.,
280 Greenberg C., Harris DJ., Arvind S., Nishant A., Prarabdh J., Rege S., Bohlman S., White

- 281 EP. 2018. A data science challenge for converting airborne remote sensing data into
282 ecological information. *PeerJ Preprints*.
- 283 Popescu SC., Wynne RH., Nelson RF. 2003. Measuring individual tree crown diameter with
284 lidar and assessing its influence on estimating forest volume and biomass. In: *Canadian*
285 *Journal of Remote Sensing*. 564–577. DOI: 10.5589/m03-027.
- 286 Pudil P., Novovičová J., Kittler J. 1994. Floating search methods in feature selection. *Pattern*
287 *Recognition Letters* 15:1119–1125. DOI: 10.1016/0167-8655(94)90127-9.
- 288 R Development Core Team. 2008. R: A Language and Environment for Statistical Computing. *R*
289 *Foundation for Statistical Computing* 1:ISBN 3-900051-07-0.
- 290 Real R., Vargas JM. 1996. The Probabilistic Basis of Jaccard's Index of Similarity. *Systematic*
291 *Biology* 45:380–385. DOI: 10.1093/sysbio/45.3.380.
- 292 Vauhkonen J., Ene L., Gupta S., Heinzl J., Holmgren J., Pitkänen J., Solberg S., Wang Y.,
293 Weinacker H., Hauglin KM., Lien V., Packalén P., Gobakken T., Koch B., Næsset E.,
294 Tokola T., Maltamo M. 2012. Comparative testing of single-tree detection algorithms under
295 different types of forest. *Forestry* 85:27–40. DOI: 10.1093/forestry/cpr051.
- 296 Yu B., Ostland IM., Gong P., Pu R. 1999. Penalized Discriminant Analysis of In Situ
297 Hyperspectral Data for Conifer Species Recognition. *IEEE Transactions on Geoscience and*
298 *Remote Sensing* 37:2569–2577.
- 299
- 300

301 **Tables captions**

302 Table 1. Task 1: overall confusion matrix. The values in the table are in square meters.

303 Table 2. Task 3: overall performances.

304 Table 3. Task 3: confusion matrix.

305

306 **Figures captions**

307

308 Figure 1. Task 1: plot level confusion matrix as a bar chart.

309 Figure 2. Task 1: Jaccard score versus crown area.

310 Figure 3. Task 1: the best 6 segmentations. Green annotations represent ground truth polygons,
311 and red annotations are predicted ones.

312 Figure 4. Task 1: the worst 6 segmentations. Green annotations represent ground truth polygons,
313 and red annotations are predicted ones.

314 Figure 5. Task 3: Accuracy and Specificity Scores (Per-Class).

315 Figure 6. Task 3: F1 Score (Per-Class).

316 Figure 7. Task 3: Precision (Per-Class).

317 Figure 8. Task 3: Recall (Per-Class).

318

Table 1 (on next page)

Task 1: overall confusion matrix. The values in the table are in square meters.

1

	Positive	Negative
True	2022.8	-
False	2416.6	1293.1

2

3

Table 2 (on next page)

Task 3: overall performances.

1

2

Performance metric	Value
Cross-entropy cost	0.8769
Rank-1 accuracy	0.8800
Classification accuracy	0.9809
Average F1 score	0.5933
Average Specificity	0.4129

3

4

5

Table 3 (on next page)

Task 3: confusion matrix.

1

2

	ACRU	LIST	OTHER	PIEL	PIPA	PITA	QUGE	QULA	QUNI
ACRU	0	0	0.	0	1.29	0	0	0	0
LIST	0	0.67	0	0.54	0	0	0	0	0
OTHER	0	0	0.74	0	0	0	0	0	0
PIEL	0	1.00	0.46	0	0	0	0	0.64	0
PIPA	0.58	0	0	0	79.14	0	0	0	0.57
PITA	0	0	0	0	2.90	0.85	0	0.50	0
QUGE	0	0	0	0	0	0	0.50	0	0
QULA	0	0	0	0	0	0	0	3.74	0
QUNI	0	0	0	0	1.39	0	0	0	19.32

3

4

5

Figure 1

Task 1: plot level confusion matrix as a bar chart.

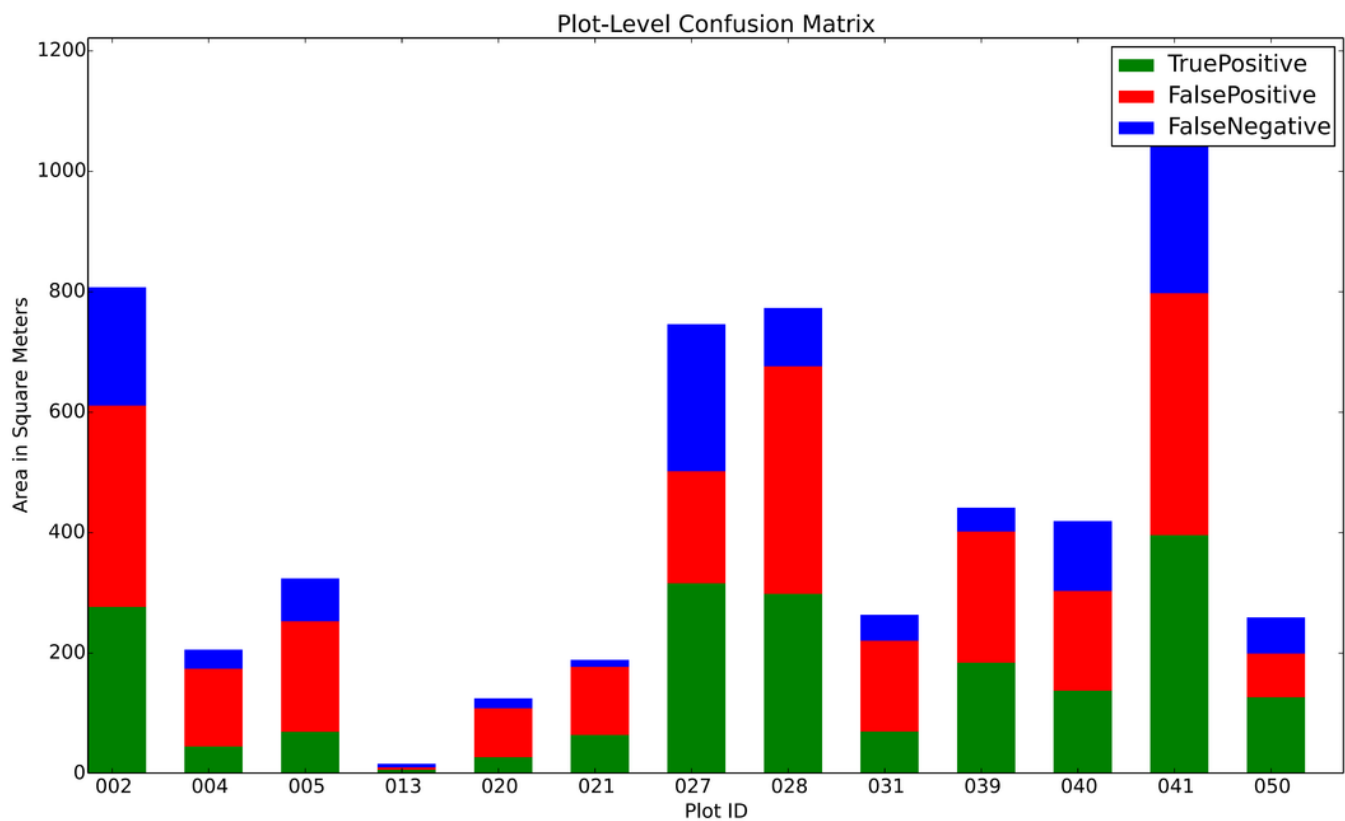


Figure 2

Task 1: Jaccard score versus crown area.

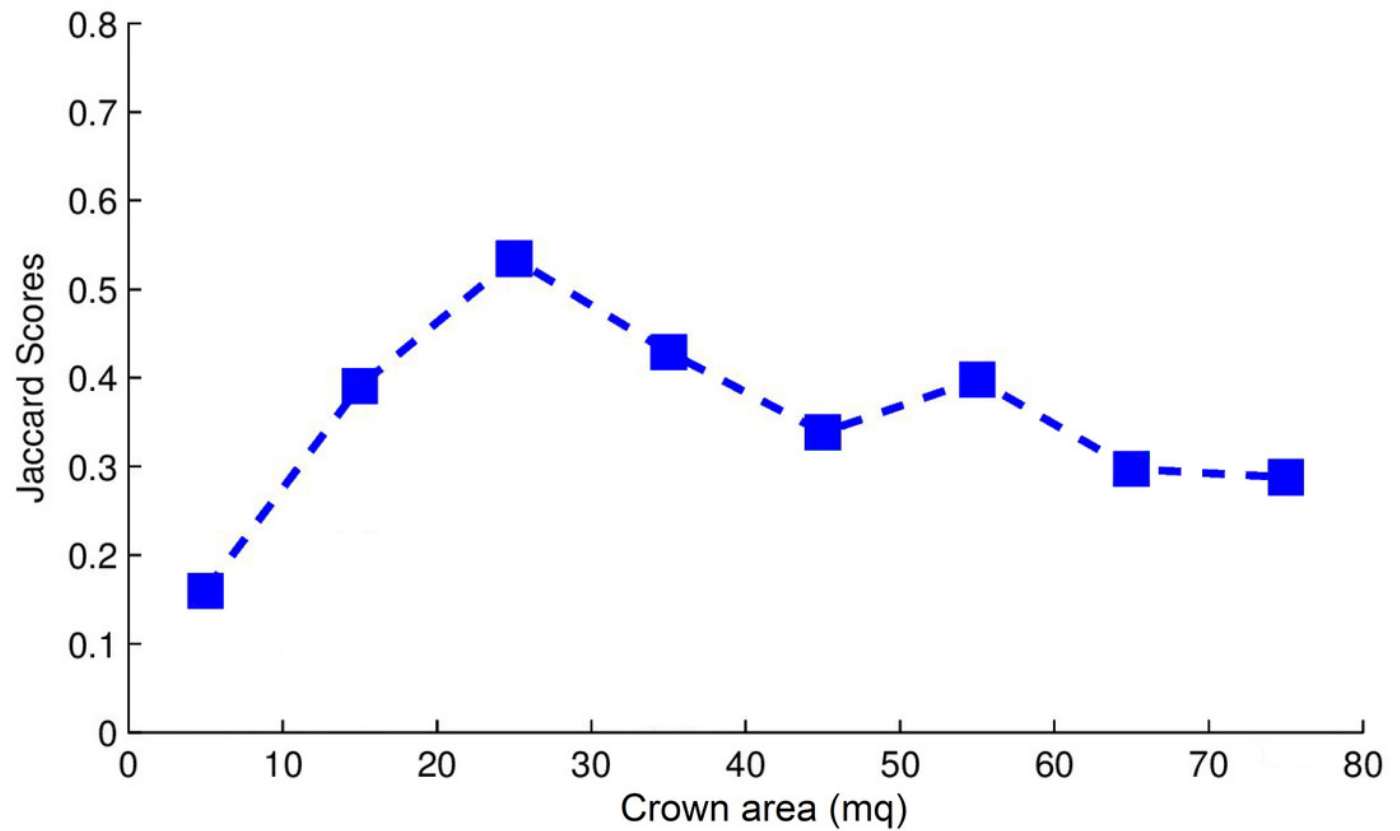


Figure 3

Task 1: the best 6 segmentations. Green annotations represent ground truth polygons, and red annotations are predicted ones.

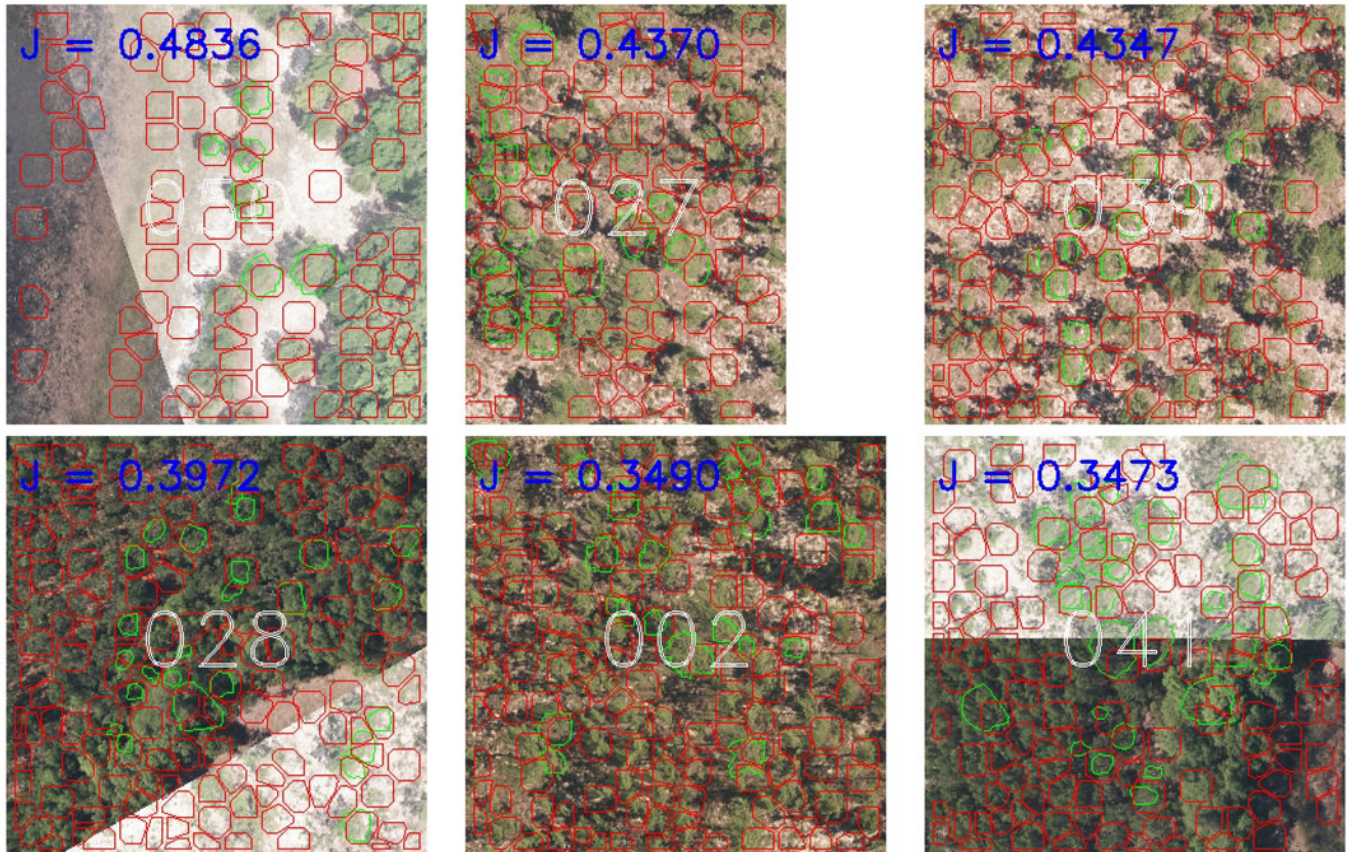


Figure 4

Task 1: the worst 6 segmentations. Green annotations represent ground truth polygons, and red annotations are predicted ones.

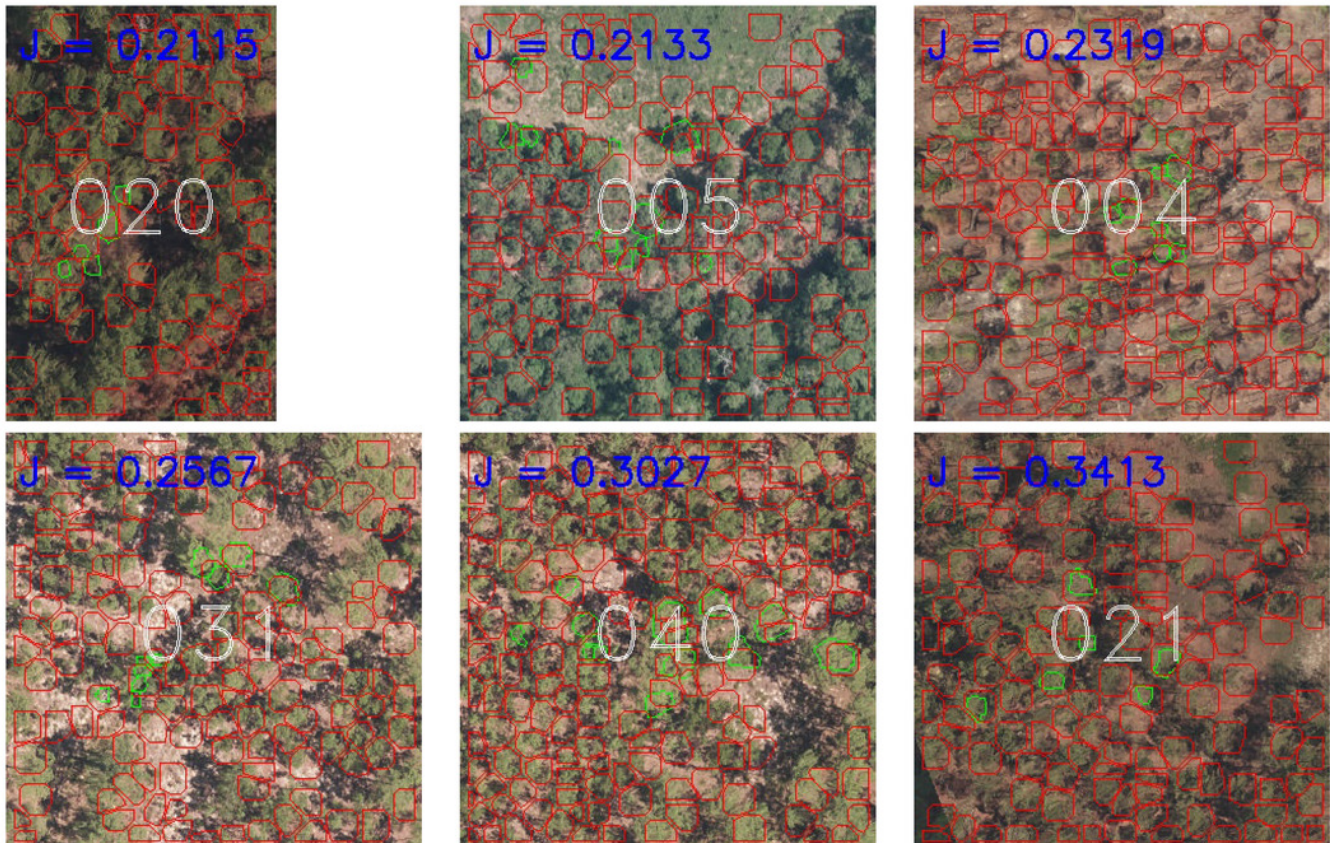


Figure 5

Task 3: Accuracy and Specificity Scores (Per-Class).

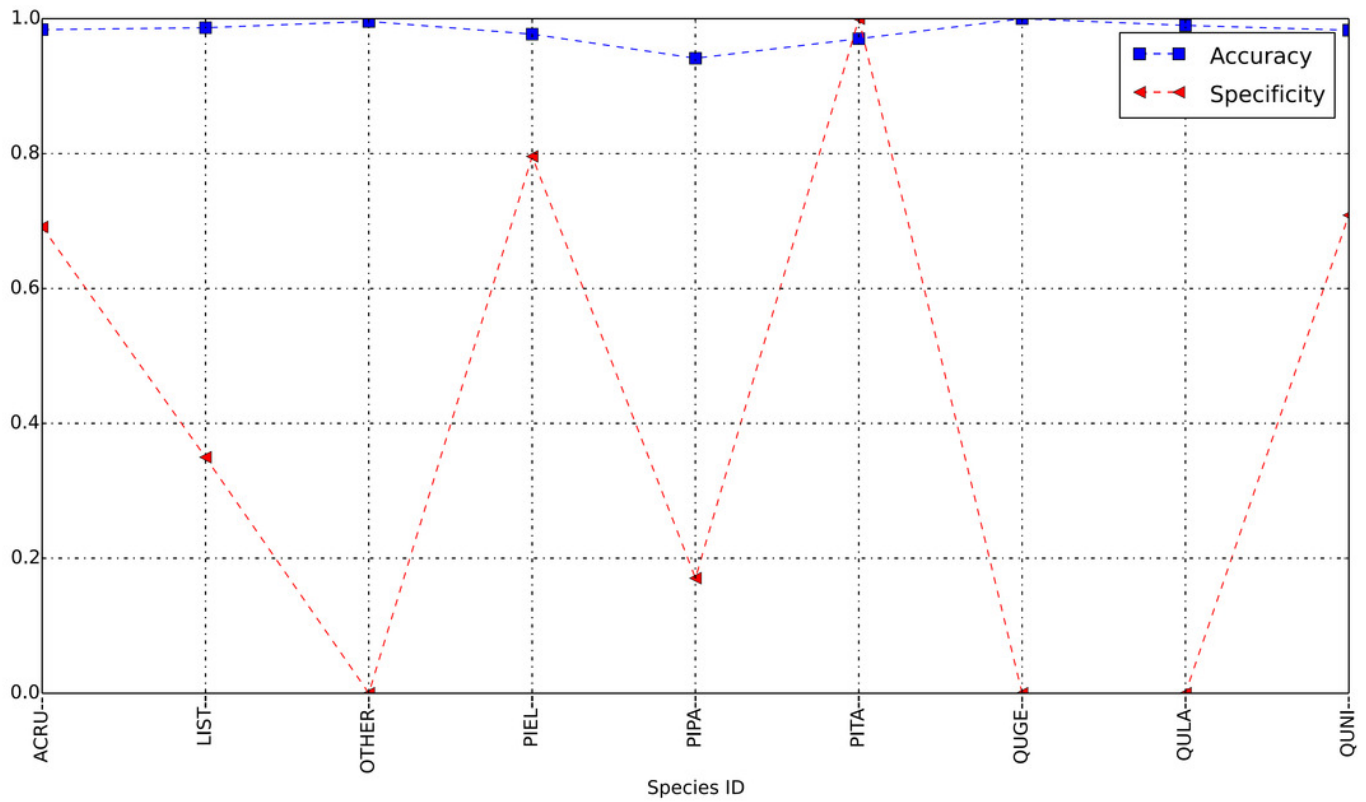


Figure 6

Task 3: F1 Score (Per-Class).

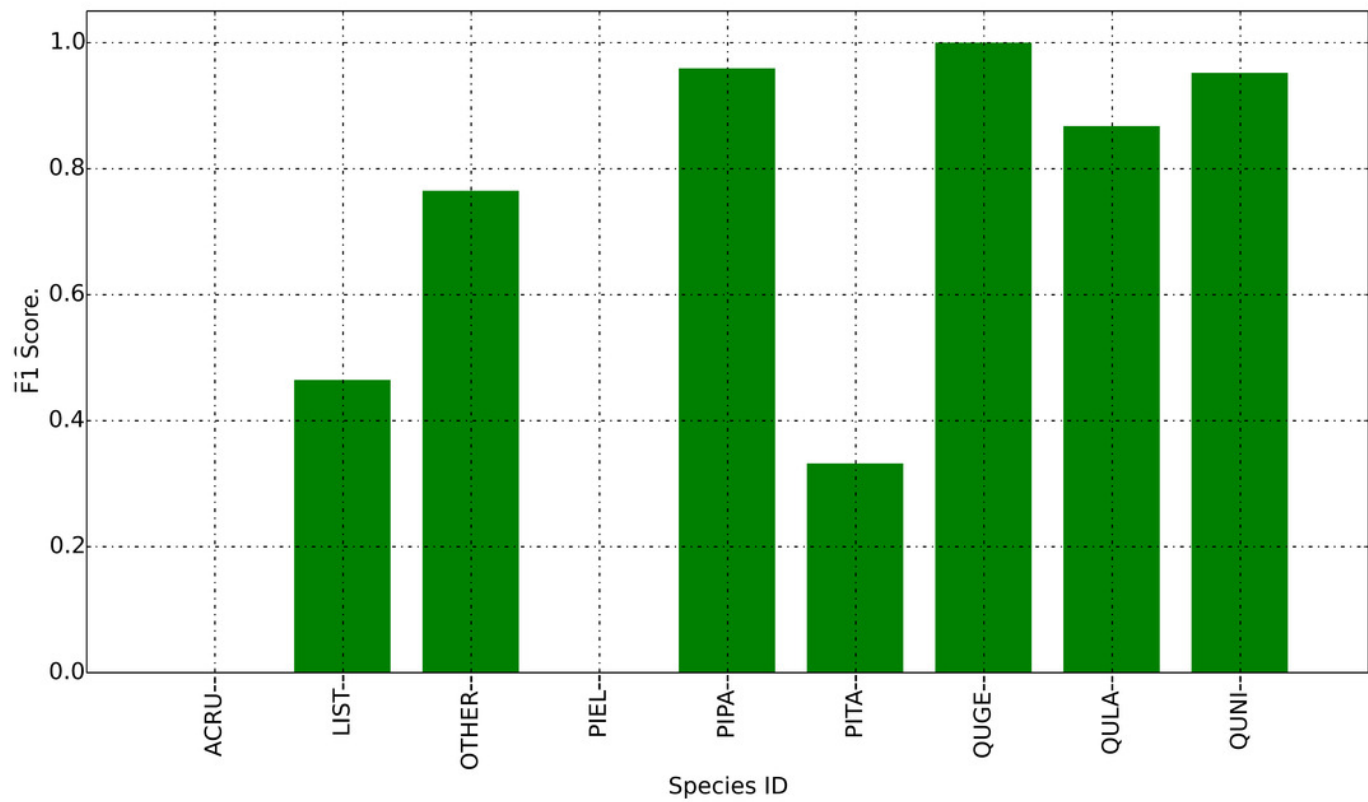


Figure 7

Task 3: Precision (Per-Class).

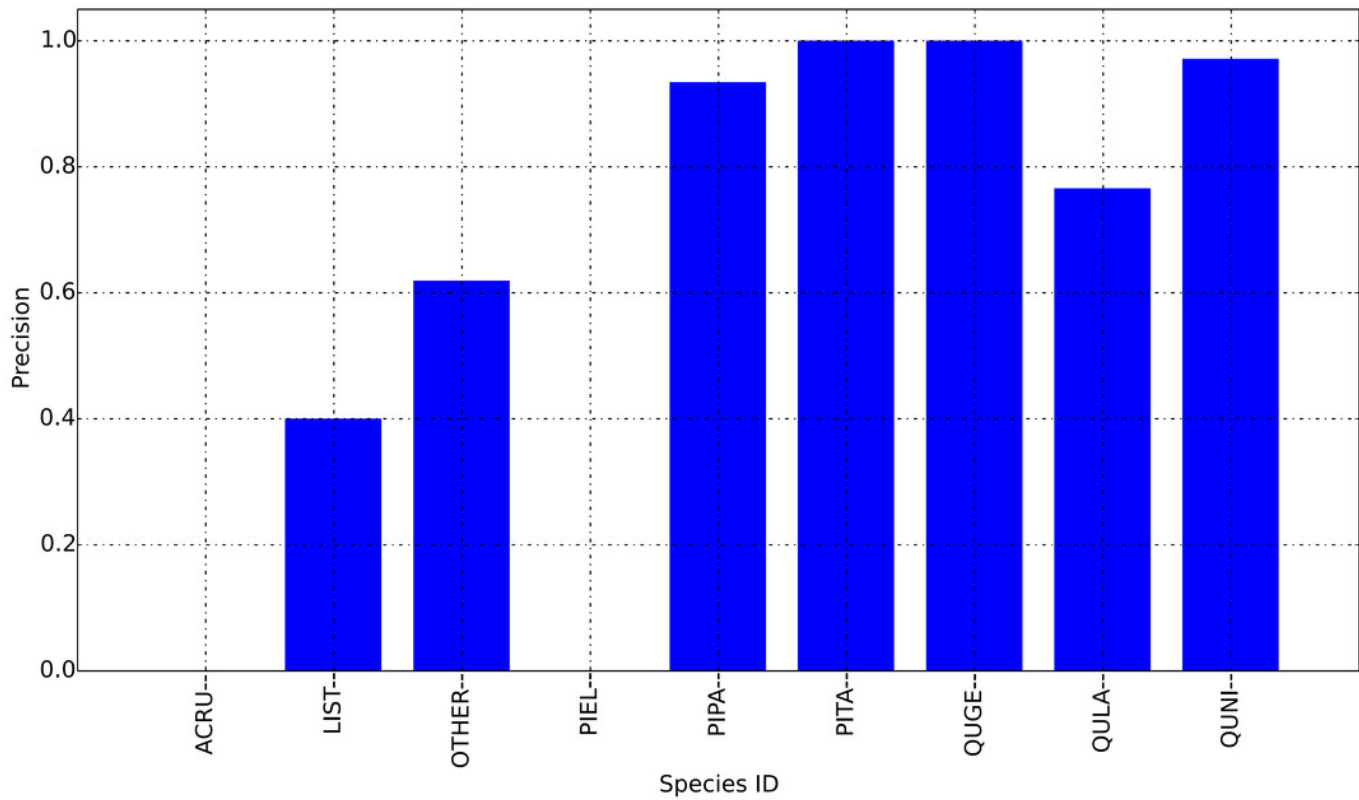


Figure 8

Task 3: Recall (Per-Class).

