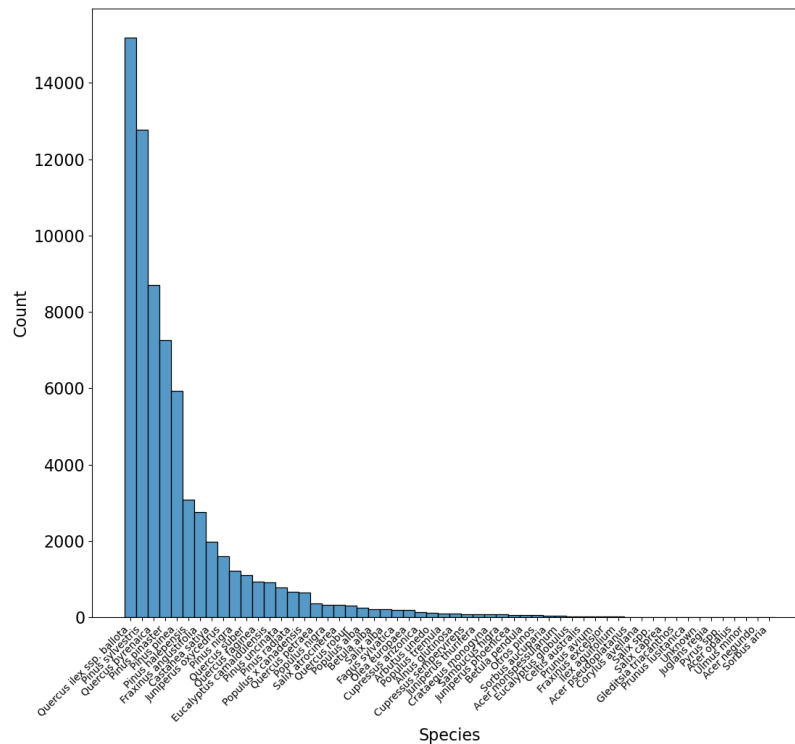# Mining Field Data for Tree Species Recognition at Scale
Supplementary Material

Dimitri Gominski[1], Daniel Ortiz-Gonzalo[1], Martin Brandt[1],
Maurice Mugabowindekwe[1], and Rasmus Fensholt[1]

University of Copenhagen, Denmark
{dg,gonzalo}@ign.ku.dk

1 Species Distribution

We indicate in Supplementary Figure 1 the complete distribution of tree species in the dataset (verified + unverified individuals). The distribution is typically long-tailed, with a few dominant species and numerous rare species.
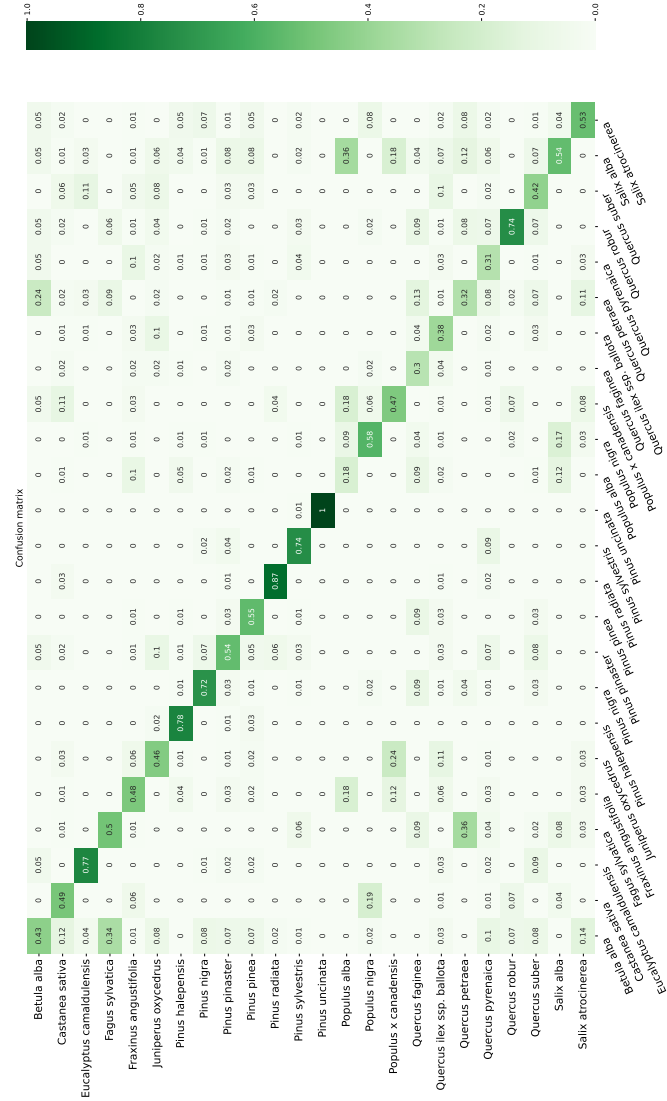

Supplementary Fig. 1: Full distribution of tree species.

2 Confusion Matrix

The confusion matrix of our best model (Supplementary Figure 2) trained on all labels + unlabeled examples with semi-supervised learning shows a relatively balanced performance accross species, with some rare species such as Salix alba having an accuracy > 50%.

3 Acknowledgments

Supplementary Fig. 2: Confusion matrix for our best model (all labels + 500k unlabeled images).