ForestInsights

Mapping New Zealand's forests through deep learning and data-centric AI

Melanie Palmer

Pearse G., Steer B., Camarretta N., Jayathunga S., Watt M.

SCION
FORESTS ▪ PRODUCTS ▪ INNOVATION

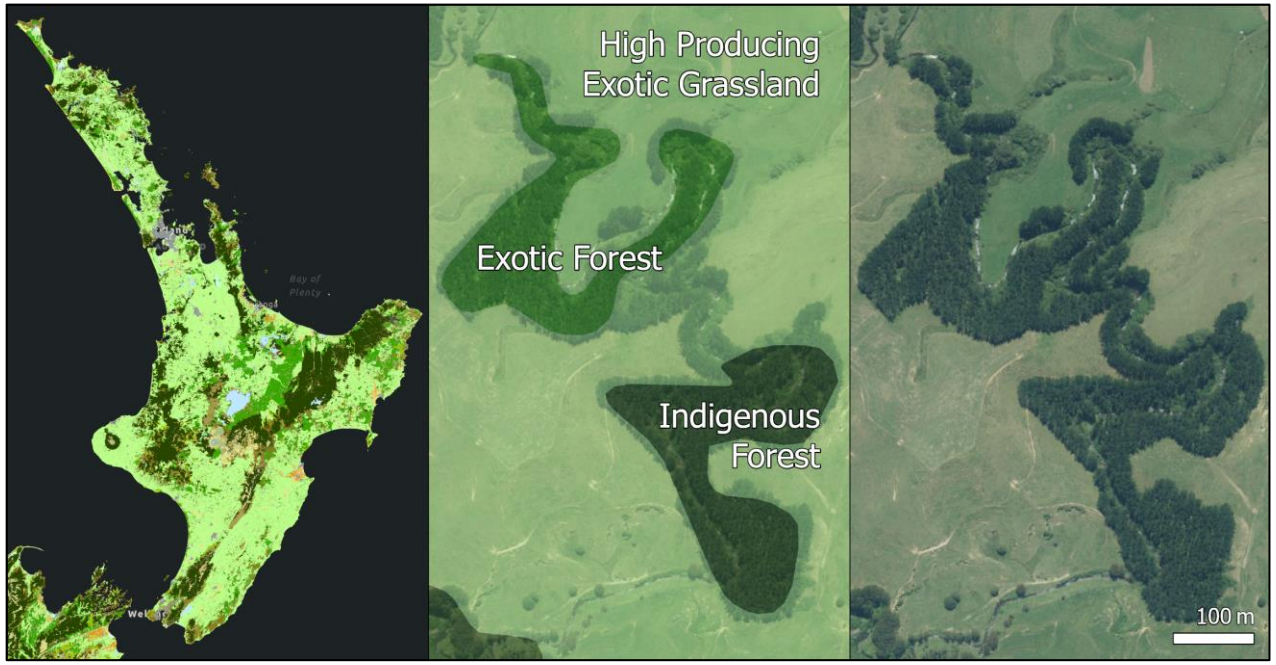
New Zealand’s goal: Net Zero 2050

- Net zero greenhouse gas emissions by 2050
- Forests are a key part of Emissions Trading Scheme
- 1.8 million ha commercial forest land
  - Around 90% radiata pine
- Significant afforestation required to meet emission reduction targets
- Growing importance of smaller growers and landowners
- Effective planning, forecasting, and policy-making relies on accurate data

SCION

Existing approaches and limitations

- National Exotic Forest Description (NEFD)
  - Survey-based
  - Limited coverage of small-scale forests
  - Lacks spatially explicit data
- Land Cover Database (LCDB)
  - Approx. 5-yearly update interval
  - Satellite-based (Sentinel-2)
  - Resolution is challenging for woodlots, seedlings, species ID



Landcover Database v5.0

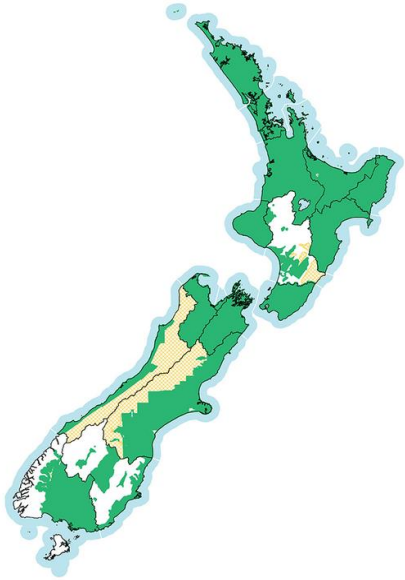
Size (ha)	Survey Frequency	Data Quality	Number of Owners	Net Stocked Area (000 ha)	Proportion of Total Area (%)
1,000+ ha	Annually	High	98	1,247	69.7
40-999 ha	Every two years	Medium	1,671	260	14.5
<40 ha	Infrequently	Low	>10,000	283	15.8

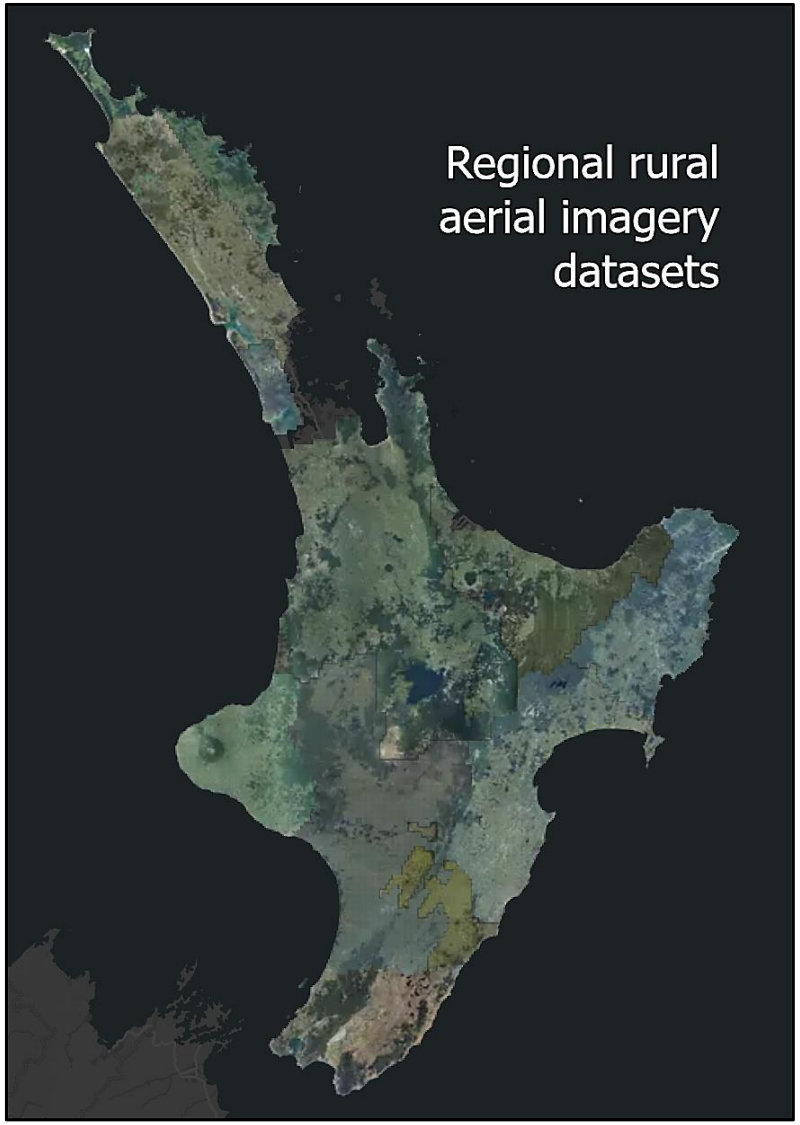
Summary of data collected for the National Exotic Forest Description by forest size. Adapted from 2023 NEFD Report

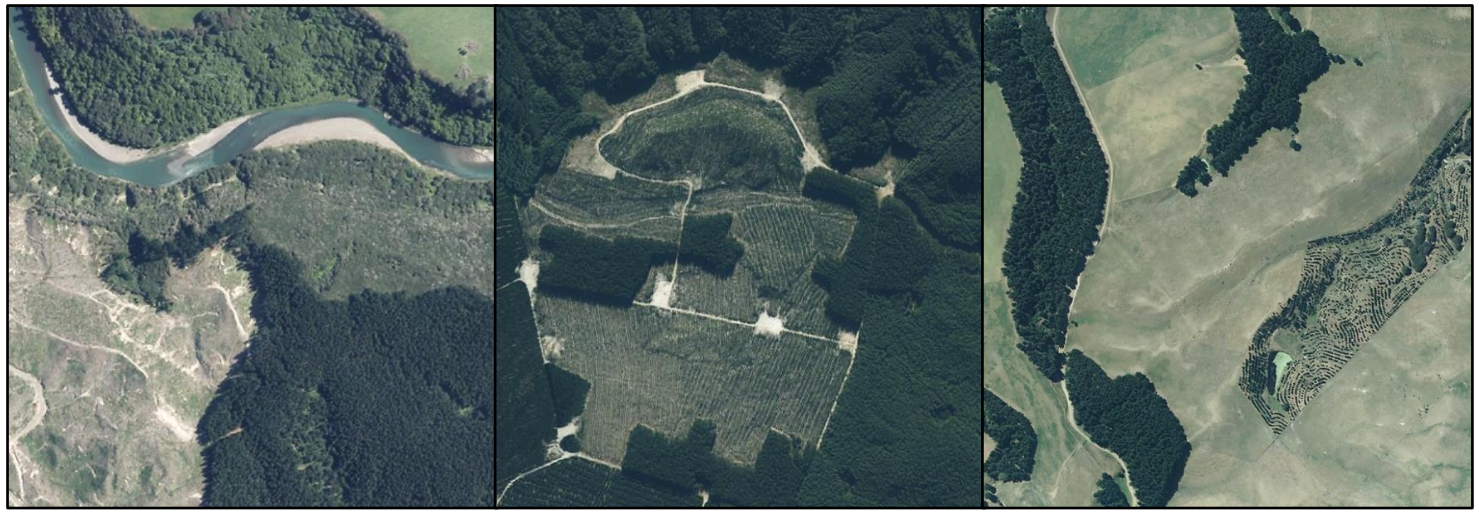
SCION

A new approach for mapping NZ's forests

- Regional fixed-wing aerial imagery
  - Routinely captured by councils
  - High resolution: 15 – 30 cm
- LINZ National Elevation Programme
  - Regional LiDAR
- **Deep learning** ***-- Lots of data!***





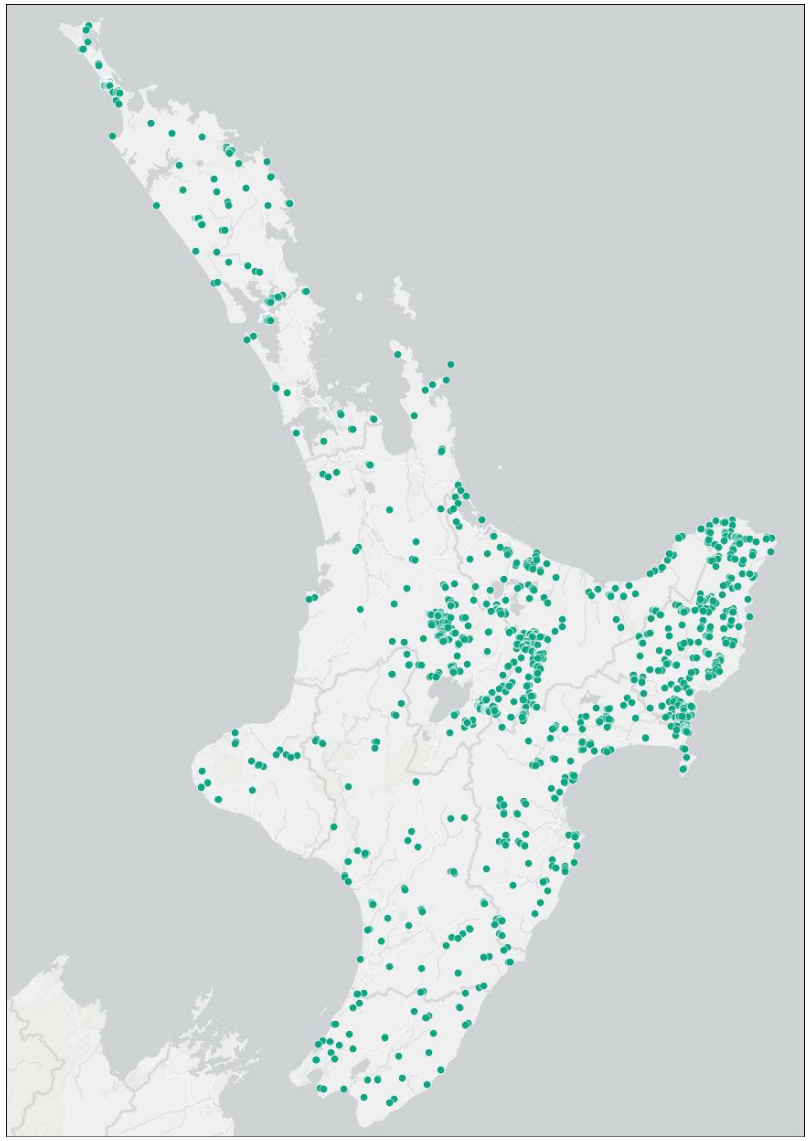


scion

High resolution exotic forest map using deep learning

Generating a training dataset

- Representative training samples gathered from across the North Island
- 1220 1:1k tiles
- 430 km^2 labelled area
- Large variety of forest examples
  - Age
  - Planting regimes
  - Different landscapes

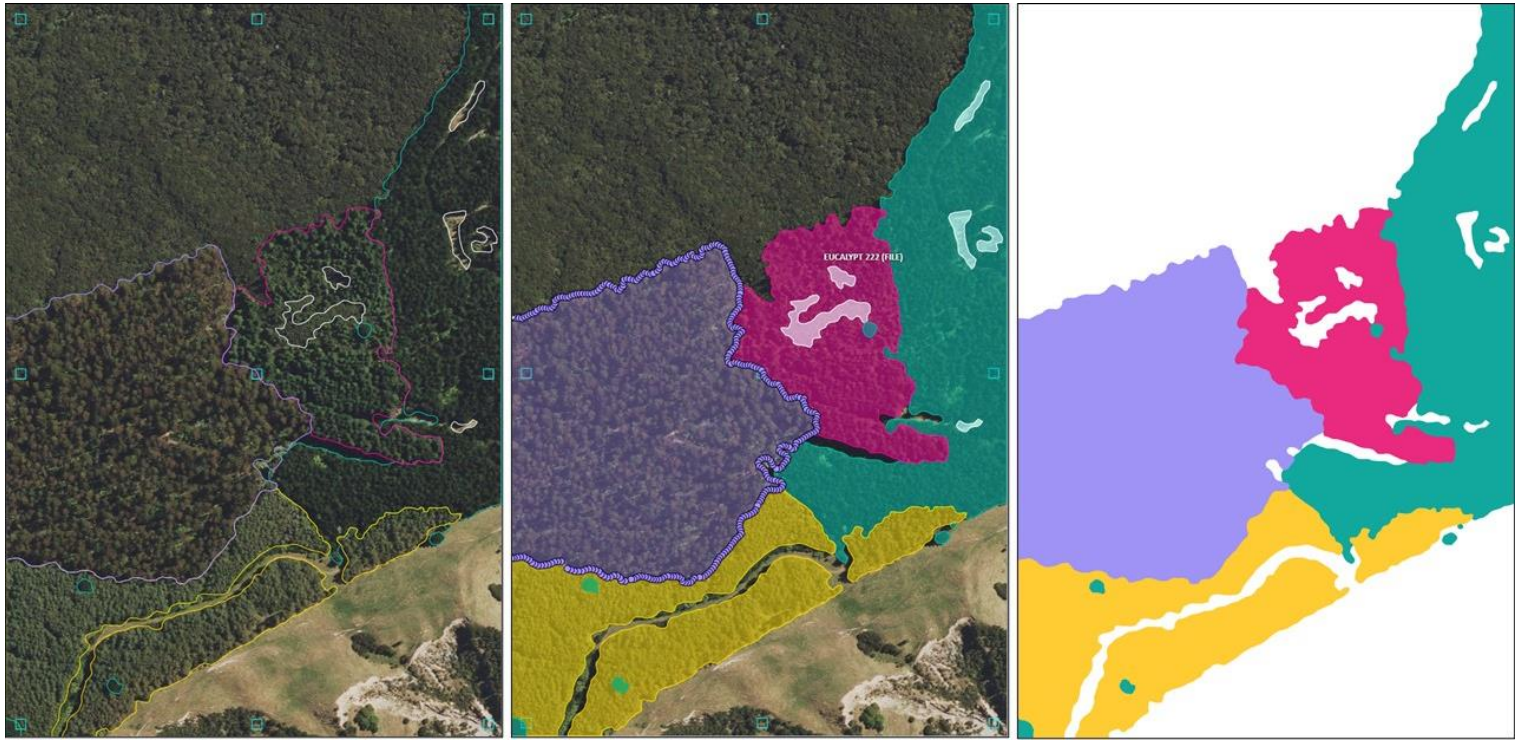


SCION

High resolution exotic forest map using deep learning

Generating a training dataset

- Hand-labelled using Computer Vision Annotation Tool (CVAT)
- Semantic segmentation
  - Each pixel is labelled
  - High level of detail
  - Exclude interior gaps



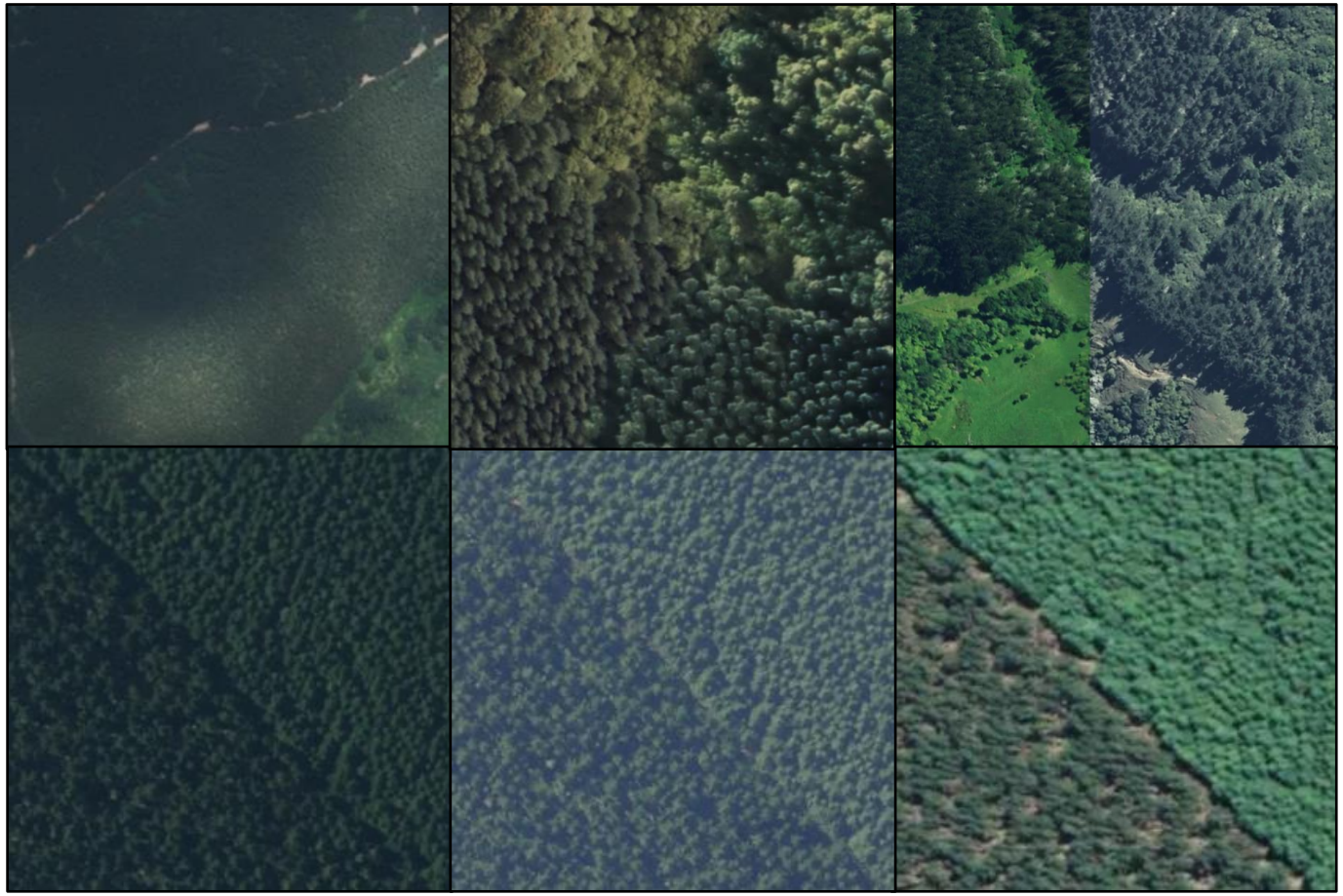
scion

High resolution exotic forest map using deep learning

Generating a training dataset

Challenges

- Diverse visual characteristics
  - Variation in camera sensors and capture conditions
  - Time of year or day
  - Clouds and shadows
- Determining species
  - Subtle differences
  - Can vary widely



scion

High resolution exotic forest map using deep learning

Generating a training dataset

Challenges

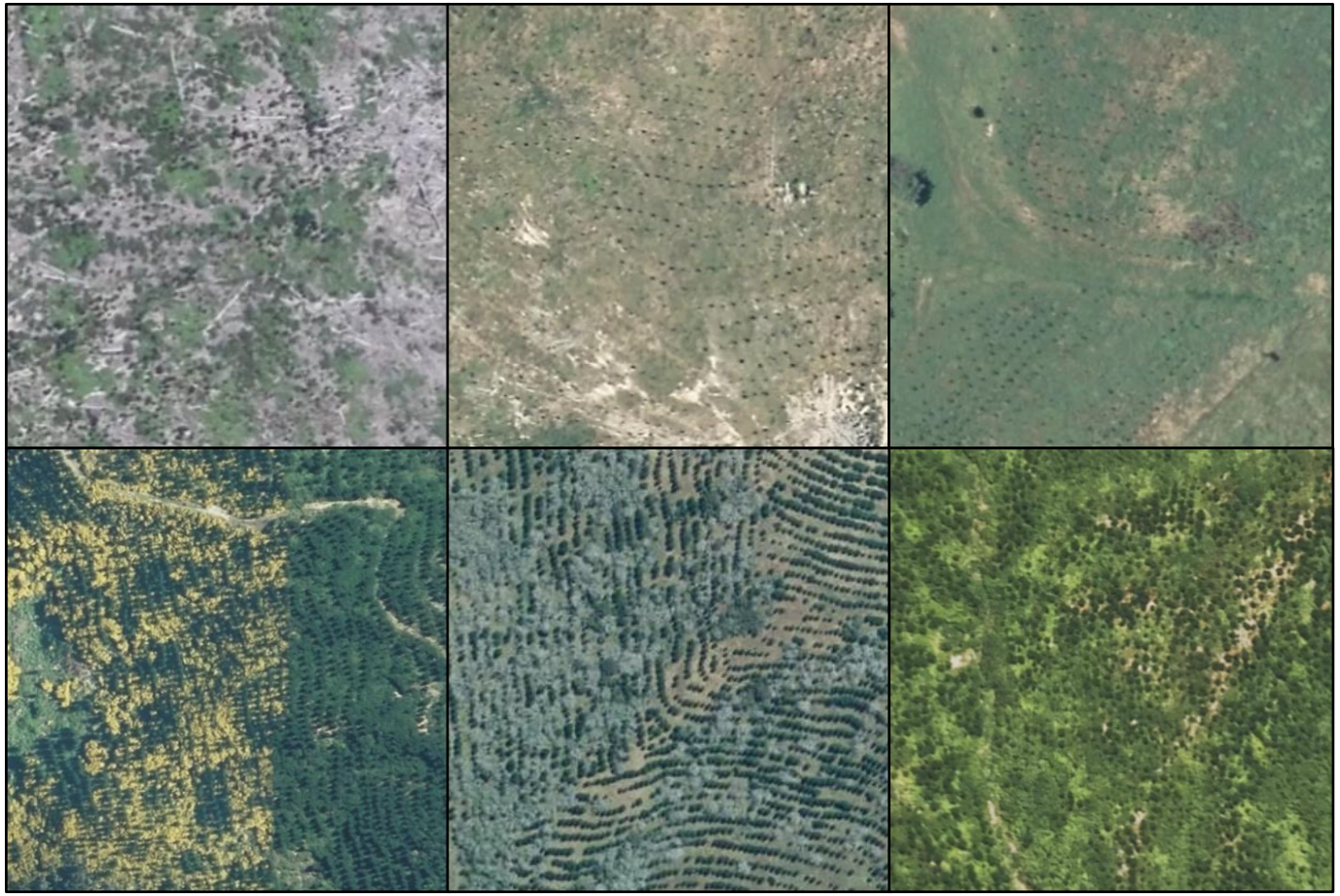
- How to define ‘forest’?
  - Dispersed
  - Wilding
  - Shelterbelts
  - Gardens

High resolution exotic forest map using deep learning

Generating a training dataset

Challenges

- Seedlings
  - Hard to see
  - Where to define boundary?
  - Weeds

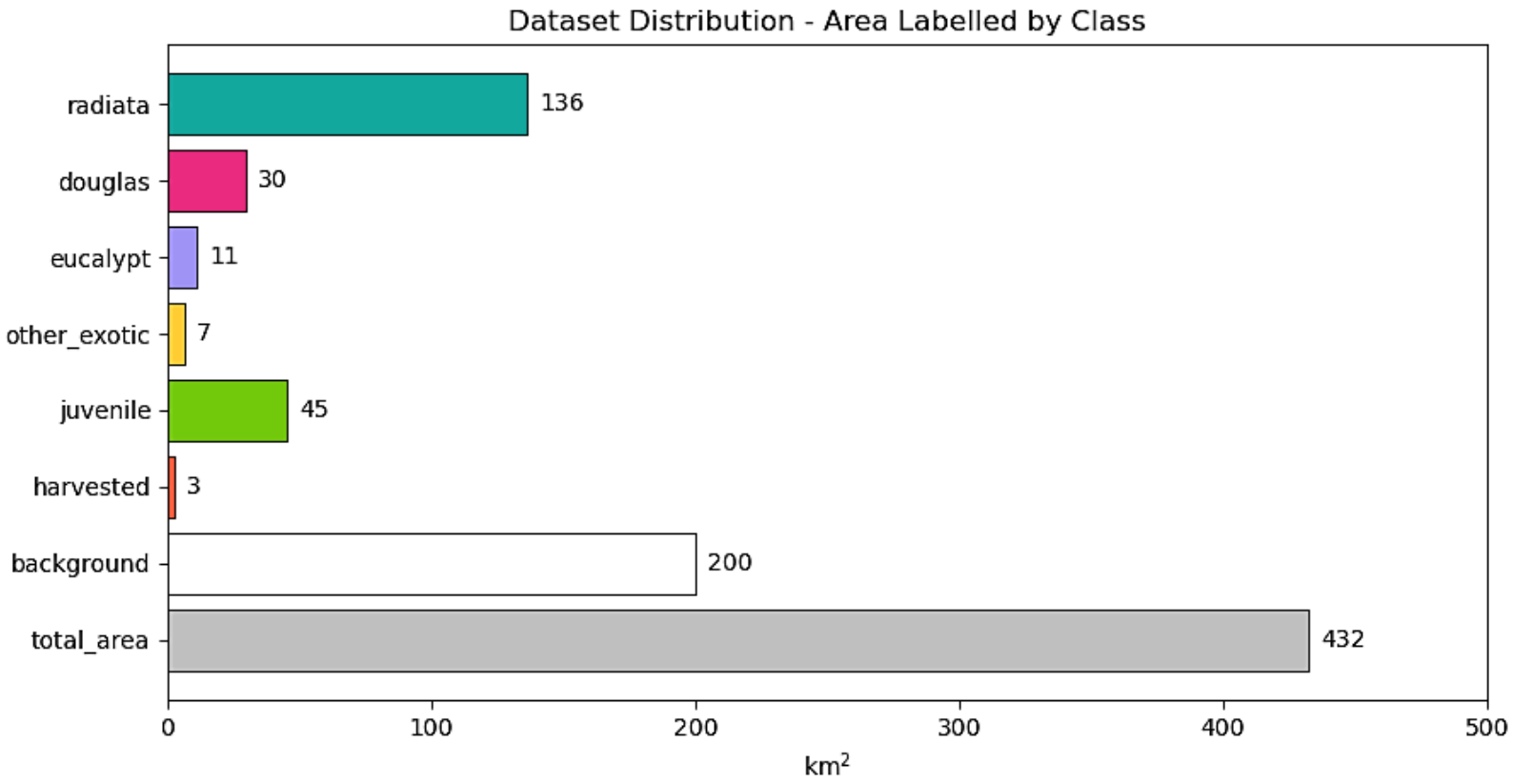


High resolution exotic forest map using deep learning

ForestInsights current dataset

Classes

- Radiata (mature)
- Juvenile (seedlings)
- Other species
  - Douglas fir
  - Eucalypt
  - Other exotic spp.
- Harvested
- Background



High resolution exotic forest map using deep learning

Training deep learning model

- Binary classification model for radiata detection

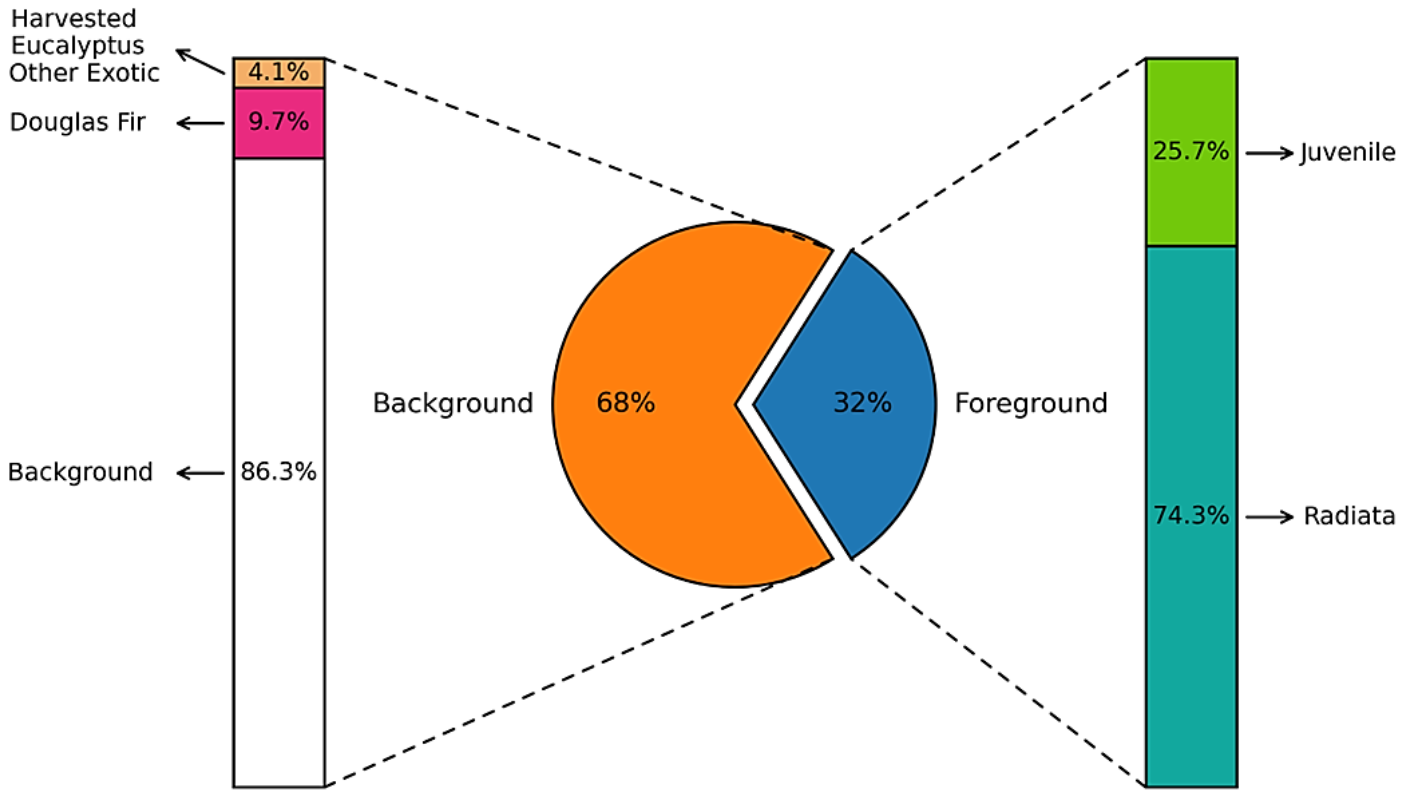
Classes

Foreground

- Radiata (mature)
- Juvenile (seedlings)

Background

- Other species
- Harvested
- Background

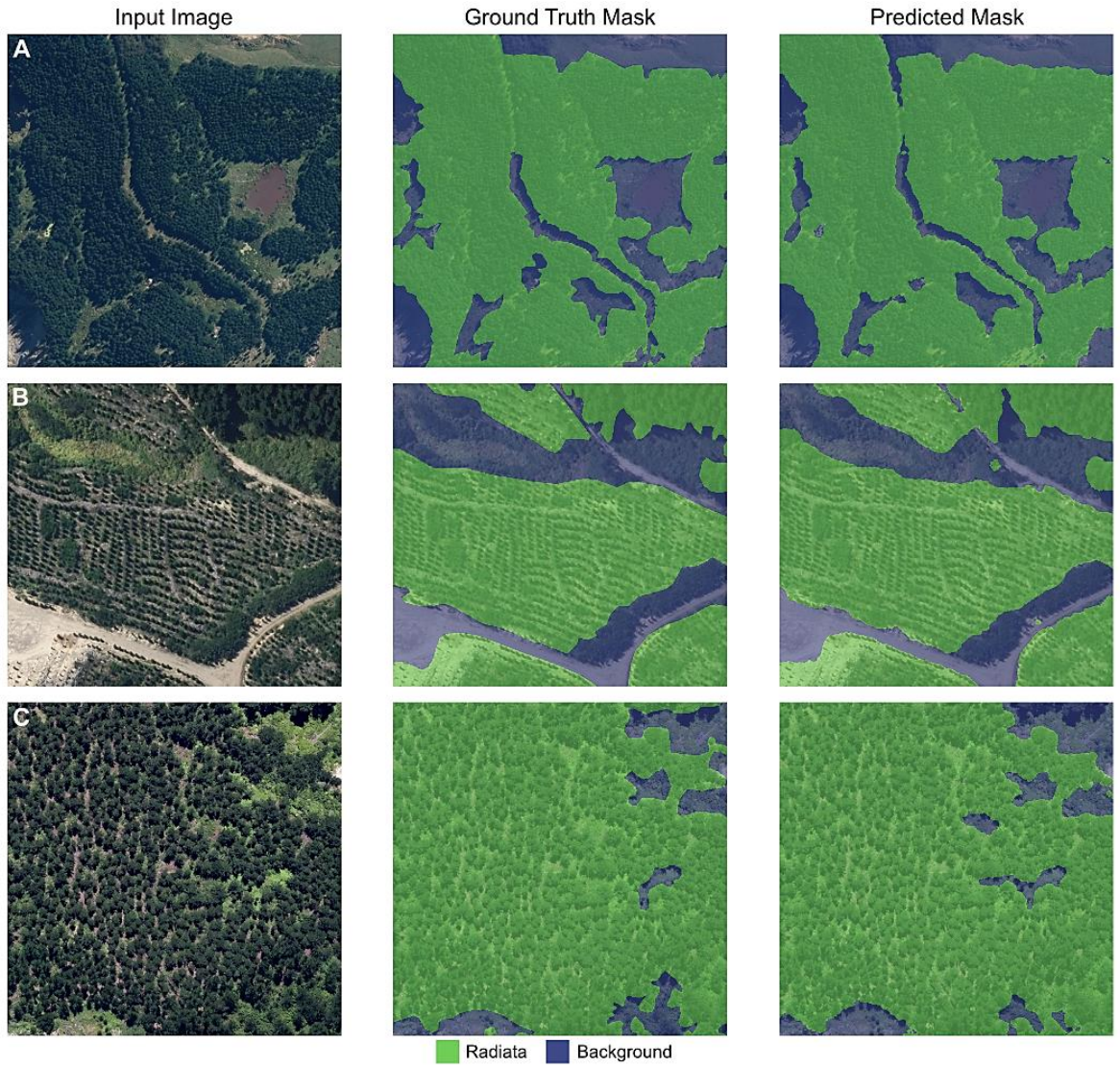


High resolution exotic forest map using deep learning

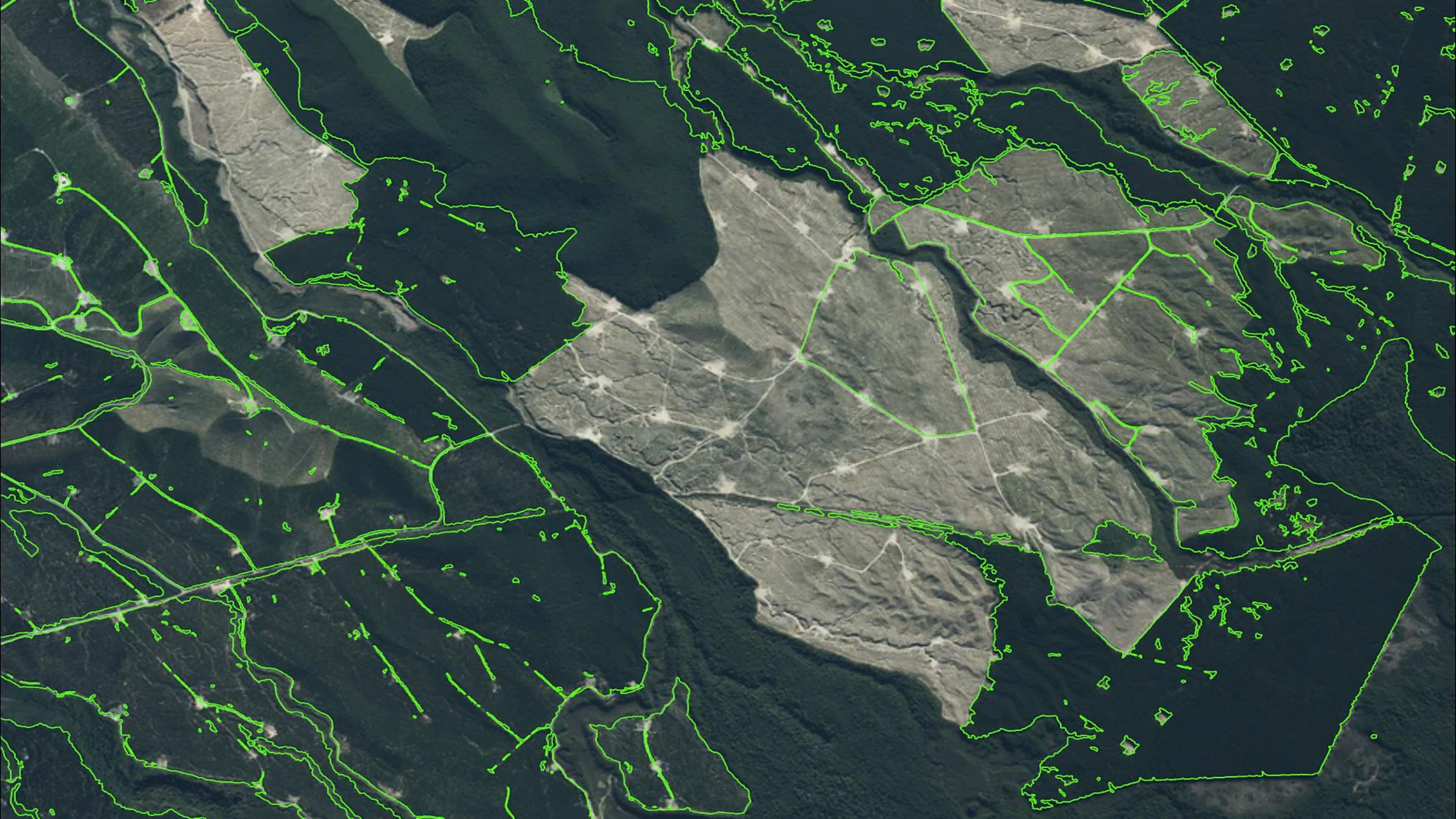
Training deep learning model

- DeepLabV3+ with ResNext-101 backbone implemented in PyTorch
- Well defined delineation of radiata across a range of contexts
- Level of detail frequently exceeded ground truth masks

Metric	Validation (15%)	Test (15%)
IoU	0.934	0.937
Accuracy	0.978	0.980
Precision	0.956	0.957
Recall	0.975	0.977
F1 Score	0.966	0.967



scion



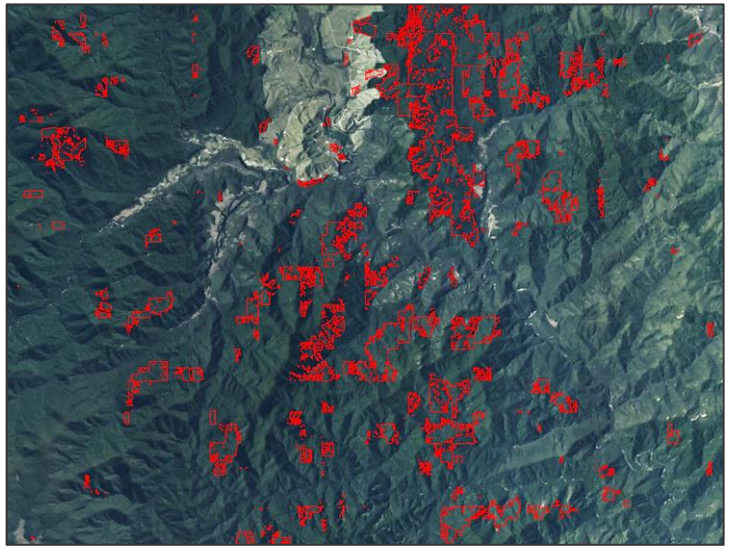
My model is doing great!

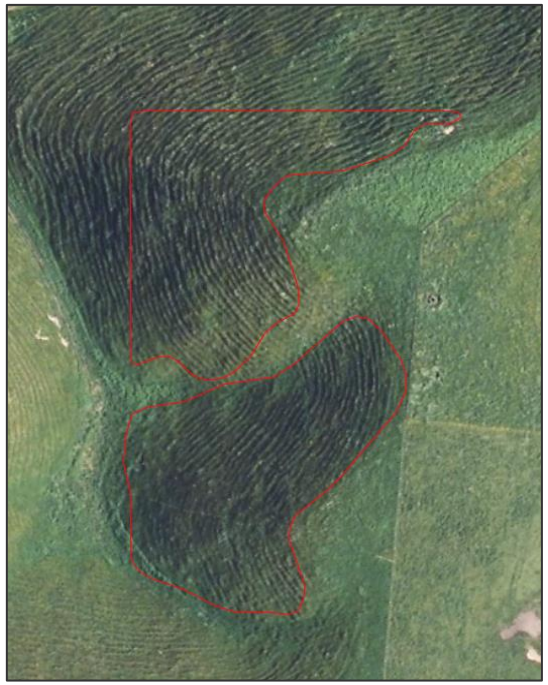
Wait… hang on…

Gartner hype curve

Model confusion

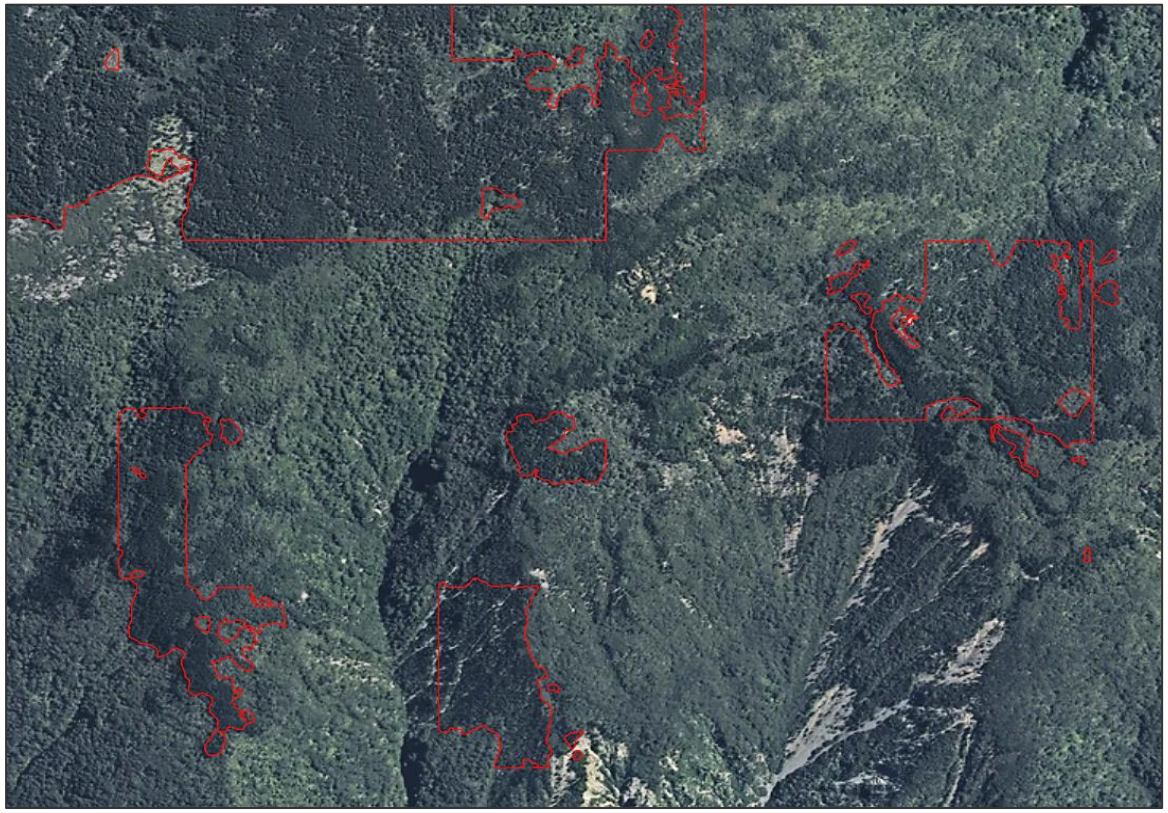
- False positives/negatives
- Unfamiliar/poor quality imagery conditions
- Unknown features





Model confusion

- Wildings
- Abandoned/unmanaged stands
- Dispersed planting

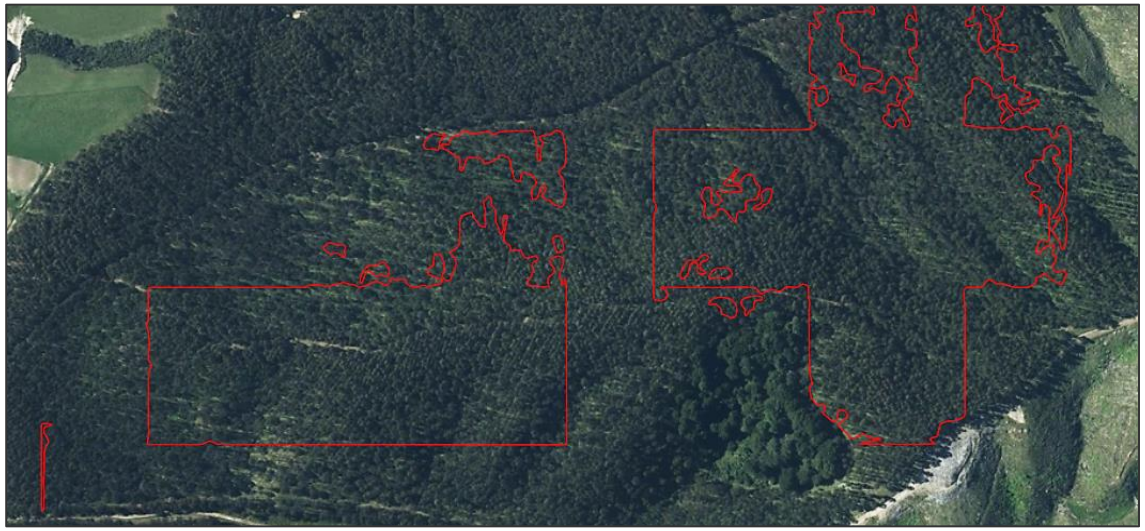


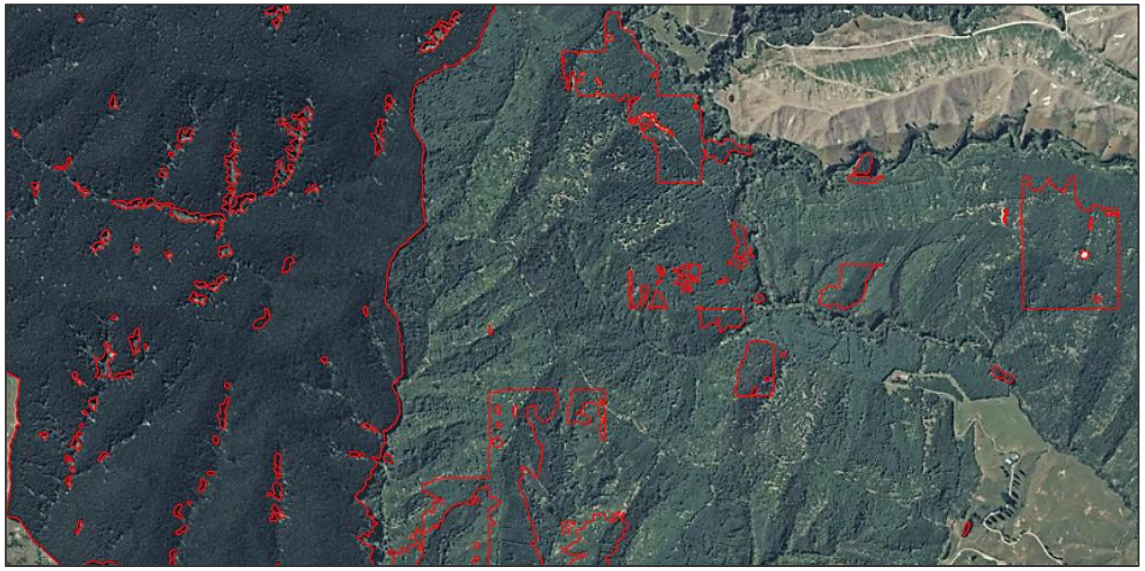
Model confusion

- Wildings
- Abandoned/unmanaged stands
- Dispersed planting

Model confusion

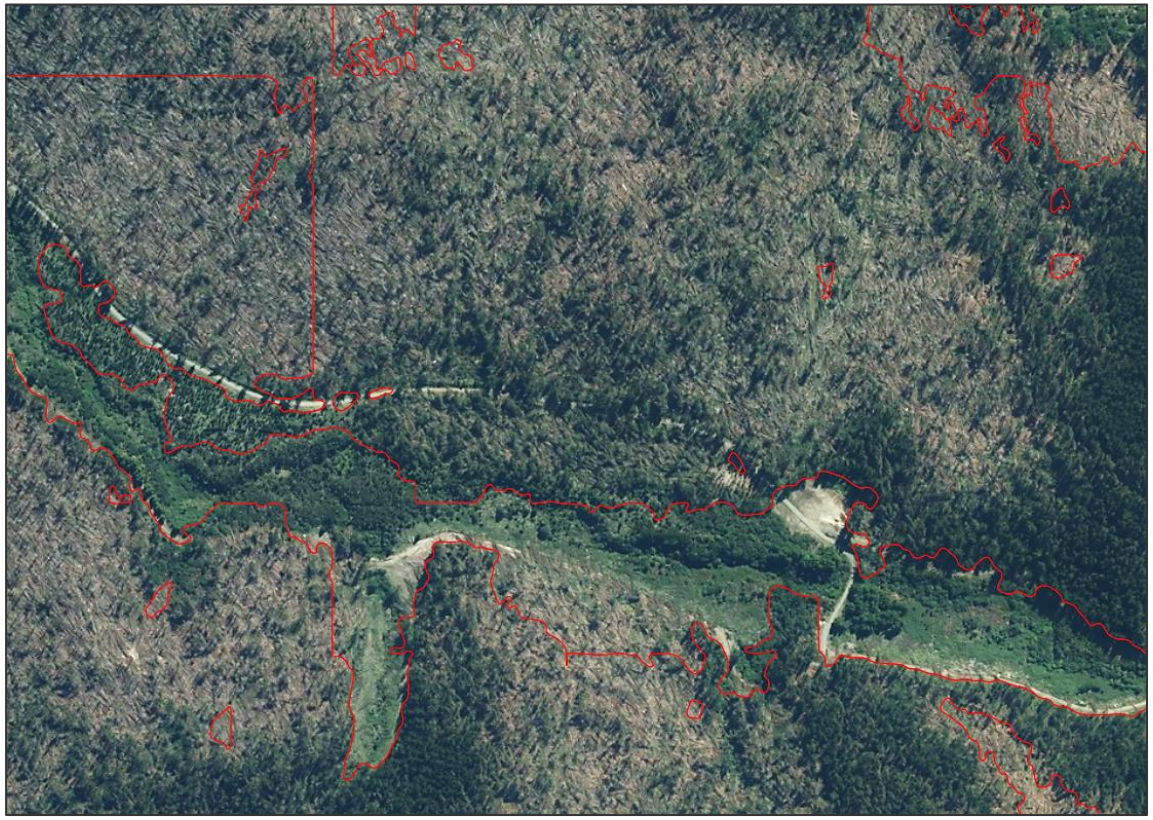
- Non-radiata species





Model confusion

- Cyclone Gabrielle
- Windthrow
- Damaged trees

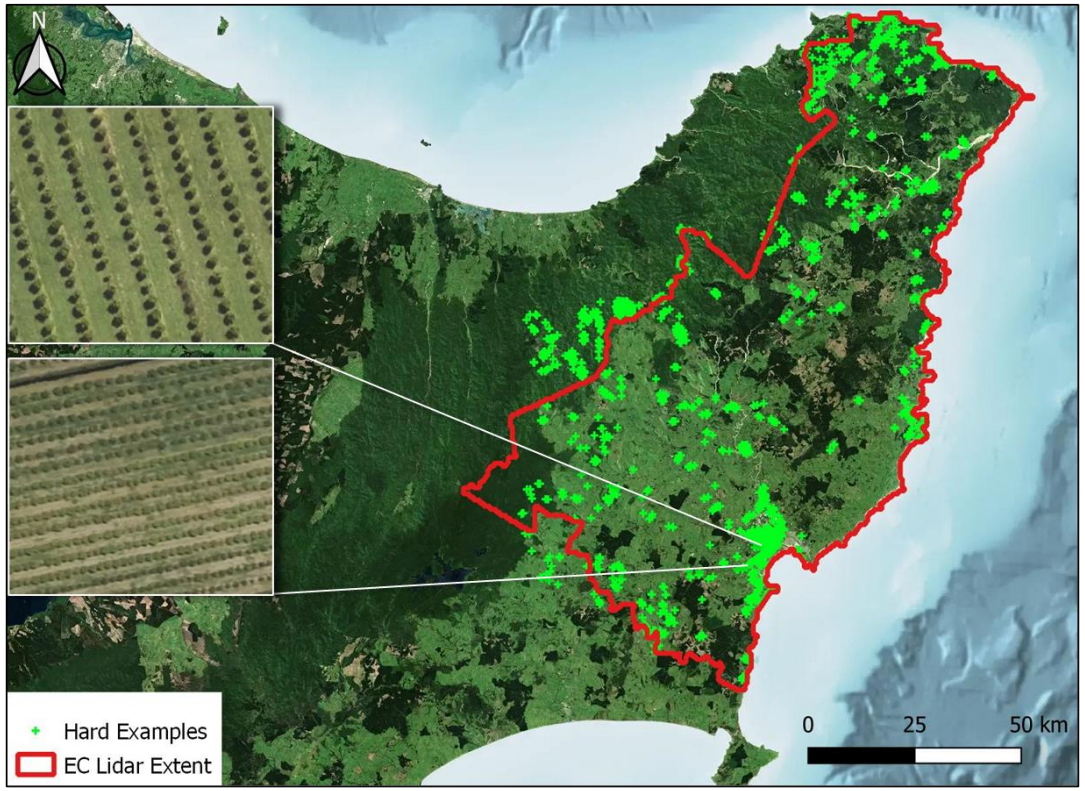


Data-centric AI

> *"The "dirty secret" of artificial intelligence is that getting the software to work well in the real world requires a large amount of high-quality data."*
>
> - Alexander Wang, Founder & CEO Scale AI in an interview with *Fortune*

Iterative dataset development

- Proof of concept → 5th iteration of the model
- Over double the size of initial dataset
- Targeted approach based on previous iteration performance
  - Inclusion of 'hard tiles'
  - Hard negative mining
- Dataset refinement



scion

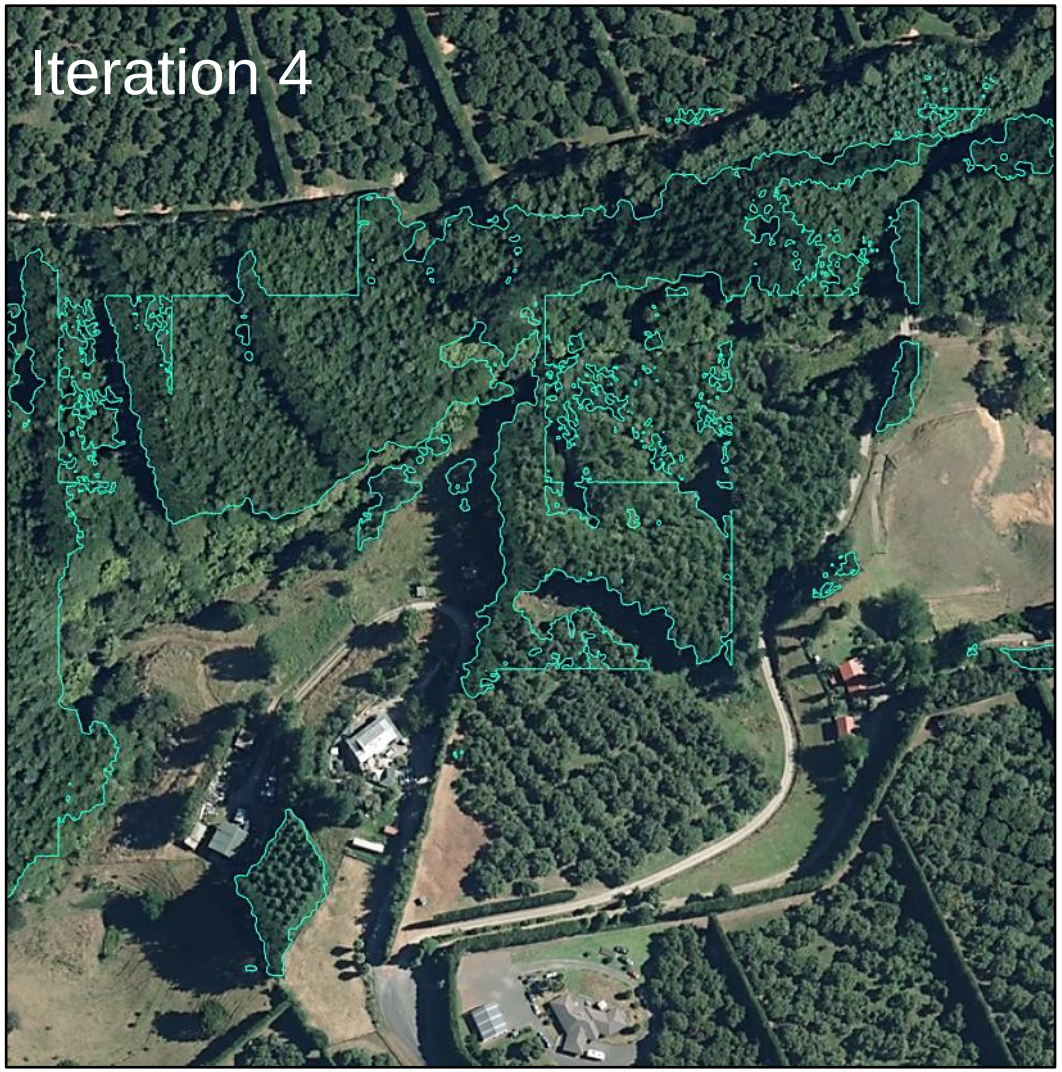
Data-centric AI

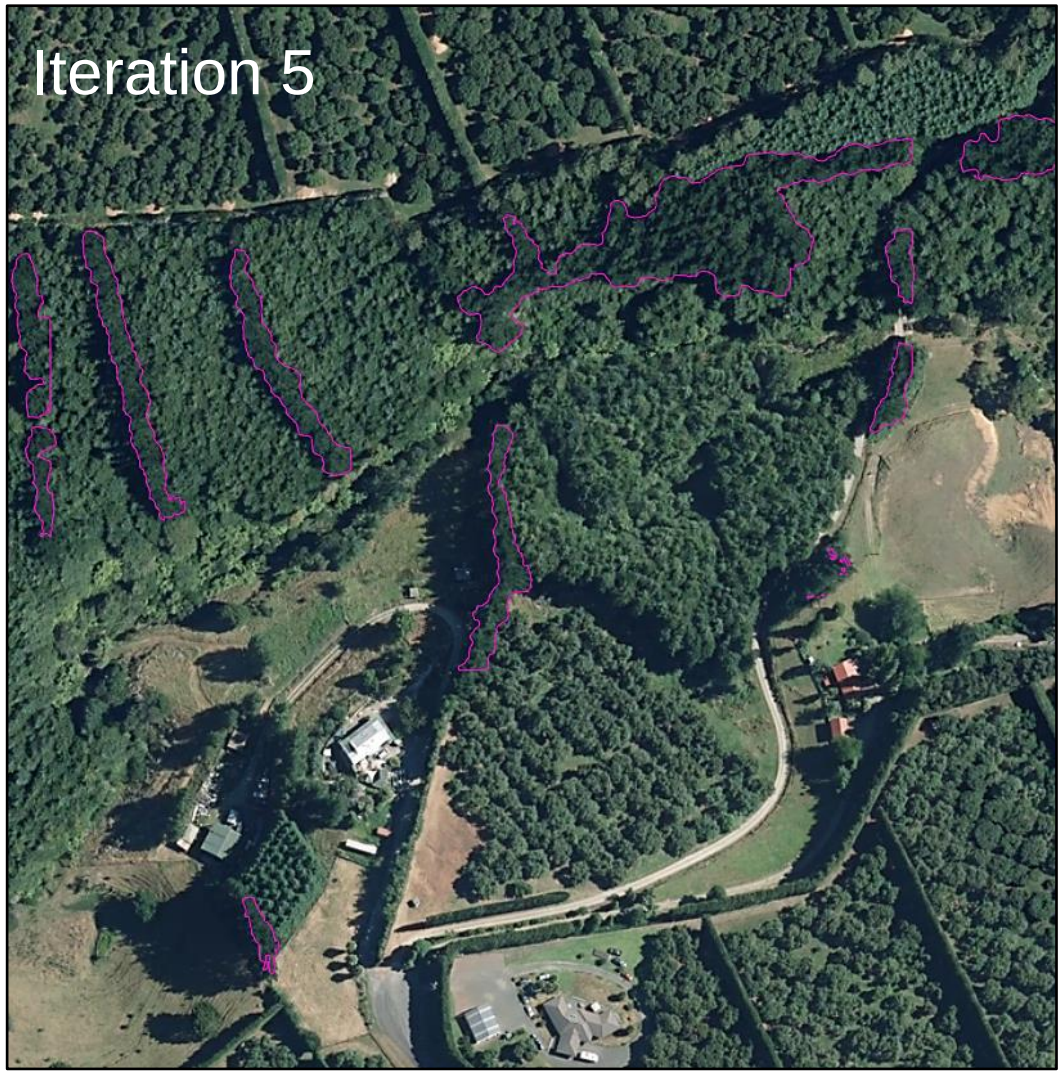
Iterative dataset development

- Image embeddings
  - Refine or re-enforce labels
  - Interrogate the model: loss vs embedding
- Error correction, mislabelling, consistency
- Find rare/under-represented examples
- Targeted labelling
  - Low-confidence areas
  - Areas of confusion
  - Address class imbalance
  - Future inference imagery

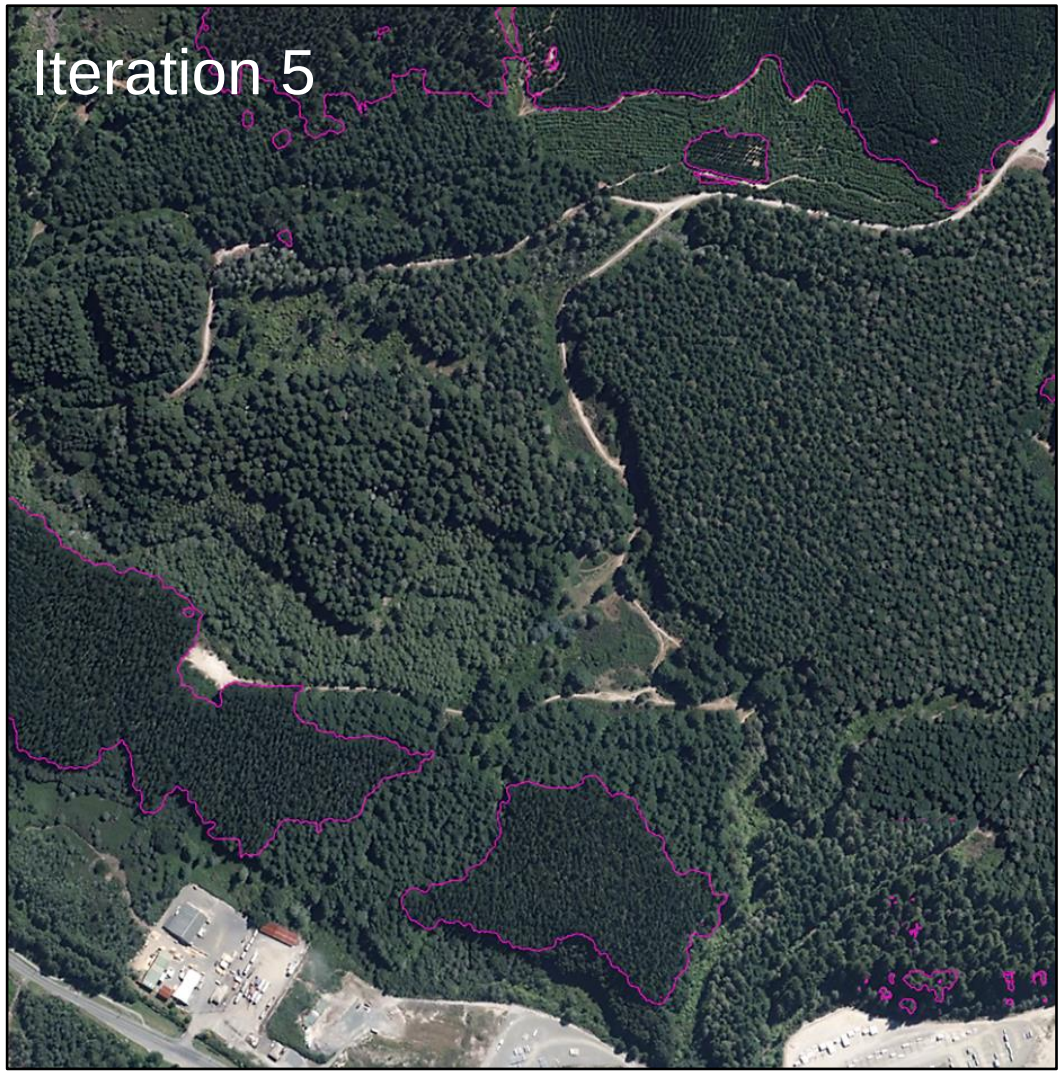
SCION

Iterative development

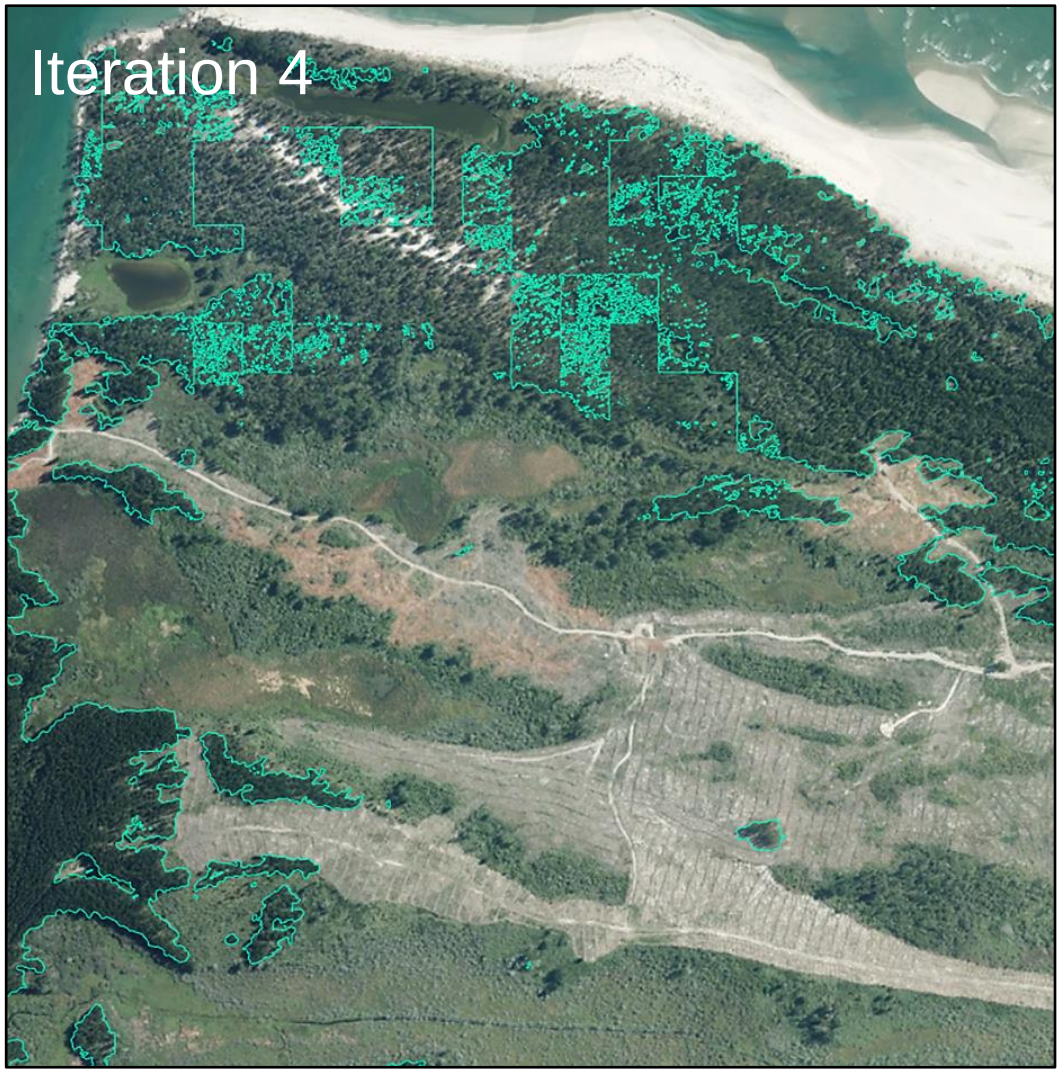


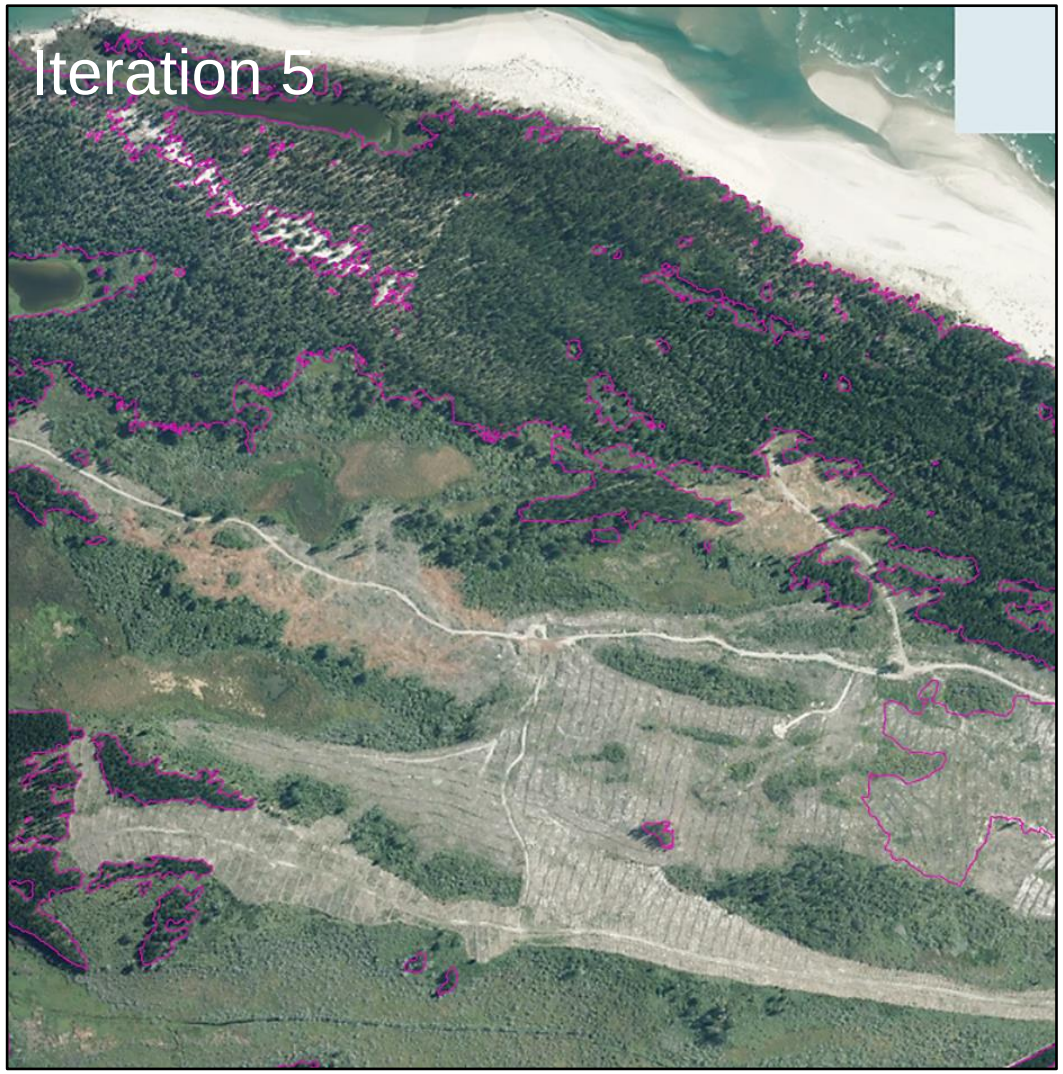


Iterative development

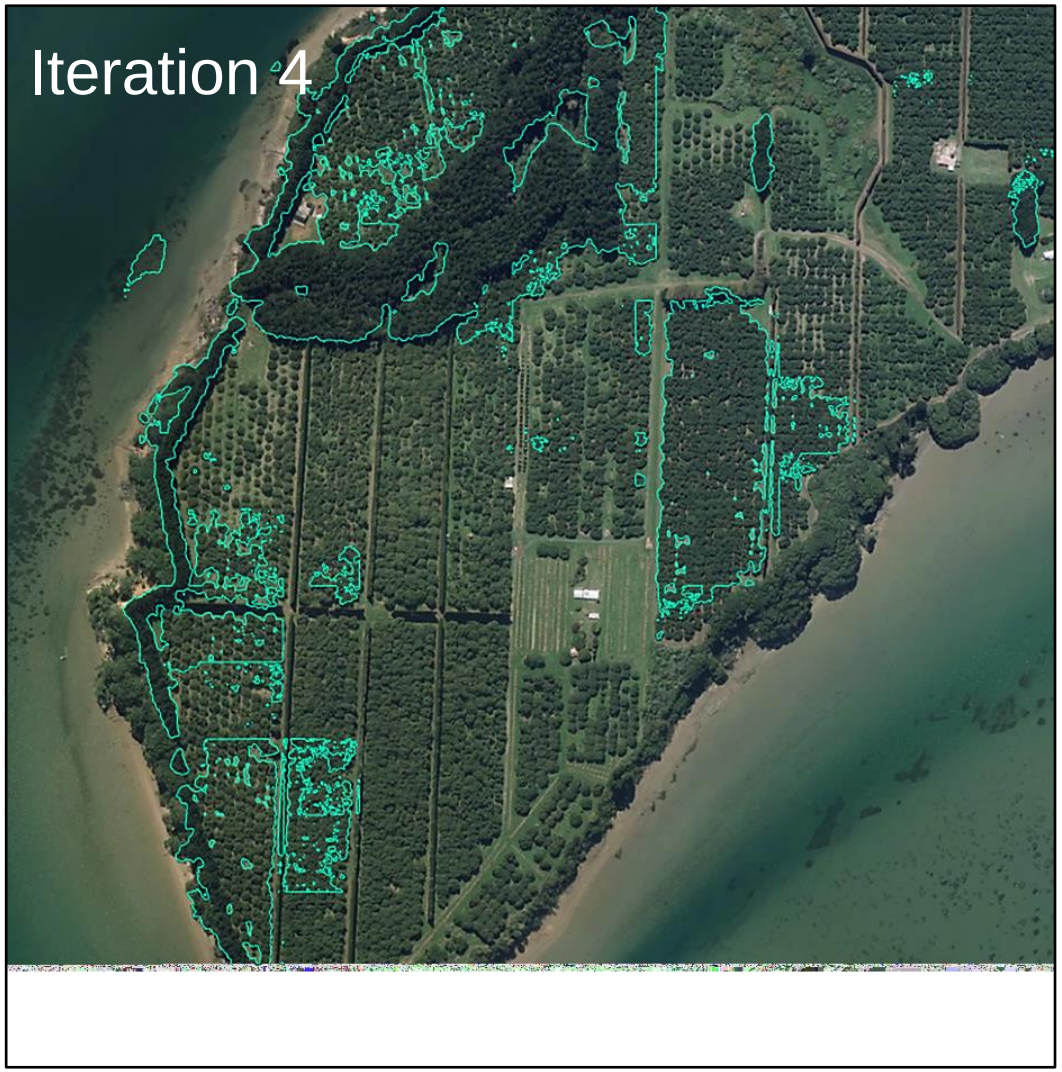


Iterative development

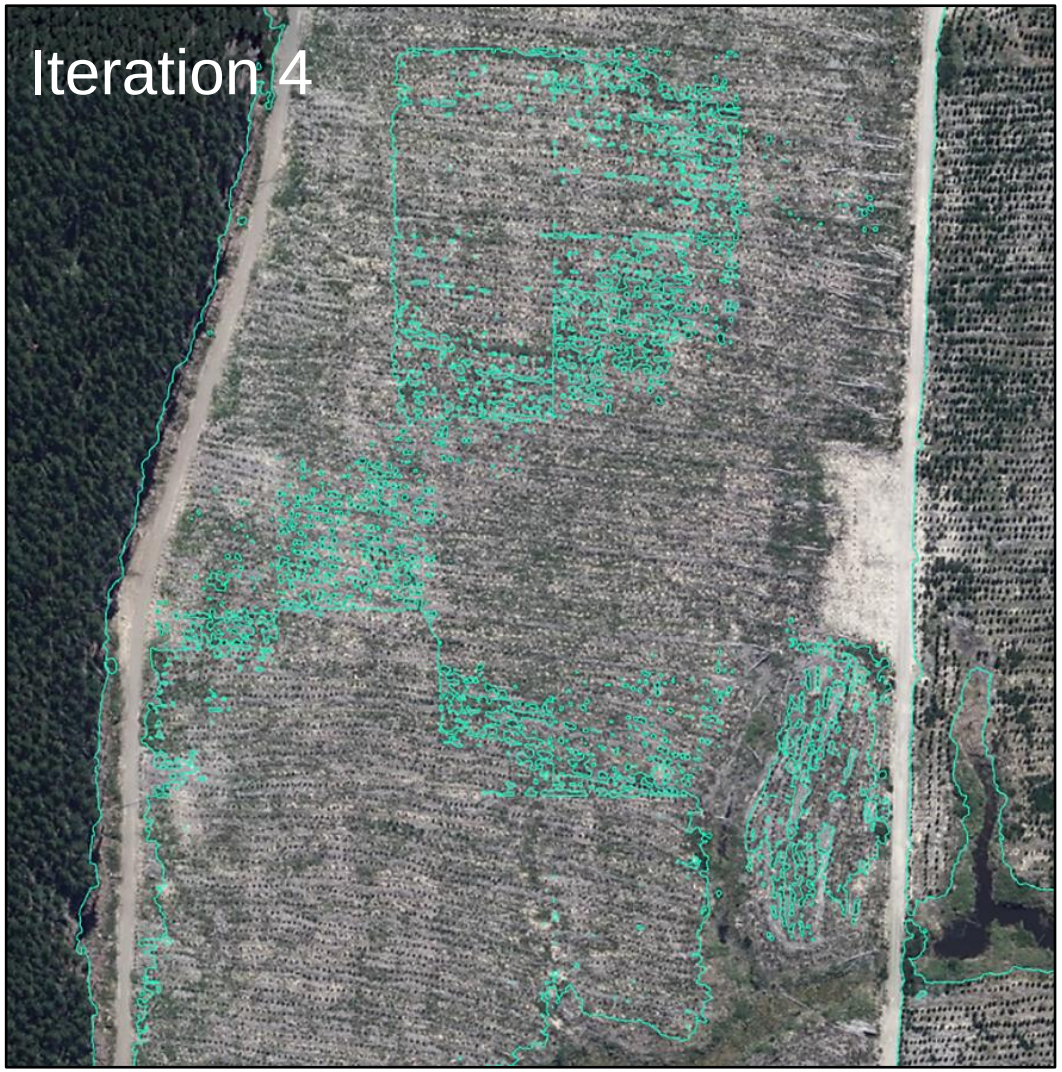


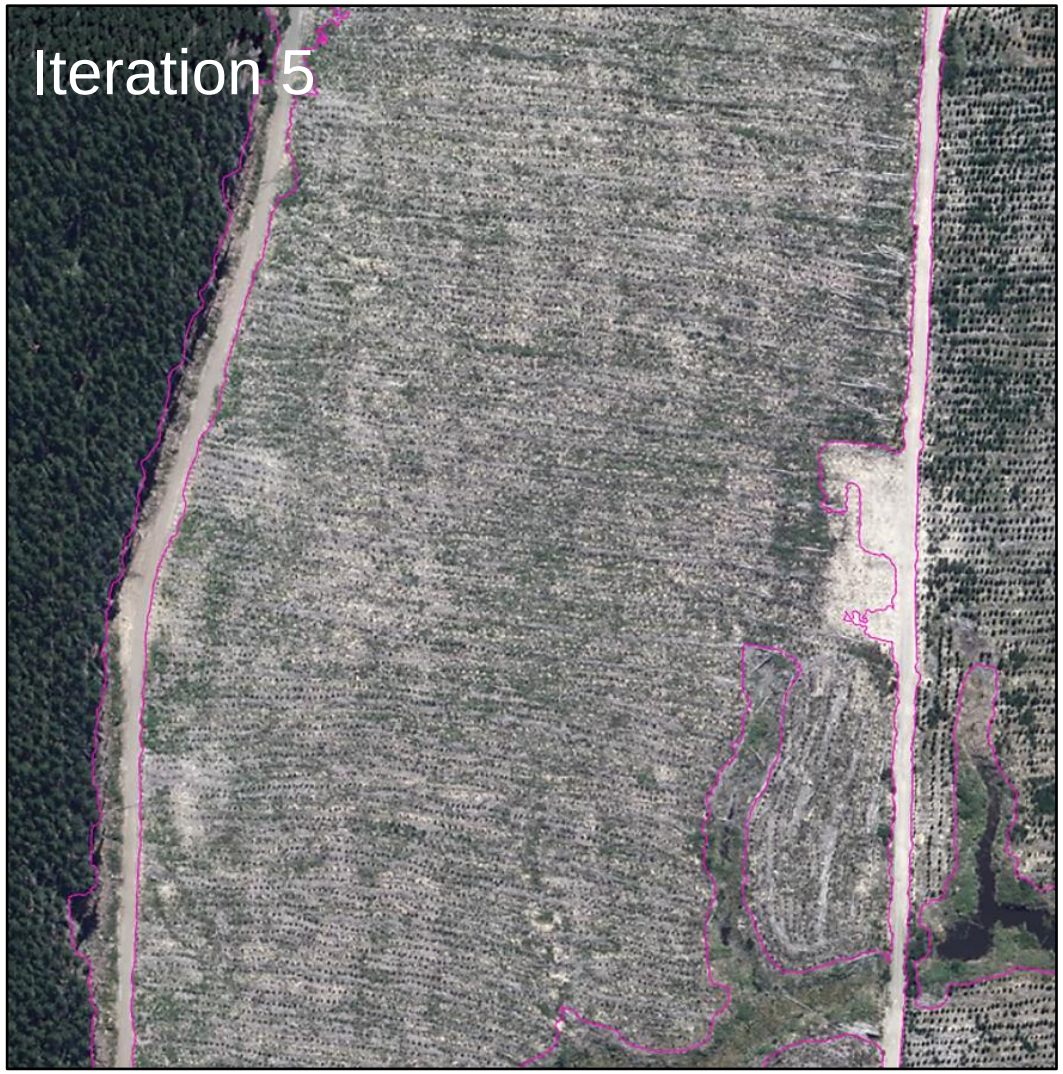


Iterative development



Iterative development

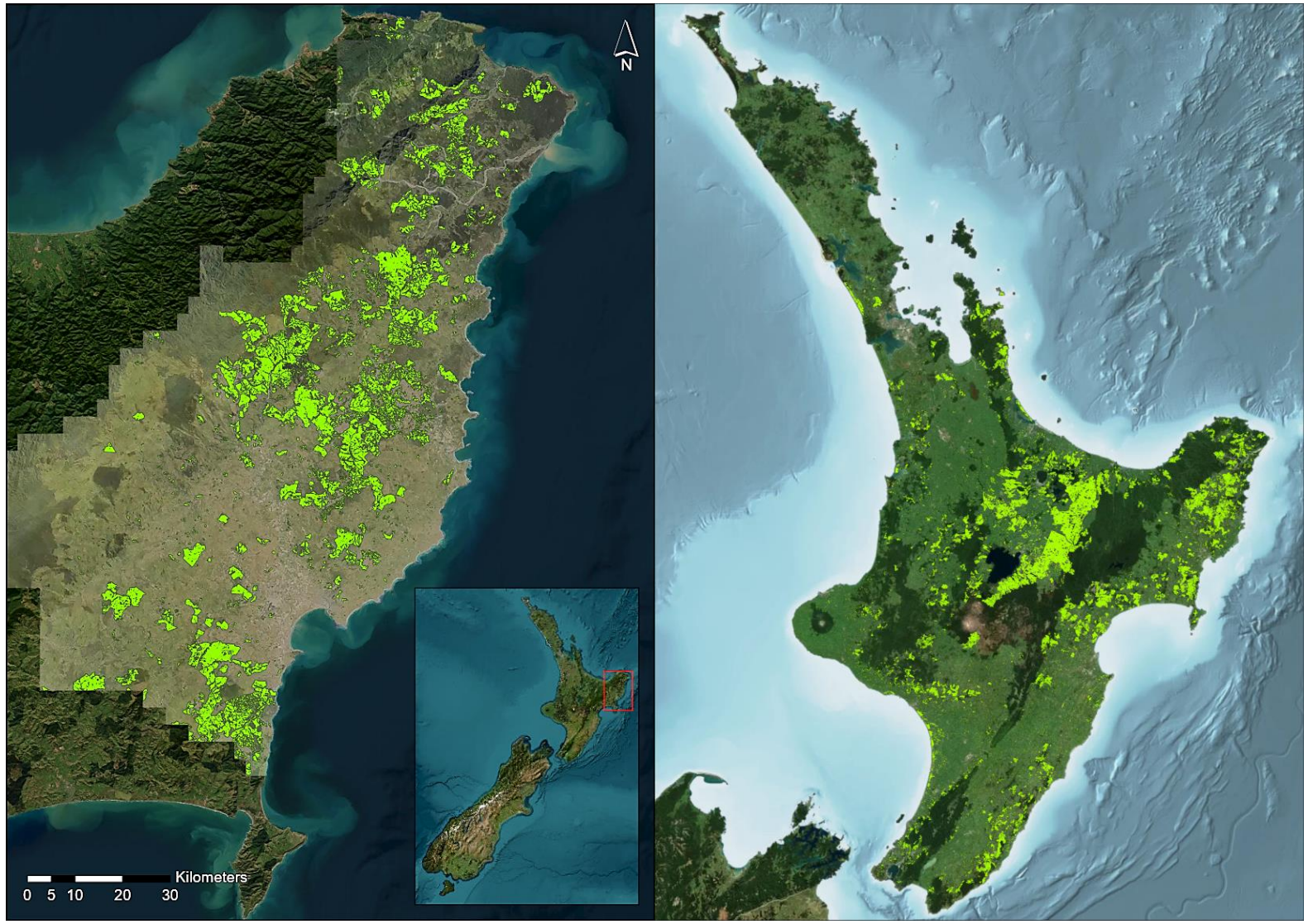




Inference

Model deployed over regional imagery

- Auckland
- Waikato
- Bay of Plenty
- Gisborne
- Hawkes Bay
- Taranaki
- Manawatū-Whanganui
- Wellington



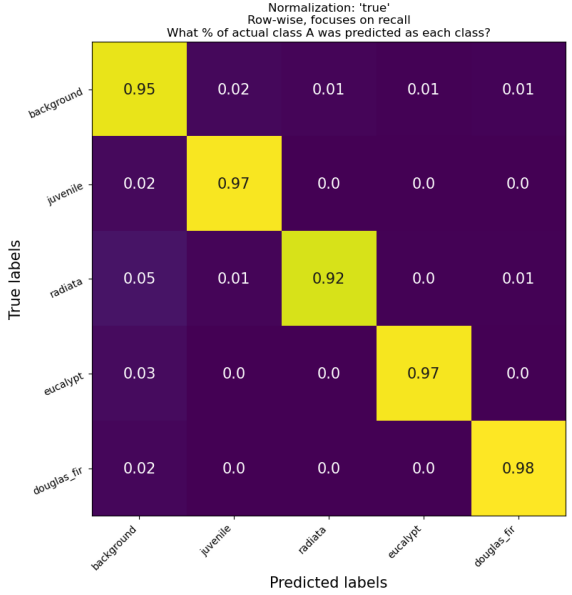
Over 1mil ha mapped!

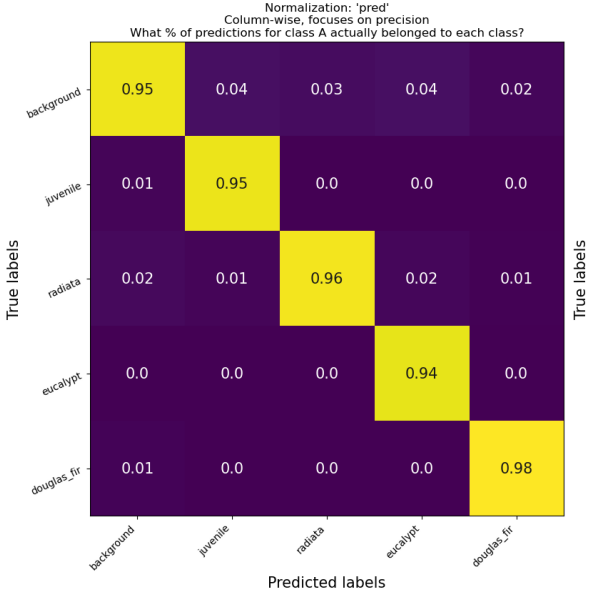
www.forestinsights.nz

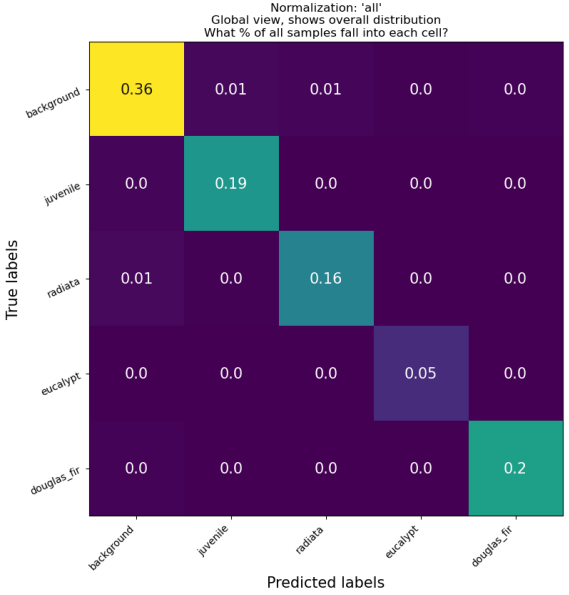
scion

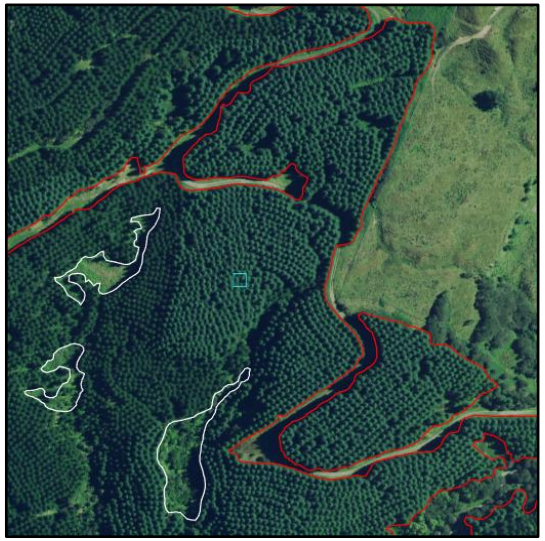
Where to next?

- **More regional inference**
  - Northland
  - South Island
- **Expand dataset**
  - Targeted improvement
  - South Island imagery
- **Other species & Multi-class model**
  - Douglas fir
  - Eucalyptus
  - Redwood
  - Cypress









Melanie Palmer
melanie.palmer@scionresearch.com
forestinsights.nz

19 November 2024

www.scionresearch.com