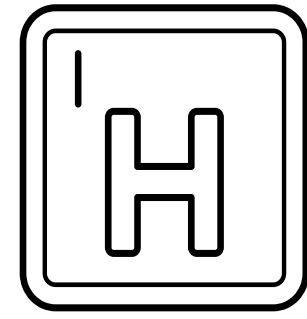# Detecting Fairy Circles from Sentinel-2 imagery in Australia

Wei Ji Leong[1], Bidhyananda Yadav[1], Adelaide Negrete[1], Ian Howat[1], and Joachim Moortgat[1]

[1]Affiliation not available

December 27, 2022

THE OHIO STATE UNIVERSITY
BYRD POLAR AND CLIMATE RESEARCH CENTER

Wei Ji Leong, Bidhyananda Yadav, Adelaide Negrete, Ian Howat, Joachim Moortgat

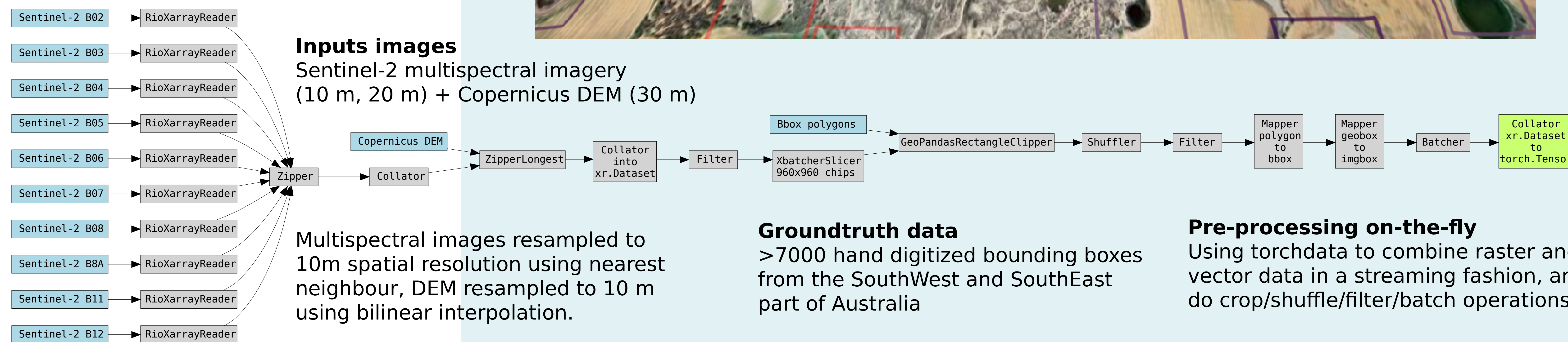# Detecting Fairy Circles from Sentinel-2 imagery in Australia

iPoster

Background

- **Ovoidal shaped** surface depressions associated with **natural hydrogen seeps** in places like Yorke Peninsula, Australia; Carolina Bays, United States; etc
- Potential **source of clean energy** for the zero carbon world
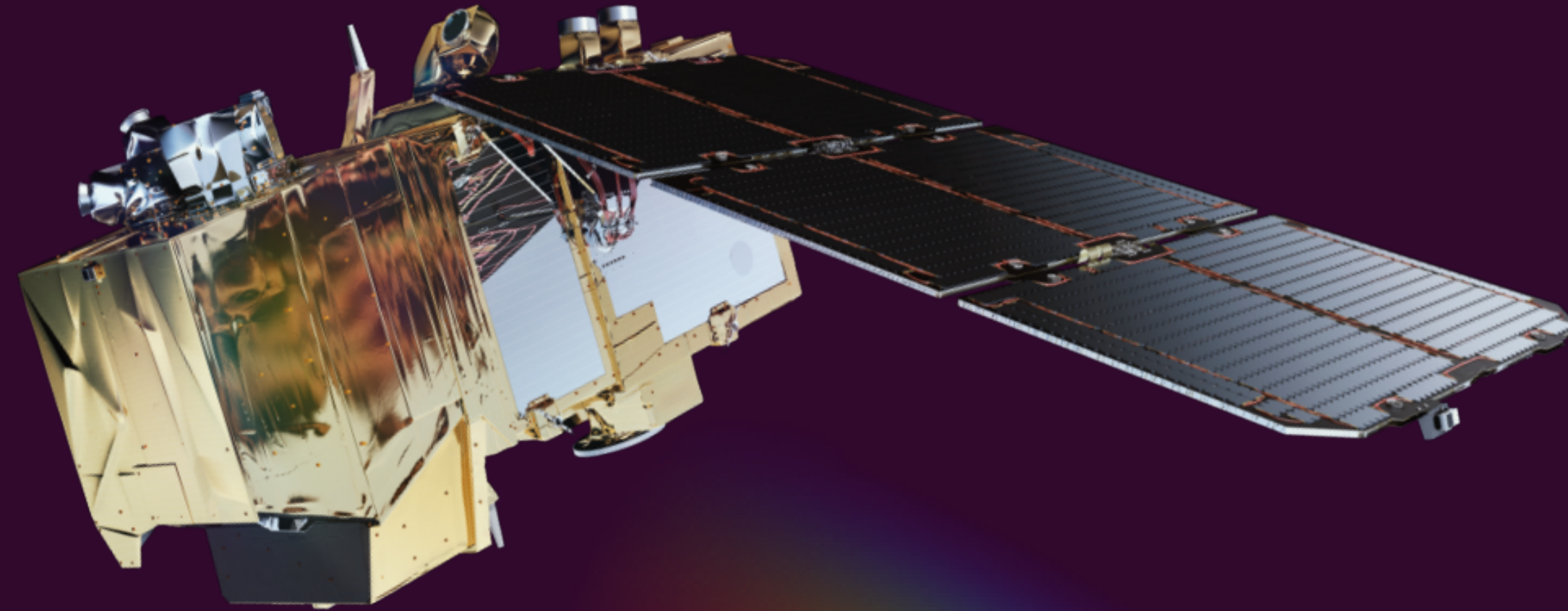- Can we **map the locations of these oval features** on a continental scale using machine learning?
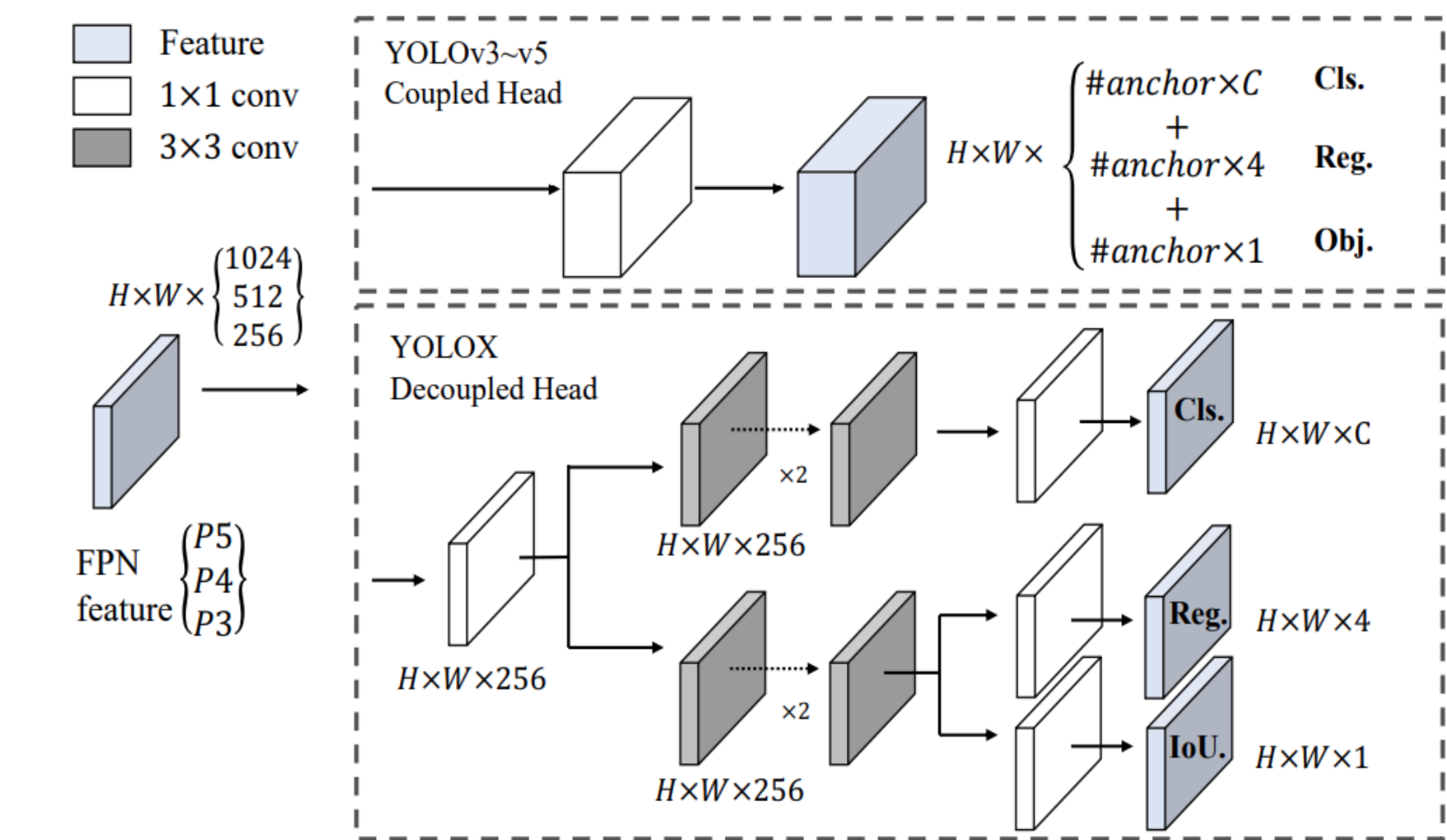
Tools used

STAC — SpatioTemporal Asset Catalog

COG — CLOUD OPTIMIZED GEOTIFF

PyTorch Lightning

MMDetection

xarray

Data pipeline

Sentinel-2 B02 → RioXarrayReader
Sentinel-2 B03 → RioXarrayReader
Sentinel-2 B04 → RioXarrayReader
Sentinel-2 B05 → RioXarrayReader
Sentinel-2 B06 → RioXarrayReader
Sentinel-2 B07 → RioXarrayReader
Sentinel-2 B08 → RioXarrayReader
Sentinel-2 B8A → RioXarrayReader
Sentinel-2 B11 → RioXarrayReader
Sentinel-2 B12 → RioXarrayReader

Copernicus DEM → Collator
Zipper → ZipperLongest → Collator into xr.Dataset → Filter → XbatcherSlicer 960x960 chips → GeoPandasRectangleClipper → Shuffler → Filter → Mapper polygon to bbox → Mapper geobox to imgbox → Batcher → Collator xr.Dataset to torch.Tensor

Bbox polygons

Inputs images
Sentinel-2 multispectral imagery (10 m, 20 m) + Copernicus DEM (30 m)

Multispectral images resampled to 10m spatial resolution using nearest neighbour, DEM resampled to 10 m using bilinear interpolation.

Groundtruth data
>7000 hand digitized bounding boxes from the SouthWest and SouthEast part of Australia

Pre-processing on-the-fly
Using torchdata to combine raster and vector data in a streaming fashion, and do crop/shuffle/filter/batch operations.
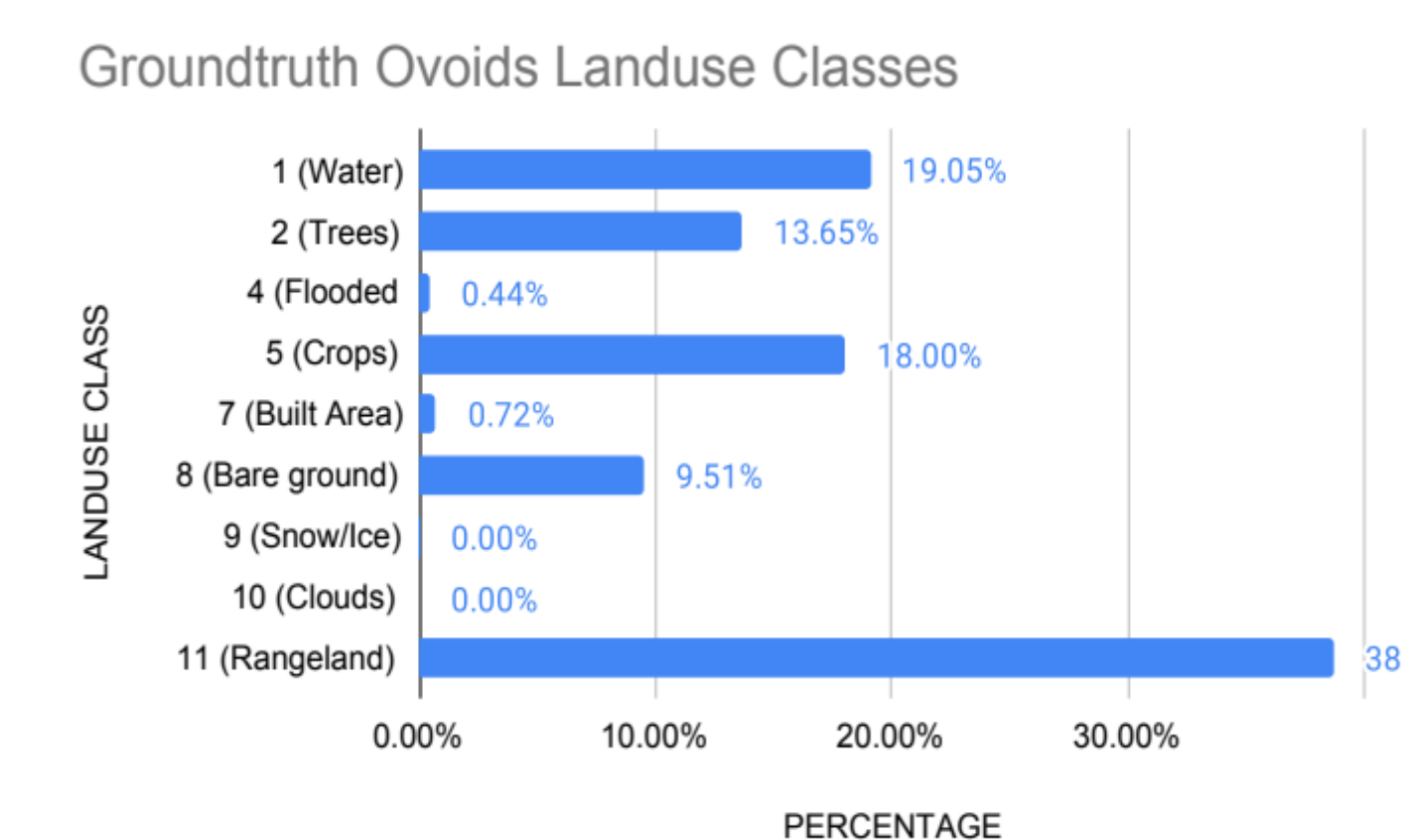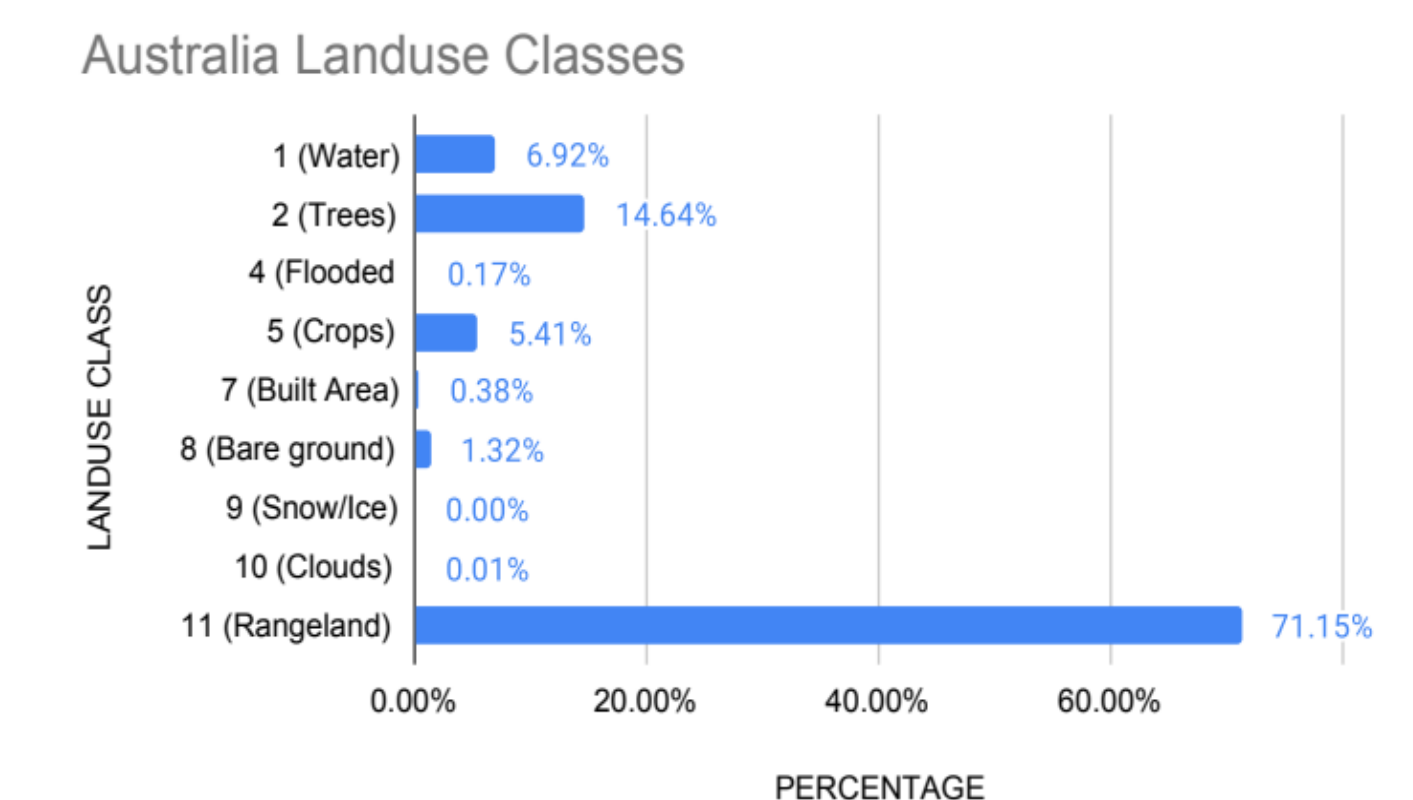
H2OvalNet with YOLOX

- YOLOX (2021) based **Convolutional Neural Network** built with Pytorch Lightning and mmdetection.
- **Light GPU memory** use = faster inference speeds (Australia in <1 week on T4 GPU).

Continental scale inference

- **Cloud-based** (Microsoft Planetary Computer) workflow to get bounding box predictions across 100s of Sentinel-2 tiles.
- Ran inference on **12 of the least cloudy Sentinel-2 images** in a year, on chips of size 9.6 x 9.6 km, with a batch size of 16.
- Resulting database of predicted bounding boxes are further postprocessed using a **custom weighted box fusion algorithm** to obtain high confidence oval feature locations.

Landuse class distribution

Australia Landuse Classes

LANDUSE CLASS	PERCENTAGE
1 (Water)	6.92%
2 (Trees)	14.64%
4 (Flooded)	0.17%
5 (Crops)	5.41%
7 (Built Area)	0.38%
8 (Bare ground)	1.32%
9 (Snow/Ice)	0.00%
10 (Clouds)	0.01%
11 (Rangeland)	71.15%

Groundtruth Ovoids Landuse Classes

LANDUSE CLASS	PERCENTAGE
1 (Water)	19.05%
2 (Trees)	13.65%
4 (Flooded)	0.44%
5 (Crops)	18.00%
7 (Built Area)	0.72%
8 (Bare ground)	9.51%
9 (Snow/Ice)	0.00%
10 (Clouds)	0.00%
11 (Rangeland)	38.

Comparing the typical land cover classification of the Australian continent versus that of oval feature locations

- Many ovoids **water filled** and over Rangeland landcover types.
- Need **more training data over vegetated (i.e. Trees and Crops) areas** like Queensland and New South Wales.