

Repurposing GIS data: Training neural networks to identify land use types

Kevan Cote
Eco-index®
kevan@moosedesign.co.nz

ABSTRACT / INTRODUCTION (Up to 200 words)

When creating neural networks for image segmentation, acquisition of large datasets for 1) training models and 2) ensuring models are reproducible, are significant challenges. One aim of the Eco-index® programme is to identify ecosystem types (e.g., wetlands, kahikatea forest) using temporally and qualitatively variable data. Our approach has been to combine available GIS shapefiles with accessible satellite and multispectral imagery to create training data for an object detection-focused neural network. Specifically, we have utilised open source Python packages (GDAL) and pre-existing Github repositories to split up the satellite imagery and train a model in a UNET framework. Our pipeline is open to a variety of data, thus providing flexibility in the types of land we can detect and the type of training data we can use (e.g., existing council/government land cover GIS data). Upcoming work includes the incorporation of additional data sources (e.g., soil moisture, elevation) alongside the satellite imagery, and the addition of a temporal element. The latter will allow for land change monitoring, and could be especially useful for tracking ecosystem restoration efforts. Adhering to the programme's Digital Public Good status, the final methodology will be packaged in a Docker container, thus making it reproducible in other applications.

ABOUT THE AUTHOR(S)

- Kevan Cote
- Kevan is an applied data scientist with the Eco-index® programme. In addition to his love of machine learning he has a unique background in manufacturing and bringing products to market which allows him to apply systematic methodology to machine learning applications. Seeing value in sharing and contributing the open source community, he strives to make code usable and easily adaptable by others.