Monitoring UK national freshwater with deep learning (Ref IAP2-18-98)

Durham University, Depts of Geography and Engineering and Computing Sciences
In partnership with Biological and Environmental Sciences, University of Stirling

Supervisory Team
Patrice Carbonneau, Durham University
Pete Hunter, University of Stirling
Toby Breckon, Durham University

Key Words
Deep Learning, Machine Learning, Drone Mapping, Lakes, Rivers

Overview
UK freshwater bodies are under increasing pressures from growing populations, increased usages and urbanisation. At a UK national scale, monitoring the evolution and health of water bodies is a considerable challenge that can be very labour intensive. Even for basic parameters such as water body size, the establishment of a comprehensive, time-dependent, database for all UK water bodies of any meaningful size remains an unattainable goal. However, given that water body dimensions can easily be established from aerial and high resolution satellite imagery, the raw data required for this task exists in very large volumes and is very often freely available. The crucial limiting factor lies in the identification of the water bodies from large data sets of imagery which fall within the remit of what is now termed ‘Big Data’. This project proposes to apply AI technology to the problem of freshwater body identification (lakes and rivers) in the UK. The project has 4 objectives:

1- Mine the DIGIMAP database for colour imagery from the years 1996 to 2018 and train an AI image classifier to identify freshwater bodies over the entire UK.

2- Ground truth and validate recent imagery with high resolution drone surveys.

3- Acquire declassified USGS (formerly from the CIA) greyscale image data of the UK, from the 1960s, 1970s and 1980s and additional imagery from the RAF archives spanning the period of 1945-1965. Train a second AI to identify surface water features in historic imagery.

4- Use the outputs of objectives 1, 2 and 3 in order to establish quantitative metrics for the evolution of UK lakes and rivers over the last 50 years.
Methodology

The project will rely on Artificial Intelligence in the form of image recognition. Both Machine learning and Deep learning will be investigated in order to develop a reliable image classifier that can approach human performance in the identification of water in national-scale time series of imagery despite changing colours and shades as a function of time. It is anticipated that this will require the usage of an advanced convolutional neural network trained with the specific features that can be seen from the air such as houses, fields, and, crucially, lakes and rivers. Several advanced AI image classifiers do exist, but these tend to be trained by the images now prevalent on social media. As a result, these networks are not well tuned to the classification of features from an aerial perspective.

Data will be sourced from the DIGIMAP national databases for the period 1996-2018 and from a mix of sources such as the RAF archive imagery on Google Earth and the USGS for the period of 1962-1984. Ground calibration and local validation of classification results, will be done with high resolution drone surveys over selected sites in England and Scotland. These low-altitude surveys will focus on water body edges and cases where there is an established seasonal change of colour, and on the question of detection limits in terms of size (i.e. the smallest detectable water body, or narrowest detectable river).

Timeline

Year 1:
- Construction of the data mining architecture
- Preliminary tests with Machine and deep learning image classifiers for colour imagery.
- Drone piloting and safety course
- Acquisition of archival spy satellite imagery from the USGS.

Year 2:
- Drone operations for validation
- Preliminary tests with machine and deep learning image classifiers for greyscale archival imagery.
- Deployment of automated data mining algorithms for 'Big Data' image acquisition at national scale.

Year 3/3.5
- Large-scale processing using Durham University Hamilton cluster.
- Development of a regional indices that summarise the evolution of waterbody area at a high temporal resolution, for the last 50 years.
- Thesis production

Training & Skills

The student will be trained for a full CAA-approved drone safety and piloting course in order to be eligible for Durham University insurance covering drone operations. Student will then be trained in advanced methods of machine learning and deep learning, including Convolutional Neural Networks. This will include funded attendance at Deep Learning summer school events that can deliver cutting-edge training in the fast-moving field of Artificial Intelligence.

References & Further Reading


Further Information

For project information, please contact: Dr Patrice Carbonneau