Project Vision

Group 5 (G5)

Index

Introduction

The NSW Spatial Services division of the DCS collects aerial and LiDAR images of the NSW landscape and is responsible for providing reliable spatial information to NSW Emergency Service Organisations. In order to better facilitate rapid responses and recovery efforts to emergency flood events, this project aims to develop a machine learning platform that will analyse aerial and LiDAR imagery to detect areas affected by floods.

The current standard of practice for DCS during a flooding event is to fly an aeroplane over flood-affected regions of interest (ROI) to capture aerial images. These images are taken as long continuous strips which must be stitched together, colour balanced, and orthorectified to produce a surface model - a representation of topographical features. This surface model is what is currently distributed to other government departments for identifying flood extent however, this does not provide a discrete boundary with coordinates. Each client of DCS performs their own processes for identifying the exact boundary. The aim of this project is to consolidate and standardise this task into a repeatable, accurate, and time-efficient process which will improve the quality of the product offered by DCS and eliminate variance and redundant workloads by DCS's clients.

Programs that are able to generalise relationships between input variables and output variables through inductive inference, whether through supervised, unsupervised, or reinforcement learning algorithms, are categorised as machine learning programs (Guo et al., 2020). Many of the difficulties in contemporary areas of machine learning research deal with defining problems that are simple for people to perform but hard to formally describe because they entail a high level of intuitive and subjective knowledge of the world, such as recognising faces in an image (Goodfellow, 2016). One class of problems that machine learning programs are proving to be effective at solving are classification problems through the use of models such as convolutional neural networks and support-vector machines (Hasan et al., 2019). If a machine learning model is capable of analysing images of landscape and classifying them as affected or unaffected by natural disaster events such as bushfires and floods, then this has the potential to greatly assist decision-making capabilities for government authorities managing and monitoring recovery and disaster-relief efforts.

By leveraging machine learning models, this project aims to develop a system capable of accepting new aerial and LiDAR images of the NSW landscape and classifying specific segments as either affected or unaffected by floods. This will provide accurate and time-sensitive information to NSW Emergency Service Organisations (ESOs).

Positioning

Problem Statement

The problem of	providing reliable and expeditious spatial information of natural disaster affected areas
affects	NSW Spatial Services and NSW Emergency Service Organisations
The impact of which is	Implementing disaster-relief and recovery plans based upon current and accurate information
A successful solution would be	A program that is able to take aerial and LiDAR images as input and classify specific sub-segments as either affected or unaffected by floods, producing a map of peak-flood extent

Product Position Statement

For	NSW DCS, Spatial Services, ESOs
Who	Need access to reliable spatial information on disaster-affected areas
The	Machine learning model trained in AWS Sagemaker
That	Will be able to take aerial and LiDAR images and classify segments as affected or unaffected by floods, classification will be performed as soon as the images are available
Unlike	Human inspection of images which has a degree of error and lack of time-efficiency
Our Product	Will provide accurate and up-to-date flood extent maps

Stakeholder Descriptions

Stakeholder Summary

Name	Description	Responsibilities
NSW Department of Customer Service - Spatial Services	Spatial Services is a division of the NSW Department of Customer Service which provides authoritative spatial and land information. Spatial Services has invited Charles Sturt University students for work experience with the Business Technology Services Team until December 20201.	Provide project scope, approach and timelines. Provide roles / responsibilities and governance arrangements. Supply the programming language definition and what tools and equipment will be required. Provide ongoing liaison for project updates.
Intellify	Intellify specialise in the delivery of Artificial Intelligence solutions. Being the second consultant in Australia and New Zealand to be awarded the Amazon Web Services Machine Learning Competency makes Intellify well placed to assist with Machine Learning projects.	Create the AWS environment to use ML / AI. Upskill Spatial Services staff in the use of ML/AI. Provide ongoing support to Spatial Services.
Dr. David Tien	Dr Tien is a Senior Lecturer in Information Technology at Charles Sturt University. He has over 34 years experience in AI related research.	Supervise and Mentor Team 5. Liaise with DCS.
Team 5	Team 5 consists of five undergraduate students studying at Charles Sturt University undertaking the ITC 303 and ITC 309 Software Development Project subjects during 2021.	Produce deliverables in line with ITC 303 / 309 requirements and project requirements as requested by DCS.

User Environment

Certain details of the user environment are yet to be specified, however some general features are established. The DCS users who will utilise the machine learning model to produce the flood extent mapping are familiar with the AWS CLI (command line interface) and ArcGIS[1], a geospatial information system that combines topographic, qualitative, and quantitative data. As outlined in the Initial Requirements document, the project is to be developed utilising AWS services to minimise workload and leverage their products to overcome technical difficulties in the machine learning development process. The user environment will therefore consist of the AWS CLI, likely utilising the AWS Lambda[2] functions as an entry point, AWS S3 Buckets[3] for storage, and AWS SageMaker[4] to host the trained machine learning model. The final output will be a surface model object for the ArcGIS software application.

The output of the system will be distributed by DCS to multiple government agencies such as the NSW RFS (Rural Fire Service) and NSW DPI (Department of Primary Industries) who will use the flood extent maps for emergency service planning. For example, NSW DPI may use the product to identify stranded herds of cattle and organise food drops. This implies there will be multiple users of the end product with different priorities.

Product Overview

Needs and Features

Need	Priority	Feature
Storage of images to be used for ML training & new images for client use case	High	Use of AWS Simple Storage Service (S3)

[1] Esri, *ArcGIS Overview*, https://www.esri.com/en-us/arcgis/about-arcgis/overview
[2] AWS, *AWS Lambda*, https://aws.amazon.com/lambda/
[3] AWS, *Simple Service Storage*, https://aws.amazon.com/s3/
[4] AWS, *AWS SageMaker*, https://aws.amazon.com/sagemaker/

Feature labelling, extraction, and reduction from images	High	Use of AWS SageMaker Ground Truth
Classification of AOI as flooded/unflooded	High	AWS SageMaker algorithm
Authentication of users	Low	Use of DCS ss-identity account for resource and service access
Generation of shapefile object for flood extent map	Medium	To be determined

Other Product Requirements

Produce Accurate Results - It is important that information produced by the system is accurate. If the output is not reliable from an accuracy point of view it will provide no benefit and will cause incorrect allocation of limited Emergency Service Organisations (ESO) resources during an emergency

Usability for non-technical users - System must be simple enough to use by a person who is considered computer literate but may not have a background in science and technology. This will enable users to get the most utility out of the system without the need for specialist training and roles to operate.

System uptime and reliability - The uptime of the system is crucial to its utility. The system may spend a lot of time idle however it could see times of extremely heavy usage during if the system was to become unstable or require regular reboots it would impact the use of the system as a means of delivering information to ESOs.

Cost - Deploying a SageMaker endpoint in the AWS Cloud environment has an ongoing cost for its time of deployment. Stakeholder meetings will need to clarify the specifics of the production pipeline and the active period required for accessing SageMaker model predictions (i.e. ongoing vs. discrete access periods).