AUSTRALIAN GEOSPATIAL-INTELLIGENCE ORGANISATION (AGO) ANALYTICS LABSV

CANDIDATE CHALLENGE DESCRIPTIONS

Please email queries to agolabs@frontiersi.com.au

Questions and further information can be provided by:

Phil Delaney, Chief Innovation and Delivery Officer, pdelaney@frontiersi.com.au
Chris Morgan, Project Manager, cmorgan@frontiersi.com.au

June 2020

Where the Streets Have No Name

Project overview

This challenge aims to test how computer vision and geospatial information processing techniques can be used to autonomously generate road transport networks. The associated need is the ability to rapidly generate mobility assessments over regions that have undergone sudden change (such as following a natural disaster) or regions where a rapid update to a transport network is required.

Problem Statement

The Australian Geospatial Intelligence Organisation (AGO) is investigating the application and integration of various computer vision solutions for one of our key source materials: high-resolution satellite-imagery. This content is used by AGO to develop various information resources, such as transport networks. These networks are vast: covering everything from road, rail and paths. Various algorithms and frameworks exist to autonomously detect and classify road and rail networks from satellite imagery. However, these techniques are typically limited to producing road or rail-segments of differing accuracy, bias and completeness, rather than networks that can be effectively utilised for mobility planning and assessment.

This challenge aims to understand the potential for autonomous transport network development and management techniques that utilise both computer vision frameworks and vector integration frameworks. Each of the following sections of this challenge describe key questions for resolution. Proposals are not required to address each and every component, though this will be an advantage.

The transport networks that may be considered in this challenge are terrestrial and include roads, rail and trails. The challenge will be expected to provide insight on applications and limitations (including accuracy) and adaptability (including environmental considerations).

Detection Performance

Through this challenge, AGO would like to understand the limitations of segment detection for roads, rail and trails. Of particular importance are to understand the detection rate (for example, for presence of a road network), the miss rate, and the overall accuracy (for example, extent of a network identified).

Various models exist for road detection and potential applications (including https://www.cosmiqworks.org/current-projects/spacenet-5/). This challenge is expected to explore and understand both Australian industry capacity to conduct this work, and also how it translates to regional conditions (more detail in the environment characterization section).

Importantly, this project will also seek to develop insight on detectable network depth. This will include distinguishing between trails, minor and major roads, and rail segments.

Transport Network Development

The development of usable transport networks is the key function being explored through this project. In this context, it is anticipated that the successful candidate would explore and assess the automated integration of network segments (derived from imagery) into a coherent network for exploitation by an existing network processing package.

AGO is interested in understanding the relevant detection and error rates, and the overall accuracy. As part of that, this challenge is expected to identify and assess various accuracy improvement strategies, such as using multiple collects to drive up probabilistic likelihood of segment and network completeness.

Input Characterisation

The primary source-data to be used in this challenge are high-resolution commercial imagery from the WorldView constellation. For each transport network type there are associated range and distribution questions, including (but not limited to):

- size/bit-depth
- obliquity
- band depth/colour mapping
- cloud cover
- file format range/distribution
- sensor type
- alignment – what happens when images e.g. from one sensor overlap those taken with another? Does the degree of overlap affect the training set?
- periodicity (uneven sampling periods?)

Understanding the interactions of these input characteristics with the potential applications will guide improved understanding of utility and limitations.

Environment Characterisation & Portability

The rapid generation of transport networks is something that AGO would want to apply across a broad range of regions. Whilst an analysis of *every* potential environment is infeasible, understanding performance (in the context of 'Detection performance' and 'Network development') in different environmental conditions is an important consideration.

In particular, if a model is trained on one or two locations, will it perform to the same degree on all locations in the same region/same country/similar regions? And if it is best at similar regions, how is 'similar' characterized and is it actually characterizable a priori?

Challenge X: Bushfire Preparation Support

Overview

This challenge aims to understand the role computer vision techniques can play in time-sensitive feature detection, in support of bush-fire preparation.

Problem Statement

Understanding the landscape in an area before a fire occurs is a key part of informing risk associated with fire management. While extensive research has been undertaken to understand landscape characteristics such as fuel load and forest structure, this challenge is interested in identifying important physical features on the ground to provide an understanding of 'what is where' in the environment. Being able to identify key features including fire breaks, hazard reduction locations, accessible sources of water and road access points from satellite imagery will provide a rich source of current information to risk managers working in these areas.

Timeliness of Information

Speed and dependability is the key feature in this challenge. Due to the dynamic and rapidly changing nature of bush fire threat, a successful trial capability will be one that can be turned around within a short time frame (for example, same-day processing and analysis) and preferable, on a fixed schedule (for example, a nominal 4pm information update cycle).

Detection Performance

Through this challenge, AGO would like to understand the limitations of feature identification, including detection and failure rates, and the overall accuracy.

Input Characterisation

The primary source-data to be used in this challenge are high-resolution commercial imagery from the WorldView constellation. For each feature there are associated range and distribution questions, including (but not limited to):

- size/bit-depth
- obliquity
- band depth/colour mapping
- cloud cover
- file format range/distribution
- sensor type
- alignment – what happens when images e.g. from one sensor overlap those taken with another? Does the degree of overlap affect the training set?
- periodicity (uneven sampling periods?)

Understanding the interactions of these input characteristics with the potential applications will guide improved understanding of utility and limitations.

Challenge X: I Can't Believe It's Not Nadir!

Overview

This research focussed challenge aims to generate insight on how object detection algorithms perform against targets in nadir and off-nadir conditions. That insight is expected to include how and when deviation from nadir starts to impact algorithm efficacy, as well as different approaches to how those problems can be measured and managed.

Problem statement

In AGOLabs2019, AGO asked Australian industry to investigate ways of producing low-cost object detection models. In response, Urbis and AIML showed that the pairing of GANs and data augmentation techniques can achieve similar results to a baseline CNN with 50% of the total training data.

AGO would like to continue investigating object detection models, this time incorporating images captured off-nadir. A typical approach to solving the problem of object detection on off-nadir imagery might be to include large quantities of labelled training data for each collection geometry, however, labelled training data for satellite imagery is uncommon and expensive to produce. Therefore, an important element of this challenge is to create a model that is robust to diverse collection geometries, while remaining low-cost with regard to labelled training data.

Although the solution to the AGOLabs2019 low-cost object detection challenge implemented GANs to meet the challenge, AGO is not prescriptive in the approaches that may be used.

Detection Performance

Through this challenge, AGO would like to understand how off-nadir image-target geometry impacts the detection performance of object detection algorithms in a variety of conditions. In this instance, performance is expected to include detection rate, miss rate and overall accuracy under different sensor/target geometry conditions.

Algorithm and Model Portability

The project will be expected to identify an approach to nadir and off-nadir object detection, and to quantify the impact of that approach. This may be in the form of training-set enhancement, model modification, or the development of new models.

Input Characterisation

The primary source-data to be used in this challenge are high-resolution commercial imagery from the WorldView constellation. For each feature there are associated range and distribution questions, including (but not limited to):

- size/bit-depth
- obliquity
- band depth/colour mapping
- cloud cover
- file format range/distribution
- sensor type

Understanding the interactions of these input characteristics with the potential applications will guide improved understanding of utility and limitations.

Challenge X: Bands: How Low Can You Go?

Overview

This research focussed challenge is concerned with quantifying advantages and limitations of multi-spectral analysis performed using machine learning algorithms.

Problem statement

The aim is to characterise the potential benefits and limitations of machine learning when applied to multispectral analysis of different phenomena. An ideal response would consider how the potential applications of multispectral analysis with machine learning could be quantified, and against which types of targets it might be most effective, for example: man-made objects, tracks, and tree-species. It would go on to categorise spectral/distribution characterisation and optimisation, for those targets.

Classification Performance

In this challenge, performance would be measured through classification fidelity (the narrowness of classification of different target-types) and accuracy (the detection and miss rates, and the overall accuracy).

Signature Libraries

AGO would like insight on the potential feasibility of using existing signature libraries and utilising them in the development of machine learning algorithms for satellite imagery-based detection.

Input Characterisation

The primary source-data to be used in this challenge are high-resolution commercial imagery. For each feature there are associated range and distribution questions, including (but not limited to):

- size/bitdepth
- obliquity
- band depth/colour mapping
- cloud cover
- file format range/distribution
- sensor type
- periodicity (uneven sampling periods?)

Understanding the interactions of these input characteristics with the potential applications will guide improved understanding of utility and limitations.

Challenge X: A Pixelating Dilemma

Overview

This challenge is concerned with finding an efficient means of maintaining distributed image libraries in network constrained environments. Specifically, rather than updating whole images, it would involve updating only pixels that have changed.

Problem Statement

The problem is mainly related to minimising the amount of data that needs to be transmitted in order to maintain common imagery baselines. These imagery baselines are consistent, orthorectified satellite imagery products across a region, and are typically used for context in decision making and battlefield preparation processes. However, our imagery baselines are relatively large, ranging from the gigabyte to the terabyte in data volumes, which presents a challenge for austere networks to accommodate.

Instead of transmitting the entire image, however, this challenge is asking the question whether only the specific pixels that have changed, could be sent, and stitched back together as an updated baseline, at the destination.

Change Detection

A pixel-to-pixel change detection algorithm would need to exist in order to support this analysis. The efficacy of that change detection would need to be measured, and include a detection rate, miss rate and a measure of overall accuracy.

Pixel-Value Normalisation

In order to effectively identify changed pixels, and also enable the destination to effectively stitch them together, the pixel-values would need to be normalised.

Small Challenge X: Space, Time and Other Probabilities

Overview

This challenge deals with the asymmetric information conditions of event-based observation and tracking data, specifically concerning the probability spaces of what happens between observations.

The Problem

Spatiotemporal data for moving entities (such as Automated Identification System – AIS, or NYC Taxi Cab data) can be understood as a sample of entity activity. Data points are spatiotemporal 'anchors' where there is quantifiable certainty about an entity's location at a given time, but between these 'anchors' are periods of uncertainty. In these periods, an entity's location is constrained by 'anchors' but precise location is unknown. This becomes increasingly problematic as data sparsity increases. AGO would like to investigate ways of accurately determining, understanding, and representing the possible movement of entities between 'anchor' points. The result of this would be 'probability spaces' for entity movements that show where entities could have been, and can be analysed to understand possible/probable pathing, entity interactions, and changes in behaviour.
An ideal solution for this challenge will involve the development of probability spaces, and use of those probability spaces to create one or more of the analytics listed above. It is important to AGO that solutions are explainable and modular, such that outputs are trusted and may be integrated into larger analytics.

System-specific Constraints

Systems present complicating factors to articulating probability spaces. These may come in the form of limitations (such as: ships may not cross land or go beyond their maximum operating speed) as well as behavioural likelihoods (such as: a ship can plausibly travel at its max speed and go to the extremity of the probability space and still arrive at the next anchor point, but it's more likely the vessel travelled between the anchor points by the shortest path and at the most economical speed).

Data Sparsity

Many processes work when used on near-perfect datasets that are complete, consistent, and high-resolution. Such data is rarely available. AGO would therefore like to understand whether the challenge can be solved under the assumption of sparse data, and at what levels of sparsity the analysis might start to degrade.

Precision-scale

Observational data that AGO contends with exist on a variety of scales that relate to the systems being analysed. Scales can include geometry (for example, elliptical error) and temporal (for example, an observation tagged to the second/minute/hour/day/etc) and tend to be system specific. An ideal investigation into this topic would involve some investigation into scale considerations for probability spaces.

Small Challenge X: Categorising Entity Movement

Overview

This challenge aims to understand the viability of automated activity categorisation and anomaly detection of entity movement data. A core component of this is understanding data requirements and whether this concept remains viable when presented with incomplete or sparse data.

Problem Statement

Entity movement data, for example AIS, is a rich source of data but is time consuming to translate into knowledge about activity. It is also difficult to identify anomalous activity without developing understanding of normal movement.

AGO would like to be able to automatically categorise movement data into activity for different types of vessels. For example, for fishing vessels, categorising the movement of those vessels into n categories, such as transit, fishing, unloading. This would allow discovery of anomalous activity conducted by vessels purporting to be fishing vessels, and would allow vessels with unknown type to be categorised as fishing vessels based on movement. It would also enable activity-oriented presentation of entity movement data.

Data Sparsity

Many processes work when used on near-perfect datasets that are complete, consistent, and high-resolution. Such data is rarely available. AGO would therefore like to understand whether the challenge can be solved under the assumption of sparse data and at what levels of sparsity the analysis might start to degrade.