AUSTRALIAN GEOSPATIAL-INTELLIGENCE ORGANISATION (AGO) ANALYTICS LAB PROGRAM

PROPOSAL BRIEFING DOCUMENT

Laura Spelbrink, Project Manager
lspelbrink@frontiersi.com.au

Phil Delaney, Chief Innovation and Delivery Officer
pdelaney@frontiersi.com.au

Document Control
Date: 16 November 2020
Version: 1.1
Reviewed By: Laura Spelbrink
Approved By: Phil Delaney

Objective

The Australian Geospatial-Intelligence Organisation (AGO) Analytics Lab Program (AGO Labs), coordinated through FrontierSI, is investigating a number of ways to better engage and work with industry. Specifically, the AGO is keen to attract a wider pool of companies and technologies to draw on for automated geospatial intelligence. The primary focus of this program is to address AGO capability challenges through a small number of short-term industry projects, with a focus upon machine learning and analytics for automated imagery analysis and product generation, including automated object classification. AGO will provided some limited commercial satellite imagery (if available) to successful participants if required.

Successful demonstrator projects may have the opportunity to progress to an additional limited operational testing phase.

What is AGO Labs?

AGO Labs is a way for AGO to rapidly assess new technology capability through a challenge-based innovation program. From this process, learnings can be gathered about the barriers and opportunities for AGO to work with an increased breadth of companies. In turn, this could lead to a long-term mechanism for AGO to access, and provide, a pipeline of activities to industry for testing new innovations and thinking in analytics.

Initial successful projects will be funded up to $100,000 for Proof of Concept projects for projects up to 6 months in length. An additional $50,000 and 3 months may be provided to some projects to undertake limited operational testing, though this process will be considered as a stage gate at the end of the Proof of Concept, and the focus on the Call for Proposals should focus on the Proof of Concept stage, rather than the operational testing stage.

There will be two rounds of AGO Labs proposal calls. One in November 2020 and one in the first half of 2021. Each is likely to fund three projects, leading to a total of at least six projects across 2020 - 2021.

At least three projects will be funded across the three Capability Challenge themes released under the first round of the AGO Labs 2020 proposal call. The three AGO Capability Challenge Topics for the first round are briefly described below. It is recommended that organisations read the full description of each challenge topic to see background, use cases and evaluation criteria, published at https://frontiersi.com.au/opportunities/agolabs/. Further details can also be found in Appendix A.

Challenge Topic 1: Where the Streets Have No Name

- 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.

Challenge Topic 2: I Can't Believe It's Not NADIR!

- This research focused 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.

Challenge Topic 3: Does This Look Fishy?

- This challenge aims to understand the viability of automated activity classification and anomaly detection over maritime track data. A core component of this is understanding data requirements and whether methodologies remain viable when presented with limited, incomplete or sparse data.

Timeline

Interested parties may submit an AGO Labs project proposal by completing a short proposal, using the template downloaded from the AGO Labs page, and emailing to the FrontierSI AGO Labs Project Manager, Laura Spelbrink, at agolabs@frontiersi.com.au **by 5:00 pm (AEDST), Friday, 11 December 2020**.

Projects will be shortlisted by a panel with representatives from both AGO and FrontierSI. Applicants may be contacted to provided further details on their proposals.

Successful Projects will be notified in December 2020. Projects are expected to start in January 2021 and delivered within a timeframe of 3 to 6 months.

Budget

At least three Proof of Concept projects will be funded across the 3 first round challenge topics, with a budget of up to $100,000. There is no requirement that additional funding will be provided to the project by applicants.

As this activity will help companies pilot technology with AGO for potential future deployment, it is expected that applicants will not operate with full commercial rates, but instead will budget the project at-cost plus 30% overheads.

For projects that are considered particularly successful and make it through a stage gate assessment between AGO and FrontierSI at the end of their Proof of Concept project will be provided up to $50,000 and an additional 3 months to undertake some limited operational testing.

The proposal response should focus on the $100,000 Proof of Concept only, operational testing will be discussed later in each successful project.

Intellectual Property

AGO may consider requests to make Background Intellectual Property in the form of data (Background IP) available to the Project, on a case by case basis, pursuant to a non-exclusive, royalty-free, worldwide licence to use the Background IP for the term of the Project and for the limited purposes of carrying out the Project.

Project Intellectual Property (Project IP) in the capability demonstrators produced during the Project will be owned by the participating partners, in agreement with the lead partner. The Australian Government is granted a perpetual licence to use any Project IP created for Defence Purposes generally (other than Commercialisation) including internal research, development, education and training.

Background IP of participating partners for the capability demonstrators is retained by the participating partners.

Technical and Data Support

Many of the challenge topics may be able to use data sources that the successful project partners already have access to. This is the preferred option for projects. Other open source data sets such as SpaceNet may also provide an option – refer to https://spacenetchallenge.github.io/datasets/datasetHomePage.html.

Satellite imagery data that have been collected by AGO may be provided to successful partners if required.

AGO will provide subject matter experts on a part-time basis to guide companies in their solution development.

Additional Resources

A webinar will be completed within two weeks of releasing the call for projects. This webinar will include a presentation of the aims of the AGO Labs program, as well as interactive Q&A with AGO and FrontierSI.

A registration link for the webinar will be provided on the AGO Labs page.

Additional questions can be directed to FrontierSI at agolabs@frontiersi.com.au

Evaluation Criteria

Projects that meet the following criteria will be considered, for further information please refer to the AGO Labs Template found at the AGO Labs website https://frontiersi.com.au/opportunities/agolabs/.

Required

- Project outputs will address the challenge.
- Evaluation of the intended approach.
- Ability of AGO to access and test outputs iteratively during the project
- Value for money
- Outcomes that can be operationalised beyond the life of the project

Constraints and Requirements

- The Australian Government contribution will be limited to $100,000 per project for the Proof of Concept
- Projects should be completed in a period of approximately a 3 to 6 months
- A final project report is required at the end of the AGO Labs project
- The lead organisation needs to be a company from Australia or New Zealand

The final project report should outline key findings and recommendations to FrontierSI and AGO. This report will include lessons learned and suggest options for industry partners to engage and work more effectively with AGO. It will also suggest options for AGO to engage and work more effectively with industry in an AGO Labs style program.

Appendix A – Challenge Topics

Challenge Topic 1: Where the Streets Have No Name

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 could potentially include paths, roads, rail, airports and ports – **however, the focus for this challenge is the current state of road networks in a rural environment**. 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. Network generation should be able to be updated when new imagery data becomes available, and the amount of time to perform this update needs to be quantifiable.

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**.

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 how it translates to regional conditions (more detail in the environment characterisation section).

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 characterization

The primary source-data to be used in this challenge are high-resolution commercial satellite imagery. For each transport network type there are associated range and distribution questions that participants may like to review, 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.

Note that the model to be developed can employ historical data incorporating new data. But this is not a guaranteed method as the real value is inferring state rather than relying on historical data (that may have been inaccurate prior to the event).

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?

Use Cases:

The following stories are provided as guidance to challenge responses in order of priority but should not be considered definitive - AGO is willing to entertain modifications to the following list or additional use cases:

1. **As a disaster relief analyst, I want to** be provided with vector road network (based on the latest imagery) that identifies and highlights transport blockages and locations (and potentially ways around blockages). Note: categorisation of blockage type and length is a stretch goal however this is a lower priority as it is expected that automatic identification of blockages will result in manual confirmation by an analyst.
2. **As a data scientist, I want to** be able to quantify the differences of road network generation between urban and rural areas. For example, the impact of obscuration and varying surface types on the successful generation of road networks.
3. **As an analyst, I want to** have the ability for new road networks to be automatically detected.
4. **As a disaster relief analyst, I need** the road network to have associated speed limits and road widths identified, to assist in mobility planning assessments. There would also be interest in categorization of road surface type or categorisation of the ability of a road to carry certain classes of vehicles - but this should be undertaken as a lower priority.

Challenge Topic 2: I Can't Believe It's Not NADIR!

Overview:

This research focused 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 AGO Labs 2019, 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 AGO Labs 2019 low-cost object detection challenge implemented GANs to meet the challenge, AGO is not prescriptive in the approaches that may be used. AGO is aware of the investigations performed under the CosmiQ Works SpaceNet 4 competition but has some key differences including:

- Goal is to generate models using sparse data sets and infrequently observed objects.
- Focus on accurate object detection rather than object segmentation.

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 characterization

The primary source-data to be used in this challenge are high-resolution commercial satellite imagery.

- This challenge has a similar design to the "Low Cost Object Models" challenge topic performed for AGO Labs 2019. The low-cost element is maintained while an additional variable (off-nadir images) is added. For the previous challenge Aircraft were selected as an object to detect. This object could be used again but the focus is on aircraft as a general category not different types of aircraft.
- Avoid the introduction of other variables in this challenge other than obliquity angle (i.e. no need to introduce weather variables).
- Based on successful vendor's proposal AGO may be able to provide suitable off-nadir imagery, noting that the intent is to work with limited data volumes.

Use Cases:

This challenge focuses at looking at the performance/other costs that obliqueness or off-nadir may have on the model. The following stories are provided as guidance to challenge responses in order of priority but should not be considered definitive - AGO is willing to entertain modifications to the following list or additional use cases:

1. **As a data scientist, I want to know** the suitable techniques to build object identification models optimised for a broad range of collection geometries and understand the associated trade-offs **so that** I can build models that are generalisable to many scenarios.
2. **As a data scientist, I want to know** the optimal degree of obliquity for my training data **so that** I understand how obliquity affects model performance.
3. As a **data scientist, I want to know** the suitable techniques to build object identification models optimised for a narrow range of collection geometries and understand the associated trade-offs **so that** I can build models that are optimised for specific scenarios.
4. **As a data scientist**, **I want to** know how a model trained on a narrow range of collection geometries will perform on new data from different collection geometries **so that** I understand the expected performance on different data.
5. **As a data scientist, I want to understand** the advantages and trade-offs between object identification models trained on a broad range of collection geometries and those trained with specific collection geometries **so that** I can select the most appropriate model development strategy to answer an intelligence question. E.g. 3 different models for nadir, off-nadir, or very-off-nadir, vs one model generalised to all of them.
6. **As a data scientist**, **I want to** know whether the quantity of data required to train a model increases or decreases when the off-nadir variable is added **so that** I understand the volume of data required to train computer vision models.
7. **As a collection strategist**, **I want to** understand how collection geometry influences detection algorithms for different objects **so that** I can advise customers on optimal collection parameters for acquiring data to train and implement automated detection models.
8. **As an imagery analyst**, I want to understand the performance of machine detections of objects in satellite imagery from different collection geometries **so that** I can better interpret and analyse different data sets. And **as a data scientist**, I want to be able to explain concepts of detection probability/accuracy to the GEOINT analysts using my models **so that** they use it properly and trust it more.

Challenge Topic 3: Does This Look Fishy?

Overview:

This challenge aims to understand the viability of automated activity classification and anomaly detection over maritime track data. A core component of this is understanding data requirements and whether methodologies remain viable when presented with limited, incomplete or sparse data.

Problem Statement:

Commonly available maritime track data is a rich source of information but is time consuming to translate into knowledge about activity. It is also difficult to identify anomalous activity without first developing an understanding of normal movement.

AGO would like to be able to automatically classify activity type and vessel type, based on maritime track data. For example, identifying fishing vessels, then classifying the movement of those vessels into several categories, such as transit, fishing, unloading. This would allow discovery of anomalous activity conducted by vessels and enable activity-oriented presentation and analysis of entity movement data.

Data and Data Sparsity

It is expected that this challenge would make use of widely accessible data such as Automatic Identification System (AIS).

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 constrained data and at what levels of constraint the analysis might start to degrade. For example, the study could assess the degradation of results if the original data set was reduced by 90%

The focus of this challenge should be the application of automatic activity classification under data constrained conditions. As such, outcomes should focus on quantifying the effects of data constraints on classification performance. For this reason, the geographic extent and category of activity has been left open-ended for this challenge.

Use Cases:

The following stories are provided as guidance to challenge responses but should not be considered definitive - AGO is willing to entertain modifications to the following list or additional use cases:

1. **As an analyst, I want to** be able to classify vessel type based on movement information alone, **so that** I can understand vessel behaviour.
    1. For example, this classification may be different to that captured in the base data when vessels have inaccurate or incomplete information.
2. **As an analyst, I want to** identify vessel sub-types (that have different 'normal activity' such as line fishing vessels vs trawlers) so that I can analyse these sub-types differently.
3. **As an analyst, I want to** identify vessels whose behaviour (e.g. speed, location, track shape) deviates from the normal activity for vessels of its type so that I can focus on vessels that are anomalous.
4. **As an analyst, I want to** classify vessel movement into different behaviours (such as transit, survey, and trawling) **so that** I can:
    1. Filter out certain types of behaviours (such as transiting) to focus/extract on other activities.
    2. Identify new or unexpected vessel routes and generate summary statistics about that activity.
    3. Delineate between behaviours that look similar but have vastly different significance (such as search and rescue vs hydrographic survey).
5. **As an analyst, I want to** use the movement of vessels to identify other events.