# Spot the GEO Satellites: From Dataset to Kelvins SpotGEO Challenge

Bo Chen[1] Daqi Liu[1] Tat-Jun Chin[1] Mark Rutten[2] Dawa Derksen[3]
Marcus Märtens[3] Moritz von Looz[3] Gurvan Lecuyer[3] Dario Izzo[3]
[1]The University of Adelaide [2]InTrack Solutions [3]European Space Agency
{bo.chen, daqi.liu, tat-jun.chin}@adelaide.edu.au mark@intrack.solutions
{dawa.derksen, marcus.maertens, moritz.von.looz, gurvan.lecuyer, dario.izzo}@esa.int

## Abstract

*The Geosynchronous Equatorial Orbit (GEO) is home to many important space assets such as telecommunication and navigational satellites. Monitoring Resident Space Objects (RSOs) in GEO is a crucial aspect in achieving Space Situational Awareness (SSA) and in protecting critical space assets. However, ground-based GEO object detection is challenging due to the extreme distance of the targets, as well as nuisance factors including cloud coverage, atmospheric/weather effects, light pollution, sensor noise/defects, and star occlusions. The Kelvins SpotGEO Challenge is designed to establish to what extent images coming from a low-cost ground-based telescope can be used to detect GEO and near-GEO RSOs solely from photometric signals that are without any additional meta-data. At the same time, the SpotGEO dataset also addresses the lack of publicly available datasets from a computer vision perspective on the satellite detection problem; by assembling and releasing such a dataset, we hope to spur more efforts on the optical detection of RSOs and enable objective benchmarking for existing and future methods. In this work, we present details of the SpotGEO dataset development, challenge design, evaluation metric, and result analysis.*

## 1. Introduction

Space technology has virtually shaped many aspects of our modern life, including communications, navigation and meteorology–to name a few. Economic losses due to the persistent disruption of space-based services will dwarf the trillions of dollars already invested into existing space assets. Unfortunately, increasing space utilisation and the worsening pollution by space debris make the prospect of disruption due to collisions realistic.

To alleviate the risk of disruption, such as collision between space assets and debris, it is crucial to develop Space Situational Awareness (SSA)[11], which is the broad task of building up-to-date information of the space environment. An important aspect of SSA is detecting and confirming the existence of tens of thousands of resident space objects (RSO) such as debris, satellites and space stations that are currently orbiting the Earth. Realistically, the detection of RSOs will be achieved using a variety of approaches such as ground-based radar [9, 20], ground-based telescopes [21, 16] and satellite-based observers [10, 14]. These approaches each have their pros and cons and complement each other in a holistic SSA system.

In this work we focus on the ground-based approach and ask the following question: can a low-cost ground-based camera be used to detect satellites orbiting the GEO? In order to answer this question, the Kelvins SpotGEO Challenge was held jointly by the Advanced Concept Team (ACT) of the European Space Agency (ESA) and the University of Adelaide, to invite the worlds' machine learning and computer vision experts to develop advanced and effective methods for this particular task. The challenge involves detecting RSOs in or close to the GEO from ground-based images. Developing SSA in GEO is of tremendous importance since that orbital belt contains some of the most critical space assets, such as satellites for communications, meteorology, navigation, etc.

### 1.1. Why is this problem difficult?

A fundamental reason behind the difficulty of this problem is the extreme distance between the observer and the target objects. In the case of the employed telescope, having an angular pixel size of about 4.5 arc seconds and the distance between the ground level and GEO, each pixel corresponds to an arc length of about 800m at GEO. Hence, targets of interest are no larger than 1 pixel in area. With atmospheric distortion and the long exposure time, the received photons from an orbiting object are then smeared over a few pixels; however, this has the effect of dimming the observed object. An example image is provided in Figure 1.

Other nuisance factors also contribute to the difficulty of the problem: cloud cover, atmospheric/weather effects, light pollution, sensor noise/defects, star occlusion (when a

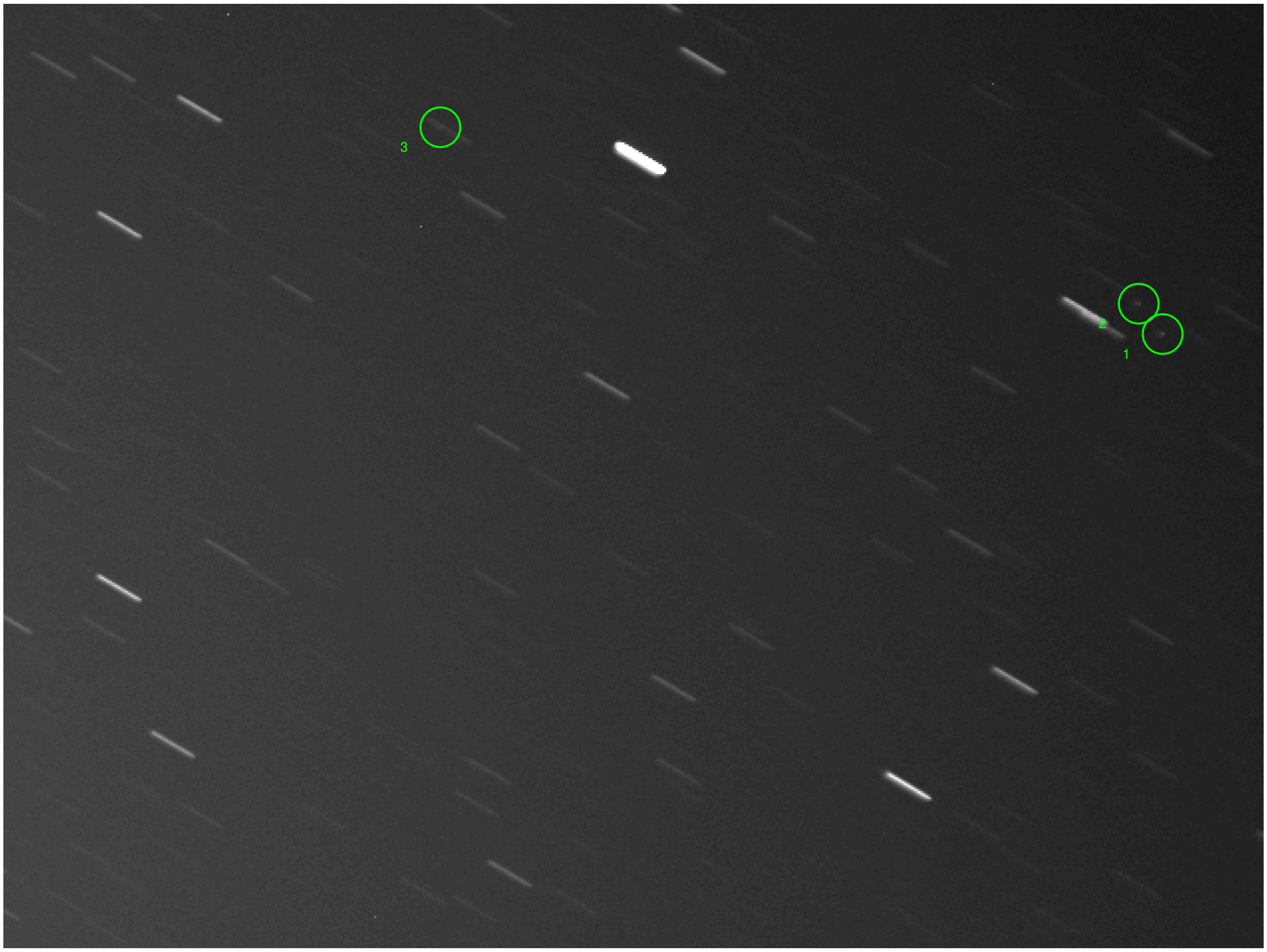


Figure 1. An example image of the SpotGEO dataset. Three GEO satellites are captured within the image, which are marked by the green circles. Note that object 3 is partially occluded by a star streak.

background star happened to cross the image coordinates of an orbiting object, as shown in Figure 1) and in rare cases, orbital manoeuvres conducted by an active GEO satellite during capture.

### 1.2. Our contributions

A number of algorithms [21, 16, 7, 13, 17, 12, 18] have already been developed for similar problems, but it is currently unclear what the best approaches are. This competition served as an invitation to researchers and practitioners in relevant fields from around the world to test our current best capability for this important problem. From the solutions submitted by the competition participants, we have summarised a 3-step pipeline that works effectively for this task and is adopted by most participants, which is discussed in Section 4.4.

At the same time, the SpotGEO Challenge also addresses the lack of publicly available datasets from a computer vision perspective on the satellite detection problem; by assembling and releasing such a dataset and developing the evaluation metric for the task, we hope to spur more efforts on automated pipelines for the optical detection of space objects and enable objective bench-marking of already existing methods.

## 2. Related works

There has been long-lasting efforts in the space domain looking into the RSOs detection problem. Earlier work by Schildknecht *et al*. [17] proposed a ground-based optical detection algorithm. This method firstly applies siderostatic tracking to align a sequence of captured images. It then uses a reference frame to either produce masks of all objects to filter other images, or subtract all other frames with the reference frame, before it scans for objects with different characteristics than stars. Yanagisawa *et al*. [21] propose a method that utilises a short image sequence with long exposure time, and image processing algorithms are used for denoising. They then exhaustively search all of the linear combinations across the image sequence using field-programmable gate array (FPGA) acceleration. However, a limitation of this technology when deployed in space is that it could suffer failures introduced in the FPGA due to the radiation environment.

More recently, Šára *et al*. [16] proposed a framework that first registers the images to a common image frame and then uses randomised heuristic for linear structure detection, followed by a scoring system to select valid tracks. Liu *et al*. [13] adopted the framework of [16] and proposed a novel deterministic topological sweep algorithm [8] which boosted the efficiency of the track extraction step.

Another line of work tackles the GEO object detection problem under the track-before-detect (TBD) treatment [5]. The key idea of TBD is to improve the Signal-to-Noise Ratio (SNR) of weak targets by accumulating spatial and temporal signals in order to increase detection confidence. Under similar settings of [16, 7, 13], Davey *et al*. [3, 4] developed a histogram probabilistic multi-hypothesis tracking (H-PMHT [19]) method for detecting RSOs. However, the limitation of their methodology is that relatively long image sequences are required in order to achieve satisfactory results.

# 3. The SpotGEO dataset

The purpose of this challenge for us was to establish to what extent images coming from a ground-based low-cost telescope can be used for detecting RSOs without utilising any information from additional meta-data. For this reason we release our dataset using PNG image format rather than the more commonly used FITS format in astronomy, as to give this competition a stronger focus on generic purpose vision algorithms and discourage the use of object catalogues.

Due to the difficulties of the problem described in Section 1.1, the detection of GEO satellites is specifically formulated to be based on a sequence of 5 images or frames, instead of a single image. This provides more information sources for the detector to better handle noisy/missing signals, as well as the possibility of utilising the geometric structure of signals in the detection, which will be described in Section 3.2.

The SpotGEO dataset consists of 6,400 sequences, each of which has 5 frames. Each frame is a grayscale image of size $640 \times 480$ pixels. To facilitate machine learning-based methodologies, we split the dataset into training and testing subsets, which makes up 20% and 80% of the 6400 sequences, respectively. Since the competition has concluded, the dataset including annotations for both training and testing subsets has been released and is publicly available on Zenodo [2].

## 3.1. Data acquisition

The dataset images were acquired using a ground-based, low-cost CMOS sensor during nighttime. The specific data acquisition approach used for this challenge is illustrated in Figure 2. Each capture instance yielded a sequence of 5 frames. For each frame, a 40-second exposure was used while the camera was kept static on the ground (equivalently, the camera was rotating at sidereal rate during exposure). To simulate a sky-sweeping scenario, after each frame in one instance had been recorded, the camera was slightly rotated to observe a different field of view (FOV) which maintains a significant proportion of overlapping with the previous FOV. This camera motion is constant between two consecutive frames within a sequence. Eventually this methodology resulted in sequences of five frames such as shown in Figure 1, where, as an example, 3 objects are also clearly marked.

Note that under the adopted capture regime, stars appear as streaks, while GEO or near-GEO objects mostly appear as blobs or shorter streaks since they are (mostly) static relative to the observer.

## 3.2. Data annotation

For each frame in each sequence, groundtruth object locations are annotated manually via careful visual inspection. The coordinates of an object within the FOV of an image is given by $(x, y)$, where

$$x \in [-0.5, 639.5], \text{ and} \quad (1)$$
$$y \in [-0.5, 479.5], \quad (2)$$

since the image sizes are $640 \times 480$ pixels and we correspond the centre of each pixel to non-negative integer coordinates.

**The definition of a valid object** We provide the exact definition of an object in accordance to the annotation process. By object we mean GEO or near-GEO orbiting objects with consistently detectable presence in the FOV of the images. As explained in Section 3.1, such objects were imaged as blobs or short streaks. Note that we are not interested in low Earth orbit (LEO) objects that occasionally appeared in the FOV of the frames and were imaged as very long streaks (longer than the star streaks). Hence, LEO objects are not considered valid objects in the dataset. Also, blob-like artefacts due to sensor noise or bright pixels due to sensor defects are not considered as objects. During our labelling procedure, an object was considered consistently detectable if it appeared in at least 3 out of 5 frames in the sequence. If an object appeared in only 1 or 2 frames in the sequence, it is not labelled as a valid object and thus is not a target of detection.

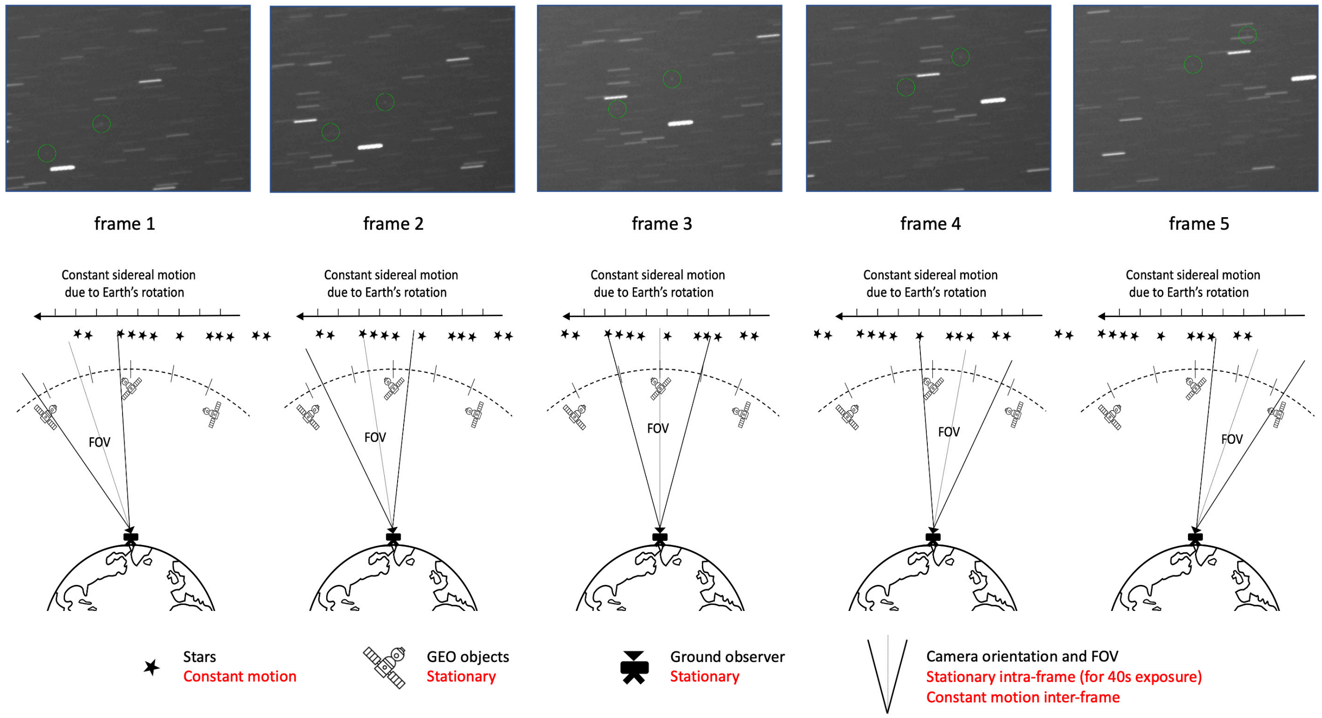


Figure 2. A conceptual illustration of the data capturing process and an example result where GEO objects are marked with green circles.

We ensured the following properties of the dataset during its development.

1. Each frame in a sequence has the same number of objects, since we have maintained that each valid object lies in the common FOVs of all frames in the sequence.
2. A valid object should form a trajectory across the sequence according to GEO orbital motion and the constant inter-frame camera motion within each sequence.

Note that the presence of a valid object in a sequence does not mean that the object is visible in all frames, since it can be occluded or is too dim to be observable in a subset of the frames.

The above properties are designed to facilitate sequence-based detection instead of frame-based detection. This prior knowledge of object trajectory across frames in a sequence can be utilised to assist filtering noisy signals and inferring hidden ones, as exemplified in Figure 3. Algorithms should thus be able to estimate the coordinates of a detected object across all frames, including frames where it is not observable.

# 4. The SpotGEO challenge

In this section we first describe the competition design and the evaluation metric for ranking participants, followed by competition results and its analysis.

## 4.1. Competition design

The competition was hosted on the Kelvins competition website, a platform created by the Advanced Concepts Team of the European Space Agency specifically for space related competitions. During the competition, participants were given the training set with groundtruth labels for developing their solutions, and the test set without labels for generating predictions.

To encourage active engagement, a leader board automatically calculates indicative scores once a participant submits a prediction, and lively ranks participants based on their best submission thus far. To ensure fairness and prevent extraction of groundtruth information from the leader board, each participant was allowed two submissions every 24 hours, and the leader board scores were based on a random half of the test set only. After the submission period ended, scores based on the full test set were calculated, which were used for the final ranking.

## 4.2. Evaluation metric

For each submission we produce two scores: an $F_1$ score and a Mean Squared Error (MSE) score. Participants are ranked based on both $1 - F_1$ and MSE, in the spirit of Kelvins' "reach the absolute zero error" motto. The official evaluation toolkit is available in [6].

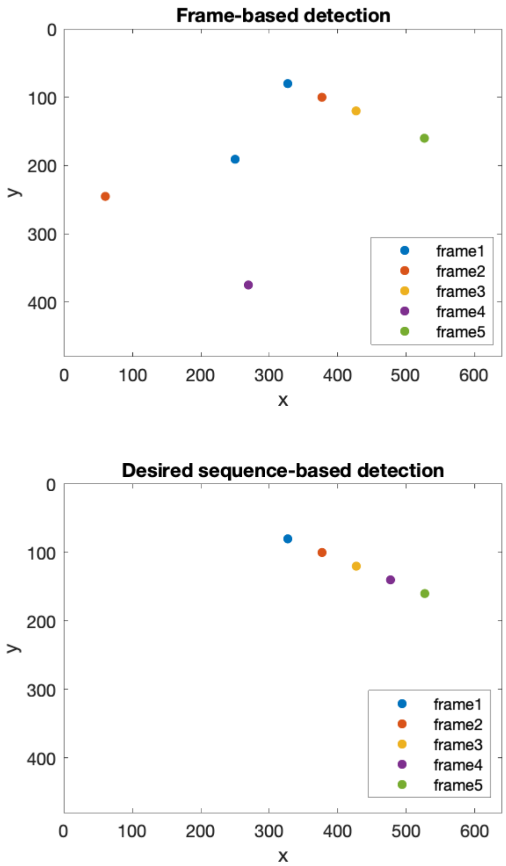


Figure 3. Examples of frame-based and sequence-based detection results. The object trajectory can be utilised to handle nuisance factors and improve detection results.

#### 4.2.1 Evaluating detection results for one sequence

To evaluate the predictions of one sequence against the groundtruth, the following inputs are required:

- Groundtruth object locations $\mathcal{Y} = \{\mathcal{Y}^f\}_{f=1}^{5}$ in the 5 frames of a sequence, where each $\mathcal{Y}^f = \{\boldsymbol{y}_j^f\}_{j=1}^{N}$ contains the coordinates of $N$ objects in frame $f$. Note that $N$ can be 0, in which case $\mathcal{Y}^f$ for all $f$, and hence $\mathcal{Y}$, are empty.
- Predicted object locations $\mathcal{X} = \{\mathcal{X}^f\}_{f=1}^{5}$ in 5 frames, where each $\mathcal{X}^f = \{\boldsymbol{x}_i^f\}_{i=1}^{M_f}$ contains the coordinates of $M_f$ objects in frame $f$. Note that $M_f$ is allowed to vary across frames. Also, $M_f$ can be 0, in which case $\mathcal{X}^f$ is empty.
- Predefined matching distance threshold $\tau$ and tolerance distance $\epsilon$, such that $0 \leq \epsilon < \tau$.

**Matching** For a given frame $f$, a *one-to-one matching* between $\mathcal{X}^f$ and $\mathcal{Y}^f$ is first obtained. Assuming for now $M_f \leq N$, the matching is encapsulated in a binary matrix

$$\mathbf{H}^f \in \{0,1\}^{M_f \times N} \tag{3}$$

with the following constraints

$$\sum_{j=1}^{N} \mathbf{H}_{i,j}^f = 1, \quad \forall i, \tag{4}$$

to the rows and

$$\sum_{i=1}^{M_f} \mathbf{H}_{i,j}^f \leq 1, \quad \forall j. \tag{5}$$

to the columns.

In words, each point in $\mathcal{X}^f$ must be matched uniquely to a point in $\mathcal{Y}^f$; not all points in $\mathcal{Y}^f$ need to be matched to a point in $\mathcal{X}^f$, but those that are matched do so uniquely.

The matching is solved via the *minimum weighted unbalanced assignment problem*

$$\underset{\mathbf{H}^f}{\text{argmin}} \quad \sum_{i=1}^{M_f} \sum_{j=1}^{N} \mathbf{H}_{i,j}^f \delta(\boldsymbol{x}_i^f, \boldsymbol{y}_j^f) \tag{6}$$

subject to the constraints (4) and (5), where function $\delta$ implements the truncated distance

$$\delta(\boldsymbol{x}_i^f, \boldsymbol{y}_j^f) = \begin{cases} \|\boldsymbol{x}_i^f - \boldsymbol{y}_j^f\|_2, & \text{if } \|\boldsymbol{x}_i^f - \boldsymbol{y}_j^f\|_2 \leq \tau, \\ \ell, & \text{otherwise.} \end{cases} \tag{7}$$

Here $\ell$ is a sufficiently large positive number, *e.g.*, the diagonal pixel length of the image. The problem can be solved efficiently via the Hungarian algorithm, maximum flow, linear programming, *etc*. Examples of the assignment problem and their one-to-one matching solutions are given in Figure 4.

If $M_f > N$, the roles of $\mathcal{X}^f$ and $\mathcal{Y}^f$ are swapped and the same problem can be solved to perform the matching. This swap is transparent to most algorithms since only simple changes are needed (e.g., adding dummy points at infinity so that $\mathbf{H}^f$ is always square). If $N = 0$ or $M_f = 0$, then $\mathbf{H}^f$ = NULL.

The matching procedure above is conducted for all frames $f = 1, \ldots, 5$.

**True positives, false negatives and false positives** Let $\mathcal{I}_f = \{1, ..., M_f\}$ and $\mathcal{J} = \{1, ..., N\}$. After the matching matrix $\mathbf{H}^f$ for frame $f$ is solved, if $\mathbf{H}^f$ is not NULL, the set of correctly detected objects for frame $f$ is

$$\mathcal{TP}^f = \left\{ (i,j) \in \mathcal{I}_f \times \mathcal{J} \;\middle|\; \mathbf{H}_{i,j}^f = 1, \; \delta(\boldsymbol{x}_i^f, \boldsymbol{y}_j^f) \leq \tau \right\}. \tag{8}$$

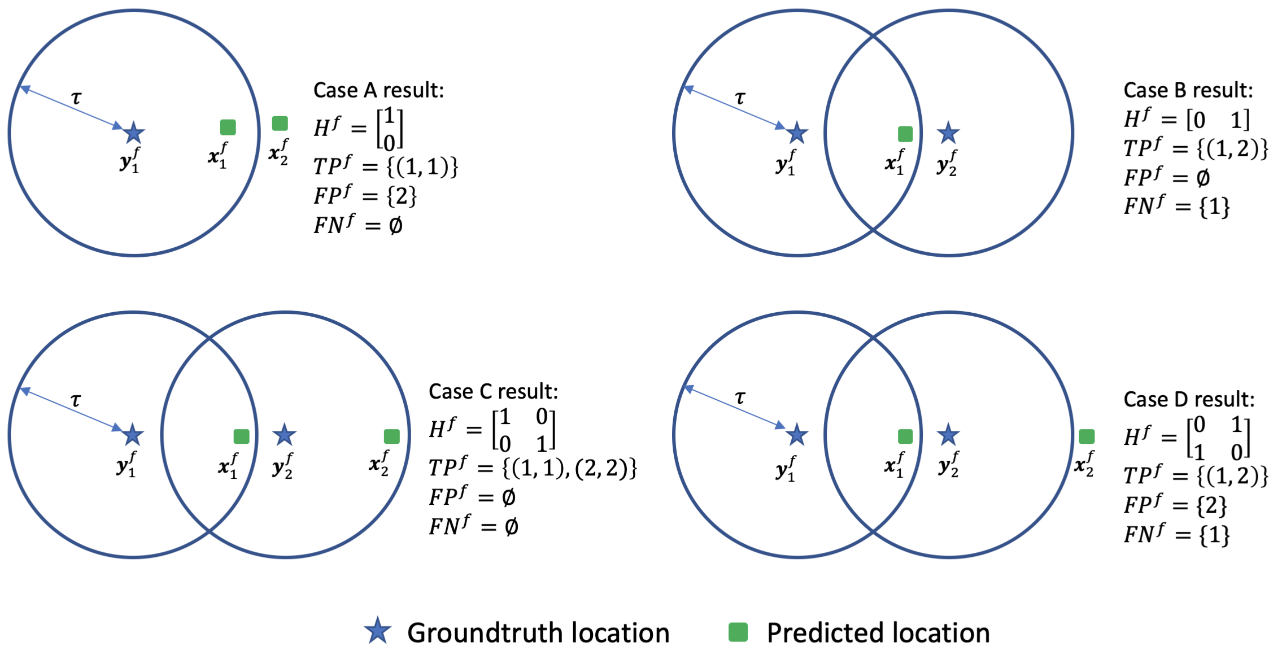


Figure 4. Examples of the assignment problem and one-to-one matching solutions. Note the different results of Case C and Case D. Because of the truncated distance defined in (7), $\boldsymbol{x}_2^f$ is no longer matched to $\boldsymbol{y}_2^f$ but to $\boldsymbol{y}_1^f$.

The set of missed objects for frame $f$ is

$$\mathcal{FN}^f = \left\{ j \in \mathcal{J} \;\middle|\; \left(\sum_{i=1}^{M_f} \mathbf{H}^f_{i,j} = 0\right) \vee \left(\sum_{i=1}^{M_f} \mathbf{H}^f_{i,j}\,\delta(\boldsymbol{x}_i^f, \boldsymbol{y}_j^f) > \tau\right) \right\}. \tag{9}$$

The set of false predictions for frame $f$ is

$$\mathcal{FP}^f = \left\{ i \in \mathcal{I}_f \;\middle|\; \left(\sum_{j=1}^{N} \mathbf{H}^f_{i,j} = 0\right) \vee \left(\sum_{j=1}^{N} \mathbf{H}^f_{i,j}\,\delta(\boldsymbol{x}_i^f, \boldsymbol{y}_j^f) > \tau\right) \right\}. \tag{10}$$

We stipulate that if $\mathbf{H}^f = \text{NULL}$, then

$$\mathcal{TP}^f = \emptyset, \tag{11}$$

$$\mathcal{FN}^f = \{j\}_{j=1}^{N}, \tag{12}$$

and

$$\mathcal{FP}^f = \{i\}_{i=1}^{M_f}. \tag{13}$$

The accounting procedure above is conducted for all frames $f = 1, \ldots, 5$. The true positive, false negative and false positive values for the sequence are then given by

$$TP = \sum_{f=1}^{5} |\mathcal{TP}^f|, \tag{14}$$

$$FN = \sum_{f=1}^{5} |\mathcal{FN}^f|, \tag{15}$$

$$FP = \sum_{f=1}^{5} |\mathcal{FP}^f|. \tag{16}$$

**Regression error** To handle the possibility of a tie in $F_1$ score, we provide another measure of detection accuracy, the MSE of localising the objects.

Given the results of the procedures above, if not all $\mathcal{TP}^f$, $\mathcal{FN}^f$ and $\mathcal{FP}^f$ are empty for frame $f$, the sum of squared error (SSE) for frame $f$ is

$$SSE^f = \sum_{(i,j)\in\mathcal{TP}^f} \pi(\boldsymbol{x}_i^f, \boldsymbol{y}_j^f) + \sum_{j\in\mathcal{FN}^f} \tau^2 + \sum_{i\in\mathcal{TP}^f} \tau^2, \tag{17}$$

where

$$\pi(\boldsymbol{x}_i^f, \boldsymbol{y}_j^f) = \begin{cases} 0, & \text{if } \|\boldsymbol{x}_i^f - \boldsymbol{y}_j^f\|_2 \le \epsilon, \\ \|\boldsymbol{x}_i^f - \boldsymbol{y}_j^f\|_2^2, & \text{otherwise.} \end{cases} \tag{18}$$

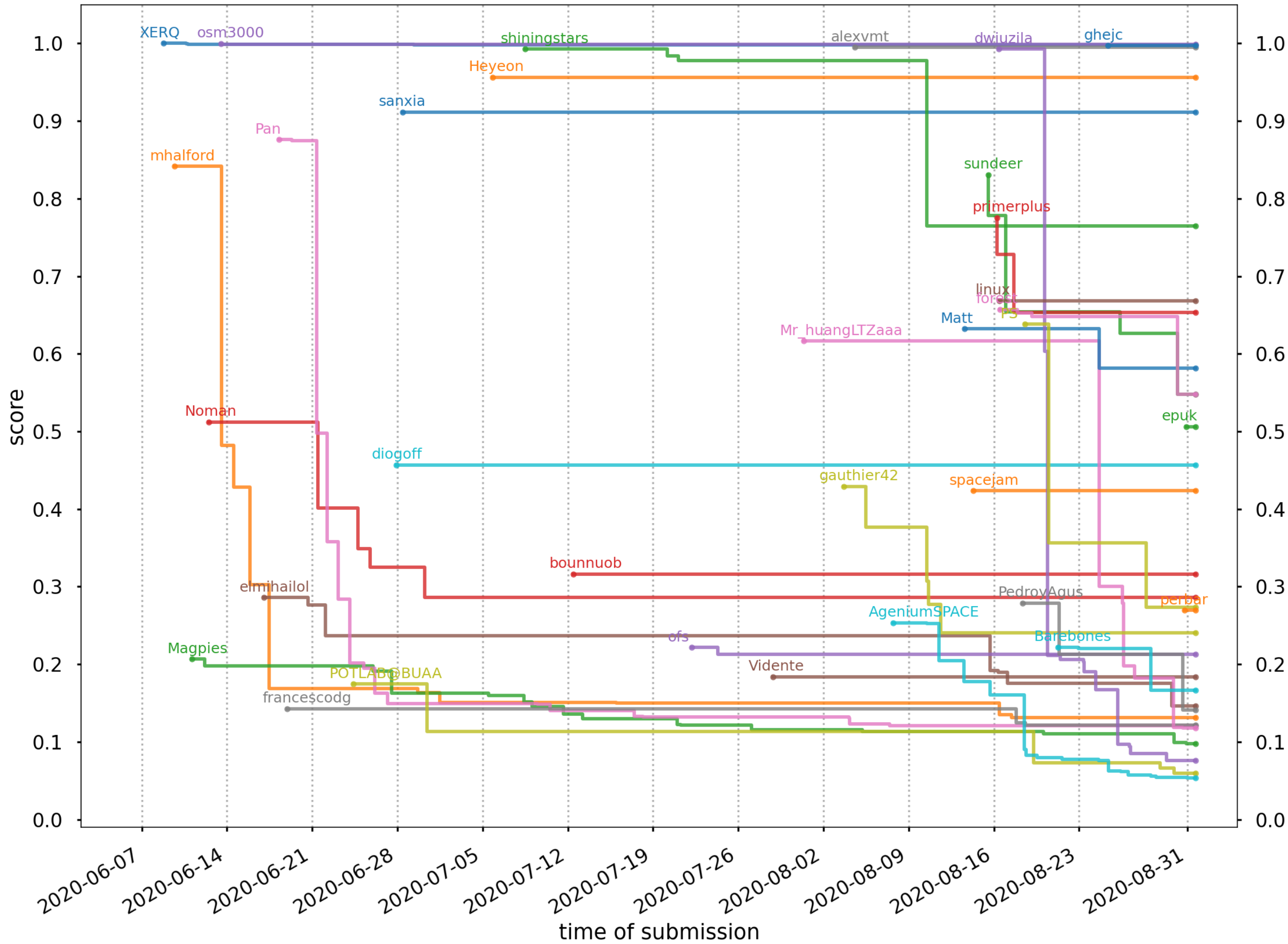


Figure 5. SpotGEO race: the evolution of best scores of each team during the submission period.

In words, $SSE^f$ accumulates the squared distance (upper bounded by $\tau^2$) between predicted object locations and ground truth object locations in frame $f$, with a tolerance of $\epsilon$ to account for inaccuracies in manual labelling. Further, each missed object and false detection respectively contribute a constant squared error of $\tau^2$ to $SSE^f$.

If $\mathcal{TP}^f$, $\mathcal{FN}^f$ and $\mathcal{FP}^f$ are all empty sets, we stipulate that $SSE^f = 0$.

The SSE for the sequence is thus

$$SSE = \sum_{f=1}^{5} SSE^f, \tag{19}$$

and the MSE for the sequence is

$$MSE = \begin{cases} \frac{SSE}{TP+FN+FP}, & \text{if } SSE \neq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

#### 4.2.2 Evaluating detection results for the whole test set

Let there be $K$ sequences of 5 frames each in the test set. Denote by $TP_k$ the true positive value for the $k$-th sequence computed according to (14) and similarly for $FN_k$ and $FP_k$. The overall precision is

$$P = \frac{\sum_{k=1}^{K} TP_k}{\sum_{k=1}^{K} TP_k + FP_k}; \tag{21}$$

the overall recall $R$ is

$$R = \frac{\sum_{k=1}^{K} TP_k}{\sum_{k=1}^{K} TP_k + FN_k}; \tag{22}$$

and the $F_1$ score is thus given by

$$F_1 = 2\frac{PR}{P+R}. \tag{23}$$

Denote by $SSE_k$ the SSE for the $k$-th sequence computed according to (19). The overall regression MSE is thus

$$MSE = \frac{\sum_{k=1}^{K} SSE_k}{\sum_{k=1}^{K} TP_k + FN_k + FP_k}. \quad (24)$$

#### 4.2.3 Ranking methodology

Participants were ranked primarily based on their submitted predictions using $1 - F_1$ (23). In the event of ties (i.e., a number of teams having the same $F_1$), regression MSE (24) would be used as a tie breaker.

### 4.3. Results

The competition attracted 54 teams globally to participate, out of which 33 teams have made valid submissions. Figure 5 shows the evolution of scores of each participant during the competition.

We asked participating teams to submit a brief description of their methodology, as a condition for their scores to enter the final ranking. This was to prevent teams solving the detection task using manual labeling. The final ranking is presented in Table 1.

| Rank | Participant Name | $1 - F_1$ | MSE |
|---|---|---|---|
| 1 | AgeniumSPACE | 0.0517 | 33838.99 |
| 2 | POTLAB@BUAA | 0.0557 | 30541.73 |
| 3 | dwiuzila | 0.0711 | 41198.46 |
| 4 | Magpies | 0.0957 | 48919.92 |
| 5 | Mr_huangLTZaaa | 0.1158 | 62021.81 |
| 6 | francescodg | 0.1211 | 65772.46 |
| 7 | mhalford | 0.1230 | 69566.91 |
| 8 | PedroyAgus | 0.1339 | 70104.97 |
| 9 | elmihailol | 0.1389 | 83172.81 |
| 10 | Barebones | 0.1634 | 105518.42 |
| 11 | gauthier42 | 0.2357 | 133979.21 |
| 12 | perbar | 0.2605 | 118436.21 |
| 13 | Matt | 0.5800 | 307932.17 |
| 14 | alexvmt | 0.9946 | 510652.52 |

Table 1. Final ranking for the SpotGEO challenge. The full list including unranked teams can be found in [1].

### 4.4. Solution analysis

Most of the submitted solutions generally follow a 3-step pipeline:

- pre-processing step - remove noise such as stars, cloud, light pollution, *etc.*;
- detection step - produce candidate objects;
- post-processing step - exploit the geometric structure to prune false detection or interpolate missing ones.

In particular, we provide brief summaries of the techniques employed by the best two teams, as they represent two distinctly different approaches yet achieved very close scores.

The solution of team AgeniumSPACE used the 3-step pipeline described above. They first perform background removal using $L_1$-spline, and estimation of star shifting using a hand-crafted star descriptor and RANSAC. In the second step, an ensemble of 10 U-Net [15] style Convolution Neural Nets (CNNs) are trained to predict object locations. Lastly, the predicted object candidates are post-processed to remove false positives and recover missing objects via line detection and trajectory filling.

The second place winning team POTLAB@BUAA tackles the problem from a quite different perspective which employs a non-learning-based approach. Their pipeline is mainly divided into two stages. In the first stage, the SNR is calculated for each pixel in each input image. Pixels with a SNR higher than a certain threshold are then selected. Lastly, candidate stars and candidate satellites are extracted via connecting adjacent selected pixels. In the second stage, they firstly estimate the inter-frame star shifts from the star candidates. This estimation is then used to filter out candidate satellites that are actually stars. The filtered candidate satellites are used to estimate the satellite shifts between consecutive frames. Finally, satellites are confirmed based on their agreement to the estimated satellite shifts.

The key difference between the two approach is the employment of CNNs for detecting candidate satellites. Albeit deep learning being a popularly employed technique amongst participating teams, the success of team POTLAB@BUAA shows that sheer thresholding and robust fitting can perform equally well as or even better than learning-based approaches in this particular problem.

## 5. Conclusion

This paper summarises the design, evaluation metric and results of the Kelvins SpotGEO Challenge, as well as the development of the SpotGEO dataset for this competition. Through the results of this competition, we learn that images obtained from low-cost ground-based cameras can be used to detect orbiting objects fairly well. This provides a direction of research to further enrich our tool box for achieving SSA.

## Acknowledgements

We gratefully acknowledge support from the Australian Research Council Linkage Project LP160100495, SmartSat CRC and Australian Institute for Machine Learning.

## References


[1] Kelvins spotgeo challenge. https://kelvins.esa.int/spot-the-geo-satellites/home/. 8

[2] Bo Chen, Daqi Liu, Tat-Jun Chin, Mark Rutten, Dawa Derksen, Marcus Märtens, Moritz von Looz, Gurvan Lecuyer, and Dario Izzo. Spotgeo dataset. *Zenodo*, https://doi.org/10.5281/zenodo.4432143, *2021.* 3

[3] *Samuel J Davey, Travis Bessell, Brian Cheung, and Mark Rutten. Track before detect for space situation awareness. In* International Conference on Digital Image Computing: Techniques and Applications. *IEEE, 2015.* 3

[4] *Samuel J Davey and Han X Gaetjens.* Track-Before-Detect Using Expectation Maximisation. *Springer, 2018.* 3

[5] *Samuel J Davey, Mark G Rutten, Neil J Gordon, M Mallick, V Krishnamurthy, and BN Vo. Track-before-detect techniques. In* Integrated tracking, classification, and sensor management: theory and applications. *Wiley Online Library, 2012.* 3

[6] *Dawa Derksen, Marcus Märtens, Moritz von Looz, Gurvan Lecuyer, Dario Izzo, Bo Chen, Daqi Liu, Tat-Jun Chin, and Mark Rutten. Spotgeo starter kit. Zenodo,* https://doi.org/10.5281/zenodo.3874368, *2020.* 4

[7] *Huan N Do, Tat-Jun Chin, Nicholas Moretti, Moriba K Jah, and Matthew Tetlow. Robust foreground segmentation and image registration for optical detection of geo objects.* Advances in Space Research, *64(3):733–746, 2019.* 2, 3

[8] *Herbert Edelsbrunner and Leonidas J Guibas. Topologically sweeping an arrangement.* Journal of Computer and System Sciences, *38(1):165–194, 1989.* 3

[9] *Joachim Ender, Ludger Leushacke, Andreas Brenner, and Helmut Wilden. Radar techniques for space situational awareness. In* Proceedings of the 12th International Radar Symposium. *IEEE, 2011.* 1

[10] *T Flohrer, J Peltonen, A Kramer, T Eronen, J Kuusela, E Riihonen, T Schildknecht, E Stöveken, E Valtonen, F Wokke, et al. Space-based optical observations of space debris. In* Proceedings of the 4th European Conference on Space Debris, *2005.* 1

[11] *John A Kennewell and Ba-Ngu Vo. An overview of space situational awareness. In* Proceedings of the 16th International Conference on Information Fusion. *IEEE, 2013.* 1

[12] *Vladimir Kouprianov. Distinguishing features of ccd astrometry of faint geo objects.* Advances in Space Research, *41(7):1029–1038, 2008.* 2

[13] *Daqi Liu, Bo Chen, Tat-Jun Chin, and Mark G Rutten. Topological sweep for multi-target detection of geostationary space objects.* IEEE Transactions on Signal Processing, *68:5166–5177, 2020.* 2, 3

[14] *A. Priest, M. Tetlow, N. Moretti, and P. Anastasiou. Small satellite constellation for space situational awareness. In* Proceedings of the International Astronautical Congress, Symposium on Small Satellite Missions, *2018.* 1

[15] *Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In* International Conference on Medical image computing and computer-assisted intervention. *Springer, 2015.* 8

[16] *Radim Šára, Martin Matoušek, and Vojtech Franc. Ransacing optical image sequences for geo and near-geo objects. In* Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference, *2013.* 1, 2, 3

[17] *T Schildknecht, U Hugentobler, and A Verdun. Algorithms for ground based optical detection of space debris.* Advances in Space Research, *16(11):47–50, 1995.* 2

[18] *P Seitzer, R Smith, J Africano, K Jorgensen, E Stansbery, and D Monet. Modest observations of space debris at geosynchronous orbit.* Advances in Space Research, *34(5):1139–1142, 2004.* 2

[19] *Roy L Streit and Tod E Luginbuhl. Probabilistic multi-hypothesis tracking. Technical report, Naval Underwater Systems Center Newport RI, 1995.* 3

[20] *H Wilden, C Kirchner, O Peters, N Ben Bekhti, A Brenner, and T Eversberg. Gestra—a phased-array based surveillance and tracking radar for space situational awareness. In* Proceedings of the International Symposium on Phased Array Systems and Technology. *IEEE, 2016.* 1

[21] *Toshifumi Yanagisawa, Hirohisa Kurosaki, and Atsushi Nakajima. Activities of jaxa's innovative technology center on space debris observation. In* Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference, *2009.* 1, 2