DATA SCIENCE CHALLENGE

[dstl] Government Office for Science SECURITY SERVICE MI5 SECRET INTELLIGENCE SERVICE MI6

Challenge Guide
Safe Passage: Detecting and Classifying Vehicles in Aerial Imagery

The following content is supplementary content to the challenge content on the website.

Data Specification

Data Files

The dataset consist of:

- **Training images** (training.zip): A set of 600 images (JPEG) at 5cm resolution, each one covering an area of 100m by 100m. [580 MB]

- **Training observations** (trainingObservations.csv): The label and location of the vehicles of interest corresponding to the training images. [252.9 KB]

- **Test images** (test.zip): A separate set of 600 images (JPEG) at 5cm resolution, each one covering an area of 100m by 100m. [567 MB]

- **Sample submission file** (sampleSubmission.csv): A sample submission file with the correct format, but with random detections. [145.5 KB]

Vehicle Classes

The training observations contain the location of observations for 9 classes of interest. These classes are:

Vehicle Class ID	Textual Name	Description	Example Image
A	Motorcycle	Any motorcycle.	
B	Light short rear	Light low-saturation coloured (e.g. white/ silver) car-sized vehicles with a short rear wind-screen and short boot. This includes car-sized vans.	
C	Light long rear	Light low-saturation coloured (e.g. white/ silver) car-sized vehicles with a long rear wind-screen or long saloon-style boot (e.g. saloons).	
D	Dark short rear	Dark low-saturation coloured (e.g. black) car-sized vehicles with a short rear windscreen and short boot. This includes car-sized vans.	
E	Dark long rear	Dark low saturation colour-ed (e.g. black) car-sized vehicles with a long rear windscreen or saloon style boot (e.g. saloons).	

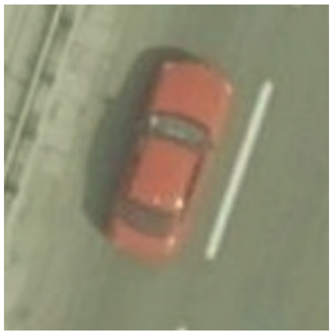
Vehicle Class ID	Textual Name	Description	Example Image
F	Red short rear	Red coloured car-sized vehicles with a short rear windscreen and short boot. This includes car-sized vans.	
G	Red long rear	Red coloured car sized vehicles with a long rear windscreen or saloon style boot (e.g. saloons).	
H	Light van	Light low saturation coloured (e.g. white/ silver) vans, larger than an average car (e.g. Ford Transit). This does not include coaches or camper vans.	
I	Red and white bus	Red and white buses. This excludes other large vehi-cles, such as red and white lorries.	

Table 1: Overview of the vehicle classes

Additional notes:

- Not all vehicles in the scenes belong to a class. For example, lorries and yellow cars do not belong to a class of interest.
- Classes B to G (cars of interest), include convertibles.
- For classes B to H (cars and light vans of interest), the colour refers to the majority colour of the vehicle (not just the parts visible in the image). For example, a short-rear black car with a red roof will be classified in class D.
- The training data provides observations with a variety of backgrounds and observations which are partially obscured (e.g. by trees) and observations that are near the edge of the image, but which are not cut off.

When generating the ground truth, observations were classified as either high or low confidence.

- **High confidence** observations are those that are certain to belong to a class.
- **Low confidence** observations are those where it is less certain which class they belong to.

The training observations which have been provided to you contain only the high confidence observations.

You are being asked to find similar high confidence observations in the test images, though any submitted observation that matches a low confidence ground truth observation will not be penalised. Low confidence observations are those where:

- The ground truth is not known because it could not be determined.

- The vehicle detected does not fall into one class, but sits between two (e.g. a light car with a medium-sized rear).

- The vehicle detected is similar to one class, but different enough not to be included (e.g. dark blue short rear).

- The vehicle is cut off by the edge of the image.

- The vehicle is a motorcycle with a cover over it.

- Observations contain image artefacts (e.g. only half a vehicle due to image stitching).

Training observations file format

The training vehicles are provided in comma separated values (CSV) format, with the first line as the headings, and the remaining rows as the entries. The headings are:

- **Id**: The unique id for this row. This is a combination of the image name (e.g. TQ2378_0_0.jpg) with file extension (.jpg) removed, an underscore (_) and the vehicle class (e.g. A). For this example the Id would be TQ2378_0_0_A.

- **Image**: The filename of the image (e.g. TQ2378_0_0.jpg)

- **Class**: The vehicle class ID (e.g. A)

- **Detections**: The complete set of detections for the image and class combination. Each detection is the centre pixel position of the observation, separated by a colon (:), e.g. xPixel:yPixel, and a pipe (|) is used to separate detections. When there are no detections, "None" is used. This provided centre pixel positon will not be exact. The pixel coordinates are integers and defined using standard convention (i.e. with the image origin at top left).

Example:

```
id,image,class,detections
TQ2379_0_0_A,TQ2379_0_0.jpg,A,None
TQ2379_0_0_B,TQ2379_0_0.jpg,B,1776:520|1824:125
TQ2379_0_0_C,TQ2379_0_0.jpg,C,1760:456
```

All combinations of image name and vehicle classes are included, even if the image does not contain any vehicles of interest.

Scoring

The submitted observation will be assessed against the ground truth for test images. When generating the ground truth, observations were classified as either high or low confidence.
The assessment will score against the high confidence observations and will ignore any observations matching a low confidence ground truth observation. Any observation that does not match a high or low confidence observation will be penalised.
The submitted results file will be scored using the Jaccard Index, defined as:

$$\frac{TP}{TP + FP + FN}$$

Where:
TP are the true positives
FP are the false positives
and
FN the false negatives.

The Jaccard index essentially rewards correct classifications (TP) and penalises incorrect or missed classifications (FP and FN). It provides a score of between 1 (perfect) and 0.

The TPs, FPs and FNs will be summed for all images and all vehicle classes and then the Jaccard index will be calculated to give the final score.

A submitted observation matches a ground truth observation if the vehicle class is the same and centre pixel position is within an acceptance boundary. An acceptance boundary is used because there is an acceptable tolerance for the exact location, and the ground truth centre position is not exact.

The following rules apply when scoring a submitted observation against the corresponding high confidence ground truth observations for the same image and class of the submitted observation:

- Any observation within the ground truth acceptance boundary will be classed as a match to that ground truth position, becoming a TP
- If there are multiple observations within the acceptance boundary, only one will be scored as a TP. The other observations will not be penalised
- Any ground truth position that does not have an observation within the acceptance boundary will be counted as a FN
- Any observation not within a ground truth acceptance boundary will be counted as an FP, unless it matches the location (using the acceptance boundary) of a low confidence ground truth observation regardless of its vehicle class (i.e. a match against a low confidence ground truth observation is based on location only, the class is ignored).

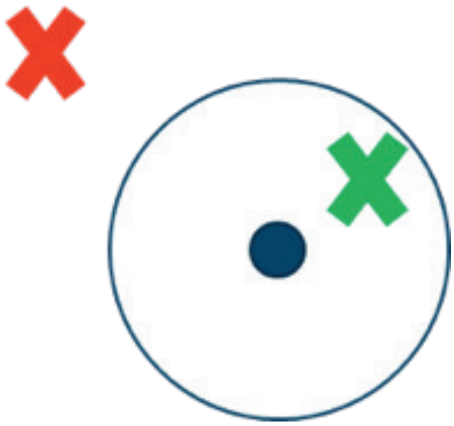


Figure 1 Ground Truth position (Blue dot) and acceptance boundary (blue circle).
A matching observation (green cross) and non-matching observation (red cross)

The acceptance boundary size is defined as the radius from the ground truth position to the edge of the acceptance boundary. This distance is provided in the table below.

Please remember that these acceptance boundaries are in relation to the ground truth centre position, which will not be exact, so you should aim to be more accurate than the distances specified.

Class	Acceptance Radius
A (motorcycle)	12 pixels (60cm)
B-G (cars)	30 pixels (150cm)
H (van)	40 pixels (200cm)
I (bus)	45 pixels (225cm)

Table 2 The acceptance boundary size per class.
Note, this is to cover the error in the ground truth and error in submitted results.

The vehicle class for low confidence ground truth observations is ignored, and therefore a match is based on location (using the acceptance boundary) only. They do not contribute to the scoring other than to reduce the false positives (FP), therefore:

1. A missed low confidence ground truth observation will not be classed as a FN (i.e. will not be penalised).

2. Any observation that matches a low confidence ground truth observation will not be classed as a TP (i.e. will not be rewarded) and will not be classed as a FP (i.e. will not be penalised).

Guidance

The aim of this section is provide some guidance to people who are new to image processing and machine learning. It is by no means the only approach and competitors are encouraged to try new and innovative solutions.

Tools that may be of interest include:

- OpenCV - http://opencv.org/

- TensorFlow - https://www.tensorflow.org/

- Theano - http://deeplearning.net/software/theano/

The following is an example approach based on a machine learning using TensorFlow. This is a semi-automated approach, although final submissions should be fully automated. The following tutorial shows how to retrain the ImageNet model for new classes of observations:

https://www.tensorflow.org/how_tos/image_retraining/

The idea is to follow this approach to retain the model to look for vehicles that match one of our nine classes of interest. In order to obtain images for our classes the training images and training observations can be used:

1. Extract a small image chip (e.g. 150 by 150 pixels) centred on the vehicle of interest and place in the corresponding folder for that class. Note, our vehicles of interest range in size from motorcycles to buses, and you will need to take this into consideration to improve the model. To start with, a standard chip size will be OK.

2. Extract a set of image chips which do not contain any vehicles of interest into a folder call 'other.' This will become an extra class that will be used to train a model to recognise chips that are not of interest.

Image chips can be extracted from the test image, and can then be classified by the retrained model to determine if they contain one of the classes of interest, or if they belong to the 'other' class and therefore can be ignored.

The Challenge Master is our expert for the challenge and will be available to offer guidance through the challenge forum.