Mafat (DDR&D) Challenge

Detection and fine grained classification of objects in aerial imagery

13.11.2017

Writing: Tzvi Diskin

Supervision: Eran Dahan

Confirmation: Shay akler

Description

In this competition the DDR&D would like to tackle the challenge of automatically exploiting information from aerial imagery data. As the volume of imagery gathered by aerial sensors is rapidly growing, we understand that the exploitation of such data could not be achieved solely by a manual image interpretation process. The following competition aim is to explore automated solutions that will enable detection and fine grained classification of objects in high resolution aerial imagery.

Participants goal is to detect and classify different objects found in aerial imagery data. The classification includes coarse-grained classification for main classes (for example: Large vehicle, solar panel) and fine-grained classification of subclasses and unique features (Such as sunroof, ladder, etc).

In addition to a comprehensive set of manually labeled objects and their respective images, in this competition the DDR&D also provides a large set of untagged images. The exploitation of those images for the task of object detection and classification, by using techniques from the area of unsupervised learning, semi-supervised learning, transfer learning, generative adversarial networks, etc., is encouraged.

Data

The dataset consists of aerial imagery taken from diverse geographical locations, different times, resolutions, area coverage and photo conditions (weather, angles and lighting). Image resolution vary between 5cm to 15cm GSD (Ground Sample Distance).

A few examples:

As can be seen, images include many different types of objects, such as: vehicles, roads, buildings, trees etc.

Participants would be asked to detect objects of the relevant classes. Every object that is detected should be classified in 4 granularity levels:

1. **Class** (example: Large vehicle) - each detection would include a single value.
2. **Subclass** (Large vehicles will include subclasses such as: Concrete mixer truck, Crane truck, Prime mover, etc.) - each detection would include a single value.
3. **Presence of Features** (has Ladder? is Wrecked? has Sunroof? etc.) - each detection can include multiple values.
4. **Object perceived color**[1] (Blue, Red, Yellow etc.) - each detection would include a single value.

d cas

Full description of general classes tagging information:

1. **Small vehicle**
   a. Subclasses - Sedan, Hatchback, Minivan, Van, Pickup truck, Jeep, Public vehicle.
   b. Features - Sunroof, Taxi, Luggage Carrier, Open cargo area, enclosed cab, Wrecked, Spare wheel.
   c. Colors - Yellow, Red, Blue, Black, Silver/Grey, White, Other.
2. **Large vehicle**
   a. Subclasses - Truck, Light Truck, Concrete mixer truck, Dedicated agricultural vehicle, Crane truck, Prime mover, Tanker, Bus, Minibus.
   b. Features - Open cargo area, Vents, Air conditioner, Wrecked, Enclosed box, Enclosed cab, Ladder, Flatbed, Soft shell box, Harnessed to a cart.
   c. Colors - Yellow, Red, Blue, Black, Silver/Grey, White, Other.
3. **Solar panel** - (No subclasses, features or colors).
4. **Utility pole -** (No subclasses, features or colors).

Participants will receive a training set, which consists of:

- **9369 tiff and jpeg images**.
- **CSV file of tagged objects** - Small vehicles, large vehicles and solar panels are described in the form of a bounding polygon, which is a set of 4 x-y (pixel) coordinates. utility poles are described by a single point (a single pair of x-y coordinates).

[1] Perceived color - object color as would be described by a person.

Only objects of small and large vehicle classes will include fine-grained classification data: subclass, features, color. please note that features are represented as boolean fields while “-1” represent a non-viable option.

- **CSV file of imagery description -** For their convenience, participants will receive a list of categorized images. available images categories:
  - **Full** - images that went through the full tagging process of all relevant classes.
  - **Partial** - images that were not tagged with solar panels, meaning solar panels a likely to appear at some images but tags of the solar panel class is unavailable for these images.
  - **No objects** - images that were positively processed in order to ensure that none of the mentioned classes objects are present in the data, and therefore they do not include tagged objects.
  - **Untagged** - meaning images that did not went through our detection pipeline and therefore do not include tagged objects, but that does not mean that no objects are present in the data.

Each image can contain more than one object of any class or even part of the object. This is valid for both the “Training Set” and the “Test Set”. The ground resolution vary between the images. The participants are requested to identify the different objects and submit the coordinate locations in the specific image pixel-grid. Having mentioned that, the training images and labels can be used in any way you decide (subject to the data terms in the Official Rules) to help detect the objects of interest in the test images.

The “Training Set” is noisy and may contain tag errors. Likewise, the set doesn’t necessarily represent the “Test Set” and the number of objects in the various classes is unbalanced between them. It should be mentioned that the majority of the “noise” comes from unlabeled objects, i.e. objects that exist in the image but were not annotated.

1	Image_id	P1_X	P1_Y	P2_X	P2_Y	P3_X	P3_Y	P4_X	P4_Y	Class	Subclass	Sunroof	Taxi	Luggage_Car	Open_cargo_	Enclosed_ca	Spare_whee	Wrecked	Flatbed	Ladder	Enclosed_bo	Soft_shell_b	Harnessed_t	Vents	Air_conditio	Color
13814	101480	724.0625	263.819489	734.381531	239.519745	808.9375	271.180512	798.618469	295.480225	small vehicle	Hatchback	0	0	0	0	0	0	0	0	0	0	0	0	0	0	silver/grey
13815	101480	644.44751	26.0853081	654.144287	1.53061104	728.55249	30.9146919	718.855713	55.4693909	small vehicle	miniVan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	white
13816	101481	177.152054	425.834259	188.68631	402.08725	258.847961	436.165741	247.31369	459.912781	small vehicle	sedan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	silver/grey
13817	101481	608.094788	408.809387	619.352234	384.929901	689.905212	418.190613	678.647766	442.070099	small vehicle	sedan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	white
13818	101481	498.088928	551.22583	509.164093	529.934021	581.911072	567.77417	570.835938	589.065979	small vehicle	sedan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	white
13819	101481	448.026764	594.874023	458.35907	575.905518	489.973236	593.125977	479.64093	612.094482	small vehicle	sedan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	white
13820	101486	477.084839	89.1953659	495.626251	50.1767006	576.915161	88.8046341	558.373718	127.823303	small vehicle	Hatchback	0	0	0	0	0	0	0	0	0	0	0	0	0	0	black
13821	101486	811.257019	170.901459	819.860291	148.496445	891.742981	176.098541	883.13971	198.503555	small vehicle	sedan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	white
13822	101486	457.39093	20.950428	457.61264	-0.6484337	525.60907	0.04956949	517.053996	37.1246223	small vehicle	sedan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	silver/grey
13823	117073	202.10513	2479.71183							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13824	117073	385.540303	2500.17774							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13825	117073	339.302512	2612.36124							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13826	117073	1259.51036	2208.34906							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13827	117073	1702.93836	1888.4745							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13828	117073	1666.55453	1884.68452							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13829	117073	1591.51287	1757.34109							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13830	117073	1372.45185	1725.50523							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13831	117073	3799.55691	1561.77797							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13832	117073	3753.31912	1608.01576							utility pole		0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Tags CSV file structure

Evaluation

Participant are asked to detect all objects that appear in the provided test set, and accurately classify them (Class, Subclass, Features, Color). The submission file should include a different list for each general class, each list should be sorted by object confidence level (high to low) for the entire test data.

We encourage participants to submit at least:

- X (To be decided) small vehicles
- X (To be decided) large vehicles
- X (To be decided) solar panels
- X (To be decided) utility poles

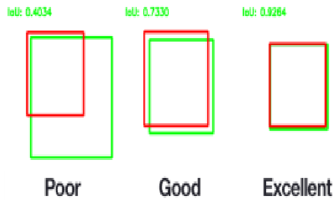
The final score will be a multiplication of two components:

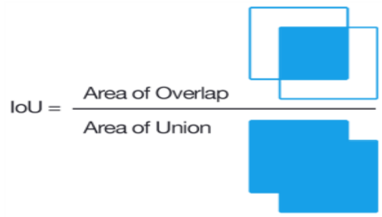
<u>Component 1 - "Coarse grained detection and classification"</u>

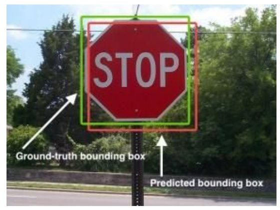
Coarse grained will be evaluated by measuring mean average precision, the first stage is deciding whether a detection is a True Positive (TP), False Positive (FP). This is done differently for objects that are bounding in bounding polygon and those that are single pixel detected. This stage will be divided into 2 sub-stages - detection by overlap area and by quality index calculation:

<u>Detection by overlap area</u>

1. True Positive (TP) will be determined when the bounding polygons (the predicted and observed) overlap with each other – intersection over union (area of overlap divided by area of union, IoU). If the IoU is larger than 0.25 and the object was detected correctly then this detection will be regarded as TP.
2. If the object is tagged by a Single Pixel, a distance metric in pixels between predicted and observed pixel should be calculated. If the distance is below a 30 pixels threshold, then this will be regarded as TP detection.
3. Otherwise, the detection will be regarded as False Positive (FP).
4. If there is more than one TP for a certain object, the first will be considered TP (provided it passed the threshold condition) and the rest will be considered FP (although they passed the threshold condition).







Quality Index

After having all the FP and TP of the predictions, the total score will be determined using Mean Average Precision (MAP)[2] for m objects out of N images. Every category in the coarse-grained classification has the same value in the score. This index varies between 0 to 1 and emphasizes correct detections with significance to confidence in each detection, meaning to distinguish between participants that detect all objects correctly, in all environmental conditions, as well as can reference their confidence in the detection.

The participant is requested to submit a list of detected objects in each image, based on the category list and sorted by the probability of of finding them in all images. The list will be sorted co that the first object will be with the highest probability for detection. The score will be calculated in each category accordingly:

$$AP(class) = \frac{1}{K}\sum_{k=1}^{K} Precision(k)rel(k)$$

N is the number of images.

M is the correct number of objects in the image.

K is the total number of objects from the class in the test data.

precision(k) is the precision calculated over the first k objects.

and $rel(k)$ equals 1 if the object k is True and 0 if the object is False

The total scope is

$$mAP = \frac{1}{N_c}\sum_{class=1}^{N_c} AP(class)$$

When N_c is the number of categories

[2] http://fastml.com/what-you-wanted-to-know-about-mean-average-precision

Component 2 - "Fine-grained classification"

As stated, for small vehicle and large vehicle classes, each object has a subclass and additional distinctive features of which the participant must detect. for example: a private carsmall vehicle can be classified as a Hatchback that have a feature of Sunroof, Taxi, Wrecked, Etc. Additionally each object color should be classified to set of predetermined colors.

Submission file will include the following information for each object: Class (single value), Subclass (single value), color (single value), the features availability of the object (multiple boolean fields). To simplify the competition, there is no need to submit the probability of the existence of a specific object, but only if it exists or not (binary form) or type (the object's color for example).

Such a file will be turned into detection matrices and the system will perform a performance appraisal for a fine-grained classification, only for the objects which were detected correctly in the coarse-grained classification (TP's).

Quality Index

Since the manner of filing the results in not probabilistic, the performance index is Multi Class Cohen's Kappa[3] which is a standard precision index for classification issues. The Kappa is calculated separately for the subclass, the color, and each feature for each main category. The total score for each main category is a simple average of the Kappa of each feature. The total score for the fine-grained classification is a weighted average between the score of the small vehicles and the large vehicles.

$$score_{small} = \frac{1}{M_{small}} \sum_{i}^{M_{small}} \kappa_i$$

$$score_{large} = \frac{1}{M_{large}} \sum_{i}^{M_{large}} \kappa_i$$

Where κ_i is Cohen's Kappa for feature i, M_{small} is the number of features of the small vehicles, M_{large} is the number of feature of the large vehicles (including the sub-class and the color)

The total score for the fine-grained classification is:

$$score = \frac{N_{large} score_{large} + N_{small} score_{small}}{N_{large} + N_{small}}$$

Final Score

[3] http://vassarstats.net/kappaexp.html

Each participant will receive three scores: coarse-grained classification score, fine-grained classification score, and a combined score which is the multiplication of both scores. The goal is to score as high as possible in the coarse-grained classification task as well as in the fine-grained classification task. A high score at one of the stages is not enough.

File Submission

Participants would be asked to submit a CSV file, each detection would be represented by a row, for every row in the dataset, submission files should contain the following:

- ImageId (string)
- Location - Small and large vehicles are described in the form of a bounding polygon, which is a set of 4 x-y (pixel) coordinates. Solar panels and utility poles are described by a single point (a single pair of x-y coordinates).
- Class (INT)
- Subclass (INT)
- Features (Booleans)
- Color (INT)

a.