

**FGAI4H-A-027**

Geneva 26-27 September 2018

**Source:** EPFL

**Title:** Running Challenges

**Purpose:** Discussion

**Contact:** Marcel Salathé E-mail: [marcel.salathe@epfl.ch](mailto:marcel.salathe@epfl.ch)

EPFL, Switzerland

# Running Challenges



## Mapping Challenge

Building Missing Maps with Machine Learning



By **Humanity & Inclusion**

Completed

24423 Views   532 Participants   718 Submissions

 118

UNFOLLOW

[Overview](#)   [Leaderboard](#)   [Discussion](#)   [Dataset](#)   [Submissions](#)   [Participants](#)

# Preamble

The Digital Epidemiology Lab has in the past developed a (largely open source) AI challenge platform called crowdAI. ([www.crowdAI.org](http://www.crowdAI.org))

Based on the experience of running 36 challenges with numerous stakeholders and in very different domains, some best practices have emerged that we assume to be valuable for benchmarking.

# Mapping Challenge

## **Problem:**

Often, accurate maps either do not exist or are outdated by disaster or conflict. Today, when new maps are needed they are drawn by hand, often by volunteers who participate in so called Mapathons.

## **Goal:**

Explore how Machine Learning can help pave the way for automated analysis of satellite imagery to generate relevant and real-time maps.

# Mapping Challenge

## Task:

In this challenge you will be provided with a dataset of individual tiles of satellite imagery as RGB images, and their corresponding annotations of where an image is there a building.

The goal is to train a model which given a new tile can annotate all buildings.

# Mapping Challenge

## Evaluation:

For for a known ground truth mask  $A$ , you propose a mask  $B$ , then we first compute  $IoU$  (Intersection Over Union) :

$$IoU(A, B) = \frac{A \cap B}{A \cup B}$$

$IoU$  measures the overall overlap between the true region and the proposed region. Then we consider it a True detection, when there is atleast half an overlap, or when  $IoU \geq 0.5$

Then we can define the following parameters :

- Precision ( $IoU \geq 0.5$ )

$$P_{IoU \geq 0.5} = \frac{TP_{IoU \geq 0.5}}{TP_{IoU \geq 0.5} + FP_{IoU \geq 0.5}}$$

- Recall ( $IoU \geq 0.5$ )

$$R_{IoU \geq 0.5} = \frac{TP_{IoU \geq 0.5}}{TP_{IoU \geq 0.5} + FN_{IoU \geq 0.5}}$$

The final scoring parameters  $AP_{IoU \geq 0.5}$  and  $AR_{IoU \geq 0.5}$  are computed by averaging over all the precision and recall values for all known annotations in the ground truth.

# Mapping Challenge

## Round 1:

We provide:

- Training data set (annotated)
- Validation data set (annotated)
- Test data set (not annotated)

You submit:

Annotations to test data set.

# Mapping Challenge

## Round 2:

You submit:

Your AI model (we containerize)

We evaluate the model on undisclosed test data set,  
provided by UNOSAT. -> final evaluation



# General

Main experience is that one needs to work with “problem owner” to define the challenge in detail - the roadblock has generally been related to technical expertise.

Whether a particular problem can be tackled by AI can be assessed relatively quickly.

# General

Data availability has been a main problem, but process of having distinct public training sets and undisclosed test sets has been a good solution to the problem.

Needs the ability to host data very securely, and to execute code close to the data, also very securely.