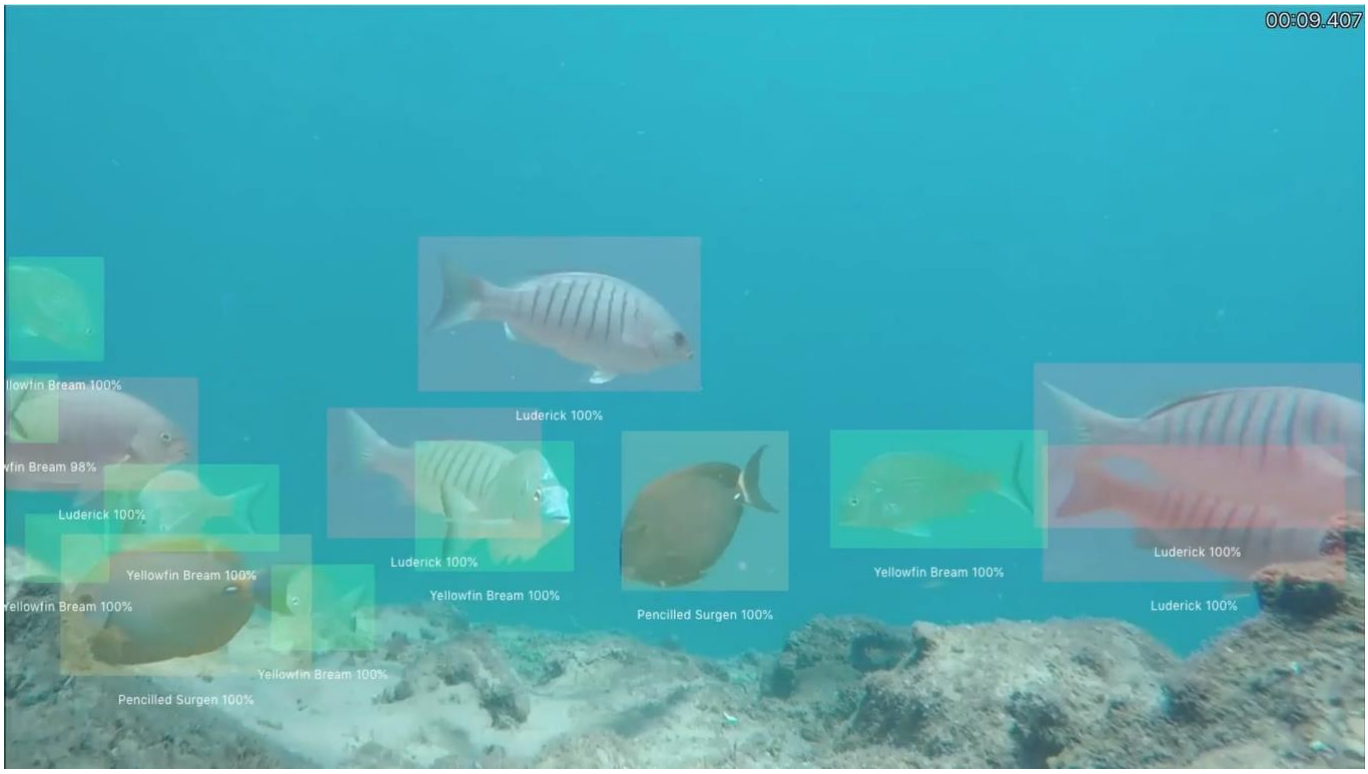
Detailed Plan – TER 2021-071



1. Identification of the project

Title: Deep Learning System for animal recognition in underwater images and videos from Mediterranean Sea

Students:

Giuseppe Alessio Dimonte	Master 2 EIT Digital (DSC)
Arturo Adan Pinar	Master 2 EIT Digital (DSC)
Yueqiao Huang	MAM5

Supervisors:
Frederic Precioso
Diane Lingrand

Type of project: Development

Date: 17/12/2021

2. Presentation of the subject

Problem: starting from the Megadetector, which is able to put bounding boxes and recognize between humans, animals and vehicles, the work to be done is to expand this model to underwater images and videos in order to put bounding boxes also for fishes and corals. After that, a UI is created allowing expert to annotate images and create a training set with the appropriate labels in order to train three Deep Learning models (object detection, tracking and instance segmentation with counting) to make them perform accurate predictions identifying, segmenting and counting different fishes and corals in the dataset.

Context: the program aim is to surveille the environment of the Mediterranean Sea.

Users: the software is intended to be used by biologist and general users interested in fishes and corals recognition.

3. Presentations of the solutions including presentation of the possible solution space, and selected solution(s)

The framework is divided in 5 main jobs:

1) Megadetector adaptation:
   - In this first step the Megadetector architecture is modified to include support for fishes and corals in underwater images and videos
2) UI creation:
   - After generating a training set with the different bounding boxes generated from the modified architecture of Megadetector, a UI is needed to allow experts in the biological field to annotate fishes and corals represented within each bounding box
3) CNN1 – Segmentation:
   - The first Convolutional Neural Network takes as input the training set with the shape of fishes and corals already segmented and the labels set from the DeepFish dataset
   - After the training, it predicts the shape of a fish or coral from an underwater image or video
4) CNN2 - Object Detection:
   - The second CNN takes as input the training set generated by the Megadetector + UI + annotation
   - After the training, it predicts the correct label for each bounding box in the input image
5) CNN3 - tracking and counting:
   - The third CNN takes as input for the training the same input as CNN2
   - After the training it can count the number of different fishes and corals in an underwater image or video

4. Positioning of the solution in relation to the existing one, see state of the art for research-oriented subjects

As pending work to do we have to study more the state of the art and the literature to position our solution compared against others.

5. Work done: the product and the process. How did you work and what you produced

The GANTT chart is defining the goals and the next actions to take, and the work produced until now.

1) Megadetector adaptation:
   - We are struggling in understanding how the Megadetector works
   - We are trying to see if some alternatives exist, like Detectron2 by Facebook
   - The SeaCLEF dataset we were requested to work with is not available anymore. Therefore, we are going to use the DeepFish dataset

2) UI creation:
   - We are studying how the Python library called "PysimpleGUI" works to implement and deploy a simple user interface written using the Python programming language
   - The completely functional UI is expected to be developed by the first week of January

3) CNN1: segmentation:
   - We wrote the code regarding this part
   - The training will be performed when we have the annotated dataset ready from experts' side

4) CNN2: object detection:
   - We wrote code regarding this part
   - We were able to write the code but not to train and evaluate it because we need to have our training dataset from the annotation part

5) CNN3: tracking and counting:
   - At the moment we did not enter in this area

6. Conclusions on the results achieved and taking a step back from the work done and what remains to be done

We are a bit late on the schedule, but we can manage it. In fact, as soon as we are able to annotate images through the UI and have our training set we can train our CNNs and test different networks types, like VGG, ResNet and EfficientNet.
After that, we can work in parallel for optimizing and improve our solution for each of the 5 macro-jobs.

7. List of bibliographic references

Our code:
https://github.com/alessiodimonte/TER
Project PNM:
https://github.com/FannySimoes/Projet_PNM
Megadetector:
https://github.com/microsoft/CameraTraps/blob/master/megadetector.md