Escola de Enxeñería Industrial

Traballo Fin de Mestrado
Resumo A3

Convolutional neural networks for object detection in subterranean environments

Pablo Santana González

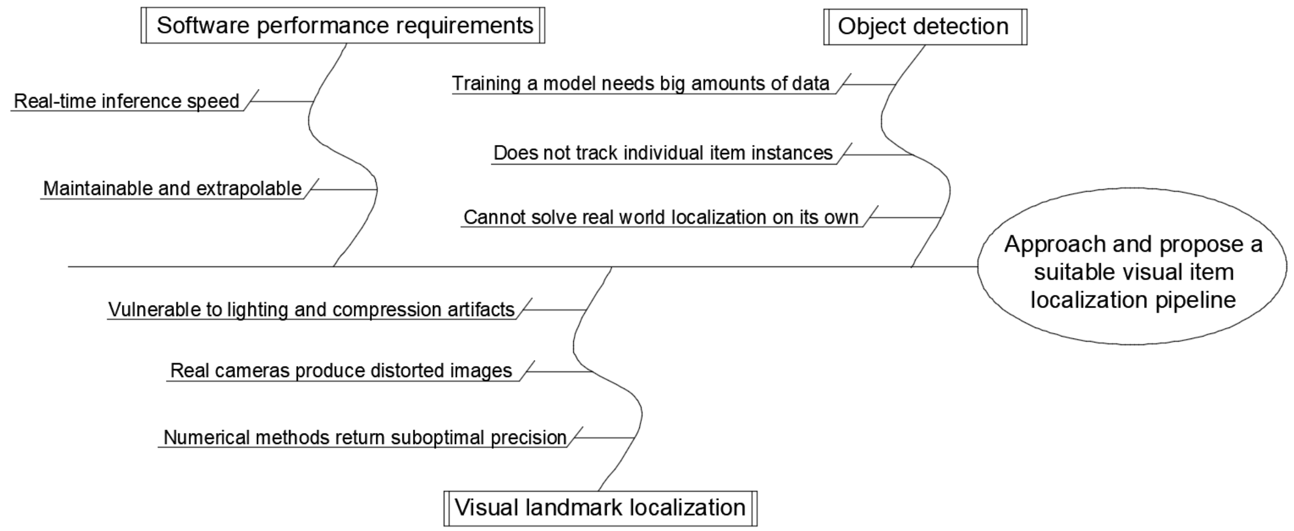
1- Problem statement

Subterranean environments are characterized by a series of hazards that vary drastically across domains and can degrade or change over time, often posing too high of a risk for having human personnel enter such places. Hence, having unmanned vehicles quickly navigate and search underground settings has been a rising interest in the world of robotics. This work aims to explore the field of applied machine vision by approaching a challenging computer vision problem consisting on the detection and localization of items within a closed environment mimicking subterranean settings, based on the input of a single camera.

2- Current scenario

Perception state-of-the-art results of item search (agent plus item localization with respect to a common coordinate frame) are nowadays achieved by the combination of different sensory such as cameras, GPS, LiDAR... however state-of-the-art computer vision techniques allow for its solution only on image input. While visual agent localization is usually based on landmarks and carried out in a deterministic manner, visual item localization needs to first identify items within pictures, for which data-driven approaches such as object detection are typically used.

3- Cause analysis



4- Target scenario

Having obtained a perception layer based on the object detection paradigm, after thourough exploration on the state-of-the art of this field by experimental benchmarking, so the position of a set of target items can be estimated from the live feed of a single camera. For this, the following goals are laid out:

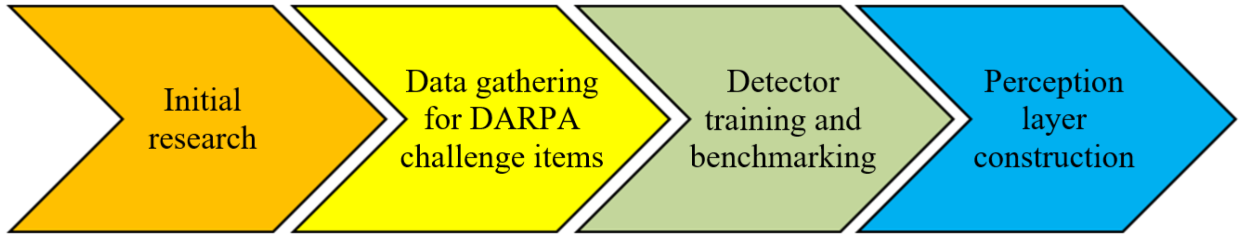
I. Approach the DARPA Subterranean challenge by using image object detection:

- Build a representative baseline dataset for a defined collection of target items.
- Perform a formal comparison of state-of-the-art object detection models using this data.

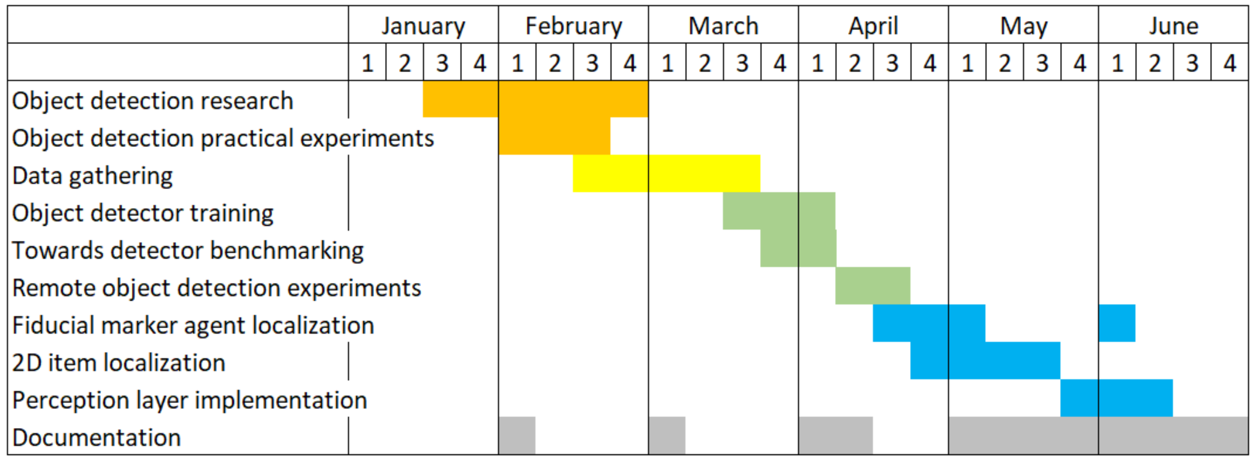
II. Design a perception layer built around an object detector for item localization:

- Achieve camera localization from fiducial markers.
- Achieve item localization and implement the complete real-time perception pipeline.

5- Course of action

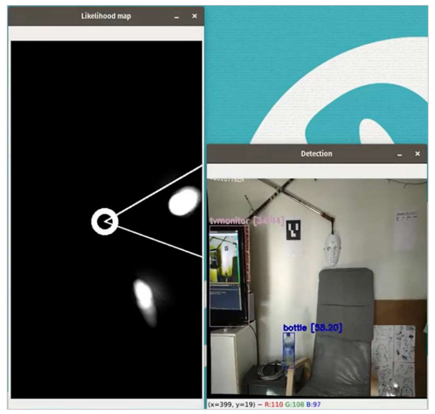


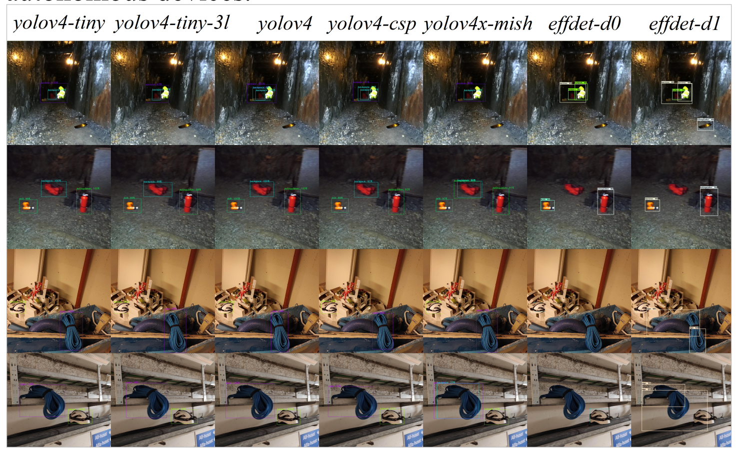
6 - Tracking



7- Results

- The performance of the most competitive real-time object detectors up to date was evaluated for a particular 6-class detection task via benchmarking on a dataset crafted towards the DARPA subterranean challenge. Additional outcomes of these tests are:
  - Two datasets towards the DARPA SubT challenge.
  - A batch of trained, competitive detectors.
- Simple computer vision techniques have proved effective for probabilistic item localization via the design and validation of the proposed monocular perception layer, its main highlights being:
  - Robustness to camera pose and object detection noise.
  - Agnosticism to the used object detection model or platform.
  - Modularity and ease of incorporation to autonomous devices.





Suggested future work includes further experimentation with synthetic data for object detection training, enhancing 3D reconstruction by addressing visual obstacles and discarding width-based localization, and expanding the current Kalman filter with inertial sensory so that position tracking can be performed when landmark visuals are interrupted.

Code, outcomes and resources available at
`github.com/pabsan-0/sub-t`