Politecnico di Torino
Dipartimento di Automatica e Informatica

PhD in Computer and Control Engineering
XXXVI cycle

ISTITUTO ITALIANO DI TECNOLOGIA

Supervisor
Tatiana Tommasi, Barbara Caputo

Efficient Object Detection Across Visual Domains

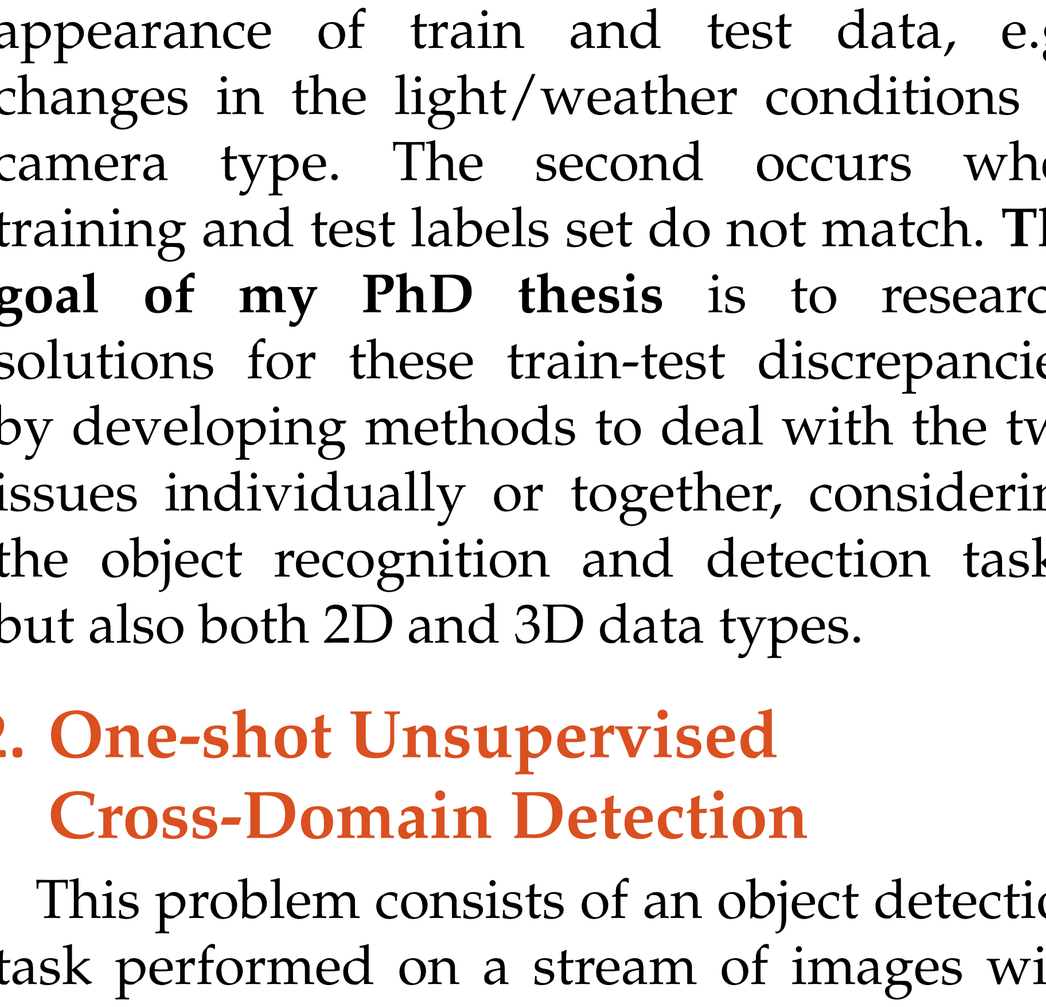
PhD Candidate: *Francesco Cappio Borlino* francesco.cappio@polito.it

1. Introduction

In the last years deep learning models have reached excellent performance in many Computer Vision tasks. Still, their use in real-world applications is often limited by some obstacles that can appear when going from lab settings to real systems. In particular, **train-test discrepancies** naturally arise and models' performances drop. There are two main types of discrepancies: **visual domain** and **semantic** distribution shifts. The first involves differences in the visual appearance of train and test data, e.g.: changes in the light/weather conditions or camera type. The second occurs when training and test labels set do not match. **The goal of my PhD thesis** is to research solutions for these train-test discrepancies, by developing methods to deal with the two issues individually or together, considering the object recognition and detection tasks, but also both 2D and 3D data types.

2. One-shot Unsupervised Cross-Domain Detection

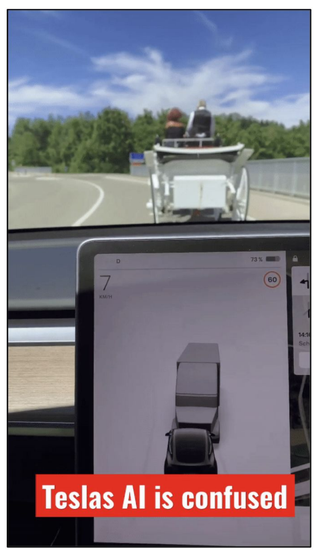
This problem consists of an object detection task performed on a stream of images with continuously varying visual domains, a situation typical of social media streams.



Our proposed solution [1] uses a multi-task training procedure involving both the main detection task and a self-supervised one. A **meta-learning** optimization strategy is further implemented in order to prepare the model for a particular inference time routine: before performing a prediction the model is adapted to the test-image's target domain by few training iterations on the **self-supervised task**.

3. Semantic Novelty Detection & Open Set Recognition

Detecting novel classes at test time may seem an easy task. Deep networks however are generally trained focusing on the **closed-world** setting, thus they are biased towards assigning test samples to known classes with high confidence. This can represent a serious problem in many safety critical applications like autonomous driving. In [2], we propose a novel paradigm for representation learning using relational reasoning designed specifically for semantic novelty detection. Moreover we have drawn [3] a picture of the state-of-the-art in Open Set Recognition on 3D point clouds to raise the attention about semantic novelty detection on this data type, which could represent a valuable asset for many safety critical applications.



4. References

1. "Self-Supervision & Meta-Learning for One-Shot Unsupervised Cross-Domain Detection", Cappio Borlino et al. In: CVIU 2022
2. "Semantic Novelty Detection via Relational Reasoning", Cappio Borlino et al. In: ECCV 2022
3. "3DOS: Towards 3D Open Set Learning - Benchmarking and Understanding Semantic Novelty Detection on Pointclouds", Alliegro, Cappio Borlino et al. In: NeurIPS 2022