# Equipment identification through image recognition

Author: Saidnassimov Darkhan

Master’s thesis Introduction chapter

2022

# 1 Introduction

### Problem statement

In recent years, computer vision algorithms have received much attention due to their potential applications in a vast variety of fields. As the technologies advance, endless possibilities arise in numerous fields, such as security monitoring, medicine and self-driving vehicles. [2] However, although over the last decade computer vision has been emerging in industrial applications, such as safety and process monitoring [2], fewer research has addressed the issue of equipment detection.

Industrial plants typically are hundreds of meters long and often it becomes frustrating to identify minor equipment parts. Identifying the parts becomes relevant once they start requiring replacement or maintenance as the plant would not be able to run at full capacity without them. Ore processing plants treat several hundreds of tons of ore per hour, and the production capacity is constant. [2]. Therefore, quite often it is troublesome to properly identify the equipment within a list of thousands of various parts in a medium to large-scaled plant.

Even though various methods have been implemented for detection of objects in various domains [2], these methods heavily rely on extensive data collection and training of models in order to work. Certain complications arise as it is often not possible to collect huge amounts of training images from the plants due to privacy and confidentiality issues. Luckily, in our scenario, the images can be collected from the 3D simulator rather easily, however, this raises restrictions on the accuracy of the models as the models will not perform as well on real images.

Hence, this thesis proposes a cross domain object detection algorithm as a solution to automatically localize and identify the equipment in a large environment and, therefore, minimize the delay before the production is online again.

### Thesis objective

The main goal of the thesis is to identify a suitable object detection method for an industrial environment and develop a minimal proof of concept application. The selected method should be able to identify an object in a real image given a labeled dataset of rendered images from a 3D equipment model and a smaller unlabeled dataset of real images.

Additionally, a method should propose a methodology to optimize the laborious data collection and labeling process. The optimization is important not only because it is a time demanding process, but also because in large plants there are thousands of objects and scalability is critical.

### Methods

In order to accomplish the objectives defined, this paper will explore multiple methods for object detection. For the purpose of training an object detection model, a vast number of training and testing pictures is required. In most traditional object detection algorithms, there is little to no difference in the environment where the training and testing images have been taken. Unfortunately, often the accuracy results are lower than anticipated when the phenomenon of domain shift takes place. [2] Domain shift occurs, for example, when the object of interest is placed under different weather or lighting conditions, such as [2]. Additionally, in an industrial environment, it often becomes challenging to take pictures due to accessibility and confidentiality regulations.

One promising approach to achieve higher performance in object detection model would be through domain adaptation. In the given scenario, a cross domain object detection model will be proposed to tackle the objectives. Domain adaptation technology’s advantages additionally include less demanding data collection process. Therefore, this project will develop a novel cross-domain object detection method based on Mean Teacher and Adversarial Training methods.[2]

The dataset utilized to implement the methods is based on TLess open-source dataset [2] due to limitations in obtaining real life data. Originally the dataset is meant for pose estimation in 3D models, therefore the dataset will be converted to formats, appropriate for the proposed methods.

The methods will be trained on rendered data from 3D models and evaluated on real images using mean average precision metrics. Finally, the selected method will be evaluated on one equipment item from a real plant.

### Scope

The thesis will be limited to proposing the solution upon analyzing multiple existing state-of-the-art methodologies. Preparing the actual real-life dataset and implementing the solution for a real plant remains out of the scope for the duration of the project. Although the proposed methods attempt to minimize the data collection and labeling process, it might take multiple months before the dataset and the model based on the real data is ready for training.

### Metso:Outotec

The issuer and the commissioner of the thesis topic is Metso Outotec Oyj. Metso Outotec is a Finnish publicly traded company that formed as a result of a merger of two companies - Metso Minerals and Outotec in June 2020. The company provides sustainable solutions and services worldwide for minerals processing, aggregates and metal refining.

Along with the focus on sustainability and strive for innovation, Metso Outotec aims on digitization throughout the equipment cycle. Therefore, the company offers various IoT solutions that attempt to optimize various production processes for customers. [2] Naturally, computer vision is of great interest to the company due to its potential in the industry and the proposed thesis topic was reported to be in high demand among the plant owners.

### Structure of the thesis

The rest of this thesis is divided into five chapters. Chapter 2 reviews the literature relevant to understand the main concepts and algorithms used during the research. Chapter 3 outlines the methodology and the analysis of the proposed computer vision solutions. Chapter 4 evaluates the solution and presents the results of the thesis using different metrics. Chapter 5 summarizes the works by discussing the accuracy of the proposed models as well as suggesting directions for future work.

### References

1. Wanyi Li, Fuyu Li, Yongkang Luo, Peng Wang, and Jia sun. Deep domain adaptive object detection: a survey. February 2020.
2. Poojan Oza, Vishwanath A. Sindagi, Vibashan VS, and Vishal M. Patel. Unsupervised domain adaptation of object detectors: A survey. May 2021.