Equipment identification through image recognition

Saidnassimov Darkhan

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 10.7.2022

Supervisor

Prof. Alexander Ilin

Advisor

Dr Christian Binder

Aalto University
School of Electrical
Engineering

Aalto University
School of Electrical Engineering

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the master's thesis

Author Saidnassimov Darkhan	
Title Equipment identification through image recognition	
Degree programme Automation and Electrical Engineering	
Major Control, Robotics and Autonomous Systems	**Code of major** ELEC3025
Supervisor Prof. Alexander Ilin	
Advisor Dr Christian Binder	
Date 10.7.2022	**Number of pages** 17+5 — **Language** English

Abstract

Your abstract in English. Keep the abstract short. The abstract explains your research topic, the methods you have used, and the results you obtained. In the PDF/A format of this thesis, in addition to the abstract page, the abstract text is written into the pdf file's metadata. Write here the text that goes into the metadata. The metadata cannot contain special characters, linebreak or paragraph break characters, so these must not be used here. If your abstract does not contain special characters and it does not require paragraphs, you may take advantage of the abstracttext macro (see the comment below). Otherwise, the metadata abstract text must be identical to the text on the abstract page.

Keywords Computer vision, domain adaptation, object detection, 3D models

Preface

I would like to thank Professor Alexander Ilin at Aalto University for his excellent guidance. Additionally, I would like to thank Dr Christian Binder for offering the opportunity at Metso Outotec and providing full support throughout the process. Finally, I would like to thank the CSC Finnish IT center for science for the computing resources that made the research possible.

Otaniemi, 10.7.2022

Saidnassimov, D.

Contents

List of Figures

List of Tables

Symbols and abbreviations

Symbols

$\mathbf{B}$	magnetic flux density
c	speed of light in vacuum $\approx 3 \times 10^8$ [m/s]
ω_{D}	Debye frequency
ω_{latt}	average phonon frequency of lattice
$\uparrow$	electron spin direction up
$\downarrow$	electron spin direction down

Operators

$\nabla \times \mathbf{A}$	curl of vectorin $\mathbf{A}$
$\frac{\mathrm{d}}{\mathrm{d}t}$	derivative with respect to variable t
$\frac{\partial}{\partial t}$	partial derivative with respect to variable t
$\sum_i$	sum over index i
$\mathbf{A} \cdot \mathbf{B}$	dot product of vectors $\mathbf{A}$ and $\mathbf{B}$

List of Abbreviations

AC	alternating current
APLAC	an object-oriented analog circuit simulator and design tool (originally Analysis Program for Linear Active Circuits)
BCS	Bardeen-Cooper-Schrieffer
DC	direct current
TEM	transverse eletromagnetic

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 possiblities arise in numerous fields, such as security monitoring, medicine and self driving vehicles. [2] Although over the last decade computer vision has been emerging in indurstial 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 need 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.

Although various methods have been implemented for detection of objects in various domains[2], these methods require an extensive dataset of images in order to work. In our scenario, the images can easily be collected from the simulator, however, this raises certain restrictions on the accuracy of the models as the models will not perform as well on real images.

This thesis proposes feasible object detection algorithms as a solution to automatically localize and identify the equipment in a large environment and minimize the delay before the production is online again.

Thesis objective

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

Additionally, a method should propose a solution to optimize the laborous data collection and labelling 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 compare multiple methods for object detection and explore the most recent developments in domain adaptation field. 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. Therefore, this paper proposes to tackle the challenge by means of domain adaptation, namely Cross Domain Adaptation, Decoupled Adaptation for Cross Domain Object Detection and Progressive Domain Adaptation for Object Detection. 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 was 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 is limited to proposing the solution upon analysing 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. Although the proposed methods attempt to minimize the data collection and labelling process, it might take multiple months before a 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 on 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.

2 Related work

TODO

Computer vision

TODO

Deep convolutional neural networks

TODO

Object detection

TODO

Domain adaptation

TODO

Domain Adaptive Object Detection

TODO

Adversarial feature learning

TODO

Pseudo-label based self-training

TODO

Image-to-image translation

TODO

Domain randomization

TODO

Mean teacher training

TODO

Graph-reasoning

TODO

3 Research Methodology

TODO

Proposed method

TODO

Implementation details

TODO

Dataset

TODO [1]

4 Validation and Results

TODO

5 Analysis and Discussion

TODO

6 Conclusion and Future Work

TODO

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.

A Appendices

Liitteet eivät ole opinnäytteen kannalta välttämättömiä ja opinnäytteen tekijän on kirjoittamaan ryhtyessään hyvä ajatella pärjäävänsä ilman liitteitä. Kokemattomat kirjoittajat, jotka ovat huolissaan tekstiosan pituudesta, paisuttavat turhan helposti liitteitä pitääkseen tekstiosan pituuden annetuissa rajoissa. Tällä tavalla ei synny hyvää opinnäytettä.

Liite on itsenäinen kokonaisuus, vaikka se täydentääkin tekstiosaa. Liite ei siten ole pelkkä listaus, kuva tai taulukko, vaan liitteessä selitetään aina sisällön laatu ja tarkoitus.

Liitteeseen voi laittaa esimerkiksi listauksia. Alla on listausesimerkki tämän liitteen luomisesta.

```
\clearpage
\appendix
\addcontentsline{toc}{section}{Liite A}
\section*{Liite A}
...
\thispagestyle{empty}
...
teksti\"a
...
\clearpage
```

Kaavojen numerointi muodostaa liitteissä oman kokonaisuutensa:

$$d \wedge A = F, \tag{A1}$$

$$d \wedge F = 0. \tag{A2}$$

B Appendices2

Liitteissä voi myös olla kuvia, jotka eivät sovi leipätekstin joukkoon: Liitteiden

Figure B1: Kuvateksti, jossa on liitteen numerointi

taulukoiden numerointi on kuvien ja kaavojen kaltainen: Kaavojen numerointi

Table B1: Taulukon kuvateksti.

9.00–9.55	Käytettävyystestauksen tiedotustilaisuus (osanottajat ovat saaneet sähköpostitse valmistautumistehtävät, joten tiedotustilaisuus voidaan pitää lyhyenä).
9.55–10.00	Testausalueelle siirtyminen

muodostaa liitteissä oman kokonaisuutensa:

$$T_{ik} = -pg_{ik} + wu_i u_k + \tau_{ik}, \tag{B1}$$

$$n_i = nu_i + v_i. \tag{B2}$$

Table C1: Table example

A	1	$e^{j\omega t}$
B	2	$\Re(c)$
C	3	$a \in \mathbb{A}$

C Temp

Latex Examples 1

Latex Examples 2

Latex Examples 3

$$D(xy) = (Dx)y + x(Dy), \qquad x, y \in \mathbb{A}. \tag{C1}$$

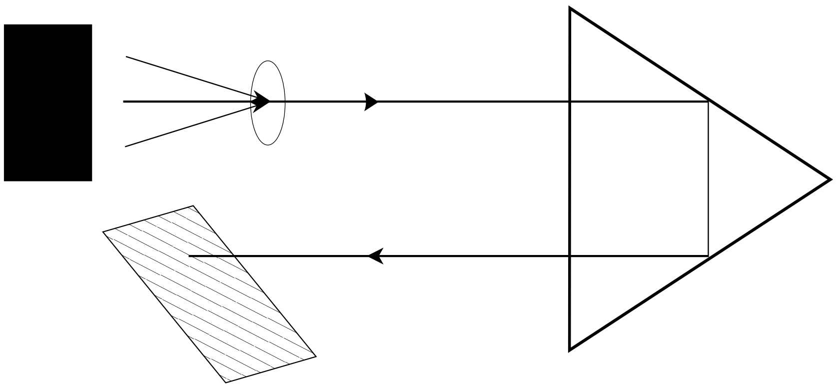

Figure C1: This is an example of a numbered caption.

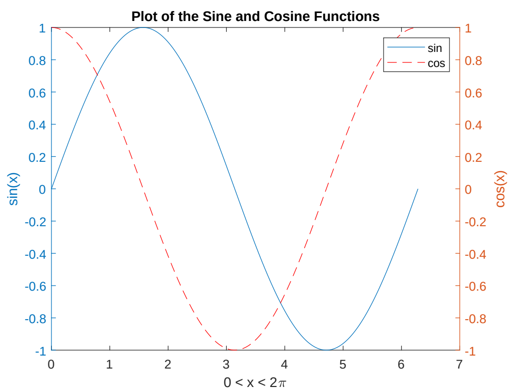

Figure C2: This is an example of a MATLAB graph.

Latex Examples 4

Footnote example [1] my footnote

Latex Examples 5

italic blabla

[2]–[2] [2, s. 83–124]

[1] My footnote ei