Title of Dissertation

**3D Object Shape Recognition and Motion tracking for Robotic pick and place application**

DISSERTATION

Submitted in partial fulfillment of the requirements of the

Degree: MTech in Artificial Intelligence and Machine Learning

By

Sai Nikhil Abinas Mohanty

2022AA05245

Under the supervision of

Arvind Raju, Principal Engineer, System SW Architecture, Intel

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE

Pilani (Rajasthan) INDIA

June, 2024

**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI**

**SECOND SEMESTER 2023-24**

DSECLZG628T **/ AIMLCZG628T DISSERTATION**

Dissertation Title : **3D Object Shape Recognition and Motion tracking for Robotic pick and place application**

Name of Supervisor : Arvind Raju

Name of Student: Sai Nikhil Abinas Mohanty

ID No. of Student : 2022AA05245

Courses Relevant for the Project & Corresponding Semester: 1. Video Analytics, Sem3

2. Computer Vision, Sem3

3. Deep Neural Network, Sem2

4. MLOps, Sem 3

## Abstract

In modern manufacturing and logistics, automation plays a critical role in enhancing efficiency, accuracy, and speed. One of the key components of automation is the ability of robotic systems to handle and manipulate objects of various shapes and sizes. However, traditional robotic arms often struggle with identifying and picking different type of objects that are moving on a conveyor belt, especially when these objects vary in shape and orientation. This challenge necessitates the integration of advanced computer vision techniques to empower robotic arms with the ability to dynamically identify and pick moving objects.

The dissertation would provide a comprehensive study and implementation on leveraging computer vision, video analytics techniques to enable a robotic arm to pick up various shapes of moving objects on a conveyor system. The project will focus on using depth camera as an input and will develop an image pre-processing and feature extraction algorithm for feature extraction and image modification best suitable for our use case.

The project will use Deep Reinforcement Learning – Self Supervised model for 3D object shape recognition and motion tracking for unlabeled videos to predict shape and location of the object. The meta data such as location and shape will be published to a ROS environment for robotic system consumption.

To make this implementation production ready, the project will also use some of MLOPs techniques during development of model such as quantization, model packaging, containerization of solution, model meta data store. Different model performance metrics will be used to evaluate the solution.

**Key Words:**

Image Processing, YOLOv5, Object shape recognition, Motion Tracking, Motion Direction, MLOPs, Object Detection, Deep Reinforcement Learning, Self-Supervision

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI**

**II SEMESTER 23-24**

DSECLZG628T **/ AIMLCZG628T DISSERTATION**

**Dissertation Outline**

**BITS ID No. 2022AA05245 Name of Student: SAI NIKHIL ABINAS MOHANTY**

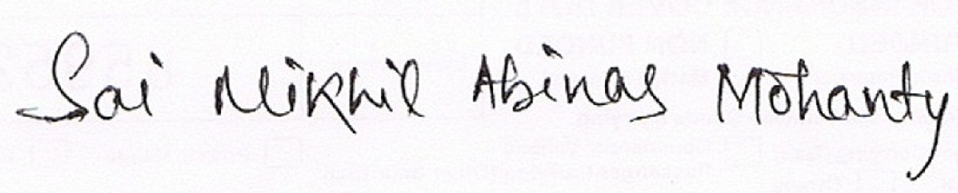
**Name of Supervisor:** Arvind Raju

**Designation of Supervisor**: Principal Engineer

**Qualification and Experience:** BE, CSE, Experience: 25+ years

**Official E- mail ID of Supervisor:** raju.arvind@intel.com

**Topic of Dissertation**: 3D Object Shape Recognition and Motion tracking for Robotic pick and place application

 A close up of a paper

Description automatically generated

(Signature of Student) (Signature of Supervisor)

Date: 07/06/2024------------ Date:--07/06/2024------

**Project Work Title**

3D Object Shape Recognition and Motion tracking for Robotic pick and place application

**1.1. Purpose:**

The purpose of this dissertation is to provide a comprehensive study and implementation on leveraging computer vision to enable robotic arm to pick up various shapes of moving objects on a conveyor system.

**1.2. Expected Outcome:**

- A Object detection and Shape recognition system that can outline shape of any type of object.

- A motion tracking system that will predict the coordinates of the object, it’s direction of motion.

- Publisher App that can publish the required predicted data to ROS environment for Robotics system consumption.

- A MLOps eco system surrounding the solution for better and seamless delivery/deployment of the solution.

**2. Literature Review:**

[1] **Cut and Learn for Unsupervised Object Detection and Instance Segmentation**

- Understood an effective method where coarse mask is drawn on the image and train the detector using loss function.

[2] **A Study on Self-Supervised Object Detection Pretraining**

- Learnt on self-supervised learning method that can be used for object detection. Learnt different object detectors, the experimental setups used and how RestNet 50 is pre-trained using R-CNN and DETR.

**3.1. Existing Process:**

Most of the industries have transformed fixed automation to programmable or flexible automation because either one of them has high adaptability and flexibility in manufacturing and automation process. Vision systems are widely used in industries especially in quality control, object sorting, and product inspection.

**3.2. Limitations:**

- For object detection problem, it requires a lot of large data set with manual annotations for learning. Hence, it becomes very time-consuming process to maintain and label such large data set.

Additionally, the model cannot predict the object which is new or not part of labelled dataset.

Most of the models are trained with a particular type of objects or shape of the object. The paper will focus on creating a model that can detect and predict any shape of the object.

**4. Justification for Methodology:**

- Image pre-processing model: This will focus on predicting shape of any type of object. So that the model can be more generic.

- The method will focus on being generic enough to predict motion with frames with different backgrounds.

- Deep reinforcement learning with self-supervised learning model will be used.

- Containerized Deployment -> This will reduce the environmental dependencies and user will be able to deploy in their environment seamlessly.

**5. Project Work Methodology:**

- Data Exploration: Collect Data sets, analyze, extract features, pre-process

- Shape Recognition: Predict shape of any object and create bounding surrounding boxes.

- Motion Tracking: Predict motion direction of the object

- Publisher App: Publish predicted meta data to ROS Environment

- MLOPs work flow development

**6. Benefits Derivable from the Work:**

- Model independent of object shape.

- Model independent of image background.

- Accuracy on object position prediction.

- Production ready solution

1. **Broad Area of Work**

The project focuses on solving a robotics arm pick and place problem by leveraging computer vision. Automation in Manufacturing and Logistics through Advanced Computer Vision and Machine Learning Techniques.

* Computer Vision Techniques
* Image Pre-processing and Feature Extraction
* Deep Learning for Object Recognition
* MLOps Techniques

1. **Objectives**

The objectives of my project are as follows:

* To pre-process the input image for making the frame suitable for analytics and to extract features.
* To use Deep Reinforcement Learning algorithms for **Object Detection and Shape Recognition**
* Motion Detection, Estimation using motion tracking and prediction algorithms.
* Publish the orientation and estimation data to a ROS environment for Robotic system consumption.
* Prepare the solution for production ready by using MLOPs.

# 3. **Scope of Work**

Scope of this dissertation is to design and develop –

* Implement image pre-processing Algorithm.
* Selection and implementation of object detection and shape recognition algorithm.
* Development of Motion Tracking Algorithm
* Implementation of predictive models to estimate object trajectories.
* Publish the prediction data to ROS environment for Robotics arm to consume.
* Containerize the solution using docker.

**4. Detailed Plan of Work** (for 16 weeks)

The plan of work should have tangible weekly or fortnightly milestones and deliverables, which can be measured to assess the adherence to the plan and therefore the rate of progress in the work. The plan of work can be specified in the table given below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Serial Number of Task/Phases | **Tasks or subtasks to be done** (be precise and specific) | **Start Date-End Date** | Planned duration in weeks | Specific Deliverable in terms of the project |
| 01 | Design of Solution | 08/06/2024-17/06/2024 | 1 Week | Work flow and block diagram design |
| 02 | Data Collection | 18/06 – 21/06 | < 1 Week | Data Collection |
| 03 | Image Pre-processing Algorithm Development | 25/06- 10/07 | 2 Week | Feature Extraction, Feature Engineering, model development |
| 04 | Object Shape Recognition Algorithm Development | 10/07- 25/07 | 2 Week | Model development, model evaluation |
| 05 | Development of predictive model for motion tracking | 22/07 – 05/08 | 2 Week | Model development, model evaluation |
| 06 | Publisher code for prediction data such as coordinates to ROS environment | 06/08 – 14/08 | 1 Week | Publisher code, preparation of ROS Env |
| 07 | MLOPs: Dockerfile creation for the solution, model meta data store into influxdb | 15/08- 30/08 | 2 Week | Containerization, DataStore implementation |
| 08 | Testing of the Solution | 31/08 – 07/08 | 1 Week | Use Case Testing, model testing |
| 09 | Final Report preparation, Documentation for Solution | 08/08 – 15/08 | 1 Week | README, User Guide, Final Report |

# Literature References

The following are referred journals from the preliminary literature review.

1. [*https://arxiv.org/pdf/2207.04186*](https://arxiv.org/pdf/2207.04186)
2. [*https://openaccess.thecvf.com/content/CVPR2023/papers/Wang\_Cut\_and\_Learn\_for\_Unsupervised\_Object\_Detection\_and\_Instance\_Segmentation\_CVPR\_2023\_paper.pdf*](https://openaccess.thecvf.com/content/CVPR2023/papers/Wang_Cut_and_Learn_for_Unsupervised_Object_Detection_and_Instance_Segmentation_CVPR_2023_paper.pdf)
3. [*https://www.semanticscholar.org/paper/The-Graph-Matching-Optimization-methodology-for-in-Willaume-Parrend/bee2f48b09a3f3d25ec9ba374f354b8cd0ee8858*](https://www.semanticscholar.org/paper/The-Graph-Matching-Optimization-methodology-for-in-Willaume-Parrend/bee2f48b09a3f3d25ec9ba374f354b8cd0ee8858)

**Supervisor’s Rating of the Technical Quality of this Dissertation Outline**

EXCELLENT / GOOD / FAIR/ POOR: \_\_\_GOOD\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Supervisor’s suggestions and remarks about the outline (if applicable).**

Date: 07/06/2024 \_\_\_\_\_\_\_A close up of a paper

Description automatically generated\_\_\_\_\_\_\_\_\_\_\_\_ (Signature of Supervisor)

Name of the supervisor: Arvind Raju

Email Id of Supervisor: raju.arvind@intel.com

Mob # of supervisor: 9986969502