Recognition POC

# Introduction

## Overview and purpose of the project

The Recognition POC application is a software tool that is designed to improve the efficiency and accuracy of the work process by reducing errors that can occur when workers take products from the wrong compartment of a bin. To achieve this, the application uses a camera placed above the bin, which is divided into compartments and processes the recording to detect the hand of the worker as they reach into the bin to take a product. After that, it can be easily determined whether he took the product from the expected compartment or not.

# System architecture

The Recognition POC application consists of two parts: a recognition system and a recognition API.

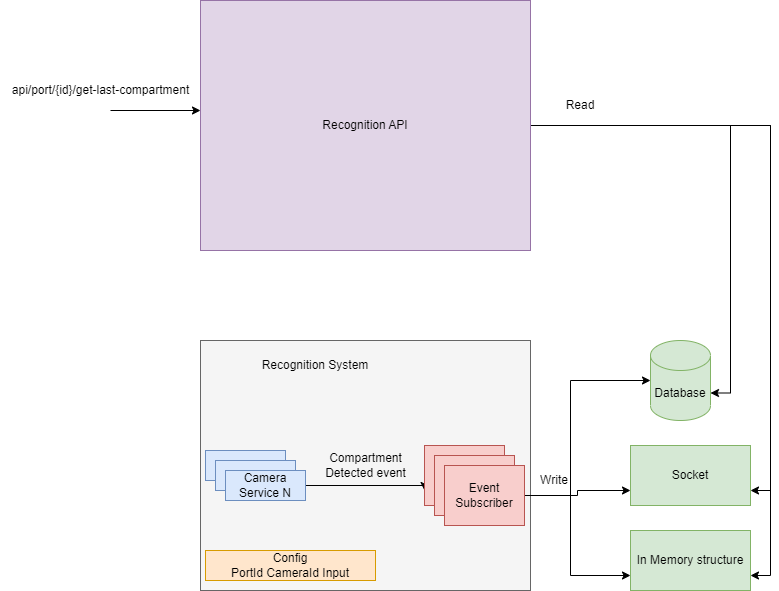


Figure System architecture

The recognition system is a combination of machine learning and computer vision technologies that are responsible for analyzing recordings and video streams and detecting when a hand has taken something from a bin. The recognition system can work with multiple recordings and cameras simultaneously. When a product is taken, an event is triggered. The application can then store information about that event in various places such as databases, sockets, and others. Currently, the system writes the event to a PostgreSQL database.

# The recognition API is a .NET application that can read from the database and provides endpoints for accessing information about product-taking events that have been recorded by the recognition system. It provides two endpoints. The first endpoint returns information about the last object taken from a specific port. The second endpoint returns whether something was taken from a specific compartment in a predefined period, based on the port and compartment specified.

# Recognition system

## Methodology

## To monitor the action, a video camera is set up and calibrated to ensure it is positioned correctly above the box. The application processes the video feed using the OpenCV library. For each frame application checks for certain conditions to determine if an object is being taken from the bin, such as the position of the hand, the speed of the hand movement, and the direction of the hand. Once it has determined that an object is being taken, the application records the coordinates of the hand and uses them to determine which compartment of the bin the object was taken from.

## Application workflow

The camera is constantly recording the bin and creating frames from the video feed. For each frame, the application checks if a hand is present. If a hand is detected, the application then checks for three conditions to determine if the hand is taking an object from the bin.

The first condition is the speed of the hand movement. It is common for the hand to move faster toward the object it wants to lift, and then slow down as it reaches the object. If the hand is slower than a predefined value, it is considered that it may have slowed down due to picking up an object.

The second condition is the position of the hand. If the hand is extended or slightly bent, it is considered that there is no object in it. However, if the hand is in a position as if it is holding or taking an object, this condition is met.

The third condition is the change in direction of the hand. If the hand is predicted to continue moving in one direction, but it starts to come back, this could indicate that the hand has changed direction because it has taken something.

If all three conditions are met, it can be concluded that the hand is taking something from the bin at that moment. The coordinates of the hand are then used to determine which compartment of the bin the object was taken from.

## Libraries

The application was created in the Python programming language, and the following libraries were used for its creation:

### OpenCV

### OpenCV (Open-Source Computer Vision) is a free and open-source library of computer vision and machine learning algorithms. In this project, it was used for the manipulation of tasks on video data, including reading video files and processing video frames.

### OpenCV provides a wide range of functionality including image processing, machine learning algorithms, and object tracking.

### One model used in the project is the Kalman Filter, which is a mathematical model that can predict the future location of an object based on its past movement. This model is used to recognize the moment when the hand has a sudden change of direction.

### Mediapipe

To detect the hand, the application uses the Mediapipe library, which is a framework for building machine learning models for multimedia processing tasks. The Mediapipe library is particularly well-suited for hand detection, as it uses machine learning algorithms to analyze the video feed and identify the location and movements of the hand.

# Possible Improvements

Some of the possible improvements are:

1. Determine what exactly is taken from the compartment using computer vision: This improvement would involve using computer vision techniques to identify and classify the specific products or items that are being taken from the compartments of the bin. This could involve training a machine learning model to recognize different types of products based on features such as shape, color, size, and other visual characteristics.
2. Detect the bin and determine its dimensions: Currently, the dimensions of the bin are predefined and hardcoded into the application. Implementing this improvement would allow the application to automatically detect the presence and dimensions of the bin in the video stream, eliminating the need to manually set these parameters.
3. More accurately determine the moment of change of direction. Throwing out the Kalman filter throws out the prediction of where the hand will be next, the idea is to create a specialized function that will track previous experiences and based on them conclude when a change of direction has occurred.

# Conclusion

The recognition POC application uses a combination of machine learning algorithms and computer vision techniques to detect the hand of a worker as they reach into a bin, track its movements, and determine which compartment of the bin is being accessed. By automating this process, the application aims to improve the efficiency and accuracy of the work process, reducing errors and improving overall productivity. It is able to capture whether the worker picked an item from expected compartment.