

# Recognising soda brands using Machine Learning

Oliwier Jaworski, Justice Lacroix

**Abstract:** This project explores real-time video stream processing using RTSP for dynamic object detection and identification, leveraging machine learning to accurately recognize and label detected objects. Seamless streaming and interaction are facilitated through WIFI connectivity. This solution is ideally suited for automation applications, such as sorting glass bottles in factories, eliminating the need for manual sorting. Overall, this approach significantly enhances efficiency in industrial processes.

## Introduction

This project aims to develop a machine learning application that can recognize at least five different objects in [figure 1] real-time video streams (RTSP) from Wi-Fi cameras. Leveraging deep learning techniques with PyTorch, we will explore the entire process from building and training the model to deploying it on new datasets. By applying advanced video analytics, the project will extract meaningful insights from live video feeds, unlocking the potential of AI-based video understanding



Figure 1: *example* objects to be recognized.

## Description of the setup

The system captures live video feeds from Wi-Fi-enabled cameras, selected for their flexibility and high-bandwidth capabilities. This connectivity allows for easy deployment and adaptability without the constraints of physical cabling.

The main input is the live video stream, enabling continuous monitoring for object detection tasks. The output consists of annotated video streams, where detected objects are labeled (e.g., "Coca Cola," "Fanta," "Ice Tea") with bounding boxes indicating their locations. This provides clear, real-time feedback to users.

Data processing will focus on object detection[1], with the model analyzing each video frame to accurately categorize objects. OpenCV[2] will serve as the graphical user interface (GUI), facilitating video display and user interactions. It allows for the overlay of bounding boxes, labels, and confidence scores directly onto the video frames, making it user-friendly and efficient.

## Overview

- **Inputs**
  - Video stream (RTSP)
- **Outputs**
  - video stream.
  - Detected objects.
- **Processing**
  - Object detection
- **Connectivity**
  - WIFI
- **User interface**
  - OpenCV

## References

- [1] Wikipedia contributors. Object Detection. Wikipedia, *The Free Encyclopedia*. **Last modified August 20, 2024**. Available online: [https://en.wikipedia.org/wiki/Object\\_detection](https://en.wikipedia.org/wiki/Object_detection)
- [2] Siromer. OpenCV GUI for YOLO Object Detection Models. *Medium*, **Aug 20, 2024**. Available online: <https://medium.com/@siromermer/opencv-gui-for-yolo-object-detection-models-ad4a8ba07e68>