**Overview of assignment:**

This project is about object detection using YOLO and image classification using Resnet for Lego pieces. The goal is to be able to tell which Lego piece by showing it to the camera therefore we do not need to struggle finding a specific piece during the buildup process.

**Description of data:**

Initial data are featured with 200 most popular Lego parts with 2000 images. Each images contains more than 100 Lego pieces. The annotation was stored in Pascal VOC format(XML).

For Resnet image classification I tried later on, I cropped the Lego parts from images and stored each to its own class folders. There were total of 12134 training images and 6014 validation images.

**Summary of methods:**

Object detection is a good way when we are trying to search for something specific from a bunch of parts. However, Object detection will be harder because it requires the model to scan through and identify the location of an object.

Classification on the other hand will be slightly easier because it is telling an object from an image where there is less misleading edges and patterns.

**Summary of model:**

For object detection I used YOLO v8n and YOLO v11n. I first tried it with YOLO v8n for 5 epochs, imagesize = 1080. Then I tried YOLO v11n for 10 epochs with a higher resolution 1440. Trying go above 1440, for example 2080 will return errors.

For image classification I used Resnet 50 and did some adjustments for final dense layers and fine tuning.

**Analysis of results**

Result for object detection did not work well. At first I consider bad predictions are due to lack of training time and low resolution setting on high resolution images. However, it was not the case, model’s accuracy stopped increasing after 8th epoch and stayed at same level for the last two epochs. Changing the model and increasing resolution only makes it slightly better.

Result for image classification is much better. I spent almost two days to figure out ray tuning and finish the work. The best results stayed about 45%. Consider it a 20 class classification with ~800 images per class, it is better than the return result of YOLO.