EE482 – Introduction to Artificial Intelligence

Semester 1 2023/2024

Title: Course project
Project 2: Distance Estimation

Name	ID
Mohammed Subhi	2035752
Talal Alqarni	2036183
Moaid Aljabri	2035724

Introduction

In this report, we are going to discuss a secondary project that we have to work on for the course EE482. The project's goal is to measure the distance between a camera and an object. Initially, we will focus on training a model to recognize an object and draw a bounding box around it. For this purpose, we will use the YOLOv8 model. After the model is trained, we will employ the bounding box drawn by the model in three distinct methods to estimate the distance between the object and the camera.

Data Set

The dataset consists of various pictures of Hot Sticks chips products taken from different distances. The data was collected by placing a camera over a scaled paper set at 0cm. The product was then positioned on the scaled paper, and multiple pictures were taken at varying distances, ranging from 18cm to 60cm. as you can see in the following figure(1).

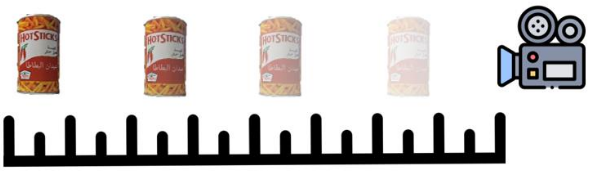

Figure 1: how data collected

These images have been divided into three groups: training, validation, and testing. All the images in the training and validation groups have been labeled to train the model.

YOLOv8

The YOLOv8 model was selected for its high performance. This model excels in accurately detecting objects, which is essential for achieving accurate results and reducing the error percentage. The effectiveness of the model improves with an increase in epochs. However, the extent of this improvement depends on the size and variety of the datasets available. The trained model will generate a bounding box that identifies the object's location within the image. This bounding box includes coordinates for the top-left corner and the bottom-right corner, which need to be as accurate as possible.

Distance Estimation

The accuracy of the estimated distance relies on the precision of the model from which data is extracted from the bounding box. If the bounding box inaccurately represents the actual size of the object's base in the image, the error in the estimated distance will increase. The object's size in the image can be calculated using the coordinates of the bounding box's corners. Specifically, the top-left corner provides coordinates X1 and Y1, while the bottom-right corner gives X2 and Y2. By subtracting X1 from X2 and Y1 from Y2, the object's width and height in the image can be determined.

By understanding the object's size in pixels on the image and its real distance from the camera, it is possible to estimate the distance. The following three methods will utilize these variables to estimate the distance.

Method 1:

This method works using a reference image to estimate distance. The reference image is selected from the testing dataset, which helps indirectly reduce the error in the trained model by including it in the calculations. The variables required for this method are the known real distance of the object in the reference image and the object's width in pixels on that image. The method operates by calculating the width of the object on the image and then using these reference variables in the following equation.

$$Estimated\ Dsitance = (reference\_width\_in\_pixels \times reference\_distance)/detected\_width\_in\_pixels$$

For optimal results, it's advisable to use an image with a moderate distance from the camera. The image used in this method is at a distance of 35cm, which is approximately midway in the desired range of 18-60cm. The accuracy of this approach is quite satisfactory, as you can see in table(1).

Method 2:

Similar to the previous method, this approach also utilizes a reference image. However, in this method, three reference images are used. After calculating the distance using each of the three references, the average of these three distances is taken to determine the final estimation. The following equations will be used.

$$Estimated\ Dsitance1 = (reference\_width\_in\_pixels1 \times reference\_distance1)/detected\_width\_in\_pixels$$

$$Estimated\ Dsitance2 = (reference\_width\_in\_pixels2 \times reference\_distance2)/detected\_width\_in\_pixels$$

$$Estimated\ Dsitance3 = (reference\_width\_in\_pixels3 \times reference\_distance3)/detected\_width\_in\_pixels$$

$$Estimated\ Dsitance = (Estimated\ Dsitance1 + Estimated\ Dsitance2 + Estimated\ Dsitance3\_)/3$$

To achieve the best outcomes, images from the beginning, middle, and end of the range (18cm, 35cm, and 60cm) have been utilized. The accuracy of this method is quite acceptable, as can be observed in the table(1).

Method 3:

In this method, a large number of images serve as references, from which data is extracted and used to determine a regression line. The real distance and the height of the object in the image were extracted from approximately 200 images to derive the following regression line equation. Additionally, the plotting of this equation is illustrated in the following Figure(2).

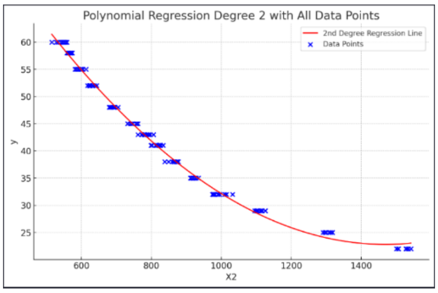

Figure 2: regression line plotting

$$Y = 0.0000424514x^2 - 0.1247301856x + 114.4322942610$$

The results at a distance of 18cm are quite poor, but they improve significantly and normalize after 25cm, as you can see in table(1).

Table 1: methods results

Real Distance	Method 1		Method 2		Method 3	
	Distance	Error%	Distance	Error%	Distance	Error%
18cm	18.37cm	2.06%	18.67cm	3.72%	29.61cm	64.50%
22cm	22.53cm	2.41%	22.91cm	4.14%	23.15cm	5.23%
25cm	24.58cm	1.68%	24.99cm	0.04%	23.83cm	4.68%
29cm	29.15cm	0.52%	29.63cm	2.17%	28.5cm	1.72%
32cm	32.84cm	2.63%	33.38cm	4.31%	32.71cm	2.22%
35cm	34.82cm	0.51%	35.39cm	1.11%	35.62cm	1.77%
38cm	37.37cm	1.66%	37.98cm	0.05%	37.65cm	0.92%
41cm	40.91cm	0.22%	41.58cm	1.41%	40.8cm	0.49%
43cm	41.99cm	2.35%	42.68cm	0.74%	42.04cm	2.23%
45cm	43.65cm	3.00%	44.37cm	1.40%	45.61cm	1.36%
48cm	48cm	0.00%	48.78cm	1.63%	49.66cm	3.46%
52cm	49.12cm	5.54%	49.93cm	3.98%	52.19cm	0.37%
55cm	54.07cm	1.69%	54.96cm	0.07%	55.65cm	1.18%
58cm	55.91cm	3.60%	56.83cm	2.02%	57.72cm	0.48%
60cm	57.89cm	3.52%	58.84cm	1.93%	58.78cm	2.03%

Discussion:

The results of methods 1 and 2 are quite satisfactory, with a reasonable error percentage. However, using the third method is not recommended, as the regression line appears to overfit the data. This overfitting might be the cause of the very poor results at 18cm.