Project: Inventory Monitoring at Distribution Centers

Domain Background

The employment of robots for stock management in logistic planning grows constantly in distribution centers daily operations. Goods are moved by robots in bins that can contain from one to several items. This project takes that into consideration and uses this trend to propose a usage of Machine Learning (ML) techniques to optimize logistic operations involving robots.

A robot that is able to count the number of items in each bin it moves would improve stock management. And the ML engineer with experience with that kind on computer vision skill can be an asset for companies.

Problem Statement

The problem I intend to solve in this project is the optimization of storage management applying computer vision. to robots used in stock operations.

Solution Statement

The chosen method to solve the proposed problem is to give robots, and consequently the controlling system, the ability to identify the number of items being transported in bins used for storage.

It is assumed that this capability could reduce time and possible mistakes in storage management processes.

Datasets and Inputs

The dataset used for this project is the Amazon Bin Image Dataset. The dataset can be found here.

This dataset holds 500,000 images and metadata from bins of a pod in an operating Amazon Center. The images were captured as the bins were being transported by robots.

The images are not divided into labelled folders. In order to use them a prior division of train images would be necessary using the metadata JSON files by Amazon in the dataset S3 folder. That step is not necessary for this project since Udacity has kindly provided a JSON file that labelled 10,441 images into 5 different folders ranging from 1 item to 5 items in each picture. Therefore, it was decided that the files to be used for training, validation and testing would be the ones listed in the forementioned JSON file.

This reduced dataset is also convenient since the AWS resources used for training in this project are limited.

Benchmark Model

Since I intend to modify the fully connected layers and the default hyperparameters of the pretrained ResNet50 network, the benchmark model for this project will be the original ResNet50 model as provided by the torchvision API.

The paper presenting the original ResNet models can be found here.

Evaluation Metrics

The main evaluation metric for comparison between the benchmark and the solution models will be result accuracy, that is, the number of correct inferences provided by the model divided by the total number of inferences made.

This metric makes sense since it encompasses the main goal of the project: to obtain a model capable of identifying the number of items in the images in the best possible way.

Project Design

My planned workflow for the resolution of this problem is:

- Creation of an ETL Pipeline in order to obtain data from the Amazon Bin Dataset, process and load it to an S3 bucket;
- Train a ML vision model using the database previously created;
- Using AWS tools, make sure the best practices are being employed and the processes are cost effective.

The ETL Pipeline will load data, separate it into train, validation and test sets and load them into separate S3 folders.

The ML model responsible for identifying the number of items in each image will be based on a pretrained ResNet50 model from torchvision that will have its fully connected layers finetuned looking to obtain better results.

After finetuning, an attempt of model improvement via hyperparameter optimization will be performed.

Finally, after arriving at a satisfactory model, the training script will be modified in order to insert debugging and profiling mechanisms. That final step aims to provide better code and instance monitoring.

All steps will be recorded in a jupyter notebook and a python training script to ensure replicability.