Machine Learning Engineer Capstone Project: Inventory Monitoring at Distribution Centers

Counting objects from the images of bins taken by an operating Amazon FC

Domain Background

Distribution centers are the foundation of a supply network which is stocked with goods to be redistributed to retailers, wholesalers or the customers. It is often defined as a warehouse which plays a key role in order fulfillment process. Amazon uses Fulfillment Center (FC) instead of "warehouse" because a standard warehouse keeps inventory until the store, whereas Amazon's warehouses ship items to the customers in addition to storage [1]. These facilities use robots to move objects from one place to another. The objects are carried in bins which can contain multiple objects. As Amazon ships ~1.6M packages a day, automation is key to optimize operations and speed up delivery [2].

Problem Statement

Inventory management allows companies to identify which and how much stocks needed at the point of a time to fulfill customers' requests. However, without using intelligence, the process is highly manual for workers to keep tracking. The main problem we focus on this project is the manual inventory monitoring to track the number of objects being sent or received. Manual classification of how many objects in each bin in the warehouse would take +20sec per bin considering the process of scanning/logging it. For Amazon, it would make 3K hours productivity loss every day (assumption of avg 3 object per bin). This is not only a problem which is related with Amazon but from all industries companies deal with inventory management.

Solution Statement

To facilitate inventory management, mostly robots are used to move objects from point A to B. If we train robots to be able count number of items in each bin, we could save workers' time with automation and intelligence. Image Classification from Deep Learning would be the method to achieve this training. As a platform, we could leverage AWS' scalability, reliability, availability and security to build, train and deploy our models. Number of correctly counted objects in each bin would be metric we will look at as an accuracy. This will give us insight on at broader scale which kind of productivity increase we could achieve. All the codes will be accessible on GitHub and with small adjustments on data preparation, this model would be able used in different occasions. As a model, I am planning to use Resnet50 which is a convolutional neural network (CNN) that is 50 layers deep with higher accuracy than Resnet34. I will utilize pretrained model and support it with few fully connected layers [3].

Datasets and Inputs

We currently have Amazon Bin Image Dataset which contains over 500,000 images and metadata from bins taken by operating Amazon Fulfillment Center. As it would be costly to train a model which has +500K images, for this project we will use 10.4K images which are provided us by Udacity team via JSON file. We have images from the bins which include between 1 to 5 objects. Normally, this dataset provides labels in the format of metadata. However, Udacity provided us in the format that we could directly get the labels. After having downloaded data, we will need

to do a split as train, test and valid. Important piece to notice is that images have different sizes in the dataset. It is important practice to transform them before training.

Benchmark Model

As a benchmark, I will leverage from "usage examples" from Amazon Bin Image Dataset as a publication [4]. Main reason is that project owner also leveraged accuracy metric and followed up a similar approach. However, the hyperparameter, dataset I will use will be more conservative such as (40epochs vs 5 or 500K images vs 10K). As it is shown below, the benchmark model achieved 55.67% accuracy. Due to the size of our dataset and cost-efficient methods we will use, I would set up a goal of 25-30% accuracy in my model.

Accuracy(%)	RMSE(Root Mean Square Error)
55.67	0.930

Evaluation metrics

The evaluation metric will be the accuracy. Given the problem we would like to solve, we need to be able to correctly classify how many objects there are in each bin. The calculation of it will be as [Correct Inferences/Total Inferences]. The benchmark model shows that it can classify number of objects in each bin with 55.67% accuracy.

Project Design

I will be using PyTorch framework with Amazon Sagemaker which is a fully managed service to build, train, and deploy machine learning (ML) models with fully managed infrastructure, tools, and workflows [5]. This service will provide us easy to create notebook instances, training jobs, endpoints with pay for what you use pricing. The project flow will look like:

1. **Data Collection** – Downloading from Amazon Bin Image Dataset and upload to S3 bucket
2. **Exploratory Data Analysis** – Identifying the volumes of each class and its distribution
3. **Training a Deep Learning model** & Hyperparameter tuning – Resnet50 will be used with fully connected layers – it will be a pre-trained neural network to add my own fully connected layer and train it.
4. **Model Evaluation** – Comparison with benchmark by keeping conservative approach in our mind
5. **Model Deployment** – Through Sagemaker endpoints, I will predict few instances to think big of replicability of this model.

References

1. "Amazon Facilities." *US About Amazon*, US About Amazon, 21 Sept. 2020, https://www.aboutamazon.com/workplace/facilities.
2. "57 Amazon Statistics to Know in 2022." *LandingCube*, 20 June 2022, https://landingcube.com/amazon-statistics/.
3. "Deep Network Designer." *ResNet-50 Convolutional Neural Network - MATLAB*, https://www.mathworks.com/help/deeplearning/ref/resnet50.html.
4. "Amazon Bin Image Dataset." *Amazon Bin Image Dataset - Registry of Open Data on AWS*, https://registry.opendata.aws/amazon-bin-imagery/.
5. Hudgeon, Doug, and Richard Nichol. "Machine Learning for Business: Using Amazon Sagemaker and Jupyter." *Amazon*, Manning Publications, 2020, https://aws.amazon.com/sagemaker/.