Project Proposal
Name: Joan Gerard

Domain Background

Distribution centers often use robots to move objects as a part of their operations. Objects are carried in bins which can contain multiple objects. These objects are then shipped to the customers who order them for instance via Amazon.

Inventory Monitoring at Distribution Centers is an important task for some companies in charge of delivering packages. When you order something from Amazon they need to make sure that they are delivering all the product items you ordered within a single package, nothing more and nothing less. A simple way to verify this is making sure that the number of objects inside the bins matches the number of objects the customer has ordered.

The problem of finding the number of objects given a picture of the product items conained in a bin is not new and many researchers and data scientists tried to resolve it. Park [1] used used the dataset to make object counting, object verification, and object quantity verification training a custom model on top of a ResNet with 34 layers for 40 epochs, a batch size of 128 and obtained an accuracy of 55.67%. Bertorello et al. [2] compared two methods using SVM and different models of CNNs such as ResNet18, ResNet34, and ResNet50 to predict the quantity of the bins obtaining an accuracy of 56.2% with ResNet34 outperforming SVM by 75%. Xu et al [3] also applied CNNs to solve the problem using data augmentation obtaining an accuracy of 54.64%.

Problem Statement

The problem is avoiding distribution centers to deliver packages with missing items. An image of a bin containing the ordered items is provided and the number of objects in the bin needs to be determined.

Solution Statement

This project uses the Amazon Bin Image Dataset that contains images of orders together with the amount of products that are in every order. The dataset contains 500K images but only a subset of them are used here to train a model using ConvNets. The objective of this project is to build a classifier that receives an image of an order with the products in it and it tries to guess how many products there are. Also, this project illustrates mainly how to train/evaluate/deploy a model in Sagemaker to be ready to be used for making predictions via an endpoint.

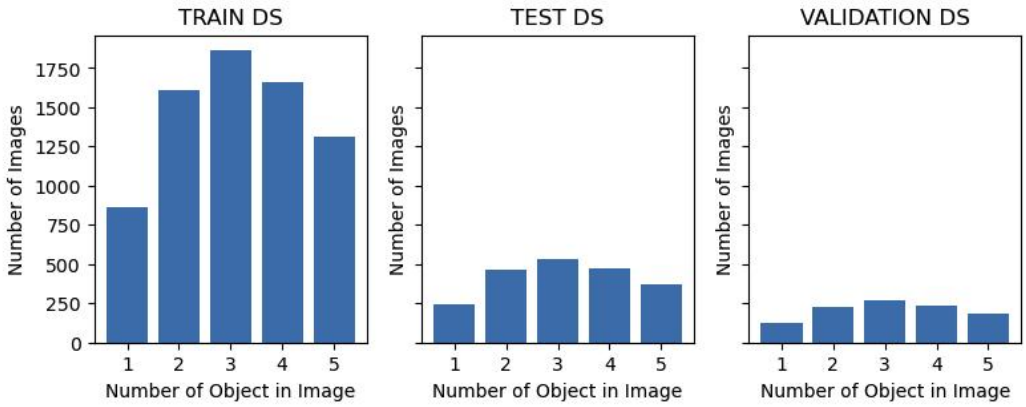
Datasets and Inputs

The Amazon Bin Image Dataset [4] contains over 500,000 images and metadata from bins of a pod in an operating Amazon Fulfillment Center. The bin images in this dataset are captured as robot units carry pods as part of normal Amazon Fulfillment Center operations.

The dataset was obtained from here and it contains images of customer's orders previously to be sent to them. Also contains labeled information about the images such as how many products it has, what the products are, their weights, etc.

The dataset is publicly available and it can be accessed either from an S3 bucket or HTTP request.

Only a portion of the dataset will be used for this project due to time and budget constrains. A distribution of the data already split in training, testing, and validation datasets is shown here:



The plan is to work with only 5 classes (number of objects in the bin) and the number of samples per class does not surpases 2000. That will help saving training time but it will unavoidably affect to the final accuracy.

Benchmark Model

There are many architectures for ConvNets for image classification in general that can be used for this project such as ResNet18, ResNet50, ImageNet, etc. Nevertheless, models cited here presented an accuracy of around 55% using ConvNets. The one that presented a higher accuracy was the one proposed by Bertorello et al. using a ResNet34.

Evaluation Metrics

The metric used for assess the classifier is the accuracy, defined as follows:

*Accuracy = Number of Correct Predictions/Total Number of Predictions * 100%*

Project Design

The project will consider the following steps:

- *Data downloading:* the dataset is downloaded locally and then upload to S3 so it can be used by other instances:
- *Data cleaning:* some of the images presented incorrect annotations of were very blurry so they were deleted from the dataset
- *Data transformation:* the images were resized to 224x224 pixels and rotated randomly before training the model
- *Model Selection:* hyperparameter tuning was performed with 4 jobs and then the best model was selected to be trained.
- *Model Training:* the best model was trained and evaluated
- *Debug and Profiling:* a complete profiler report was obtained with relevant information such as GPU/CPU utilization and debugger information such as training loss was reported.
- *Model Deployment:* it deploys the best model to a Sagemaker endpoint.
- *Model Inference:* once the model is deployed, it can be used to make predictions for different images never seen before.

Conclussion

In conclusion the project objective is to create a classification model that will predict the number of objects found on a bin picture, for that purpose the Amazon Bin Images Dataset will be used to train the model. The process of creating such model will include the downloading, cleansing transformation of the data and then model selection, training and evaluation. Additionally, the model will be deployed for making further model predictions.

Bibliography

[1] Park, Eunbyung, Amazon Bin Image Dataset (ABID) Challenge GitHub Repository, https://github.com/silverbottlep/abid_challenge. July, 2017.
[2] Rodriguez Bertorello, Pablo; Sripada, Sravan; and Dendumrongusup, Nutchapol. Amazon Inventory Reconciliation Using AI. December, 2018.
[3] Xu, Zihao and Salloum, Mariam. Deep Neural networks for Object Enumeration. 2018. DOI:10.1109/BigData.2018.8622191
[4] Amazon Bin Image Dataset was accessed on July 2023 from https://registry.opendata.aws/amazon-bin-imagery