Udacity AWS Machine Learning Engineer Capstone Project

Oluwafunmilola Esther Egbekunle |19-02-2022

I. DEFINITION

Project Overview

Distribution centres are integral to the smooth running of a supply chain network. It is often a large warehouse where products/Goods are stocked to be redistributed to wholesalers or retailers. This calls for increased movements of goods in and out for an effective supply chain management system. The goods transported needs to be handled in a timely, efficient and prompt manner.

Distribution centers often employ the use of robots to move objects as a part of their operations. These Objects are carried in bins which can contain multiple objects. It is crucial that the Bins contains the correct of items for a smooth running and increased efficiency.

The efficient processing of a distribution center has a direct impact on the final price of the product delivered to the end user and an indirect impact on the cost of goods through reduced inventory.

This project focus on focuses on the use of Machine Learning to build a model that can count the number of objects in each bin. Which can be used to track inventory and make sure that delivery consignments have the correct number of items.

Problem Statement

The act of Tracking the records of sales and inventories in a Distribution Centre manually can be very cumbersome, prone to errors and inefficient most especially for companies operating more than one Distribution centers and stores selling thousands of products per month. It's critical to manage the process accurately, efficiently, and cost-effectively devoid of error. Else, one loses the risk of failing to meet customer demand which ultimately degenerates into increased costs and reduction in revenue.

This process can be largely augmented by the application of Computer vision and Machine learning to detect and classify various images present in a particular bin and computes its resulting count which ultimately can be useful in predicting stock level. Managing inventory strategically has never been more important, and it only gets more challenging as product development evolves.

Metrics

Being a classification exercise, the metrics to be used in this project is the Accuracy(precision)

$$\frac{1}{N}\sum_{i=1}^{N} 1[p_i == g_i]$$

II. ANALYSIS

Data Exploration

The lifeblood of any organization is Data. In this project, we made use of the Amazon Bin Image Dataset (Amazon Bin Image Dataset). The Datasets contains over 500,000 images in jpg format alongside a corresponding metadata of the items contained in each bin in json format.

The bin images datasets consist of multiple/random object categories and various instances and the metadata consists of the like the associated Image name, expected quantity of items, dimensions and units of each item object category identification (Amazon Standard Identification Number, ASIN), dimensions and units of each item, weights, expected quantity of items e.t.c.

In the below sample image, there are 2 object contained in the bin which corresponds to the information in the metadata file.

- Bin Image- image(600.jpg)

- Metadata - [json file (1000.json)]

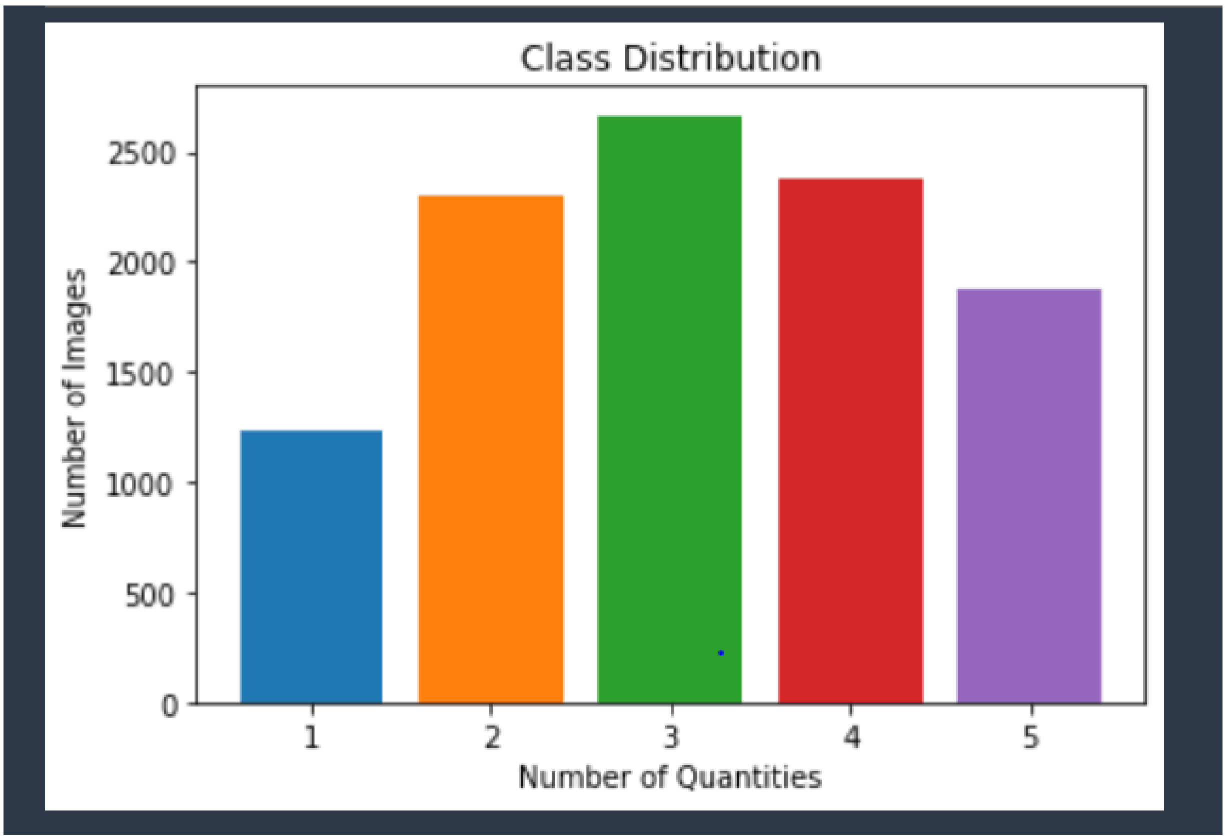

```
{
  "BIN_FCSKU_DATA": {
    "B000I4AQHA": {
      "asin": "B000I4AQHA",
      "height": {
        "unit": "IN",
        "value": 3.8
      },
      "length": {
        "unit": "IN",
        "value": 9.1
      },
      "name": "Pedigree Little Champions 12 Pouch Variety Pack Dog Food With 4 Beef, 4 Chicken, 4 Beef And Chicken, 3.97 lb Carton",
      "quantity": 1,
      "weight": {
        "unit": "pounds",
        "value": 4.3
      },
      "width": {
        "unit": "IN",
        "value": 6.6
      }
    },
    "B00TPSF6G6": {
      "asin": "B00TPSF6G6",
      "height": {
        "unit": "IN",
        "value": 4.4
      },
      "length": {
        "unit": "IN",
        "value": 12.5
      },
      "name": "EBP8386 Full Set of 4 Turn Signal Blinker Winker Indicators for Suzuki GS DR GSF",
      "quantity": 1,
      "weight": {
        "unit": "pounds",
        "value": 1.8999999999999997
      },
      "width": {
        "unit": "IN",
        "value": 5.0
      }
    }
  },
  "EXPECTED_QUANTITY": 2,
  "image_fname": "600.jpg"
}
```


The tapes in front of the bins is so as to prevent the items from falling out of the bins, and sometimes it might make the objects unclear.

Exploratory Visualization

In this Project, only a subset of the Data was used. The original Dataset contained over 500,000 images, however the subset used contained a total of 10,441 Images. The Data was chosen in such a way that the number of objects in a BIN ranges from 1 to 5.

Below is a bar chart of the distribution of the data used.



Algorithm and Technique

The algorithm used in the project is based on Convolution Neural Network architecture (CNNs). A major benefit to CNNs is the ability to make use of pretrained models and further fine-tuning it to implement it for other tasks (also known as Transfer Learning). Of all the pretrained models available like the Inception, VGG, Resnet e.t.c, the pre-trained model fine-tuned for use in this project is "**RESNET34**".

The following steps were performed in the fine-tuning of the model.

- The RESNET34 Pretrained model was chosen and created
- All the convolutional layers were freezed (No training)
- Two Fully Connected layer was added.

During training of the model, the Forward pass was performed throughout the whole network. However, for the backward pass, only the weight of the fully connected network was updated since the convolution layer was frozen.

The following hyperparameters are used during tuning and training;

• batch size (number of images to look at once during a single training step)

• learning rate (how fast to learn)

As this is a multi-class classification problem, the Cross-entropy loss function is used as the cost function, and Adaptive Moment Estimation (Adam) method is used as an optimizer.

Benchmark

The project was embarked on with an aim to attaining an accuracy score of about **55.67%**

III. METHODOLOGY

Data Preprocessing

In the implementation of this project, only a subset of the Amazon Bin images dataset was used which contained a total of 10,441 images which is available in an S3 bucket named (**aft-vbi-pds**) in the US-east-1 AWS region.

There is a unique numerical identifier for each images and the corresponding metadata pair which contains information like the associated Image name, expected quantity of items, dimensions and units of each item.

As part of the starter files, A Json file that containing the 10,441 subset of the metadata was made available and arranged in folders where the folder number corresponds to the number of objects in the Bin.

The subset of data created is arranged in subfolders. Each of these subfolders contains images where the number of objects is equal to the folder's name. For instance, all images in folder 3 have ideas with three things. The number of objects will serve as the categories for classifying bin images.

Since the unique identifier of the metadata corresponds to the images, for example the image named 1000.jpg will have its metadata named 1000.json. the images were then downloaded by replacing ". json" with ".jpg" into a folder and split into the training, testing and validation Data in ratio 8:1:1.

Some of the images had poor quality and some objects were hardly visible. For this reason, the Data was augmented, It was also resized to 224 x 224 and normalized

Implementation

The project was implemented using AWS Cloud services (majorly Sagemaker, S3 and cloudwatch logs).

A notebook instance with the ml.t2.medium' instance type which has 2vCPUs and 4Gb memory was set up. The necessary packages were installed such as Torch, TorchVision, Pytorch, Boto3. E.t.c.

The Downloaded Datasets was then uploaded to Amazon Simple storage service(S3). A training script was created to fine-tune the 'RESNET34' pretrained model that takes in the model training hyperparameters as arguments and reads, loads, and preprocesses the train, test and validation data.

During the Model training, A batch size of 64 and learning rate of 0.05 was chosen. In a bid to increase the performance of the model and hyperparameter tuning was done. As many parameters for the learning rate and batch size was chosen.

model debugging and profiling was also used to better monitor and debug the model training job. The model was thereafter deployed to an endpoint.

Refinement

In the initial training carried out, I specified the below hyperparameter values

```
hyperparameters = { "test-batch-size": "100",

                    "batch-size": 64,

                    "lr": 0.05}

Hyperparameters
```

This produced the below test los and accuracy

```
Test set: Average loss: 96.0
Testing Accuracy: 15.0
```

During the hyperparameter tuning, the below hyperparameter ranges were specified

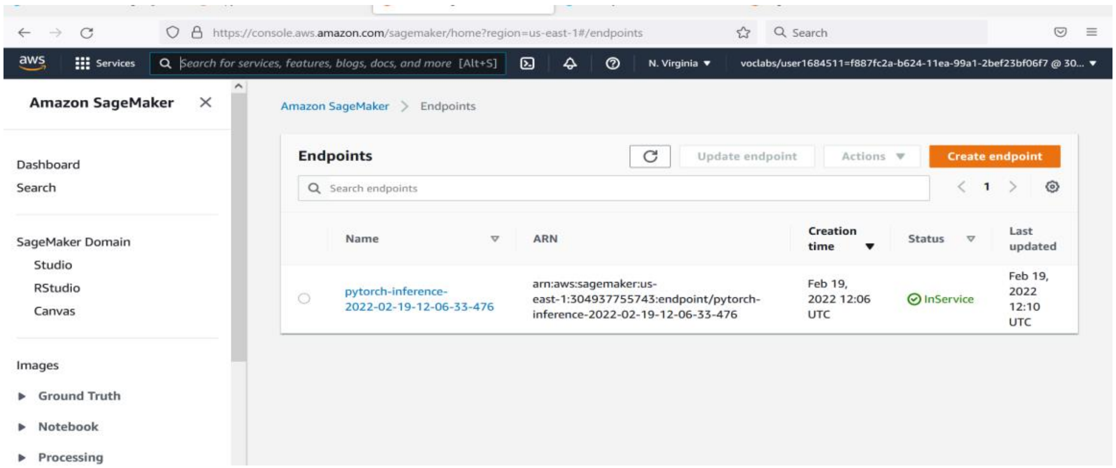
```
hyperparameter_ranges = {

    "lr": ContinuousParameter(0.001, 0.1),

    "batch-size": CategoricalParameter([32, 64, 128, 256])

}
```

However, I kept getting the error "Best training job not available for tuning job" eand when I checked the sagemaker console, I got the info "Best training job summary data is available when you have completed training jobs that are emiting metrics". even when my metric definition matched the test log.
This made me result into using the fixed hyperparameters values specified in the initial training Job for the model debugging and profiling.

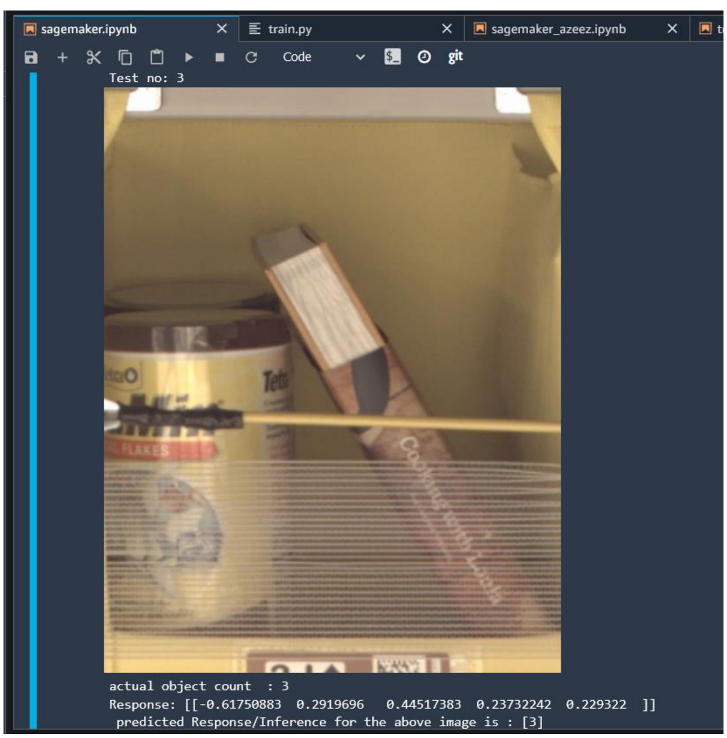
IV. RESULTS

Model Evaluation and Validation

The model was deployed with pytorch using the fixed hyperparameters values to an endpoint. The AWS lightweight runtime solution was used to invoke an endpoint. Below is a screenshot of the deployed endpoint



Below is a screenshot of the endpoint result



Justification

Only a sybset of the available Dataset was used (10,441 of 535,234 images). With this small subset of images dataset the model's accuracy reached the testing accuracy of 15.0. Although, it is below the benchmark accuracy of 55.67%, it is still a significant progress considering that not all the dataset was used. Perhaps training with a larger Dataset and more Data preprocessing can further improve the model past the benchmark, and is below the bench mark result mentioned earlier.

V. CONCLUSION

Amazon SageMaker is a fully managed service that provides a robust platform for developing machine learning models.

The goal of this project was to build a Machine learning model that can count the number of objects in each bin. Being a deep learning project, a pretrained model was fine-tuned to implement the project.

The project encompasses the full cycle of Machine learning; Data Preprocessing, Data Modelling, Model Training, Model testing, Model deployment and Inference.

A system like this can be used to track inventory and make sure that delivery consignments have the correct number of items.

VI. REFERENCES

1. Inventory Management: Definition & Processes [2021] (shipbob.com)
2. Amazon BIN Image dataset Amazon Bin Image Dataset - Registry of Open Data on AWS
3. Documentation https://github.com/awslabs/open-data-docs/tree/main/docs/aft-vbi-pds
4. Amazon BIN Image Dataset challenge silverbottlep/abid_challenge: Amazon Bin Image Dataset Challenge (github.com)