MACHINE LEARNING ENGINEER CAPSTONE

Inventory Monitoring at Distribution Centers

By Eric Moreno

Domain Background

A supply chain is a network between a business and the suppliers that make and distribute the products, supplies, or services the business needs to operate. There can be many players in a supply chain, including warehouses, trucking companies, retailers, distribution centers, and producers.

Regardless of a company's supply chain being large or small, inventory management is vital to a company's health because it balances supply with demand by ensuring that product is available at the right time by tracking product up and down the supply chain. Too much stock costs money and reduces cash flow and too little stock could lead to unfilled customer orders and lost sales.

Problem Statement

The largest operator of distribution centers is Amazon, which operates more than 175 fulfillment centers worldwide, with more than 150 million square feet of space. Amazon sells more than 12 million products and relying on manual processes to manage inventory for such a large operation would not be feasible and would be extremely costly. The advent of digital transformation has allowed many industries to benefit from artificial intelligence and this project aims to discover how artificial intelligence can be incorporated to improve inventory monitoring.

Datasets and Inputs

The dataset used for this project will be the Amazon Bin Image Dataset, which consists of 500,000 images of bins containing one or more objects from Amazon Fulfillment Centers. Each image is in JPEG format and contains corresponding JSON metadata files which describe the items in each bins.

Solution

This project will produce a Convolutional Neural Network (CNN) which is a deep learning algorithm that can take in an input image, assign importance (learnable weights and biases) to various aspects / objects in the image, and be able to differentiate one from the other. AWS SageMaker will be used to train the CNN model to classify the number of objects in each bin.

Benchmark Model

The Amazon Bin Image Dataset (ABID) Challenge was conducted in 2017 and serves as a benchmark. Below are the results on the validation split performed by the initial challenge.

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

Bin Quantity	Per class accuracy (%)	Per class RMSE
0	97.7	0.187
1	83.4	0.542
2	67.2	0.710
3	54.9	0.867
4	42.6	1.025
5	44.9	1.311

Evaluation Metrics

Counting the number of items in a bin will be evaluated by two metrics: accuracy (precision) and root mean squared error (RMSE).

$$\text{Accuracy: } \frac{1}{N}\sum_{i=1}^{N} 1[p_i == g_i] \qquad \text{RMSE: } \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - g_i)^2}$$

Project Design

A pre-trained ResNet50 will be used for this project. It is a 50-layer deep convolutional network which beats the limitation of the VGG-16 model by solving the issue of diminishing gradient. The ResNet architecture uses the concept of skip connections, allowing inputs to skip some convolutional layers and the result is a significant reduction in training time and improved accuracy.

Below is an outline of the project design:

Data Preparation

- Download data and store in S3 bucket
- Split training data into train, test, and validation sets

Exploratory Data Analysis

- Explore the distribution of classes of images
- Identify any class imbalance

Model Training

- A training script will be created for data pre-processing and hyperparameter tuning, and will load a pre-trained ResNet50 model
- Data augmentation will be performed on the training set
- Hyperparameters for consideration are learning rate and batch size
- A training estimator will be set up

Model Evaluation

- Model performance will be evaluated using accuracy and RMSE

Deployment

- The model will be deployed to an endpoint
- A lambda function will be created which will connect to the endpoint and make inferences using a test image

---

Sources

- https://www.bigrentz.com/blog/amazon-warehouses-locations
- https://www.netsuite.com/portal/resource/articles/inventory-management/inventory-management.shtml
- https://registry.opendata.aws/amazon-bin-imagery/
- https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
- https://datagen.tech/guides/computer-vision/vgg16/