Inventory Monitoring at Distribution Centers Report

AbdelKarim Eljandoubi

August 19, 2023

1 Domain Background

A Distribution Center serves as a spacious warehouse where products and goods are frequently kept in stock, ready for redistribution to wholesalers or retailers. These centers play a vital role in ensuring a smooth operational flow within a company's supply chain network, connecting the company, its suppliers, and ultimately the consumers.

Within the Distribution Center lies an expansive area containing numerous goods arranged in storage containers referred to as bins. These centers buzz with constant activity as goods are continuously moved in and out. Hence, it is of utmost importance that this movement transpires efficiently, promptly, and within specified time frames. The effective management of these processes directly influences the final pricing of products delivered to end users and indirectly affects overall inventory costs by minimizing stock holdings.

2 Problem Statement

The manual task of keeping track of sales and inventory records within a Distribution Centre can become exceedingly burdensome, susceptible to errors, and notably inefficient—particularly for companies operating multiple Distribution Centers and stores that handle a substantial volume of products on a monthly basis. Achieving accurate, efficient, and cost-effective management of this process is crucial. Otherwise, the potential consequence is the inability to meet customer demand, leading to escalated costs and diminished revenue.

This challenge can be significantly alleviated through the application of Computer Vision and Machine Learning. These technologies enable the identification and categorization of diverse images found within specific storage bins. This, in turn, facilitates the calculation of item quantities, which holds the potential to aid in forecasting stock levels. Strategic inventory management has assumed heightened significance, and its complexity continues to grow as product development advances.

3 Solution Statement

The project will be addressed by capitalizing on advanced techniques in deep learning, specifically Convolutional Neural Networks (CNNs), along with machine learning approaches. The dataset at hand comprises a substantial collection of approximately 535,000 images depicting bins containing varying quantities of objects. Each image is accompanied by a metadata file that furnishes crucial details such as the object count, dimensions, and object types.

The primary objective of this undertaking is to create a model capable of classifying the number of objects within each bin. To execute this classification task effectively, the chosen approach involves employing a pre-trained Convolutional Neural Network (CNN). This CNN model will be subjected to further training and fine-tuning procedures using Amazon SageMaker.

The central purpose of this project is to establish robust evidence that machine learning techniques can indeed be harnessed to devise a system capable of inventory tracking. This system holds the potential to ensure the accuracy of delivery consignments by verifying that the correct number of items is included.

4 Datasets and Inputs

Data serves as the lifeblood of any organization. To successfully undertake the projects outlined below, it is imperative to gather and utilize specific dataset components from sources like the Amazon Bin Image Dataset, AWS, and the internet.

Dataset Details:

1. This dataset encompasses approximately 535,000 images in jpg format, representing various bins. Each image is paired with corresponding metadata detailing the items present within it. The metadata is presented in json format. This dataset is accessible through an S3 bucket named ”aft-vbi-pds” within the US-east-1 AWS region.

2. Each image is assigned a unique numerical identifier, linked to its corresponding metadata. For instance, an image labeled as 1111.jpg will have a metadata file named 1111.json. This one-to-one correspondence facilitates data organization.

3. The quantity of items within each bin exhibits variability across different bins. For the project, only a subset of this data will be utilized. The selection process aims to include bins with item counts ranging from 1 to 5. Notably, the distribution of classes within this subset is not evenly balanced. A visual representation of the subset is provided below.

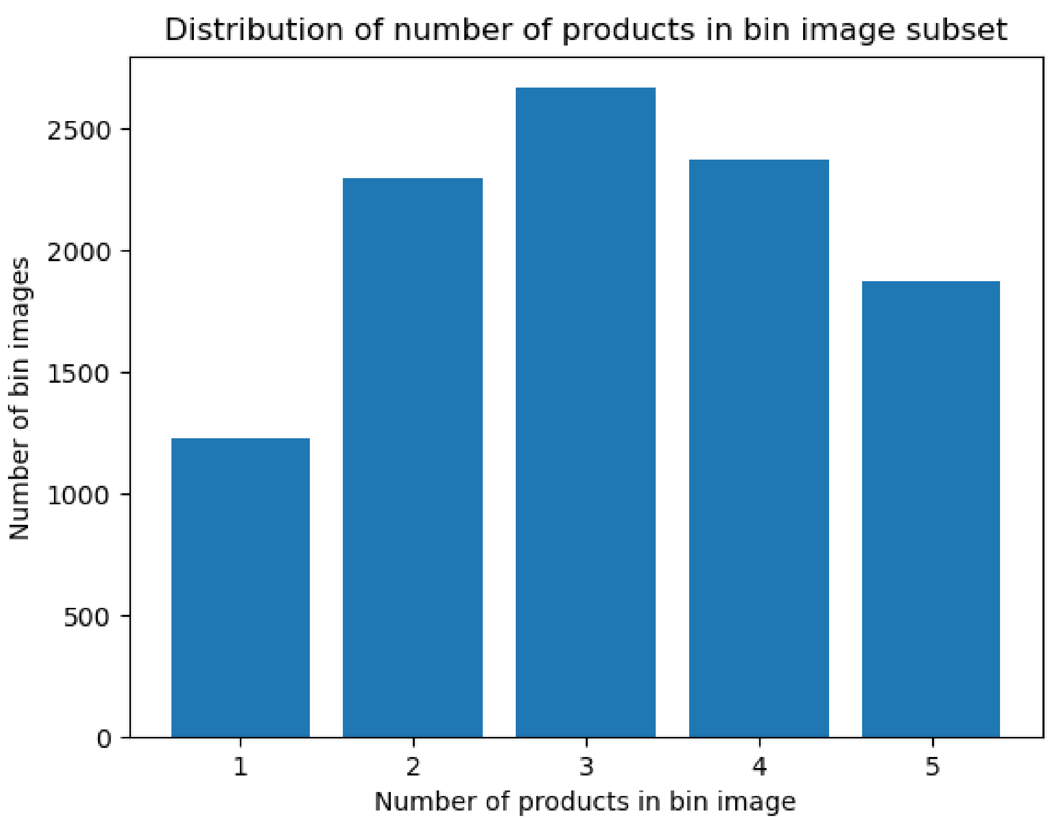

Figure 1: Distribution of number of products in bin image subset.

4. Within the metadata (json) files, essential details are included, such as the associated image’s name, the anticipated item quantity, along with dimensions and units of measurement for each individual item. These elements contribute to a comprehensive understanding of the dataset.

4.1 Data preprocessing

The dataset is partitioned into training (80%), validation (10%), and testing (10%) subsets. Preprocessing steps involve operations such as resizing, random flipping, conversion to tensor format, and normalization before forwarding the data to the model

5 Benchmark Model

During the hyperparameter tuning process, all variants of the ResNet architecture will be assessed. Subsequently, the variant that attains the highest testing accuracy score will be identified and chosen for implementation. This selected variant will then be utilized for the intended purpose.

6 Evaluation Metrics

Given that this constitutes a classification problem, the general accuracy of the classification holds value as a means to assess the effectiveness of the trained model's performance.

$$\text{accuracy}(\mathbf{g}, \mathbf{p}) = \frac{1}{n} \sum_{i=1}^{n} 1_{g_i = p_i} \tag{1}$$

Assigning $1_{x=y}$ as the indicator function, and denoting $\mathbf{p}$ as the prediction and $\mathbf{g}$ as the ground truth.

7 Hyperparameter Tuning

The hyperparameters chosen are :

1. The learning rate $\text{lr} \in [10^{-7}, 10^{-2}]$. The learning rate for the Adam optimizer is recognized as a hyperparameter that holds the potential to enhance the training process beyond relying solely on the predetermined default value.

2. The model name is selected from a list of ResNet variants. Selecting an appropriate pretrained CNN, proficient in extracting pertinent features, holds essential importance for accurately counting objects.

3. The number of layers (x) follows an exponential pattern of two 2^x where $x \in [|1, 6|]$. It signifies the depth of the network, where greater depth yields a more precise approximation; however, this also entails grappling with the challenge of gradient vanishing.

4. The number of neurons (y) follows an exponential pattern of two 2^y where $y \in [|8, 12|]$. It denotes the width of the network, where increasing width results in a more accurate approximation, but it also introduces the concern of overfitting.

8 Model Evaluation

The culmination of extensive hyperparameter tuning efforts led to the attainment of a peak test accuracy score, which settled at 33%. This result reflects the optimization of various parameters to enhance the model's performance, demonstrating a notable achievement within the scope of this experimentation. The selected benchmark [Par17] model showcases an approximate accuracy of 56%. Regrettably, the experimental model fell short of achieving comparable results to this benchmark. Moving forward, it's imperative to invest further effort into refining the data, exploring potential transformations, and reevaluating the suitability of the chosen models. These steps are essential to bridge the performance gap and advance the overall effectiveness of the approach.

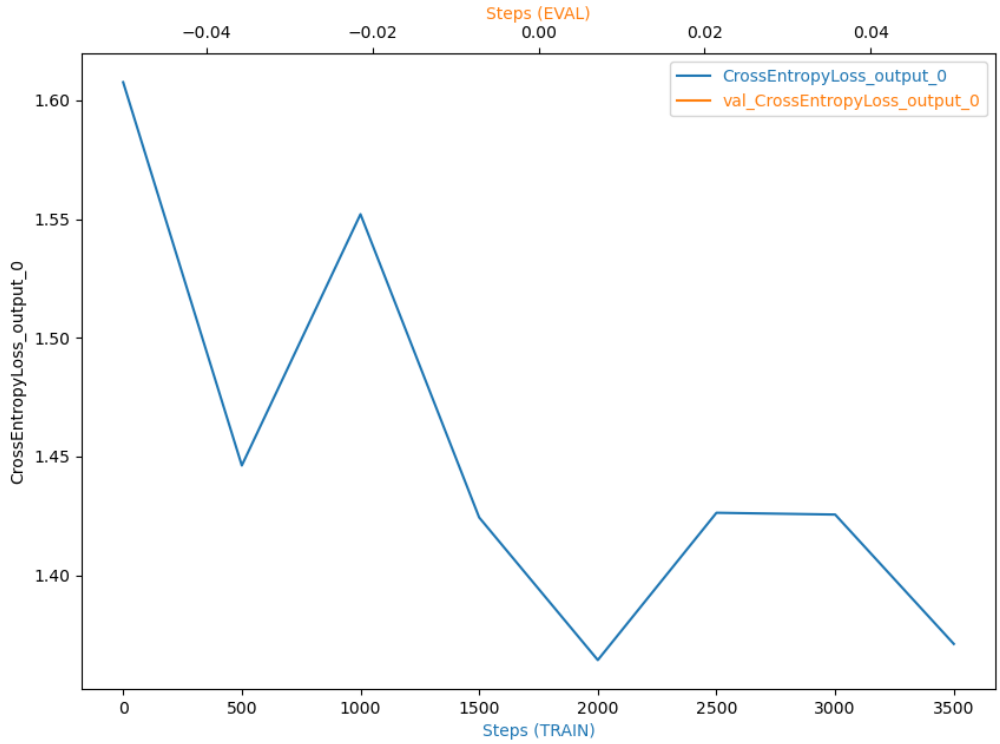

Figure 2: Cross Entropy Loss.

9 Improvements

To enhance performance, a potential approach could involve leveraging AWS Ground Truth for the labeling process, aiding in annotating our dataset intended for object detection tasks. Subsequently, we could refine pre-existing object detection models, encompassing both CNN-based and transformer-based architectures. To ensure accuracy, the application of the Non-Maximum Suppression algorithm could then be employed to determine the count of valid objects post filtration.

References

[Par17] Eunbyung Park. Amazon bin image dataset(abid) challenge. *Github*, 2017.