AWS Machine Learning Engineer Nanodegree

Capstone Proposal

Lorna Yen 11th Nov, 2024

Domain Background

Inventory monitoring and automation in distribution centers is an essential component of supply chain management. With the increasing use of robotics and automated systems, efficient and accurate inventory tracking has become a priority. Traditionally, inventory counting relies on manual labor, which is prone to errors and inefficiency. In this project, we explore the application of computer vision techniques to automate object counting in bins using image data, which can significantly streamline the operations in distribution centers, enhance accuracy, and reduce labor costs.

Problem Statement

The objective of this project is to build a machine learning model that can accurately count the number of objects in a bin based on image data. This task is challenging because bins may contain different quantities and configurations of objects, and object colors may be ambiguous, making it difficult to perform precise counts manually or through basic image processing techniques. Automating this process with a trained model can improve the efficiency of distribution centers by providing real-time and accurate object counts for better inventory control.

Datasets and Inputs

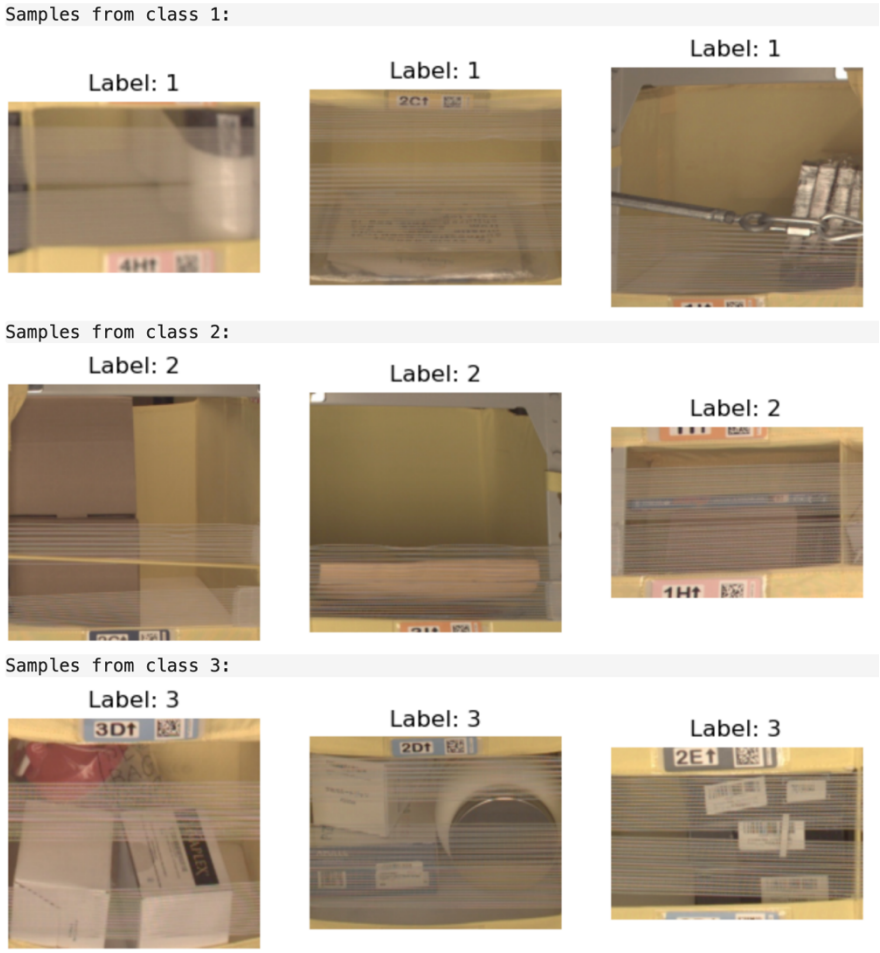
This project utilizes the **Amazon Bin Image Dataset**[1], which consists of approximately 500,000 images of bins, each containing a variable number of objects. To manage resource constraints, we will work with a subset of this dataset (10,441 images) provided by Udacity for educational purposes. Each image is annotated with metadata, including the count of objects in the bin, which serves as the label for the classification task. This subset contains images with object counts ranging from 1 to 5, providing a variety of configurations for training the model.

We utilize a data subset list, `file_list.json`, provided by Udacity, to download images from the Amazon Bin Image Dataset's S3 storage. This file includes labels and file URLs, structured as follows:

```
{ "1": ["data/metadata/100313.json", "data/metadata/09915.json",
 ...], "2": ["data/metadata/100894.json", ...], ... }
```

Each image is accompanied by metadata for object count and other attributes.

Sample Image Data:



Solution Statement

The proposed solution is an end-to-end machine learning pipeline built on AWS, leveraging Amazon SageMaker to train, tune, deploy, and monitor a convolutional neural network (CNN) model. A ResNet50-based architecture and the Amazon Bin Image Dataset will be used for the task of counting objects in bins, utilizing transfer learning to accelerate model convergence. Hyperparameter tuning will be conducted to optimize the model for accuracy and reduce the risk of overfitting. The final model will be deployed as an inference endpoint on SageMaker, providing a scalable solution for real-time object counting in bin images.

Benchmark Model

This project references a benchmark accuracy of 55.67%, which was achieved in the Amazon Bin Image Dataset Challenge using a ResNet34 model[2]. While the primary focus of this project is to build a scalable machine learning pipeline on AWS rather than enhancing model accuracy, this benchmark provides a useful reference point for

evaluating model performance. The benchmark model used accuracy and RMSE (Root Mean Square Error) as evaluation metrics, which can guide the validation metrics for this project's CNN-based approach.

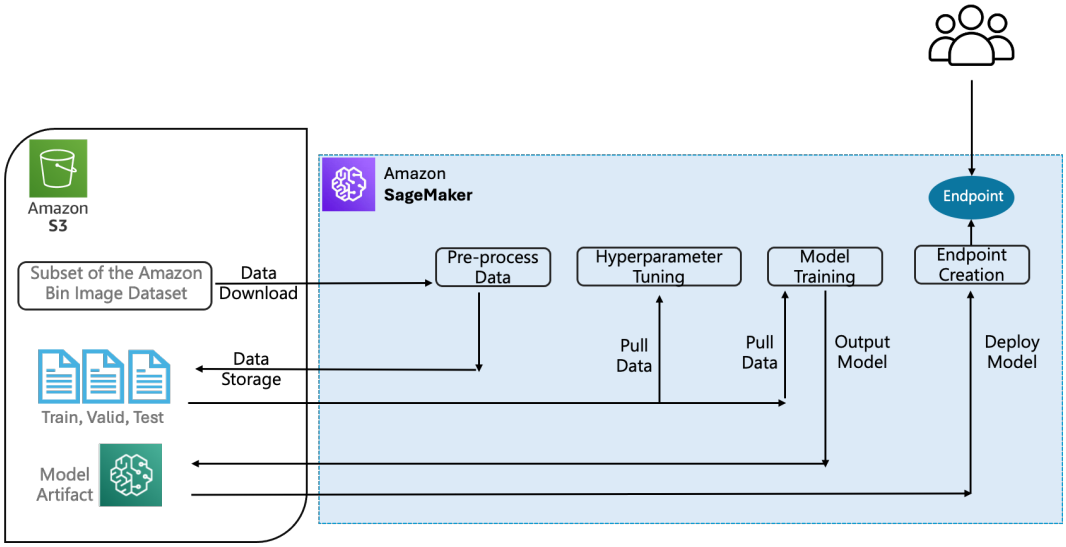
Evaluation Metrics

The primary evaluation metric for this project will be accuracy in predicting the correct object count in each bin. Additionally:

- **Cross-Entropy Loss** will be tracked to measure the error during training and validation.
- **Confusion Matrix** will provide insights into specific misclassifications. To optimize model convergence and detect potential issues, Amazon SageMaker's SMDebug will be used to monitor metrics such as training loss, validation loss, and resource utilization.

Project Design

Here is the high-level architecture of the project workflow:



- Data Ingestion and Preprocessing: The subset of the Amazon Bin Image Dataset will be ingested and split into training, validation, and testing sets. Images will be resized and normalized as part of the preprocessing pipeline.
- Model Selection and Training: A ResNet50 model is fine-tuned specifically for object counting. Training is conducted on SageMaker, with checkpointing to manage training progress and prevent overfitting.
- Hyperparameter Tuning: Key hyperparameters like learning rate, batch size, and number of epochs are optimized using SageMaker's Hyperparameter Tuner to enhance model performance.

- Debugging and Profiling: SMDebug will be integrated to monitor key metrics, track training and validation losses, and identify potential issues such as vanishing gradients, poor weight initialization, and underutilized resources.
- Model Deployment and Inference: The final model will be deployed to a SageMaker endpoint, allowing real-time inference on new bin images to predict object counts.

Reference

[1] Amazon, "https://registry.opendata.aws/amazon-bin-imagery/," 2024. [Online]. Available: https://registry.opendata.aws/amazon-bin-imagery/.

[2] silverbottlep, "Amazon Bin Image Dataset(ABID) Challenge," 20 Jully 2017. [Online]. Available: https://github.com/silverbottlep/abid_challenge.