Project Proposal

Domain Background

The modern world is increasingly relying on automation, especially in the logistics and distribution industry. Automation not only improves efficiency but also reduces the possibility of human errors. An essential aspect of this automation is maintaining the inventory at distribution centers. These centers often use robotic systems to manage bins filled with multiple items. As a consequence, there's a growing need for a system that can accurately count the number of objects within each bin to maintain inventory records accurately and ensure the correct number of items in each delivery consignment.

Problem Statement

The primary problem to solve is the development of an accurate model that can count the number of objects in each bin based on bin images. This model will be trained using the Amazon Bin Image Dataset, consisting of half a million images of bins with varying numbers of objects. The metadata files associated with each image provide details like object count, object type, and image dimensions.

Solution Statement

To tackle the problem, a machine learning model using a Convolutional Neural Network (CNN), specifically ResNet50, will be built. This model is ideal for the task at hand due to its efficiency in image classification tasks. The model will be trained using the AWS SageMaker service, which provides the infrastructure for training deep learning models at scale. After training, the model will be evaluated on a separate testing dataset and subsequently deployed to an endpoint for real-time predictions.

Datasets and Inputs

The Amazon Bin Image Dataset will be used for training the model. This dataset comprises 500,000 images of bins with varying numbers of objects, along with associated metadata files for each image detailing the object count, object type, and image dimensions. This dataset will be divided into training, validation, and testing sets. The training set will be used for training the model, the validation set for tuning the hyperparameters, and the testing set for evaluating the model's performance.

Benchmark Model

A simple Convolutional Neural Network (CNN) will serve as the benchmark model. The CNN model is known for its performance in image classification tasks and serves as a good starting point. Comparing the performance of the custom ResNet50 model to this simple benchmark will validate the effectiveness of a more complex network architecture.

Evaluation Metrics

The performance of the model will be evaluated using two metrics: accuracy and sparse categorical accuracy. These metrics were chosen because they offer different perspectives on the model's performance, ranging from a general overview (accuracy) to a more nuanced understanding (sparse categorical accuracy) of how well the model can predict the number of items in a bin.

Project Design

The project will follow the following steps:

1. **Data Preprocessing:** The images and their corresponding labels will be loaded, and the data will be split into training, validation, and testing sets.
2. **Model Building:** A deep learning model based on the ResNet50 architecture will be built. The model will be compiled with appropriate loss functions and metrics.
3. **Model Training:** The model will be trained using the training set, with the validation set used for hyperparameter tuning.
4. **Model Evaluation:** The performance of the model will be evaluated using the testing set and the chosen evaluation metrics.
5. **Model Deployment:** The trained model will be deployed to an AWS SageMaker endpoint for real-time predictions.
6. **Prediction Testing:** The endpoint will be tested using sample images to ensure it's working as expected.
7. **Performance Monitoring:** The performance of the endpoint will be monitored to ensure the model continues to function as expected over time.