AWS Machine Learning Engineer Nanodegree Capstone Project Proposal.

Inventory Monitoring at Distribution Centers.

09.04.2023

—

Henok Wubet.

Domain Background,

Computer vision research dates back 60 years, with computer scientists continuously exploring novel approaches to enable computers to extract meaningful information from input images.

Computer vision algorithms involve various techniques for acquiring, processing, analyzing, and comprehending digital images to extract data from the real world. This interdisciplinary field aims to enable computers to fully understand digital images, emulating human vision.

In recent years, there has been a surge in the development of deep learning techniques and technologies, especially since 2010. With high-performance computers and GPUs, deep learning models can be trained and improved over time, providing businesses with powerful tools to analyze images.*[5]*

The retail industry has seen rapid advancement in recent years as automation and machine learning are increasingly integrated into various operations. One such area is optimizing warehouse product storage and retrieval, essential for efficient order fulfillment. The Amazon Bin Image Dataset is a valuable resource for exploring machine-learning solutions to object counting and identification within bins.

The research on counting objects in images has significant practical implications across various domains. Convolutional Neural Networks (CNNs) have shown outstanding performance in detecting and counting objects. This project aims to leverage the progress made in this field and use it to count things inside Amazon bins. The ultimate aim is to improve product storage and retrieval accuracy and efficiency in warehouse settings.

Problem Statment

Distribution centers often use robots to move objects as part of their operations. Things are carried in bins that can contain multiple objects. Accurately tracking inventory and ensuring that delivery consignments have the correct number of items is crucial for the smooth operation of the distribution centers. To solve this problem, we propose building a Classifier ML model that counts the number of objects in each bin (items per bin is always between 1 and 5). This project will use the Amazon Bin Image Dataset and AWS SageMaker to preprocess data and train the model. The project will demonstrate end-to-end machine learning engineering skills acquired during the nano degree.

Datasets & Inputs

To accomplish this project, we will utilize the Amazon Bin Image Dataset, which comprises 500,000 images depicting bins that contain one or multiple objects. Each image in the

Dataset is accompanied by a metadata file containing information such as the number of objects, dimensions, and the object type. Our objective is to classify the number of objects present in each bin.*[1]*

- This Dataset consists of about 535,000 images in jpg format of Bins alongside corresponding metadata of the items contained in each Image in JSON format. This data is available in an S3 bucket (aft-vbi-pds) in the US-east-1 AWS region.
- Each image has a unique numerical identifier and the corresponding metadata pair. For example, the image named 1000.jpg will have its metadata named 1000.json
- The number of items in each bin varies across the bins.
- The metadata(JSON) file contains information like the associated Image name, expected quantity of items, dimensions, and units of each item.

Inputs

1. Items information (SKUs, Dimension, Number of Items) which, in the case of these projects, can be obtained from the metadata file
2. Images of the storage containers (BINS)
3. BIN image Information, e.g., Number of Objects

Solution Statment

The project's primary objective is to develop a machine-learning model that can accurately count the number of objects in each bin in distribution centers.*[3]*

The solution to this problem involves training a deep learning model using the ResNet-50 architecture to predict the object counts in an image from 1-5. Using transfer learning with pre-trained Torch models, the ResNet-50 model will be fine-tuned on the given dataset. We will train the model and tune the hyperparameters using Amazon Sagemaker. The model will be evaluated on the validation set using metrics such as accuracy and RMSE. The final model will be deployed and tested on a test set to evaluate its performance.

Benchmark Model

This project aims to evaluate the precision of the model's output and achieve a score of 55.67% or higher. *[2]*

The previous project, which utilized the resnet 34-layer architecture and was trained from scratch, achieved an accuracy score of 55.67%. This score will serve as the benchmark for the current project.

Evaluation Matrics

The evaluation metric for this problem is the Accuracy Score and Root mean square error (RMSE).*[2]*

Accuracy measures the proportion of correctly predicted object counts to the total number of object counts in the validation set. It is computed as

$$Accuracy = \frac{Number\ of\ Correct\ prediction}{Total\ number\ of\ prediction}$$

$$accuracy(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(\hat{y}_i = y_i)$$

RMSE measures the average difference between the predicted and accurate counts. It is computed as

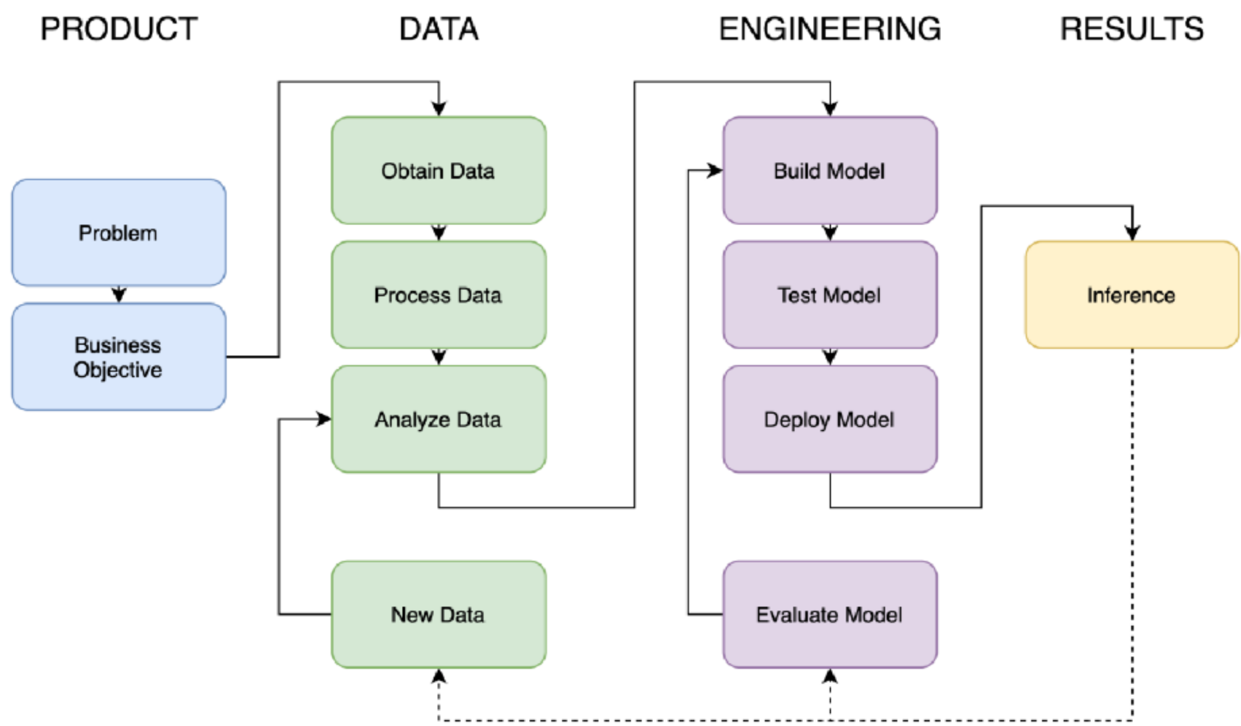
$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (p_i - g_i)^2}$$

Project Design

1. Data Preprocessing: The ABID dataset comprises millions of images categorized into different classes. This project will only use moderate data with classes from 1-5. This data will be extracted from the ABID metadata. Then the data will be preprocessed using AWS SageMaker, Which includes resizing, cleaning, and fixing wrong values to ensure the model's effectiveness
2. Data Splitting: Split the selected subset into 70% training, 15% validation, and 15% test datasets.
3. Hyperparameter Tuning: After preparing the data, we will utilize Amazon SageMaker's Hyperparameter Tuning service to tune the hyperparameters of the CNN and achieve the best possible performance. Hyperparameter tuning involves finding the best set of hyperparameters to achieve the highest model accuracy. SageMaker's Hyperparameter Tuning service runs multiple training jobs and tests different hyperparameter combinations. We will choose the best combination of hyperparameters based on the highest validation accuracy.
4. Model Training: After determining the best hyperparameters, we will train the CNN on the prepared dataset using the chosen hyperparameters. In this project, we will base the CNN architecture on a pre-trained model called ResNet50 that has undergone training on the ImageNet dataset. This approach will enable us to use the feature extraction capabilities of the pre-trained model and fine-tune the last few layers to enhance performance on our specific task.
5. Model Evaluation: We will use accuracy, and root mean squared error (RMSE) metrics to evaluate the trained model. Accuracy, a standard classification metric, measures the percentage of correctly classified images. Meanwhile, RMSE is a regression metric gauges the difference between the predicted and ground truth counts.

   To guarantee that the proposed model surpasses the benchmark model, we will compare the accuracy and RMSE scores of the two models.

6. Model Deployment: Once the model is trained & evaluated, it will be deployed to a SageMaker endpoint, and we can query an image for a prediction.
7. Multi-Instance Training: The same model will be trained, but the training workload will be distributed across multiple instances.
8. Report: A README file will be created, describing the project, explaining how to set up and run the code, and describing the results.



Conclusion

This document outlines a project to develop a machine-learning model that can accurately count the number of objects in each bin in distribution centers. The project will use the Amazon Bin Image Dataset, which consists of approximately 535,000 images of bins alongside corresponding metadata of the items contained in each image in JSON format. The ResNet-50 architecture will be used to predict the object counts in an image from 1-5, and the model will be fine-tuned on the given dataset using transfer learning with pre-trained Torch models. The project will be evaluated on the validation set using metrics such as accuracy and RMSE, with a benchmark accuracy score of 55.67% to be achieved. The project design includes data preprocessing, splitting, hyperparameter tuning, and model training. This project aims to improve product storage and retrieval accuracy and efficiency in warehouse settings.

References

[1] https://github.com/awslabs/open-data-docs/tree/main/docs/aft-vbi-pds

[2] https://github.com/silverbottlep/abid_challenge/tree/master

[3]https://towardsdatascience.com/comprehensive-guide-on-multiclass-classification-metrics-af94cfb83fbd

[4] https://www.fast.ai/

[5] Pablo Rodriguez Bertorello, Sravan Sripada, & Nutchapol Dendumrongsup. "Amazon Inventory Reconciliation Using AI." https://github.com/pablo-tech/AI-Inventory-Reconciliation/blob/master/ProjectReport.pdf