Amazon Bin Image Dataset(ABID) Challenge

Domain Background

The Huge companies like Amazon all over the world usually move objects from place to another and those companies got many objects to be moved so they move the objects using bins which have many objects inside and each object has a number so the company could keep track of the object and make sure that everything is going well .

The Amazon Bin Image Dataset contains 50,000 images and metadata from bins of a pod in an operating Amazon Fulfillment Center. The bin images in this dataset are captured as robot units carry pods as part of normal Amazon Fulfillment Center operations. This dataset can be used for research in a variety of areas like computer vision, counting genetic items and learning from weakly-tagged data.

The main goal of this project is to build a ML model that can count the number of objects in each bin in order to track inventory and check that bins have the appropriate number of items in order to reduce stock mismatches.

The project aims to build Machine learning model which is able to count the objects in each bin and make sure that each bin has the specific amount of objects inside it .

http://cs229.stanford.edu/proj2018/report/65.pdf

Problem Statement

The Amazon Bin Image Dataset contains images and metadata from bins of a pod in an operating Amazon Fulfillment Center. The bin images in this dataset are captured as robot units carry pods as part of normal Amazon Fulfillment Center operations

when we want the bins double-checked. The task is sometimes very challenging because of heavy occlusions and a large number of object categories.

We would like to open a new challenge in order to attract talented researchers in both academia and industry for these tasks. As a starting point, we provide baseline methods and pre-trained models for two tasks, counting and object verification tasks.

This is a simple task where you are supposed to count every object instance in the bin. We count individual instances separately, which means if there are two

same objects in the bin, you count them as two.

As we could see, the problem is an Image Classification problem as the image is provided to build the ML/DL model to identify the number of objects in each bin.

I would use AWS SageMaker for training and showing the result.

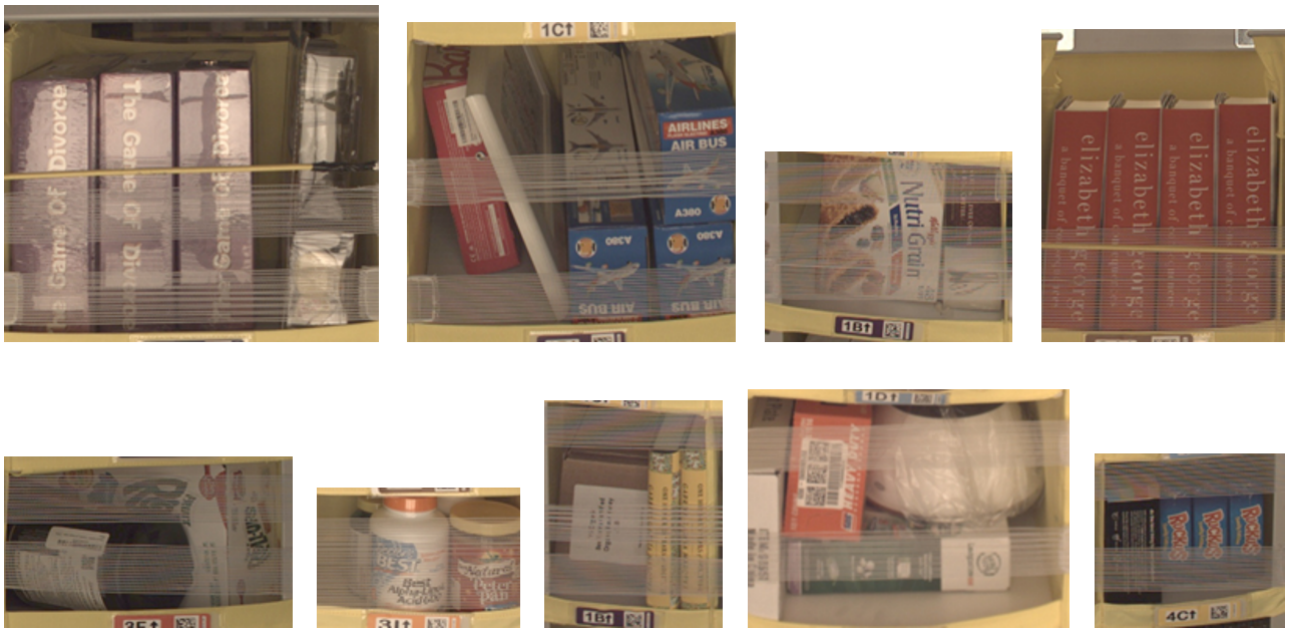
Datasets and Inputs

These are some typical images in the dataset. A bin contains multiple object categories and various number of instances. The corresponding metadata exists for each bin image and it includes the object category identification(Amazon Standard Identification Number, ASIN) which contains more than 500,000 image and metadata, quantity, size of objects, weights, and so on. The size of bins are various depending on the size of objects in it. The tapes in front of the bins are for preventing the items from falling out of the bins and sometimes it might make the objects unclear. Objects are sometimes heavily occluded by other objects or limited viewpoints of the images.

https://www.kaggle.com/datasets/dhruvildave/amazon-bin-image-dataset

The dataset contains 6 classes which identify each object in the picture .

Picture :



MetaData:

```
{
    "BIN_FCSKU_DATA": {
        "B00CFQWRPS": {
            "asin": "B00CFQWRPS",
            "height": {
                "unit": "IN",
                "value": 2.399999997552
            },
            "length": {
                "unit": "IN",
                "value": 8.199999991636
            },
            "name": "Fleet Saline Enema, 7.8 Ounce (Pack of 3)",
            "normalizedName": "(Pack of 3) Fleet Saline Enema, 7.8 Ounce",
            "quantity": 1,
            "weight": {
                "unit": "pounds",
                "value": 1.8999999999999997
            },
            "width": {
                "unit": "IN",
                "value": 7.199999992656
            }
        },
        "ZZXI0WUSIB": {
            "asin": "B00T0BUKW8",
            "height": {
                "unit": "IN",
                "value": 3.99999999592
            },
            "length": {
                "unit": "IN",
                "value": 7.899999991942001
            },
            "name": "Kirkland Signature Premium Chunk Chicken Breast Packed in Water,
12.5 Ounce, 6 Count",
            "normalizedName": "Kirkland Signature Premium Chunk Chicken Breast Packed
in Water, 12.5 Ounce, 6 Count",
            "quantity": 1,
            "weight": {
                "unit": "pounds",
                "value": 5.7
            },
            "width": {
                "unit": "IN",
                "value": 6.49999999337
            }
        },
        "ZZXVVS669V": {
            "asin": "B00C3WXJHY",
            "height": {
                "unit": "IN",
                "value": 4.330708657
            },
            "length": {
                "unit": "IN",
                "value": 11.1417322721
```

```
            },
            "name": "Play-Doh Sweet Shoppe Ice Cream Sundae Cart Playset",
            "normalizedName": "Play-Doh Sweet Shoppe Ice Cream Sundae Cart Playset",
            "quantity": 1,
            "weight": {
                "unit": "pounds",
                "value": 1.4109440759087915
            },
            "width": {
                "unit": "IN",
                "value": 9.448818888
            }
        }
    },
    "EXPECTED_QUANTITY": 3
}
```


Solution Statement

Solution would be to build a Deep Learning model which would help to count the objects in each picture by using pre pre-trained model like Resnet as we used before in the previous project .

Benchmark model

There are many who worked on same dataset and shows great result I would comparing results with them.

https://github.com/silverbottlep/abid_challenge
http://cs229.stanford.edu/proj2018/report/65.pdf

As we could see the results on the repository acc = 55.67 , RMSE = 0.93 :

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

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

I would use the evaluation at the end each epoch and see the progress of the model would use both Accuracy and RMSE for Evaluation Metrics .

Project Design

First : Download the Dataset from S3 .

Second: Put each class of the dataset in a folder which identify number of the objects in each image .

Third : Split the dataset to train and test .

Forth : Upload the splitting Dataset to S3 bucket

Fifth : Creat train_model.py file to train the model so it could call Resnet model

Sixth : Launch the training job and track the result for each epoch

Seventh : Tuning the Hyperparameter to get the best Hyperparameter which shows the best result .

Eighth : Deploying our model and use the end point to predict new images then delete the end point