05/05/2022

CAPSTONE PROPOSAL

Inventory Monitoring at Distribution Centers

AWS Machine Learning Engineer Nanodegree Program

UDACITY

Table of Contents

Introduction.

In this capstone project proposal, prior to completing the following Capstone Project, we will leverage what we have learned throughout the Nanodegree program to author a proposal for solving a problem of our choice by applying machine learning algorithms and techniques. A project proposal encompasses seven key points:

- The project's **domain background** — the field of research where the project is derived.
- A **problem statement** — a problem being investigated for which a solution will be defined.
- The **datasets and inputs** — data or inputs being used for the problem.
- A **solution statement** — the solution proposed for the problem given.
- A **benchmark model** — some simple or historical model or result to compare the defined solution to.
- A **set of evaluation metrics** — functional representations for how the solution can be measured.
- An outline of the **project design** — how the solution will be developed, and results obtained

1 Domain Background

Monitoring inventory means developing an understanding of what your best-selling products are, how often you turn over that inventory and what you need to have on hand to satisfy orders. Having the right products in stock allows you to fulfill orders in a timely manner and develop a loyal customer base.

Inventory monitoring and management is vital to a company's health because it helps make sure there is rarely too much or too little stock on hand or in a particular bin, limiting the risk of stockouts and inaccurate records.

There is an old business axiom that says, “**Nothing happens until somebody sells something**." With inventory monitoring, which could be changed to "**Nothing happens until somebody counts something**."

Inventory can be anything from boxes of ice cream cones in the storeroom at a sweet shop to a million-square-foot warehouse full of goods for a big box chain. But in either case, accurate inventory monitoring is a key to that company's success.

Simply put, inventory is the goods that a business owns that it plans to sell. If your company is an apparel retailer, products become inventory when you take possession of shirts, dresses, suits, and accessories from your suppliers. Those products leave the stock when they are sold to customers. Inventory can be stored on premises or at warehouses, distribution centers and other facilities.

2 Problem Statement

Keeping track of your Inventory is fundamental to your success in retail. At the most basic level your job is to supply the products to meet consumer demand. You cannot do that without effective inventory management.

Inventory systems suffer from several problems. Some of the more common issues include stock outs, excess inventory, misplaced inventory, and employee errors

In other hands, manual Inventory Management requires a huge workforce, and it is also error prone. So, there is a need for an effective inventory management.

The problem faced by the company is they do not have (or at least they do not have yet) any systematic system to record and keep their inventory data. It is difficult for the administrator to record the inventory data quickly and safely because they only keep it in the register and not meticulously organized. Therefore, it is difficult for the admin to estimate their profit.

With the new system developed, companies can manage their inventory data easily, quickly, and more secured. We could solve the above problem with machine automated tasks, Computer Vision Process aids us in solving the defined problem.

3 Evaluation or performances metrics

Our input is a dataset that consists of multiple bin images. A bin contains multiple object categories and various number of instances, each labeled with one of 5 different classes (number of objects categories, denote by **EXPECTED_QUANTITY** in each image's metadata).

Consequently, our goal is to predict the number of objects categories in each bin images considering i.e., to predict the **EXPECTED_QUANTITY.**

Label	No of Training Samples used
1	1228
2	2299
3	2666
4	2373
5	1875

Figure 1 - Dataset repartition

Given that the data is highly imbalanced, and we are facing a multi-class classification problem, **accuracy** will not be a useful metric. We can look at other metrics like **ROC-AUC**, **precision-recall curve, F1 score**, etc.

- AUC considers the true positivity rate and false positivity rate (thus at individual class level) and hence helps us assess the performance better than accuracy.
- F1 score is also a good metric since both precision and recall are useful for our problem and F1 score is a harmonic mean of precision and recall (Higher the better)

4 Solution Statement

Our solution to this problem is to create for example a responsive web application (also viewable on mobile) that would give real-time updates like real time update to hospital room supplies. Using video monitoring of the shelves, machine learning algorithms would recognize the types and quantities of available supplies.

For the sake of simplicity, we will (in this project) collecting data first.

Our solution consisted of an automated object detection machine learning algorithm recognizing supplies and their quantities plus a responsive front-end. The user (for example) or someone else can click on an image to view the live photo feed of the supplies, as well as a table of its supplies and quantities.

One of the areas where computer vision has made huge progress is image classification and object detection. A neural network trained on enough labeled data will be able to detect and highlight a wide range of objects with impressive accuracy.

To do this, thousands of images of objects needed to be generated in some sort of container.

5 Datasets and Inputs

In order to train a computer vision system for inventory monitoring, a method for creating a dataset is required. Those method are sometimes quite enough difficult to obtain.

Hopefully, Amazon provides us the **Amazon Bin Image Dataset** that we can use to train, test and validate our model.

5.1 Description of the Image Dataset

The **Amazon Bin Image Dataset** contains over 500,000 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. Bin images dataset provides the metadata for each image from where number of items can be derived.

The dataset is under the license **Creative Commons Attribution-NonCommercial-ShareAlike 3.0 United States (CC BY-NC-SA 3.0 US)** https://creativecommons.org/licenses/by-nc-sa/3.0/us/

5.2 Resources on AWS

Description: Over 500,000 bin JPEG images and corresponding JSON metadata files describing items in bins in Amazon Fulfillment Centers.

Resource type: S3 Bucket

Amazon Resource Name (ARN): arn:aws:s3:::aft-vbi-pds

AWS Region: us-east-1

6 Benchmark Model

The object counting using convolutional neural network has been introduced with connected component analysis from the following research paper is used as a benchmark to achieve the results. The authors had produced relatedly satisfactory results.

N. K. Verma, T. Sharma, S. D. Rajurkar and A. Salour, "Object identification for inventory management using convolutional neural network," 2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), 2016, pp. 1-6, doi: 10.1109/AIPR.2016.8010578.

The neural network chosen for this experiment uses the **ResNet** architecture. **ResNet** was chosen because it stacked some additional layers in the Deep Neural Networks which results in improved accuracy and performance. The intuition behind adding more layers is that these layers progressively learn more complex features.

ResNet are also chosen in order to prevent/solve the problem of the vanishing/exploding gradient, this architecture introduced the concept called Residual Network.

7 Set of Evaluation Metrics

4 Major Problems Faced during Inventory Control

- The Classification Problem.
- The Order Quantity Problem.
- The Order Point Problem
- Safety Stock

For the sake of simplicity, we will only focus on **the Classification Problem and the A/B analysis.**

This widely used classification technique recognizes different items of inventory for the purpose of inventory monitoring which assumes that **a firm should not exercise equal attention on all items of inventory** since a firm has to maintain several types of inventories. It, therefore, should pay maximum attention to those items that are

- most costly, and /or
- slow moving.

On the contrary, inventories which are less expensive should be given less control effort. Thus, the firm should be selective in its approach towards the inventory control management. This analytical approach called ABC Analysis Approach, tends to measure the relative cost significance of each component of inventories.

The overall accuracy of the classification and F1 score will be used to evaluate the performance of the trained model.

8 Workflow or Project Design