Project Proposal:

Domain Background:

Deep learning has revolutionised the operations of many industries. Advancements in computer vision, natural language processing and speech recognition have been beneficial to many companies that are using this technology to help or change their ways of working.

In computer vision the process of identifying and localising objects in an image is called object detection. One application is in distribution centres where the use of robotics to carry out tasks is prevalent. For example , the research carried out by Devalla et al demonstrates a system where robotics can pick and place desired coloured objects from one location to another improving the efficiency in packaging and logistics industries [1]. The ability of these robots to accurately select objects and choose the correct number of items for delivery enables fast operations and helps track inventory. Further research by Bahaghighat et al explain their methodology of counting blister cards within a drug package while moving in a conveyor belt. They use a layered approach which include object detection , feature extraction and classification to decide if among four defined classes Null, One, Two, and Three - these are the number of blisters cards may be found in a drug package [2] .

Problem statement:
Using deep learning methodologies, an object detections algorithm will be built to classify the number of objects in bins.

The datasets and inputs:
To complete this project we will be using the Amazon Bin Image Dataset. The dataset contains 500,000 images of bins containing one or more objects. We have been provided with a subset of this data. Within this there are 5 classes referring to the number of objects in each image. For each image there is a metadata file containing information about the image like the number of objects, it's dimension and the type of object. The subset of the dataset provided has imbalanced classes, the number of objects being:

| Class | Number of images |
|-------|------------------|
| 1 | 1228 |
| 2 | 2299 |
| 3 | 2666 |
| 4 | 2373 |
| 5 | 1875 |

Therefore we can observe that the classes are very imbalanced, class 1 have the lowest sample size.  This will impact the evaluation methodology.

Solution statement :
A convolutional network will be employed , we will use a pretrained network then apply transfer learning to create a bespoke solution to this problem.

Benchmark model :
There are many network architectures that have been developed for the purpose of object detection including VGGNet, ResNet, Inception, MobileNet and many more. .The pretrained model we would like to initially pursue is ResNet-18 . The benchmark will be the results from the selected pretrained model without applying transfer learning. Any performance increases will be attributed to transfer learning

Evaluation metrics :
Due to the nature of the imbalanced classes we will use an confusion matrix to determine the performance of the model using a confusion matrix. This will look at the precision, recall and F1 score of each class.

Project design:
1. Download and process the data and upload to S3 into train and test folders. The processing will include the following:
    a. We may need to resize the images
    b. Scale and apply normalisation to images
Apply Data augmentation such as horizontal flip
We will also split the data into training and test 80:20 ratio.
Use a pretrained network to observe benchmark evaluation metrics. We will initially select ResNet-18.
Apply transfer learning and train the model.
 Evaluate the updated model .

Reference:

[1] Devalla, Vindhya, et al. "Design and development of object recognition and sorting robot for material handling in packaging and logistic Industries." International Journal of Science and Advanced Technology 2.9 (2012).
[2] M. Bahaghighat, L. Akbari and Q. Xin, "A Machine Learning-Based Approach for Counting Blister Cards Within Drug Packages," in IEEE Access, vol. 7, pp. 83785-83796, 2019, doi: 10.1109/ACCESS.2019.2924445.