OSA Project

Presented By:

Joe BAROUK

4th year CCE Student and AI Engineer Intern at CME

Enterprise :	**CME Offshore, Beirut, Lebanon**
Supervised by :	**Dalia Jaber, AI Engineer**
	Kamal Youssef, Senior Solution Architect
Duration :	**From 05/06/2023 To 28/07/2023**

Content

Datasets

- Roboflow Dataset
- Spinneys Dataset
- Combined Dataset (Roboflow Dataset + Spinneys Dataset)
- Test0 Dataset (I gathered and labeled images found from datasets on Roboflow that could be used for further testing)

List of figures

List of tables

1 Introduction

In the fast-paced world of retail, maintaining well-stocked shelves is a critical aspect of ensuring customer satisfaction and maximizing sales. Supermarkets often face the challenge of keeping track of product availability on their shelves, leading to occasional stockouts and missed sales opportunities. To address this issue, an innovative solution is presented - an Automated Shelf Monitoring System powered by cutting-edge AI technology.

This project proposes the deployment of a camera-based monitoring system strategically positioned within the supermarket. With state-of-the-art AI algorithms, this system will continuously analyze the shelves, detecting empty spaces where products have run out or are insufficiently stocked. The AI model, trained on a diverse dataset, can accurately detect empty sections on a shelf, providing precise and timely feedback to supermarket management.

The primary goal of this system is to streamline the restocking process, reducing response times to stockouts, and enhancing overall operational efficiency. By leveraging real-time data and instant notifications, supermarket staff can proactively address shelf gaps, ensuring that customers find the products they desire promptly. This improved inventory management not only boosts sales but also enhances customer experiences, fostering loyalty and positive word-of-mouth.

2 Project Requirements

- Scan a Shelf using a Camera
- Detect Empty Spaces using AI (Object Detection)
- Install the model on a Raspberry Pi

3 Project Constraints

- Small Size AI model
- Edge AI model for Image Processing
- Recognition should even happen:
    - Under poor and strong lighting
    - In presence of many shelves

4 Solution Design

- Resources Used:
    - YOLOv8
    - Roboflow
- Resources to be used:
    - Camera
    - Raspberry Pi or any alternative

5 Results

- ***Solution 1*: Labeling Empty and Semi-Empty Spaces (Roboflow Dataset + some (~40) images from Spinneys Dataset)**

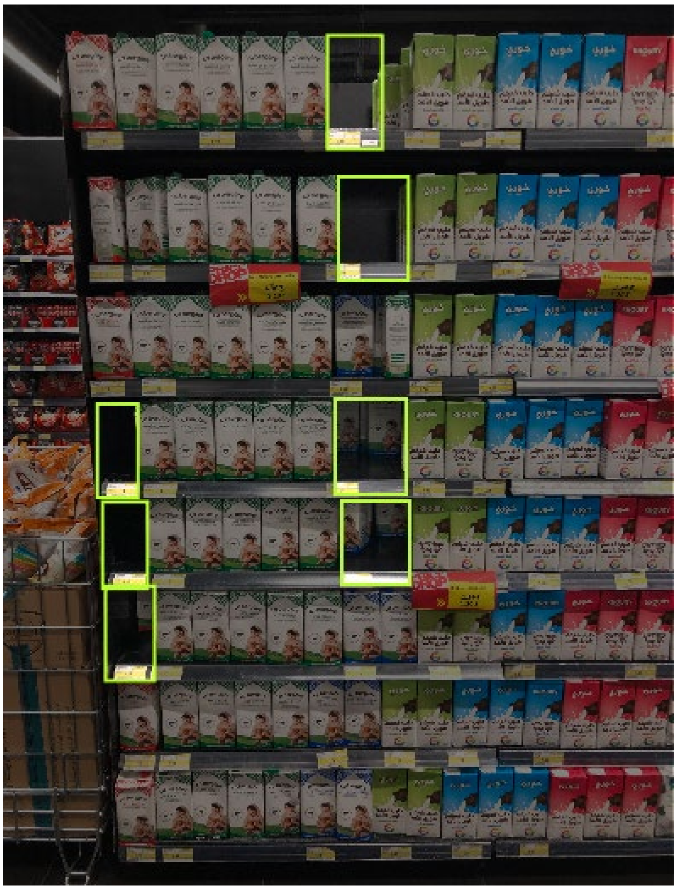

Figure 1: Example of empty and semi-empty labeling

```
Solution 1:
-----------
train: no augmentation - nano
train1: with augmentation: brightness:30%, shear, rotation - small
train2: with augmentation(x3): brightness:30%, shear, rotation - small
train3: with augmentation(x3): brightness:30%, shear, rotation - nano
train4: with: optimizer:AdamW, dropout:0.5 - with augmentation(x3): brightness:30%, shear, rotation - nano
train5: train4 + some more epochs
train6: dropout:0.5 - with augmentation(x3): brightness:30%, shear, rotation - medium
train7: dropout:0.5 - with augmentation(x3): brightness:15%, shear, rotation - small
train8: optimizer: AdamW, dropout:0.5 - with augmentation(x3): brightness:15%, shear, rotation - nano

train9: dropout:0.5 - with augmentation(x2): brightness:15% - nano
train10: with augmentation(x2): brightness:15% - nano

train11: dropout:0.5 - with augmentation(x2): brightness:18% - nano
train12: dropout:0.5 - with augmentation(x3): brightness:18% - extra large
train13: dropout:0.8 - with augmentation(x2): brightness:18% - nano

train15: dropout:0.5, weight_decat:0.5, cos_lr:True, deterministic:False, dropout:0.9, with augmentation(x2): brightness:18% - nano
train16: epoch: 200, dropout:0.5, weight_decay:0.5, cos_lr:True, deterministic:False, dropout:0.9, with augmentation(x2): brightness:18% - nano
train17: epoch: 200, dropout:0.5, weight_decay:0.001, cos_lr:True, deterministic:False, dropout:0.9, with augmentation(x2): brightness:18% - nano
```

Figure 2: Models tried for Solution 1

VALID	P	R	mAP50	mAP50-95
train5	0.917	0.857	0.918	0.656
train4	0.923	0.839	0.908	0.627
train3	0.907	0.857	0.911	0.651
train2	0.936	0.873	0.94	0.713
train6	0.93	0.863	0.933	0.717
train7	0.951	0.905	0.957	0.742
train8	0.908	0.833	0.901	0.609
train9	0.942	0.881	0.935	0.698
train10	0.942	0.881	0.935	0.698
train11	0.943	0.862	0.929	0.681
train12	0.952	0.89	0.957	0.769
train13	0.916	0.883	0.937	0.682
train14	0.608	0.486	0.467	0.124
train15	0.881	0.854	0.91	0.618
train16	0.907	0.862	0.926	0.669
train17	0.933	0.859	0.929	0.678

Table 1: Results on Validation Set (Solution 1)

TEST	P	R	mAP50	mAP50-95
train5	0.947	0.867	0.942	0.677
train4	0.949	0.851	0.934	0.645
train3	0.946	0.878	0.932	0.667
train2	0.947	0.921	0.96	0.734
train6	0.946	0.909	0.946	0.728
train7	0.974	0.929	0.956	0.732
train8	0.938	0.833	0.924	0.612
train9	0.933	0.866	0.949	0.722
train10	0.933	0.866	0.949	0.722
train11	0.926	0.833	0.89	0.652
train12	0.95	0.816	0.91	0.736
train13	0.946	0.822	0.898	0.659
train14	0.681	0.48	0.481	0.124
train15	0.91	0.779	0.876	0.596
train16	0.907	0.801	0.88	0.644
train17	0.909	0.819	0.886	0.651

Table 2: Results on Test Set (Solution 1)

SpinneysTest	P	R	mAP50	mAP50-95	F1-score
train	0.546	0.343	0.366	0.159	0.42
train5	0.651	0.363	0.386	0.142	0.47
train6	0.624	0.444	0.448	0.181	0.52
train7	0.679	0.393	0.439	0.192	0.5
train8	0.543	0.423	0.409	0.163	0.48
train9	0.643	0.484	0.522	0.217	0.55
train10	0.643	0.484	0.522	0.217	0.55
train11	0.653	0.491	0.557	0.245	0.56
train12	0.639	0.581	0.638	0.28	0.61
train13	0.691	0.564	0.606	0.262	0.62
train14	0.699	0.602	0.652	0.266	0.65
train15	0.74	0.543	0.638	0.267	0.63
train16	0.714	0.551	0.583	0.244	0.62
train17	0.636	0.489	0.516	0.211	0.55

Table 3: Results on Spinneys Dataset (Solution 1)

Analysis: The performance of the models on the validation and test sets appears to be satisfactory. However, when evaluating the models on the Spinneys Dataset, which consists of images taken specifically from the Spinneys supermarket, the results are significantly worse. This suggests that the models are suffering from overfitting and weren't trained on enough images, as they were primarily trained on images from the Roboflow dataset with only a limited number of Spinneys images. Consequently, the models' ability to generalize to new, unseen data is compromised.

A notable issue observed during visual inspection is that the models are significantly better at detecting empty spaces within the supermarket, but they struggle to accurately identify semi-empty spaces. This limitation led to the consideration of a potential and easier solution, which we'll refer to as "Solution 2".

- ***<u>Solution 2</u>: Labeling Empty Spaces only (Roboflow Dataset + some (~40) images from Spinneys Dataset)***

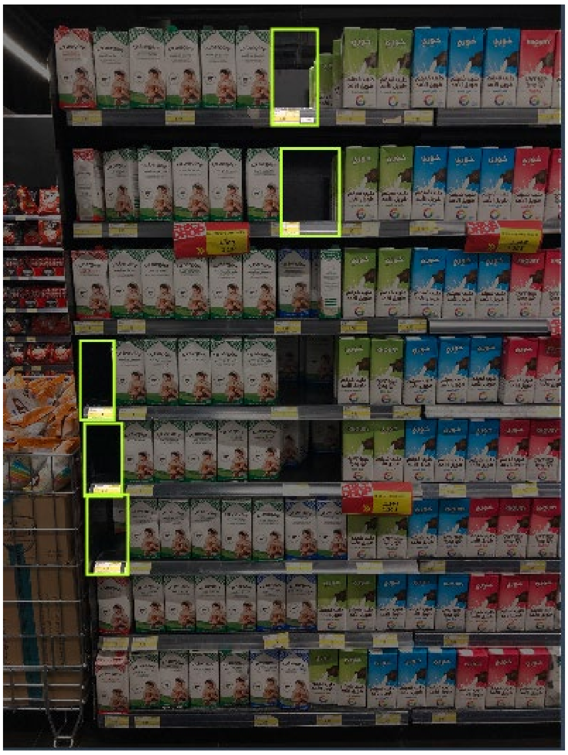

Figure 3: Example of empty only labeling

```
Solution 2:
-----------
train: v1, no augmentation, dropout=0.7, optimizer=auto, epoch=200 - nano
train1: v1, no augmentation, dropout=0.7, optimizer=auto, epoch=200 - small
train2: v2, no augmentation, dropout=0.7, optimizer=AdamW, epoch=200 - nano
train3: v2, no augmentation, dropout=0.7, optimizer=Adamax, epoch=200 - nano
train4: v2, no augmentation, dropout=0.7, optimizer=RAdam, epoch=200 - nano
train5: v3, with augmentation(x2):brightness:20%, dropout=0.7, optimizer=auto, epoch=100 - nano
```

Figure 4: Models tried for Solution 2

VALID	P	R	mAP50	mAP50-95
train	0.932	0.885	0.931	0.692
train1	0.924	0.9	0.947	0.735
train2	0.873	0.855	0.911	0.637
train3	0.915	0.865	0.927	0.677
train4	0.898	0.871	0.928	0.698
train5	0.918	0.861	0.931	0.69

Table 4: Results on Validation Set (Solution 2)

TEST	P	R	mAP50	mAP50-95
train	0.921	0.879	0.93	0.688
train1	0.937	0.931	0.96	0.749
train2	0.96	0.845	0.915	0.642
train3	0.898	0.883	0.936	0.694
train4	0.928	0.879	0.939	0.707
train5	0.875	0.885	0.923	0.688

Table 5: Results on Test Set (Solution 2)

SpinneysTest	P	R	mAP50	mAP50-95	F1-score
train	0.684	0.584	0.617	0.303	0.63
train1	0.754	0.564	0.616	0.309	0.65
train2	0.732	0.551	0.609	0.306	0.63
train3	0.78	0.614	0.671	0.386	0.69
train4	0.747	0.589	0.668	0.368	0.66
train5	0.86	0.548	0.687	0.394	0.67

Table 6: Results on Spinneys Dataset (Solution 2)

Analysis: At this stage, adjustments were made to the dataset, by refining the labels to focus solely on empty spaces. The performance on the validation and test sets is also promising in this solution. However, despite some improvement on the Spinneys Dataset, it falls short of the desired level. Upon careful analysis, it was evident that the models' overfitting persisted and the dataset was still lacking images, hindering their ability to generalize effectively.

As a result, the training data was expanded by incorporating the full Spinneys Dataset with the Roboflow Dataset. This will be "Solution 3".

By combining these two datasets, greater diversity and complexity were introduced to the training samples. This diverse training data would expose the models to a wider range of scenarios, encouraging better generalization and improved performance on unseen data.

- ***<u>Solution 3:</u> Combining the full Spinneys Dataset with the Roboflow Dataset while still considering empty spaces only***

```
Solution 3:
-----------
train: v2, no augmentation, dropout=0.7, optimizer=auto, epoch=200 - nano
train1: v2, no augmentation, optimizer=auto, epoch=200 - nano
train2: v2, no augmentation, dropout=0.7, optimizer=AdamW, epoch=200 - nano
train3: v2, no augmentation, dropout=0.9, optimizer=auto, epoch=200 - nano
train4: v2, no augmentation, dropout=0.7, cos_lr=True optimizer=auto, epoch=200 - nano
train5: v3, augmentation(x2): brightness:20% - bounding box rotation:10%, optimizer=auto, epoch=115 - nano
train6: v4, augmentation(x3): flip:horizontal - crop:15-25%, optimizer=auto, epoch=115 - nano
train7: v4, augmentation(x3): flip:horizontal - crop:15-25%, optimizer=auto, epoch=115 - small
```

Figure 5: Models tried for Solution 3

VALID	P	R	mAP50	mAP50-95	F1-score
train	0.933	0.864	0.936	0.678	0.9
train1	0.933	0.864	0.936	0.678	0.9
train2	0.894	0.866	0.92	0.64	0.88
train3	0.933	0.864	0.936	0.678	0.9
train4	0.918	0.863	0.93	0.67	0.89
train5	0.921	0.88	0.935	0.674	0.9
train6	0.907	0.905	0.944	0.696	0.91
train7	0.929	0.894	0.953	0.723	0.91

Table 7: Results on Validation Set (Solution 3)

TEST	P	R	mAP50	mAP50-95	F1-score
train	0.923	0.844	0.908	0.672	0.88
train1	0.923	0.844	0.908	0.672	0.88
train2	0.878	0.802	0.88	0.617	0.84
train3	0.923	0.844	0.908	0.672	0.88
train4	0.913	0.818	0.907	0.648	0.86
train5	0.941	0.834	0.936	0.686	0.88
train6	0.924	0.853	0.925	0.682	0.89
train7	0.91	0.906	0.939	0.721	0.91

Table 8: Results on Test Set (Solution 3)

Analysis: In this approach, the full Spinneys Dataset was integrated directly into the overall Roboflow dataset, eliminating the need for a separate table. The performance on the validation and test sets demonstrated positive outcomes, indicating that the models are functioning well. However, there is still room for further improvement.

Upon visual comparison with the two previous solutions, this current approach showcased superior results. It effectively detected nearly all empty spaces, though occasionally with low confidence scores. This suggests that while the models are generally successful in identifying empty spaces, there are instances where it may be uncertain about its predictions.

Various attempts were made to enhance performance, including data augmentation. Unfortunately, the impact of augmentation was limited, and increasing the model's size only marginally improved detection scores. The larger models exhibited slightly higher confidence scores, but it did not increase the number of empty spaces detected compared to the smaller models.

Overall, the current solution presents the best performance thus far, surpassing the previous attempts. While it successfully detects a substantial number of empty spaces, addressing the occasional low-confidence detections remains a priority for further refinement.

- Testing Solution 3 on a new dataset (Test0 Dataset)

Test0	P	R	mAP50	mAP50-95	F1-Score
train	0.837	0.8	0.863	0.453	0.818
train1	0.837	0.8	0.863	0.453	0.818
train2	0.841	0.8	0.847	0.438	0.820
train3	0.837	0.8	0.847	0.453	0.820
train4	0.775	0.786	0.825	0.428	0.780
train5	0.876	0.717	0.843	0.447	0.789
train6	0.881	0.704	0.831	0.443	0.783
train7	0.83	0.774	0.863	0.484	0.801

Table 9: Results on Test0 Dataset (Solution 3)

Analysis: Upon analyzing the metrics, it may appear that the models' performance is not optimal based solely on numerical indicators. However, when visually inspecting the results, they can be considered quite promising. The models exhibit the ability to detect almost all the empty spaces, although not always with a high confidence score.

Considering these findings, these images can be included into the Combined Dataset for further training of the model. Unfortunately, due to time constraints, this task was not undertaken.

The visuals can be accessed in the designated folder named "Train Solution3 Predict Test0" on the GitHub repository.

6 Testing

In this section, a series of visuals showcasing the outcomes of the "Train6" model from "Solution 3" on the Spinneys Test Set is presented. The full visuals can be accessed in the designated folder named "Train6 Solution3 Predict" on the GitHub repository.

Figure 6: Prediction 1

Figure 7: Prediction 2

Figure 8: Prediction 3

Figure 9: Prediction 4

7 Limitations

The models encounter (rarely) some limitations:

- Not detecting all empty spaces

The model didn't detect the space delimited by the green box.

- Detecting spaces that do not correspond to a shelf

The model outputted the bounding boxes (confidence 0.64 and 0.41) as empty spaces.

- Detecting small empty spaces that do not correspond to a place where a product will or can be placed

The model outputted the bounding box (confidence 0.87) as an empty space but there isn't enough space to place a Dettol bottle.

- Detecting with low confidence

8 Future Work

The next phase of the project involves evaluating the performance of Solution 3 on a fresh and new dataset taken from a supermarket in Lebanon other than Spinneys to assess even more its generalization capabilities with new examples. This step is critical to ensure that the model can adapt well to unseen scenarios and maintain consistent accuracy. Once a robust and reliable model is established, the implementation of the system in supermarkets can be proceeded with. To achieve this, strategically positioned cameras will be installed to capture the store's layout effectively. The pre-trained model will be deployed on a Raspberry Pi or a similar device, allowing real-time processing of the camera feed.

A crucial aspect of this implementation is conducting multiple tests to validate the model's performance under various supermarket conditions. This iterative testing process will help identify any potential challenges and fine-tune the system to ensure optimal performance.

Upon successful completion of these tests and ensuring the model's reliability, the camera and model can be seamlessly integrated into the existing supermarket system. This integration will enable efficient monitoring and analysis of empty spaces, facilitating improved store management and customer experience.

As an addition to the system later on, the model can be trained to detect the misplacement of a product and which products belong to the observed shelf.

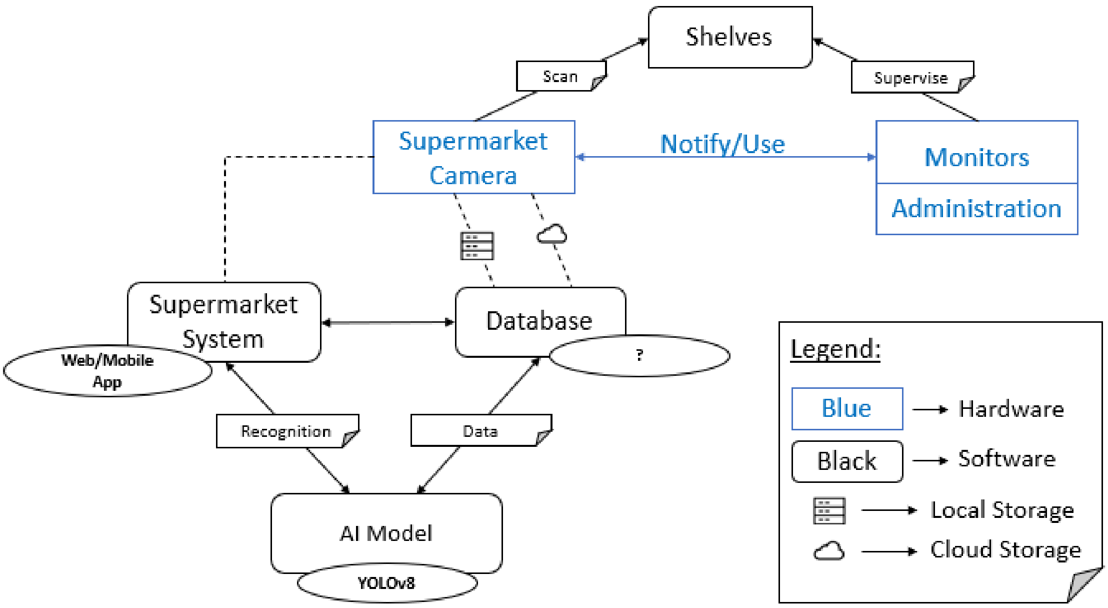

Figure 10: Technical Schema for Supermarket System

9 Conclusion

In conclusion, the Automated Shelf Monitoring System marks a significant leap forward in the realm of supermarket inventory management. By harnessing the power of AI and real-time data analysis, a solution that addresses the perennial challenge of maintaining well-stocked shelves in a fast-paced retail environment was created.

Throughout the course of this project, an advanced AI model capable of accurately detecting empty spaces on supermarket shelves was successfully developed.