CMPE 258: Project Milestone

BrailleCart: AI-Powered Grocery Assistance for the Visually Impaired

Team 6:
Sudip Das, 017549415, sudip.das@sjsu.edu
Alekhya Vaida, 017550039, alekhya.vaida@sjsu.edu
Prithvi Elancherran, 017518709, prithvi.elancherran@sjsu.edu

1. Achievements So Far

1.1 Grocery Object Detection Using YOLOv8n (Fully functional)

- Implemented object detection for various grocery items using the YOLOv8n model, achieving a precision of 99.52%.
- Training & Validation Results: The model achieved a mean average precision of 99.52% indicating high accuracy in detecting and classifying grocery items. The precision-confidence curve showed a precision of 1.00 at a confidence level of 0.878, suggesting that the model is highly reliable at this confidence threshold.
- Training & Validation Metrics: The training and validation losses (box loss, classification loss, and distribution focal loss) showed a consistent downward trend over 10 epochs, mages: indicating effective learning and reduced errors. Precision and recall values approach 1.0, suggesting that the model is both accurate and comprehensive in its detections.

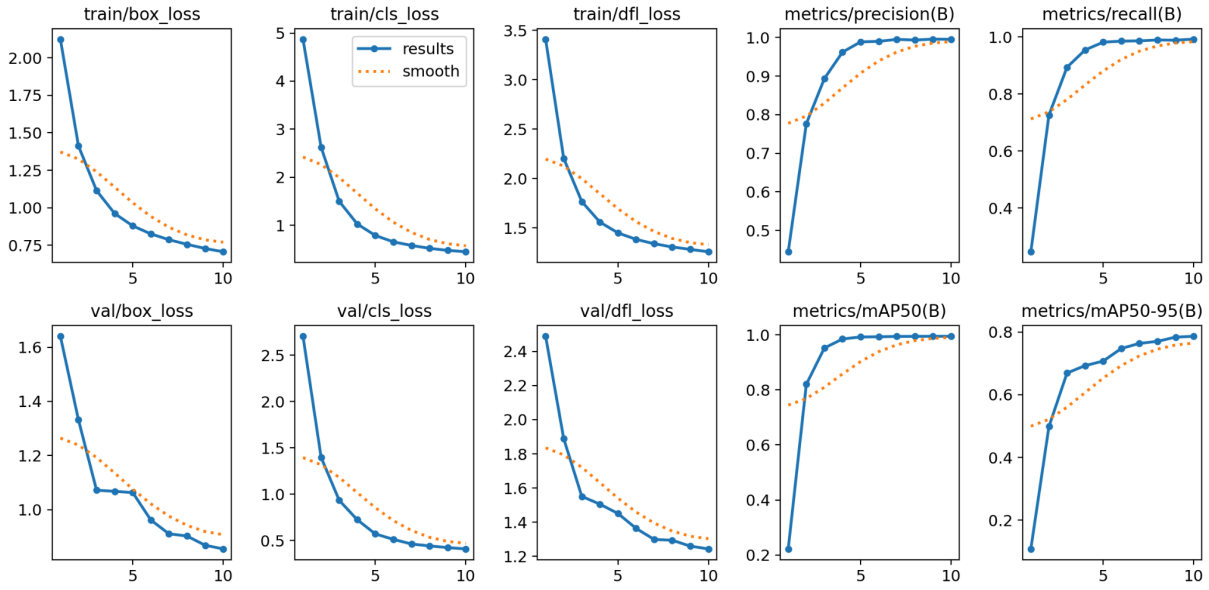

Figure 1: Model training results and metrics

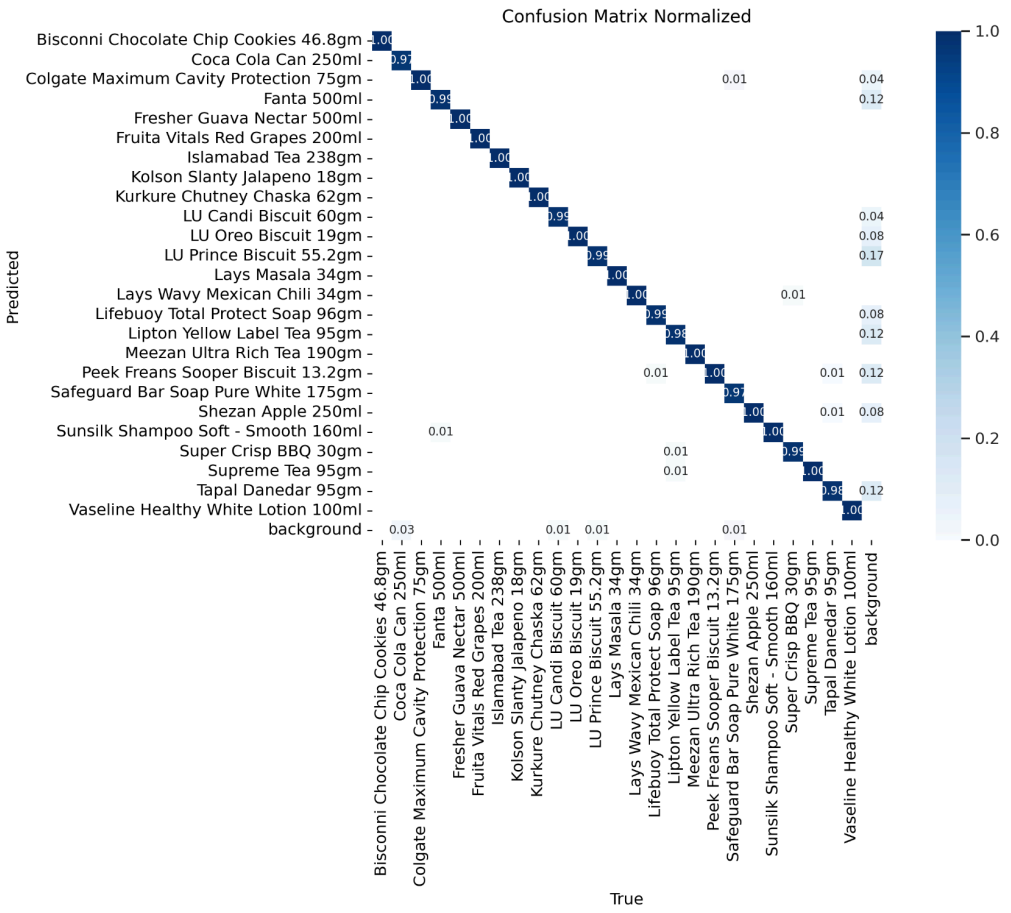

Figure 2: Confusion matrix

1.2 UI Development with Streamlit (Partially functional)

- Created a basic user interface using Streamlit to facilitate interaction between the visually impaired user and the BrailleCart system.
- This UI currently allows users to upload images of the products which triggers object detection and receive audio description of the product.

1.3 Planned for future development

- OCR integration
- LLM-based text-to-speech

1.4 Code Link: https://github.com/PrithviElancherran/BrailleCart

2. Baseline Modules Description

2.1 YOLOv8n Object Detection Module

- This is the core component of our system, responsible for detecting grocery items from real-time camera input.
- YOLOv8n was chosen for its accuracy and speed, which is crucial for providing visually impaired users with timely information.

- This module forms the baseline for all subsequent modules, as accurate item detection is key to the rest of the system functioning effectively.

2.2 UI (Streamline-based)

- The user interface provides a bridge between the visually impaired user and the BrailleCart system.
- The UI allows users to initiate object detection and receive feedback in an accessible format.
- This module's partial functionality currently supports basic interactions, but we plan to enhance it with additional features in upcoming milestones.

3. References

- Roboflow Universe, "Grocery Dataset," Link
- Ultralytics, "SKU-110K Dataset," Link

4. Challenges Encountered

4.1 OCR Integration

- The initial attempt at OCR integration faced issues in detecting labels under different lighting and angle conditions.
- The quality of the images varied, and the text was often partially obscured, making it difficult for the OCR to extract meaningful information.

5. Plans to Overcome Challenges

5.1 Improving OCR Integration

- Plan to experiment with additional OCR tools such as Tesseract and implement preprocessing techniques to enhance image quality.
- Aiming to make this module more robust by addressing specific issues with text detection under varied lighting conditions.

5.2 Optimizing Module Integration

- To address latency, we plan to optimize communication between modules using asynchronous programming and improve computational efficiency.
- This will involve profiling each component to identify bottlenecks and addressing them appropriately.

5.3 Future Developments

- To Enhance the UI with features such as voice commands and better flow.
- To Refine the LLM response generation to provide a more natural and conversational experience.