Logo Detection in Videos

Objective

The aim of this project is to establish a machine learning (ML) pipeline capable of identifying Pepsi and CocaCola logos in video files. The pipeline is designed to capture frames from the video, use a YOLOv7 model to recognize logos, and generate a JSON file with the timestamps of detected logos.

Dataset Used

Pepsi and CocaCola Images dataset consists of 400 images designed for computer vision tasks, focusing on logo recognition and classification. [1]

Approaches Considered

- YOLOv3
- YOLOv5
- YOLOv7 (Selected)

Detailed Analysis

1. YOLOv3

Pros:

- Well-established and widely used in various applications.
- Good balance of speed and accuracy.

Cons:

- Older architecture with limitations in performance compared to newer models.
- Larger model size and slower inference times compared to more recent versions.

Decision: Not considered due to being an older solution with less optimal performance.

2. YOLOv5

Pros:

- Improved performance over YOLOv3 with faster and more accurate detections.
- Actively maintained and widely supported in the community.

Cons:

- Though newer than YOLOv3, it is not the latest model available.
- Model size and complexity can be higher.

Decision: Considered initially but not chosen because YOLOv7 offers further improvements.

3. YOLOv7

Pros:

- State-of-the-art performance in terms of speed and accuracy.
- Optimized for real-time applications and edge devices.
- Advanced features like anchor-free detection and dynamic training mechanisms.

Cons:

- Newer and less documented compared to older models.

Decision: Chosen due to its superior performance and optimizations, making it ideal for the project's requirements.

Figure 1: Enter Side by side comparison of YOLOv5 (left) and YOLOv7 (right)

Final Decision

YOLOv7 was selected as the primary model for the logo detection project. The decision was based on its cutting-edge performance, optimizations for real-time and edge applications, and the ability to handle complex detection tasks efficiently.

Initial Steps

- **Familiarization:** Started with YOLOv7 due to prior experience with YOLO models and the compelling performance benefits of the latest version.
- **Implementation:** Integrated YOLOv7 into the project for logo detection, leveraging its advanced capabilities to meet the project requirements.

Conclusion

Selecting YOLOv7 is consistent with the project's objectives of attaining excellent performance, efficiency, and real-time detection abilities. Choosing the newest developments in the YOLO series ensures that the project uses cutting-edge technology to accurately detect logos in different situations, leading to strong and precise results.

Results

Observations

- **True Positive Rates (Diagonal Elements)**:
  - **Coke**: The model correctly identifies coke 83% of the time.

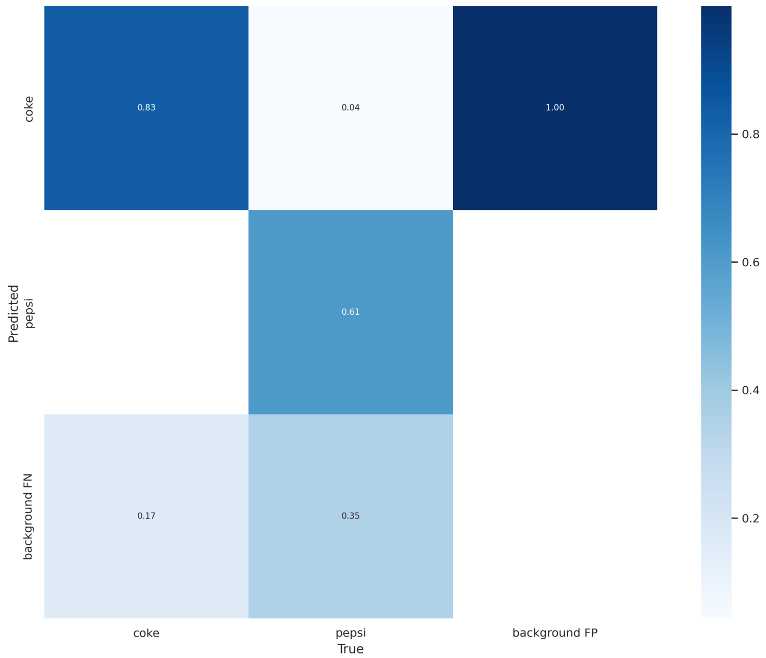

Figure 2: Confusion Matrix

- **Pepsi**: The model correctly identifies pepsi 61% of the time.
- **Background FP**: The model has a high rate (100%) of correctly identifying background false positives.

- **False Positive and False Negative Rates (Off-diagonal Elements)**:
  - Coke predicted as Pepsi: 4%
  - Coke predicted as background FN: 17%
  - Pepsi predicted as background FN: 35%
- **Misclassification Insights**:
  - The model has some confusion between coke and pepsi, with a small percentage of coke being predicted as pepsi.
  - A significant portion of the pepsi is misclassified as background false negatives (35%), indicating the model struggles more with identifying pepsi compared to coke.
  - The high accuracy in identifying background FP (100%) suggests the model is very effective at distinguishing non-relevant items from the relevant classes.

References

[1] My workspace. Pepsi and cocacola images dataset. https://universe.roboflow.com/my-workspace-7m1hi/pepsi-and-cocacola-images, Dec 2022. Visited on 2024-07-08.