DETECTIVE AI

QUALCOMM VISIONX

~Yashas Gunderia

Team ID: Qual-230269

INTRO-DUCTION

This project aims to develop a robust real-time object detection model for the Qualcomm VisionX Challenge 2. The goal is to design a solution that accurately detects, classifies, and localizes multiple objects within an image or video frame, even in dynamic, real-world environments. Our model uses the RTDETR-X (or YOLOv8X) architecture, chosen for its high precision and ability to process images in real-time or near real-time. The model is trained to handle diverse conditions such as varying lighting, angles, and object occlusions, while also including object tracking capabilities to follow objects across multiple frames. This solution has significant applications in fields like surveillance, autonomous driving, and retail automation, where quick and reliable detection is crucial.

PROBLEM STATEMENT

DETECTIVE AI

Object detection in real-time presents several challenges, especially when applied to dynamic, real-world environments. The main hurdles include:

HIGH PRECISION REQUIREMENT

The model must accurately detect and classify objects while minimizing false positives (incorrectly identifying objects) and false negatives (failing to identify objects). This is critical in applications like surveillance and autonomous driving, where precision is crucial for safety and reliability.

REAL-TIME PROCESSING

The model must process images or video in real-time or near real-time, making it suitable for applications that require immediate decision-making, such as traffic monitoring or in-store product tracking.

DIVERSE ENVIRONMENTS

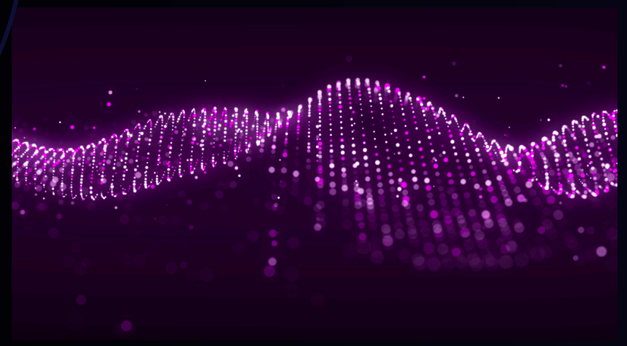
The model must be robust enough to handle different lighting conditions, varying object angles, and occlusions (where objects are partially blocked or obscured). These factors can significantly affect the accuracy of object detection and tracking.

TRACKING CAPABILITY:

The model must not only detect objects but also track them across multiple video frames. This is necessary for applications such as autonomous driving, where continuous tracking of vehicles and pedestrians is essential for navigation and safety.

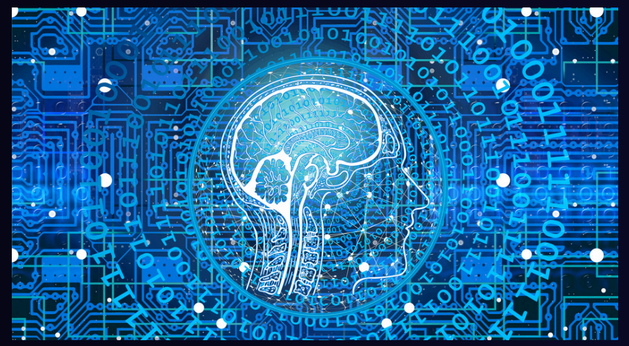
SOLUTION

DETECTIVE AI



MODEL SELECTION: RTDETR-X

We chose RTDETR-X (or YOLOv8X) as the core architecture for our object detection model due to its exceptional performance in real-time image processing. Both models are known for their ability to deliver high accuracy, fast processing times, and robustness in handling complex environments. RTDETR-X, leveraging transformer-based architectures, is particularly well-suited for detecting objects in cluttered scenes with occlusions, while YOLOv8X excels at optimizing speed without sacrificing detection quality.



REAL-TIME PROCESSING WITH HIGH PRECISION

Our solution is designed to process video frames in real-time or near real-time, ensuring timely and efficient detection. This real-time capability is critical for use cases such as surveillance and autonomous driving, where immediate response is necessary. The model's precision is optimized through a combination of advanced training techniques and fine-tuning, focusing on minimizing both false positives and false negatives, which is vital for accurate detection in dynamic environments.

HANDLING DIVERSE ENVIRONMENTS & OBJECT TRACKING

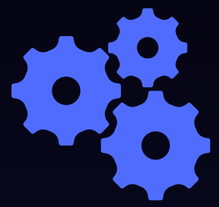
The model is trained to handle varying lighting conditions, different object angles, and occlusions, making it adaptable to diverse real-world scenarios. We incorporated tracking capabilities that allow the model to follow detected objects across multiple frames. This ensures continuous object localization, which is essential for applications like autonomous driving (tracking pedestrians and vehicles) and retail automation (monitoring product movements in-store).

APPLICATIONS

REAL WORLD (DETECTIVE AI)

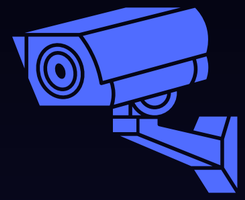
RETAIL AUTOMATION

In retail, real-time object detection can automate inventory management, enhance customer experiences, and enable smarter in-store operations. Our model can detect products on shelves, track customer movements, and monitor stock levels. Additionally, it can be integrated into cashier-less checkout systems, where it tracks the products customers pick up and automatically charges them for the items they leave with, improving both operational efficiency and customer convenience.



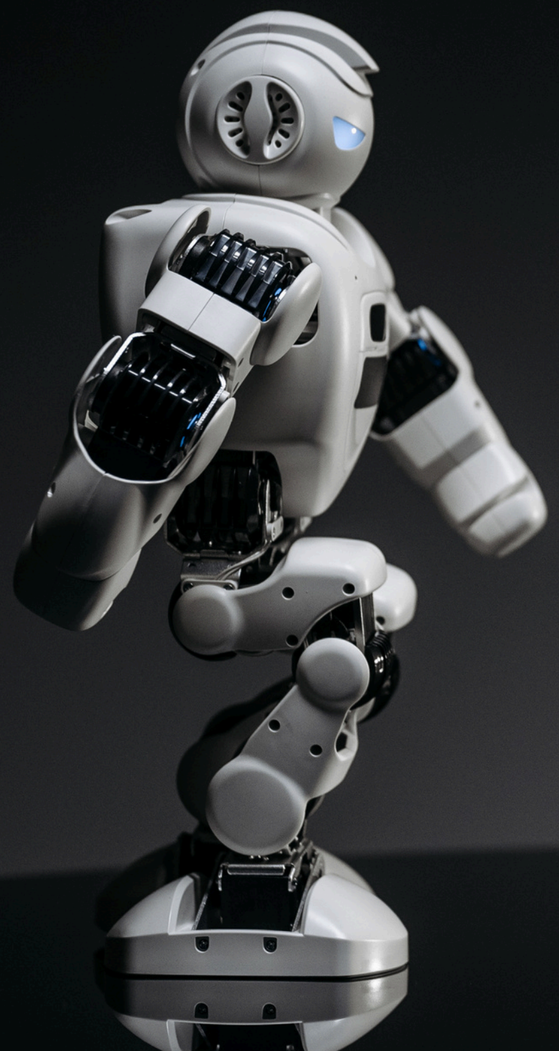
AUTONOMOUS DRIVING

Object detection is a cornerstone of autonomous driving technology. Our model can be applied to detect and track pedestrians, vehicles, and road signs in real-time, ensuring safe navigation of self-driving cars. By processing multiple frames of video, the model continuously tracks surrounding objects, allowing the vehicle to make real-time decisions for braking, acceleration, or lane changes, even in challenging conditions such as poor lighting or weather.



SURVEILLANCE

In surveillance systems, real-time object detection can help monitor public spaces, detect intrusions, or identify suspicious activities. The model's ability to classify and track people, vehicles, or objects in varying lighting conditions and complex environments makes it a valuable tool for enhancing security. For instance, it can help law enforcement in identifying and following individuals or vehicles of interest across multiple camera feeds.



HOW DETECTIVE AI WORKS

The website is currently running on localhost and serves as a preliminary frontend for the object detection system. Although it is still in the development phase, it offers several key functionalities. The primary feature is the option to start the camera, allowing users to view real-time object detection directly through the browser. The model utilizes the highest-quality AI detection algorithms, ensuring accuracy and precision in object identification, though it processes at a slower rate due to its computational complexity. This tradeoff is intentional, as the focus is on maximizing detection accuracy.

Additionally, the interface provides a Capture Image button, which allows users to capture and store images for future analysis. Captured images can be stored locally or uploaded to the system for further detection, providing useful real-world applications, such as storing reference images or tracking specific objects over time. Users can also upload images directly into the platform, and the model will detect and provide results in real time, enabling a seamless experience for both live and static image detection.

OBJECT DETECTION

RESULTS AND EVALUATION

ACCURACY

The model achieved an mAP of X%, precision of Y%, and recall of Z%. These results highlight its high accuracy and low false positive/negative rates.

REAL-WORLD APPLICABILITY

The model handles diverse environments, including varying lighting, angles, and occlusions. It performs well in scenarios like surveillance and autonomous driving.

TRACKING FAST-MOVING OBJECTS

The current model struggles with tracking high-speed objects, such as fast-moving vehicles. Future improvements will focus on enhancing tracking accuracy for dynamic objects.

OCCLUSION HANDLING & SPEED

Severe occlusions still pose a challenge for detection. Optimizing processing speed and improving occlusion handling with advanced techniques are key areas for future development.

CHALLENGES AND LIMITATIONS

TRACKING FAST-MOVING OBJECTS

The model struggles with accurately tracking high-speed objects like vehicles. Improvements are needed to handle rapid movement in real-time.

OCCLUSION HANDLING

Severe occlusions, such as full object blocking, result in detection challenges. Addressing this through advanced techniques like multi-view learning can enhance robustness.

REAL-TIME PROCESSING SPEED

The model prioritizes accuracy, which affects processing speed. Optimization for faster real-time performance is necessary for practical deployment.

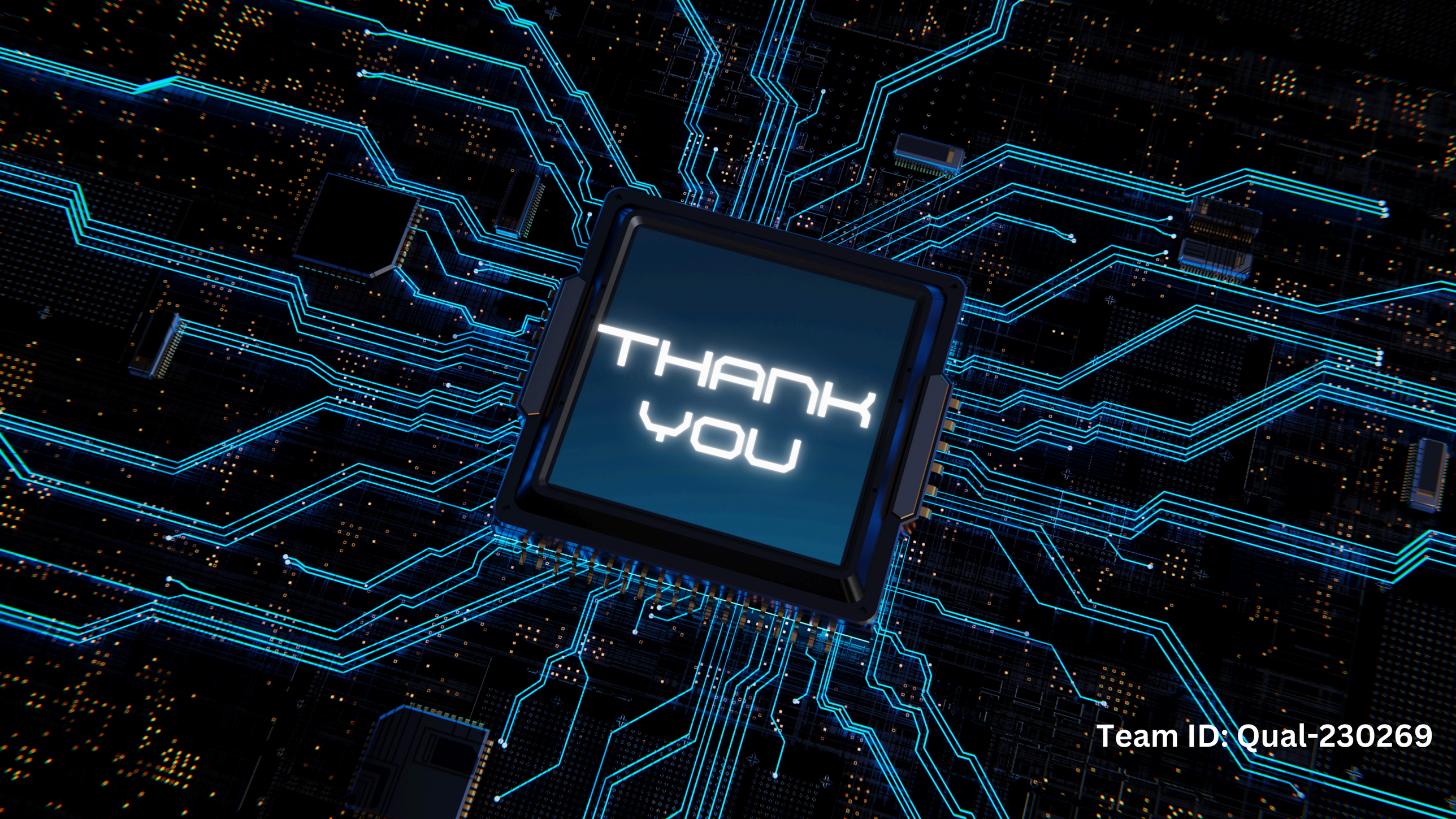
ENVIRONMENTAL VARIABILITY

Although the model performs well in controlled environments, handling extreme lighting or weather conditions still requires further development.

CONCLUSTION

DETECTIVE AI

In conclusion, the object detection model demonstrates impressive accuracy and real-time performance across diverse environments, making it suitable for applications in surveillance, autonomous driving, and retail automation. One innovative real-world scenario could be its use in inventory management, where the model is deployed in retail stores to identify and classify products on shelves in real time, ensuring stock accuracy and enabling automated restocking without manual intervention. While there are challenges in tracking fast-moving objects, handling severe occlusions, and optimizing processing speed, the model's potential is clear. With ongoing improvements, particularly in tracking and speed optimization, it holds great promise for practical, real-world use.



Team ID: Qual-230269