Overview :

The Vision Drive project was an exploration into using **computer vision for real-world perception and decision-making**. The goal was to build a pipeline that can take visual input, extract meaningful information from it, and use that information to guide actions (or simulate them). This project helped me connect theoretical concepts of vision with practical implementation challenges.

What I Did :

Built the vision pipeline

- Used OpenCV for image preprocessing (resizing, filtering, color space conversion).
- Implemented detection logic using deep learning models (YOLO-based approach).
- Extracted relevant features from frames for further decision-making.

Model integration and testing

- Integrated pretrained models and tested them on real and sample data.
- Tuned thresholds and parameters to reduce false detections.
- Compared performance under different lighting and background conditions.

Experimentation and iteration

- Tried multiple preprocessing techniques to improve detection stability.
- Experimented with different frame rates and resolutions to balance speed vs accuracy.
- Debugged issues related to inconsistent predictions and latency.

Structuring the project

- Organized code into modules (data input, processing, inference, output).
- Added documentation and comments to make the pipeline understandable and reusable.

What I Learned

- Computer vision is **less about perfect models and more about good pipelines**.
- Preprocessing and data quality often matter more than model complexity.
- Real-world vision systems require constant tradeoffs between **accuracy, speed, and robustness**.
- Debugging vision models requires both visual inspection and systematic testing.
- Small design choices (like frame size or color space) can drastically affect results.