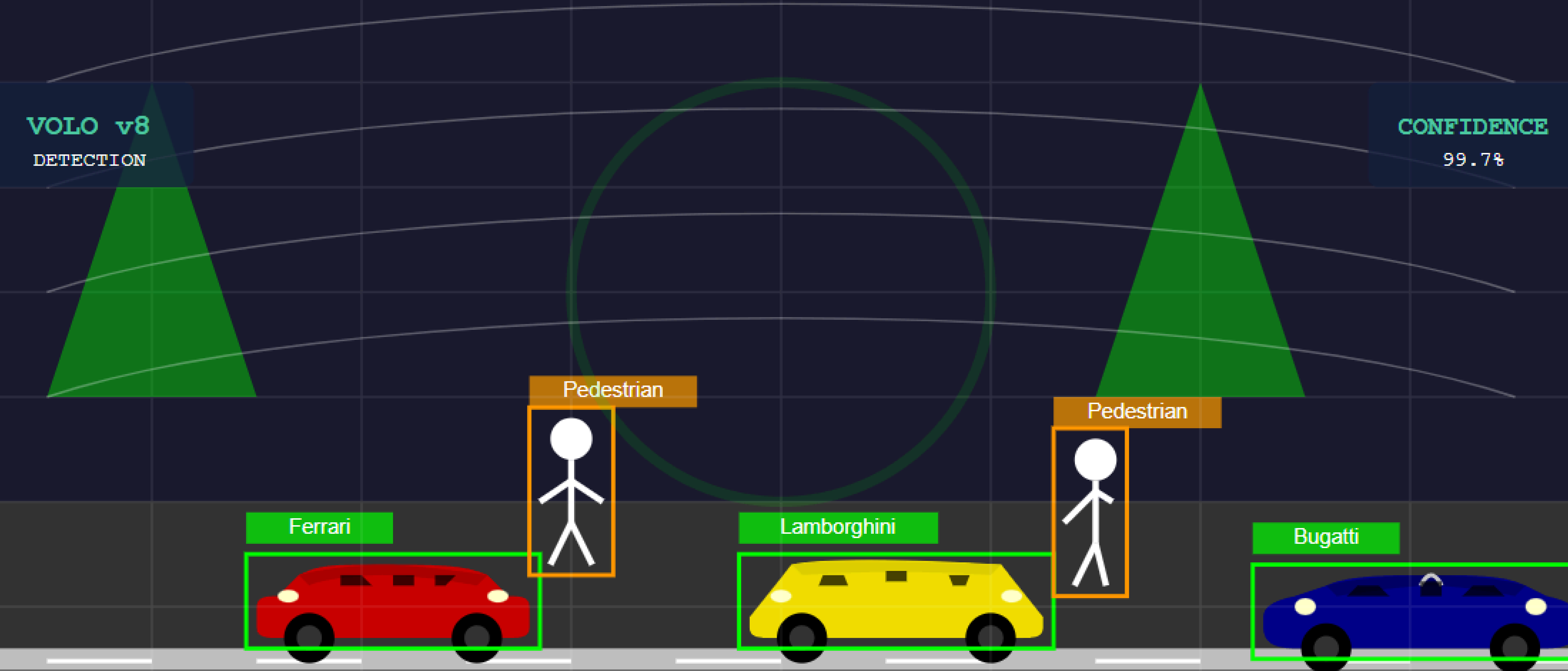
Smart Vision

Car & Pedestrian Detection using VOLO v8



BDD100K: Real-World Driving Dataset for Object Detection

01 Real-World Diversity:

Contains driving scenes from multiple U.S. cities with variations in weather (sunny, rainy), time (day/night), and environments (urban, rural, highway), making it ideal for training robust models.

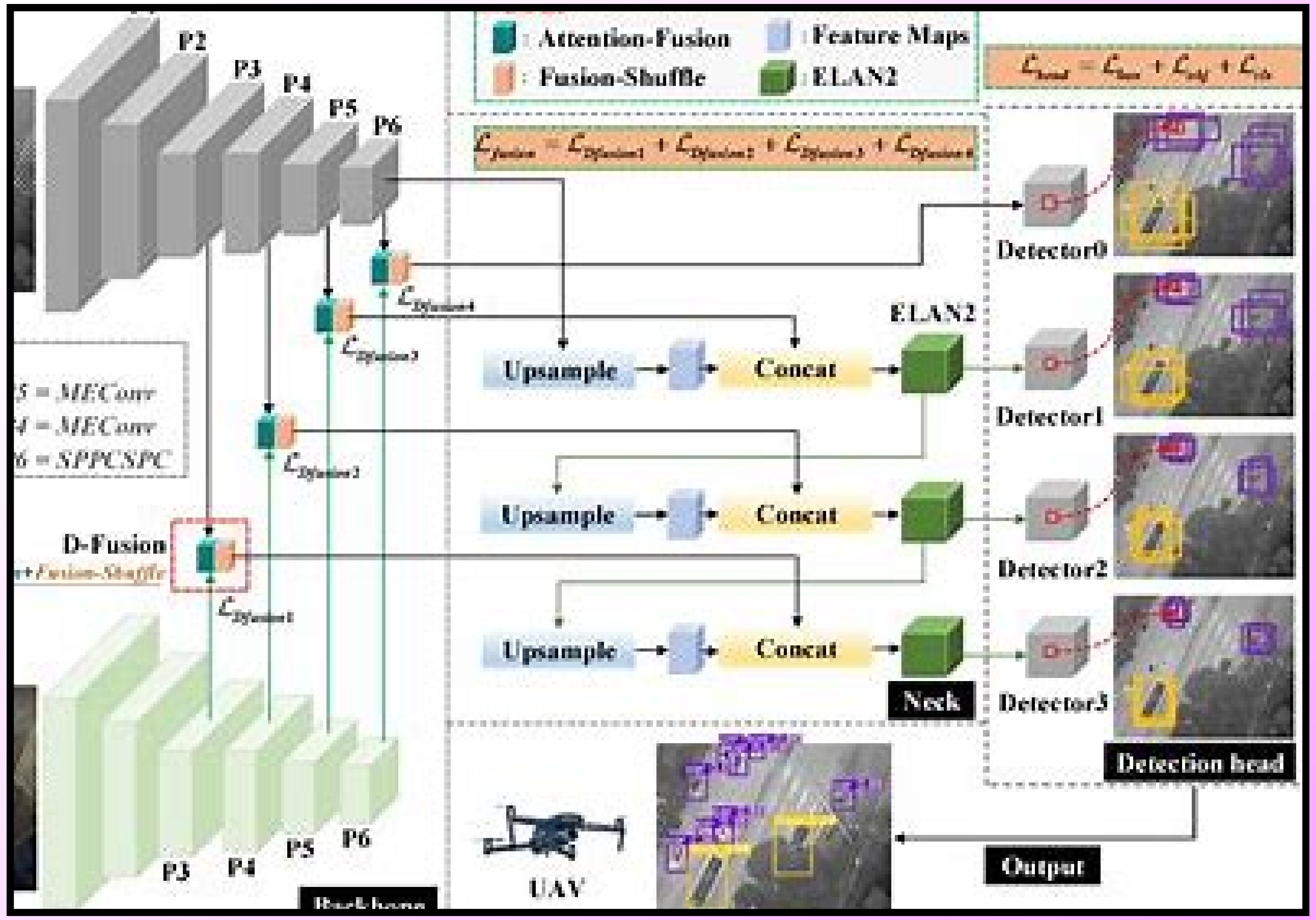
02 Rich Annotations

Includes detailed labels for 10 object classes (e.g., car, person, traffic light), with bounding boxes, lane markings, and drivable areas — useful for both detect.ion and segmentation tasks

03 Ideal for Autonomous Tasks:

Tailored for autonomous vehicle research, supporting multiple tasks like object detection, lane detection, and tracking in realistic traffic scenarios.

MODEL ARCHITECTURE: YOLO V8



1.Backbone – CSPDarknet:

Utilizes a Cross Stage Partial (CSP) backbone to extract rich features from the input image while reducing computation and maintaining accuracy.

2. Neck – PAN (Path Aggregation Network):

Enhances feature fusion from different scales, improving object detection for both small and large objects by combining low-level and high-level features.

3. Head – Anchor-Free Detection:

YOLOv8 adopts an anchor-free head design for simplified object prediction, directly regressing object centers and bounding boxes, leading to faster inference and easier training.

4. Modular & Lightweight Design:

Optimized for real-time applications with a balance between speed and accuracy; supports easy integration with ONNX, TensorRT, and deployment on edge devices.

TRAINING YOLOV8 ON BDD100K

01 Data Preparation

Converted BDD100K annotations to YOLO format (class, x_center, y_center, width, height). Split the dataset into training, validation, and test sets for structured evaluation.

02 Training Configuration

Used Ultralytics' YOLOv8 repo with custom data.yaml. Set hyperparameters like image size, epochs, batch size, and pretrained weights to fine-tune on car and pedestrian classes.

03 Monitoring & Evaluation

Tracked training using metrics like mAP, loss curves, and precision-recall. Evaluated model on validation set to avoid overfitting and ensure generalization.

EVALUATION METRICS & MODEL PERFORMANCE

01 Mean Average Precision (mAP)

Measures the average precision across all classes and IoU thresholds. Higher mAP reflects the model’s ability to detect objects accurately and consistently. In our case, mAP@0.5 showed strong performance for the 'car' class and decent results for 'pedestrian'.

02 Intersection over Union (IoU):

IoU = Area of Overlap / Area of Union between predicted and ground truth boxes. A common threshold is IoU > 0.5 for a correct detection. Our model maintained good IoU values on large objects, with minor drops on occluded or small-scale pedestrians.

03 Precision & Recall:

Precision: Ratio of correct detections to total predicted detections (low false positives). Recall: Ratio of correct detections to all actual objects (low false negatives).

04 Model Performance Summary:

Consistently detected cars with high accuracy and low false positives. Pedestrian detection had slightly lower recall due to motion blur, occlusions, and small size. Overall, the model generalizes well across diverse BDD100K scenarios.

EVALUATION METRICS & MODEL PERFORMANCE

01 Inference Speed:

YOLOv8 is optimized for real-time detection, achieving fast inference speeds even on edge devices like Jetson Nano or mobile GPUs. This makes it highly suitable for live video feeds in traffic or surveillance settings.

02 Model Size & Variants:

Offers different variants (YOLOv8n, s, m, l, x) depending on the trade-off between speed and accuracy. Smaller models were tested for mobile deployment, while larger ones provided higher accuracy in evaluations.

03 Export & Deployment:

The trained model was exported to ONNX and TorchScript formats for compatibility with other platforms. Deployment-ready with integration into camera pipelines or web-based dashboards.

04 Post-Processing Improvements:

Applied confidence thresholding and non-maximum suppression (NMS) to refine predictions. Improved output stability and reduced overlapping boxes in crowded urban scenes.

RESULTS & VISUALIZATIONS

Detection Samples:

Included sample images from the validation set showing accurate bounding boxes around cars and pedestrians. Visual outputs demonstrate strong performance even in varied lighting and occlusion.

Bounding Box Confidence:

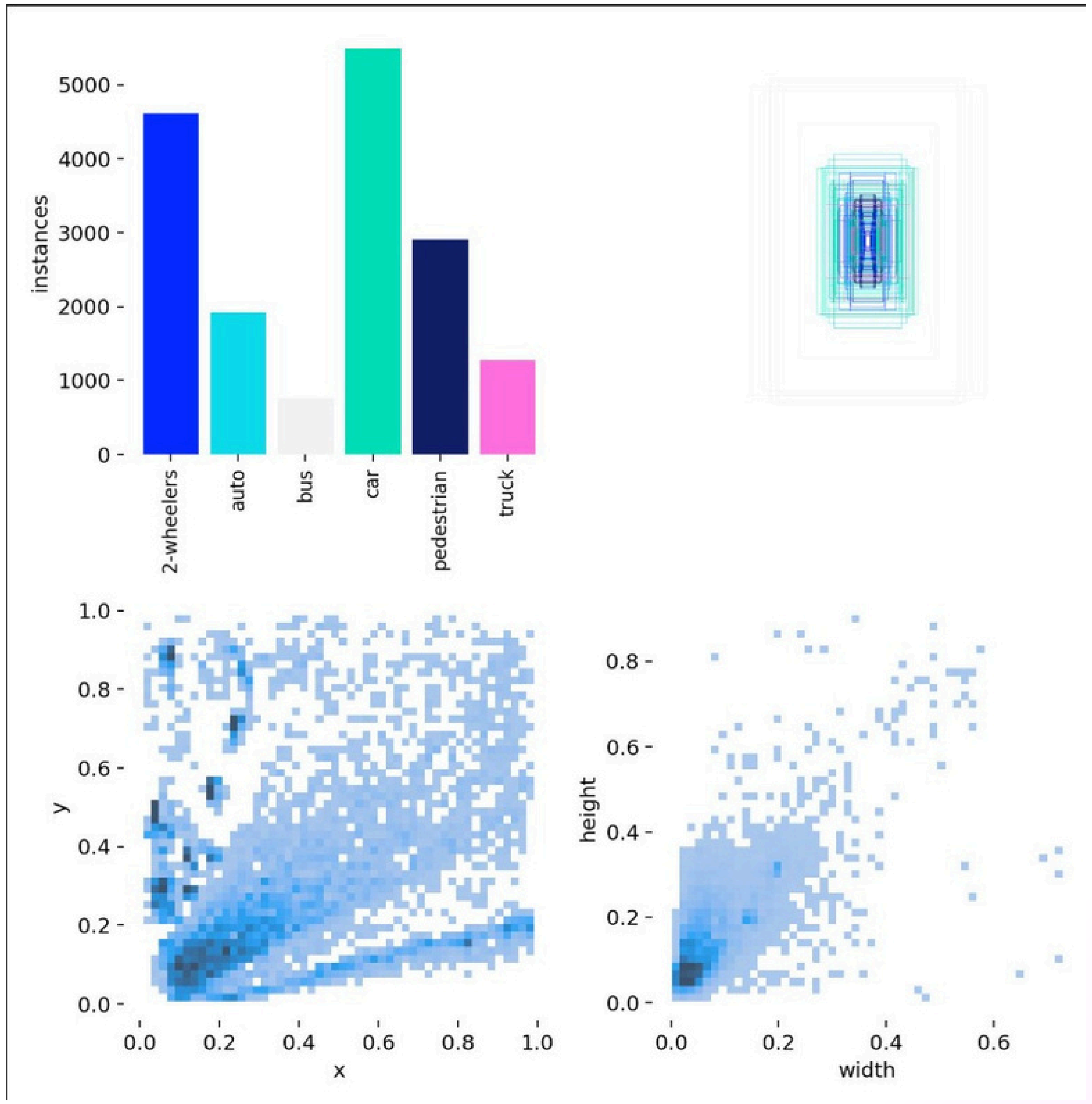
Boxes are color-coded based on class, with label and confidence scores. High-confidence detections (>90%) are visible on most car instances; pedestrian confidence varies slightly in cluttered scenes.

Before vs After Training:

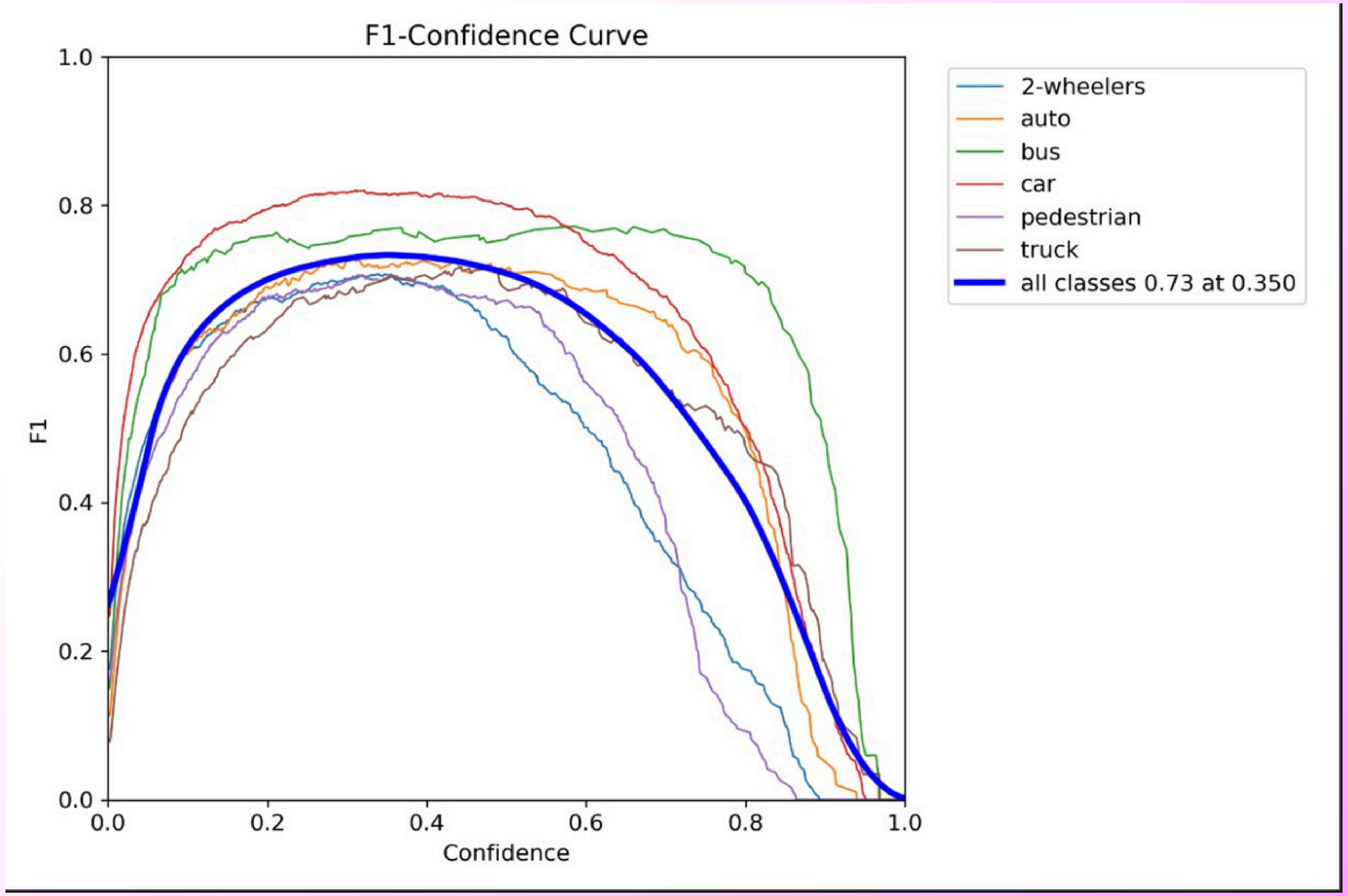
Side-by-side comparison of predictions using pretrained weights vs. fine-tuned YOLOv8 model shows significant improvement in recall and reduced false positives.

imes

VISUALIZATIONS



Label Encoding Graph



F1 Confidence Curve

CONCLUSION & FUTURE WORK

Project Summary:

Successfully trained YOLOv8 on BDD100K to detect cars and pedestrians with strong mAP and real-time inference performance.

Impact:

This system can be integrated into smart traffic monitoring, ADAS (Advanced Driver Assistance Systems), or pedestrian safety applications.

Future Improvements:

Expand to include more classes (e.g., traffic lights, bicycles). Train with more night-time and foggy images. Deploy on edge devices using TensorRT or OpenVINO.

THANK YOU

22BAC10004 - Sreeram MV
22BAC10026 - Sourabh Shrivastava
22BAC10027 - Saaransh Gupta

22BAC10032 - Badrikanath Praharaj
22BAC10047 - Nishant Kumar
22BEC10022 - Ayush Billore