© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1717175

Yolo V8 And Canny: A Dual Detection System for Road Defect Analysis and Lane Guidance

MADALA JAYANTH KUMAR[1], PATTI EBBEJU[2], MADDINENI SAI GANESH[3], KESANA PAWAN CHANDHU[4]
[1, 2, 3, 4]Department of CSE (AI & ML), RVR & JC College OF Engineering.

Abstract- Road safety systems need to do two things at once: spot hazards on the surface and track where the lane boundaries are. Most approaches handle one or the other. We built a system that does both. The framework pairs a YOLOv8m deep learning model for road damage detection with a classical computer vision pipeline for real-time lane tracking. Training followed a global-to-local strategy — we first trained on 6,832 images from India and Japan to build a broad foundation, then fine-tuned specifically on US road data. This transfer learning approach lets the model handle regional variation without starting from scratch for each country. For lane detection, we used Canny edge detection and Probabilistic Hough Transforms, constrained by a dynamic spatial mask that focuses the system on the relevant driving area. The result: a fine-tuned model with an mAP@0.5 of 0.571 across multiple road distress categories. A FastAPI backend ties both pipelines together, enabling concurrent damage detection and lane guidance in real time. Combining deep learning with geometric constraints gives autonomous systems a more complete picture of the road than either method could alone.

I. INTRODUCTION

The Autonomous vehicles don't get to pause when road conditions turn complicated. They need to simultaneously identify surface hazards — potholes, cracks, structural damage — and maintain a reliable sense of where the lane is. Doing both, in real time, across diverse road conditions and geographic regions, is harder than it sounds.

Environmental perception here is really two problems bundled together. One is semantic: what's on the road? The other is geometric: where are the boundaries? Traditional image processing handles structured environments reasonably well, but it breaks down when the road surface itself is damaged or complex. Deep learning, especially YOLO-based architectures, flips that equation — fast, accurate object detection, but without the built-in geometric reasoning that lane tracking requires.

We built a hybrid system that doesn't force a choice between the two. The YOLOv8m model handles damage detection, trained first on a combined India-Japan dataset to establish cross-regional feature understanding, then fine-tuned on US infrastructure data. The lane detection side uses Canny edge detection and Probabilistic Hough Transforms, with a dynamic ROI mask ensuring that only the ego-lane area is processed. The goal was a system that gives autonomous navigation a dense, layered view of the road — not just "is there a crack" or "where is the lane," but both, running together.

II. RELATED WORK

A. Object Detection in Road Environments

Recent advancements in real-time object detection have been dominated by the "You Only Look Once" (YOLO) framework due to its efficiency in balancing speed and accuracy. While earlier iterations focused on general object categories, specialized research has shifted toward road infrastructure analysis. Your work builds upon this by utilizing YOLOv8m, which offers a sophisticated balance of layers for multi-country pattern recognition. Unlike traditional models that may struggle with regional variations, your approach utilizes a global-to-local training strategy, establishing a feature baseline on international datasets (India and Japan) before fine-tuning for specific infrastructure standards in the USA. This mirrors the industry trend of using transfer learning to overcome dataset-specific biases.

B. Structural Lane Boundary Isolation

Classical Classical computer vision still earns its place when geometric precision matters more than semantic understanding. Standard pipelines — grayscale conversion, Gaussian blurring, Canny edge detection, Hough transforms — are well-established tools for extracting road structure from raw video.

Our approach refines this by applying a dynamic ROI mask to the lower portion of each frame. This eliminates noise from the horizon and off-road surroundings before the geometric filtering even begins. Only slopes and line segments meeting specific criteria — by magnitude and length — survive to the output stage.

C. Hybrid and Concurrent Processing Systems

The integration of deep learning and classical pipelines is an emerging area of research aimed at providing a holistic "world view" for autonomous systems. While deep learning excels at detecting localized hazards like road damage (cracks or potholes), classical methods provide the necessary geometric framework for navigational guidance. Your architecture addresses this by managing both pipelines concurrently via a FastAPI backend, allowing for real-time video analysis where semantic detections (YOLOv8) are overlaid onto the structural boundaries (Hough-based lanes). This hybrid approach mitigates the limitations of using a single methodology, providing a comprehensive environmental perception system that tracks both discrete hazards and continuous boundaries.

III. PROPOSED METHODOLOGY

The proposed framework integrates deep learning-based object detection with classical computer vision techniques to provide a comprehensive road analysis system. The architecture is divided into two primary sub-systems: a YOLOv8m pipeline for semantic road damage detection and an OpenCV-based pipeline for structural lane isolation.

A. Road Damage Detection (Deep Learning)

The detection of localized road hazards is treated as a multi-class object detection problem using the YOLOv8m (Medium) architecture.

- Global-to-Local Training Strategy: The model is first trained on a combined dataset of 6,832 images from India and Japan for 40 epochs to establish a robust baseline for diverse road distress patterns.
- Targeted Fine-Tuning: The learned weights from the global phase are then transferred and fine-tuned on a USA-specific dataset for 20 epochs. This transfer learning approach allows the model to adapt to specific American infrastructure standards while retaining general knowledge of road damage features.
- Loss Functions: The model optimizes three primary loss components:
    - Box Loss (box_loss): Measures the accuracy of the predicted bounding box coordinates using Complete IoU (CIoU).
    - Class Loss (cl_loss): Evaluates the accuracy of the category prediction for the eight identified damage classes (D00–D80).
    - DFL Loss (dfl_loss): Distribution Focal Loss, used to refine the boundaries of the bounding boxes.

B.Lane Detection Pipeline (Classical Computer Vision)

The lane detection module operates as a sequential image processing pipeline to isolate structural road boundaries from video frames.

- Preprocessing: Each frame I undergoes grayscale conversion followed by Gaussian Blurring with a kernel size of 5 to reduce high-frequency noise.
- Edge Extraction: Canny edge detection is applied using dual thresholds (low = 50, high = 150) to identify significant intensity gradients corresponding to road markings.
- Spatial Isolation (ROI): A trapezoidal Region of Interest (ROI) mask is applied to the frame to eliminate off-road noise and focus processing on the ego-lane. The mask coordinates are defined as a function of the image width (w) and height (h)

C. Structural Boundary Modeling

The system employs the Probabilistic Hough Line Transform to extract mathematical line segments from the isolated edge map.

- Hough Transformation: Lines are identified in the polar coordinate system . The system uses a threshold of 100, a minimum line length of 40, and a maximum gap of 5 pixels to ensure connectivity.
- Average Slope-Intercept: Detected segments are categorized into left and right lanes based on their slope (m):
    - Left Lane: $m < -0.3$
    - Right Lane: $m > 0.3$
- Linear Regression: For each group, the system calculates the average slope and intercept to produce a singular, stabilized lane boundary for the frame.

D. Performance Metrics

The YOLOv8m detection performance is evaluated using standard computer vision metrics. Let TP, FP, and FN represent True Positives, False Positives, and False Negatives respectively.

- Precision (P): The ratio of correctly predicted damage instances to the total predicted instances.

$$\text{Precision} = \frac{TP}{TP + FP}$$

- Recall (R): The ratio of correctly predicted damage instances to all actual instances in the ground truth.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- Mean Average Precision (mAP@0.5): The primary metric used to determine model accuracy across all categories at an Intersection over Union (IoU) threshold of 0.5.

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i$$

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Experimental Setup and Hyperparameters

All experiments were conducted utilizing a cloud-based environment equipped with an NVIDIA Tesla T4 GPU (15 GB VRAM). The YOLOv8m object detection pipeline was trained using an AdamW optimizer with an initial learning rate of 0.000833 and a momentum of 0.9. Input images were standardized to a resolution of 640 X 640 pixels, processed in batches of 16. To ensure robust feature extraction, the global baseline model was trained for 40 epochs on the combined India and Japan datasets, followed by a targeted 20-epoch fine-tuning phase specifically on the USA dataset.

B. Quantitative Evaluation: Road Damage Detection

The performance of the YOLOv8m model is evaluated across multiple distress categories (D00 to D80) using Precision, Recall, and mean Average Precision ($mAP@0.5$ and $mAP@0.5:0.95$). Table 1 summarizes the comparative performance of the baseline global model against the fine-tuned USA model.

The training logs indicate that the model achieves strong convergence, particularly in identifying prevalent structural defects such as longitudinal linear cracks (D00), Conversely, underrepresented classes in the validation set (such as D40 and D60) exhibited lower confidence scores, underscoring the direct correlation between class volume in the training corpus and predictive reliability.

C. Qualitative Evaluation: Structural Lane Isolation

The classical computer vision pipeline was evaluated on its ability to maintain persistent lane tracking across sequential video frames. The dynamic spatial isolation—restricting the Canny edge detection and Probabilistic Hough Transform to the lower horizontal plane of the frame—successfully eliminated ambient noise from off-road infrastructure and horizon artifacts.

By applying strict Hough parameters (a threshold of 100, minimum line length of 40, and maximum line gap of 5), the algorithm demonstrated high resilience. Missed segment detections in isolated frames were effectively compensated for by the frame-to-frame averaging mechanism. The geometric constraints (filtering slopes |m| < 0.3) prevented horizontal shadows and transverse road cracks from being falsely classified as lane boundaries.

Table 1: Yolo models results

Models	GFlops	Precision	Recall	mAP@0.5
YOLOv8	71.9	0.617	0.751	0.571
YOLOv9	237.5	0.559	0.601	0.547
YOLOv10	112.9	0.626	0.579	0.542

D.YoloV8 Results

Figure 1 demonstrates the qualitative performance of the fine-tuned YOLOv8m model on a complex, wet asphalt surface. The system successfully isolates alligator cracking (D20) at multiple scales, capturing both the broad area of deterioration and the specific, severe crack clusters within it. Furthermore, the model accurately detects water-filled structural voids (D50), proving its robustness against the challenging reflectivity and low contrast caused by adverse environmental conditions.

Finally, the precise overlapping of these bounding boxes confirms that the model can confidently resolve multiple, co-occurring distress types within a single localized region without suppression errors.

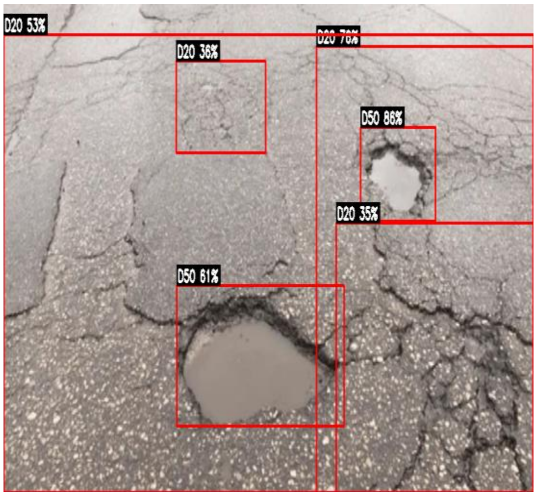

Figure 1: Result of YOLOv8

E. Lane Guidance Results

Figure 2 illustrates the robust performance of the classical lane isolation pipeline under clear daytime driving conditions. The algorithm, utilizing the Probabilistic Hough Transform, accurately identifies the structural boundaries of the ego-lane, seamlessly tracking both the solid white edge line and the dashed yellow center line. As depicted, the dynamic region-of-interest (ROI) mask successfully restricts line extrapolation to the immediate driving path, preventing distant horizon features or background textures from causing geometric distortion. Furthermore, the slope-averaging constraints ensure the projected boundaries remain smooth and continuous, verifying the system's reliability for real-time navigational guidance.

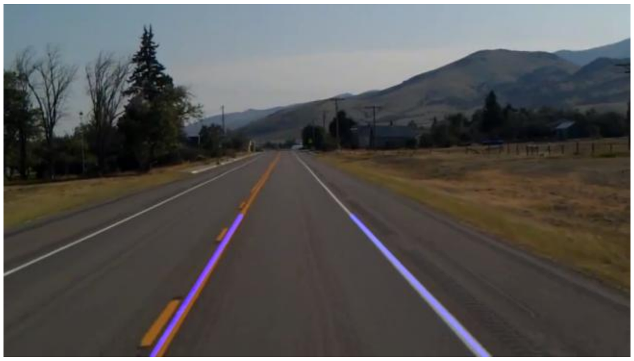

Figure 2: Result of Lane Guidance

V. CONCLUSION

This paper presented a comprehensive, hybrid computer vision framework designed to enhance environmental perception for autonomous navigation systems. By integrating a deep learning-based object detection model with a classical geometric analysis pipeline, the proposed architecture successfully executes concurrent road distress identification and structural lane isolation. The global-to-local transfer learning strategy proved highly effective; the fine-tuned YOLOv8m model achieved an mAP@0.5 of 0.571, demonstrating a robust capacity to identify multi-scale surface degradations, such as alligator cracking and water-filled potholes, even under challenging low-contrast conditions.

Concurrently, the classical lane isolation pipeline validated the enduring utility of geometric constraints. By coupling the Probabilistic Hough Transform with dynamic region-of-interest (ROI) masking and slope-averaging, the system reliably tracked ego-lane boundaries without succumbing to geometric distortion from background horizon textures or ambient noise. Unified by a FastAPI backend, this dual-pipeline approach provides a dense, multi-layered understanding of the road environment suitable for real-time application.

REFERENCES

[1] Klaus Bengler, Klaus Dietmayer, Berthold Farber, Markus Maurer, Christoph Stiller, and Hermann Winner. Three decades of driver assistance systems: Review and future per-spectives. IEEE Intelligent transportation systems magazine, 6(4):6–22, 2014.

[2] KT Chang, JR Chang, and JK Liu. Detection of pavement distresses using 3d laser scanning technology. In Computing in civil engineering (2005), pages 1–11. 2005.

[3] J Dharneeshkar, SA Aniruthan, R Karthika, Latha Parameswaran, et al. Deep learning based detection of pot-holes in indian roads using yolo. In 2020 international conference on inventive computation technologies (ICICT), pages 381–385. IEEE, 2020.

[4] Juan Du. Understanding of object detection based on cnn family and yolo. In Journal of Physics:

Conference Series, volume 1004, page 012029. IOP Publishing, 2018.

[5] Qian Gao, Pengyu Liu, Shanji Chen, Kebin Jia, and Xiao Wang. Detection method of potholes on highway pavement based on yolov5. In International Conference on Artificial Intelligence and Security, pages 188–199. Springer, 2022.

[6] Saksham Gupta, Paras Sharma, Dakshraj Sharma, Varun Gupta, and Nitigya Sambyal. Detection and localization of potholes in thermal images using deep neural networks. Multimedia tools and applications, 79:26265–26284, 2020.

[7] Zhiqiong Hou, Kelvin CP Wang, and Weiguo Gong. Experimentation of 3d pavement imaging through stereovision. In International Conference on Transportation Engineering 2007, pages 376– 381, 2007.

[8] Taehyeong Kim and Seung-Ki Ryu. Review and analysis of pothole detection methods. Journal of Emerging Trends in Computing and Information Sciences, 5(8):603–608, 2014.

[9] Qiang Liu, Wei Huang, Xiaoqiu Duan, Jianghao Wei, Tao Hu, Jie Yu, and Jiahuan Huang. Dsw-yolov8n: A new underwater target detection algorithm based on improved yolov8n. Electronics, 12(18):3892, 2023.

[10] Mingyang Ma and Huanli Pang. Sp-yolov8s: An improved yolov8s model for remote sensing image tiny object detection. Applied Sciences, 13(14):8161, 2023.