Integration and Comparison of Vision Models for Smart Inspection Cell

Case Study

Intelligent Systems in Production

Technische Hochschule Deggendorf
Campus Cham

Date : 15.01.2026

Group members:

Sai Prasanth Parnambedu	(Mt no : 12502673)
Yaswitha Pallela	(Mt no : 12504195)
Rakshith Thatikonda	(Mt no : 12503565)
Prarthana Shenoy	(Mt no : 12504810)
Chandrika Tirukkovalluri	(Mt no : 12502673)

Guided by

Prof. Hamidreza Heidari

Contents

Problem Statement

- Surface defects in automotive spur gears (scratch, dent, crack, pitting) can reduce drivetrain reliability.
- Manual and rule-based inspection methods are subjective and do not scale well for high production volumes.
- There is a need for a data-driven vision system that ensures accurate defect detection under real-time inference constraints.
- This work evaluates deep learning models for defect classification and detection, focusing on accuracy, robustness, and inference latency.

Objectives

- Train an object detection model (YOLOv8) to detect defects and localize them with bounding boxes.
- Train an image classification model (MobileNetV2) to classify images into: good, scratch, dent.

Compare both approaches using:

- Accuracy / Precision / Recall / F1-score (classification)
- mAp50 / mAp50-95(detection)
- Inference latency for real-time suitability
- Productivity rate

Workflow

Overview:
This project follows an end-to-end engineering methodology covering data creation, model development and performance validation to integrate into an industrial AI-based inspection system.

1. Data Preparation

- CAD-based synthetic spur gear dataset
- Defects added: scratches, dents, voids
- Variations in lighting, orientation, camera angle
- Dataset split: 80% train | 10% val | 10% test

2. Annotation & Quality Assurance

- Manual annotation in YOLO format
- Bounding box verification using Python scripts
- Ensured clean, reliable training data

Workflow

3. Model Development

YOLOv8 → Defect detection & localization, MobileNetV2 → Lightweight defect classification
Training strategy:

- Transfer learning
- Hyperparameter tuning
- Data augmentation
- Early stopping

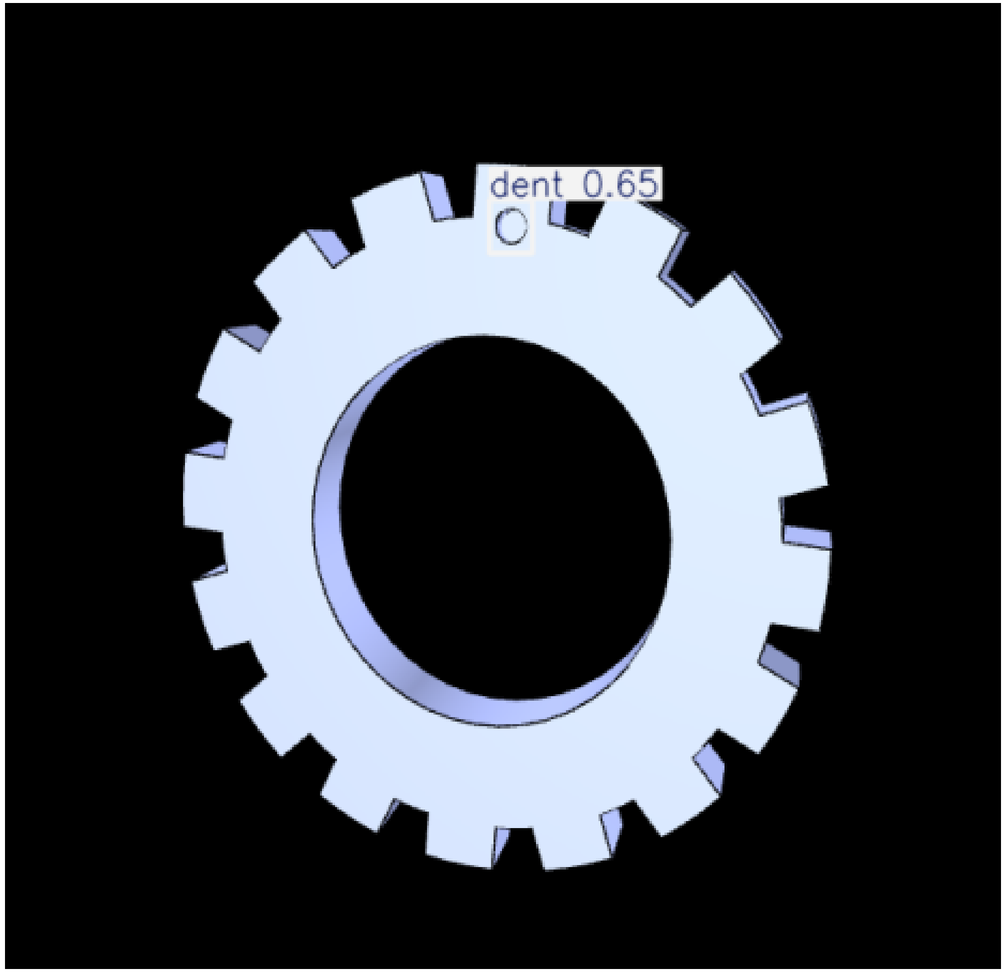
4. Evaluation & Validation
- Tested on unseen dataset
- Cross-validated using both models for reliability
- Metrics: Accuracy, Precision, Recall, F1, mAP@50, Latency, Productivity Rate

5. Future Work - Hardware Integration

- RoboDK for robotic inspection cell simulation
- SimPy for inspection workflow logic
- Workflow:
  Conveyor → Image Capture → AI Inference → Accept/Reject Sorting

Image classification results

Detected Defects in gear_dent_Gear_dent_34.png

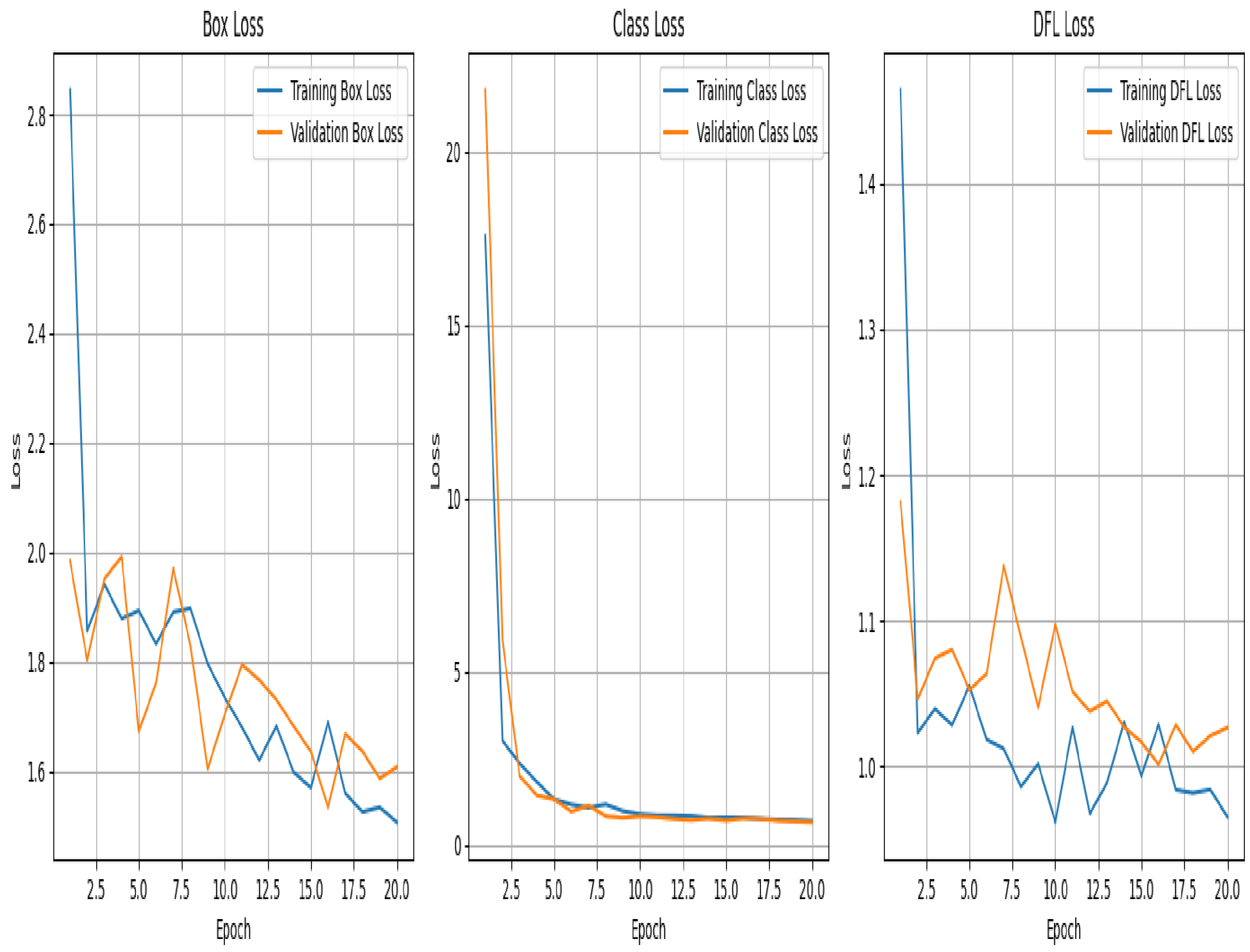


True Label: good
Predicted Label: good

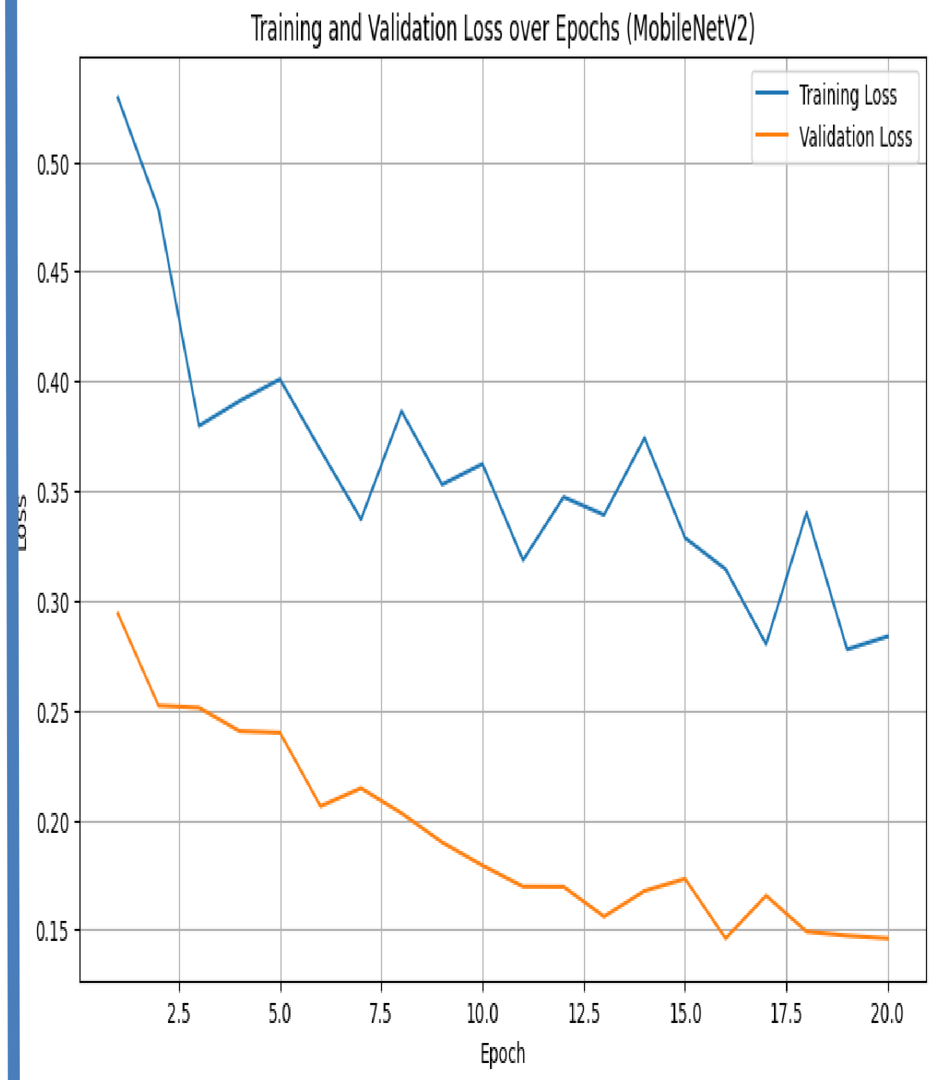
Results from YOLOv8

Results from MobileNetV2

Loss Curves

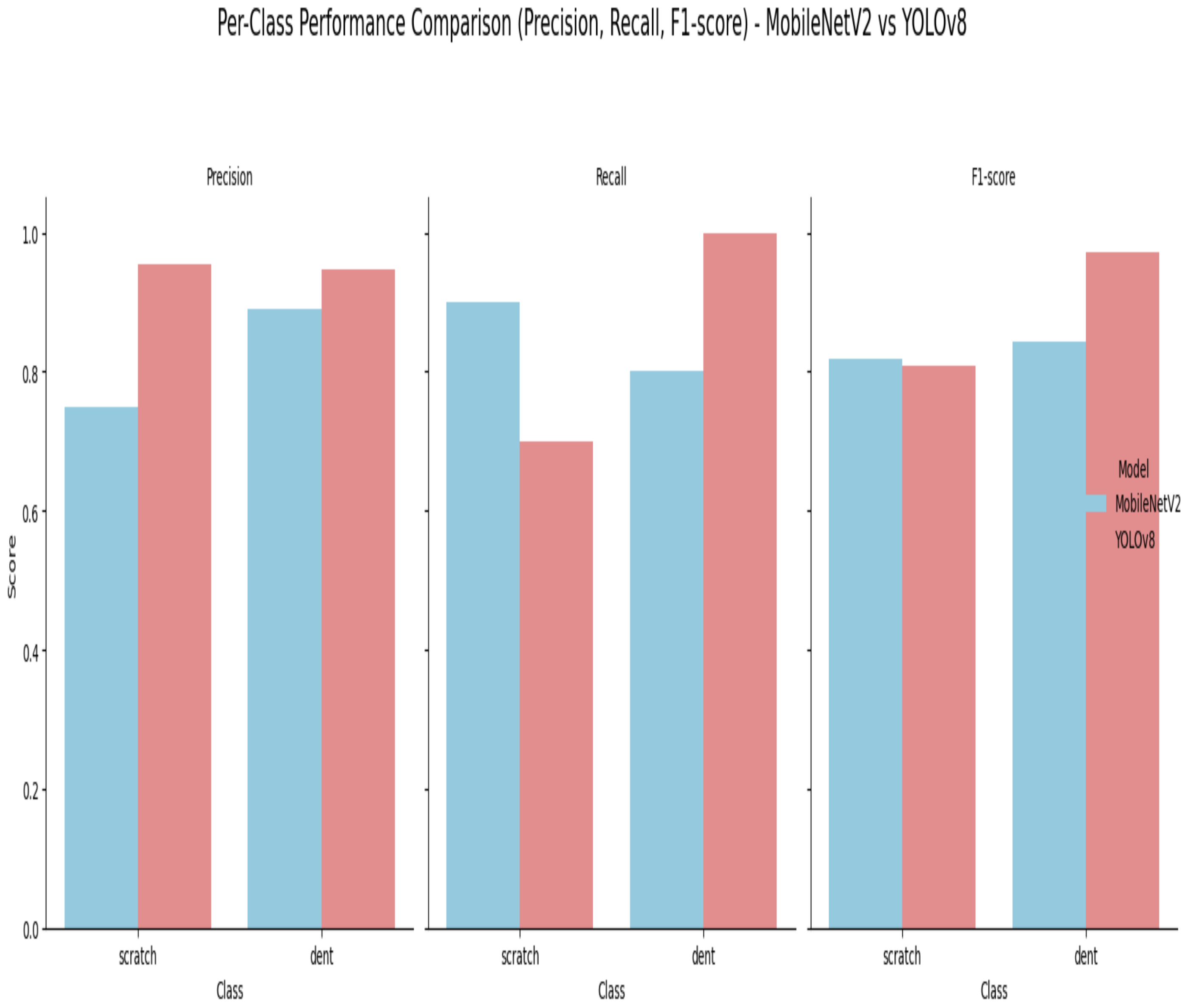


Loss curve for YOLOv8



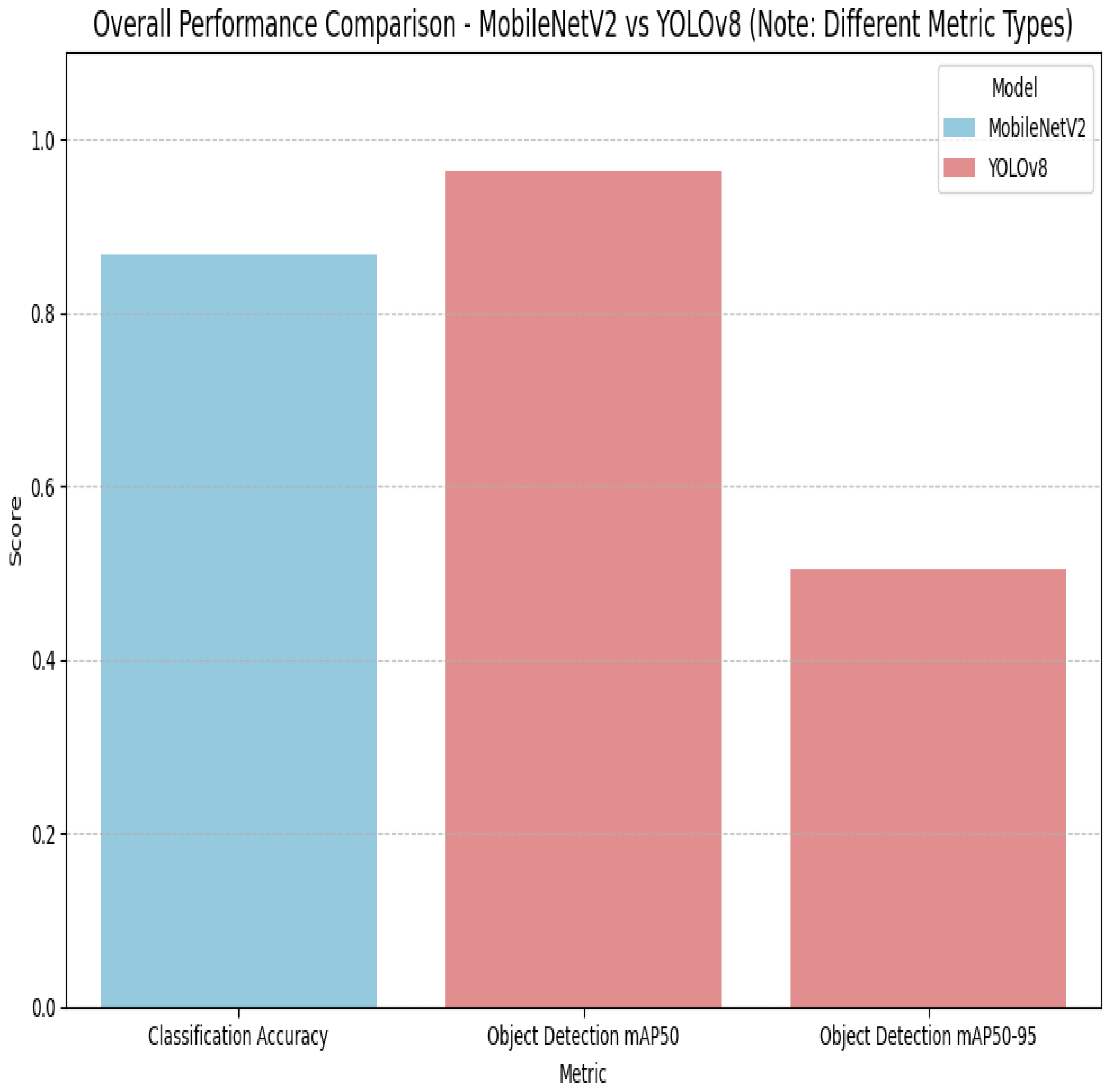
Loss curve for MobileNetV2

Model Performance Comparison



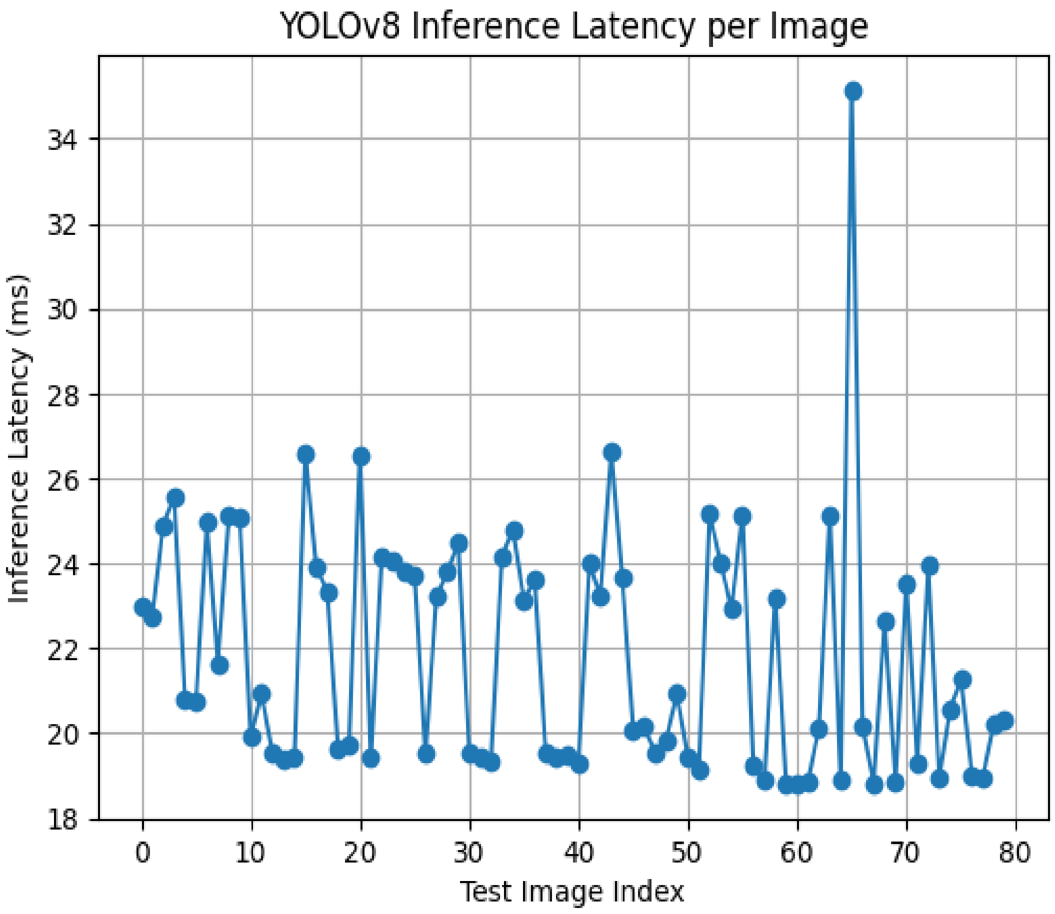
Model	Class	Precision	Recall	F1-score
Mobile NetV2	Scratch	.75000	0.9	.81818
Mobile NetV2	Dent	.88888	0.8	.84210
YOLOv8	Scratch	.95400	0.7	.80749
YOLOv8	Dent	.94600	1.0	.97225

Model Performance Comparison



Model	MobileNetV2	YOLOv8
Test Accuracy	.86666	NaN
mAP50	NaN	.964
MAP50-95	NaN	.505

YOLOv8 Inference Latency

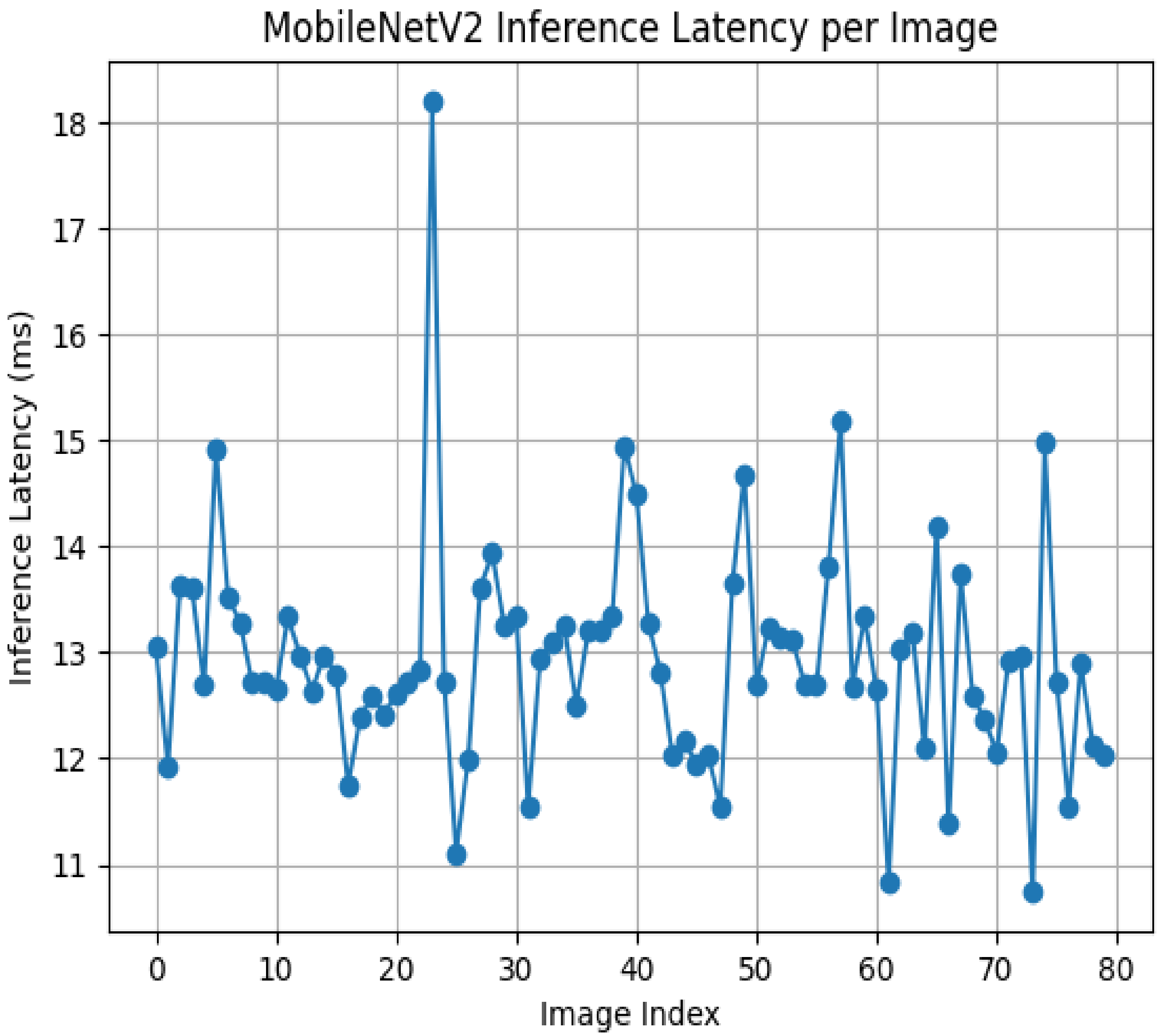


- Latency measured on GPU (Tesla T4, Colab) with warm-up runs, then multiple timed inferences.

Result :

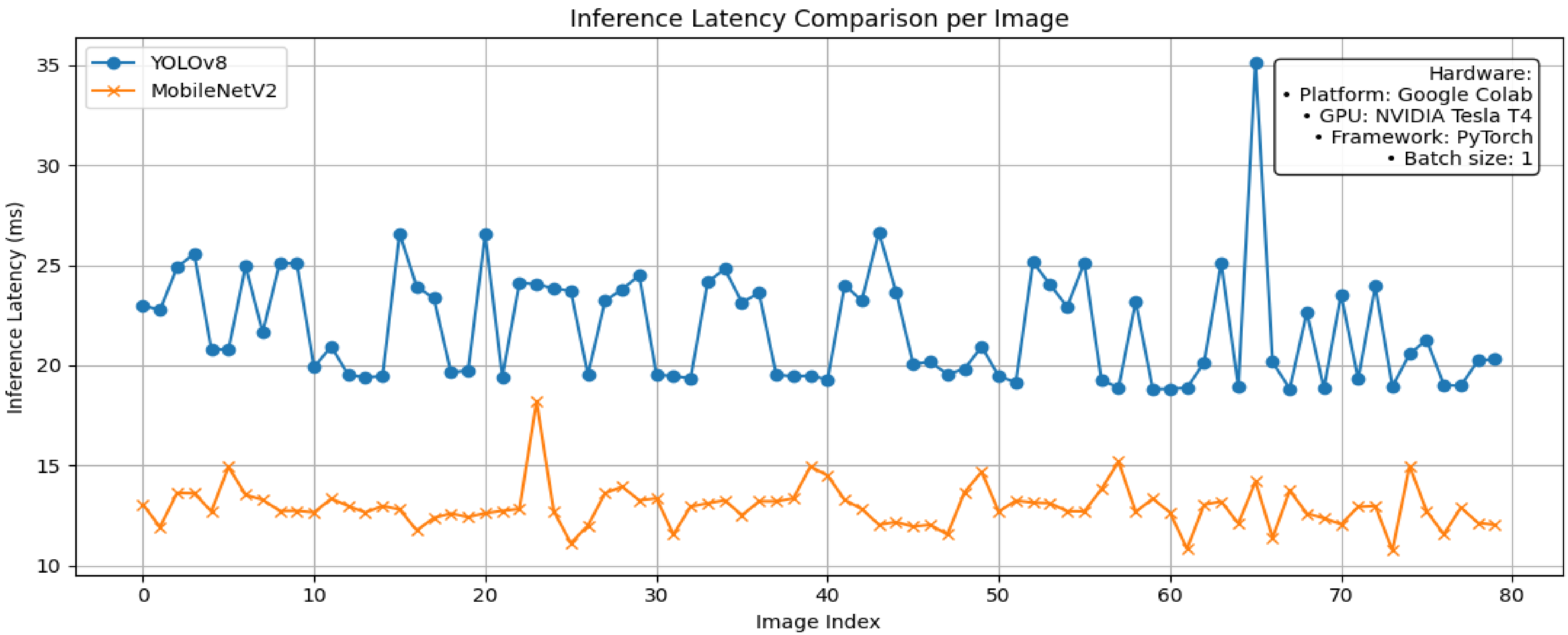
- Avg Latency: **21.87 ms** ;
- Std Latency: **2.83 ms**

MobileNetV2 Inference Latency



- Latency stays consistent across test images with small variations.
- Most predictions fall around **12–13 ms**, with rare spikes.
- Indicates efficient deployment potential for edge/production use.

Inference Latency Comparison



- MobileNetV2 is consistently faster (~12–13 ms/image) than YOLOv8 (~20–25 ms/image)
- Productivity of MobileNetV2: 80 images/sec, YOLOv8: 45 images/sec
- YOLOv8 shows higher variance and occasional spikes, due to detection + post-processing overhead.

Conclusion

- YOLOv8 achieves higher detection quality (mAP50 ≈ 0.964, mAP50–95 ≈ 0.505) and provides bounding-box localization.
- MobileNetV2 gives strong classification performance (accuracy ≈ 0.867) and is faster (~12–13 ms/image) than YOLOv8 (~22 ms/image).
- Trade-off: YOLOv8 is best when defect location matters; MobileNetV2 is best for fast screening.

Future Scope

- Validate models on real camera data
- Optimize inference latency for edge devices
- Extend to multi-view and multi-defect inspection
- Enable real-time integration with automated systems

Reference

Xiang Wan, Xiangyu Zhang and Lilan Liu, "An Improved VGG19 Transfer Learning Strip Steel Surface DefecRecognition Deep Neural Network Based on Few Samples and Imbalanced Datasets", https://doi.org/10.3390/app11062606

K. Khurana, R. Kumar, and S. Singh, “Molo: Hybrid model using mobilenetv2 and yolov8 for edge devices,” Discover Artificial Intelligence, vol. 5, no. 1, pp. 1–15, 2025

R. Yan, H. Liu, Q. Zhao, and Y. Chen, “Stms-yolov5: A lightweight algorithm for gear surface defect detection,” Sensors, vol. 23, no. 12, p. 5581, 2023

H. Nguyen, “Mobilenetv2 + enhanced feature pyramid for fast object detection,” Journal of Theoretical and Applied Information Technology, vol. 98, no. 21, pp. 3341–3352, 2020

M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4510–4520

Thank You

https://github.com/Pallelayaswitha1/Integration-and-Comparison-of-vision-models-for-smart-inspection-cell/tree/main