Automated Railway Crack Detection Using Machine Learning: Analysis of Deep Learning Approaches

2024 IEEE 15th Annual Information, Technology and Mobile Communications Conference

24 October 2024

ANDREW D'ARMS, HWAPYEONG SONG, HUSNU S. NARMAN, NEVZAT C. YURTCU, PINGPING ZHU, AND AMMAR ALZARRAD

Outline

- Introduction
- Methodology
- Results
- Conclusion

Problem

- Finding cracks and gaps in railroad tracks is very time consuming when done manually
- Doing so automatically can be slow depending on implementation, so better solutions were needed

Previous Solutions

- Machine vision
- Convolutional neural networks
- YOLOv3
- RetinaNet
- FCN and U-Net
- YOLOv5

Limitations

- Detect unrelated defects, like scars
- Some models are slower and unable to perform real-time
- Labels are inconsistent, including box sizes
- Existing datasets too small in scope

Background

- YOLO, or You Only Look Once, is a real-time object detection model using a CNN known for speed and accuracy
- It does this by placing bounding boxes around relevant areas of the image depending on what classes it was trained on

Models

- YOLOv3 incorporated a feature pyramid network
- YOLOv5 started using Darknet in its backbone
- YOLOv6 focused on hardware optimizations
- YOLOv8 featured an anchor-free split head, offering better balance between accuracy and speed
- YOLOv9 improved accuracy using programmable gradient information and a generalized efficient layer aggregation network
- YOLOv10 incorporated non-maximum suppression and overall optimizations
- ResNet101 skip connections and residual blocks and is commonly used in the field, but slower

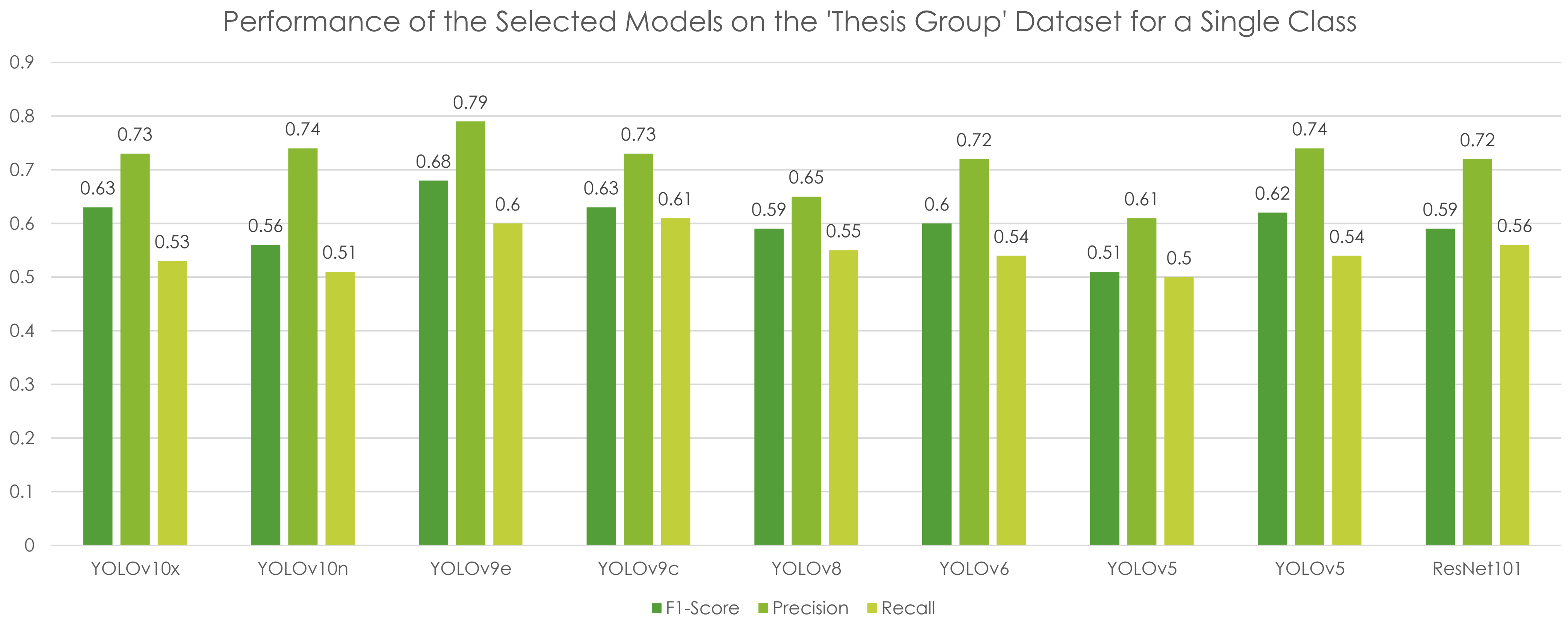
Methodology

- Nvidia Geforce RTX 3090 used for all testing (24GB of VRAM)
- Ultralytics package used for most testing, compiles most YOLO models into one Python package
- Original packages used for YOLOv5 and YOLOv9
- Precision, recall, and F1-scores recorded (harmonic mean of precision and recall)

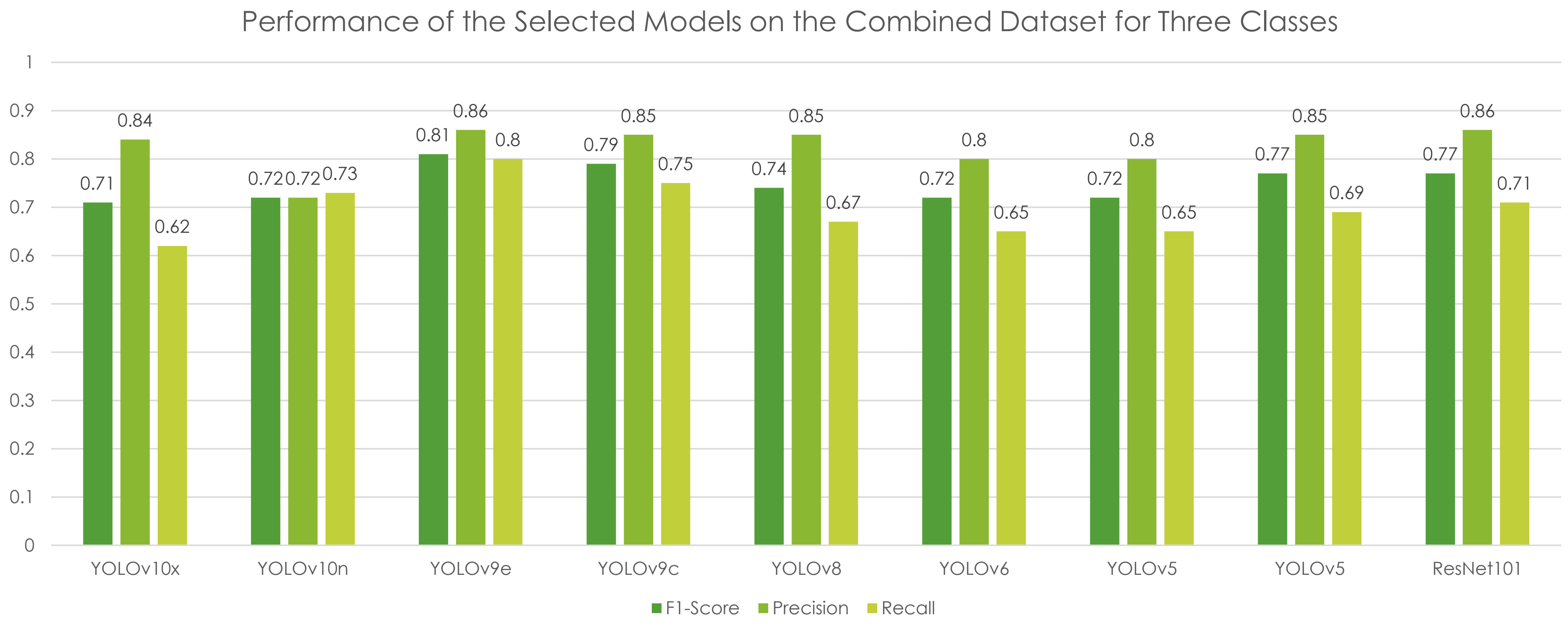
Datasets

- Single-class dataset made by 'Thesis Group' on Roboflow, ~1,000 images of a variety of cracks and gaps
- Combined dataset combined original and 5 other similar datasets
- Combined dataset includes 2,000 images, 5,000 after augmentation

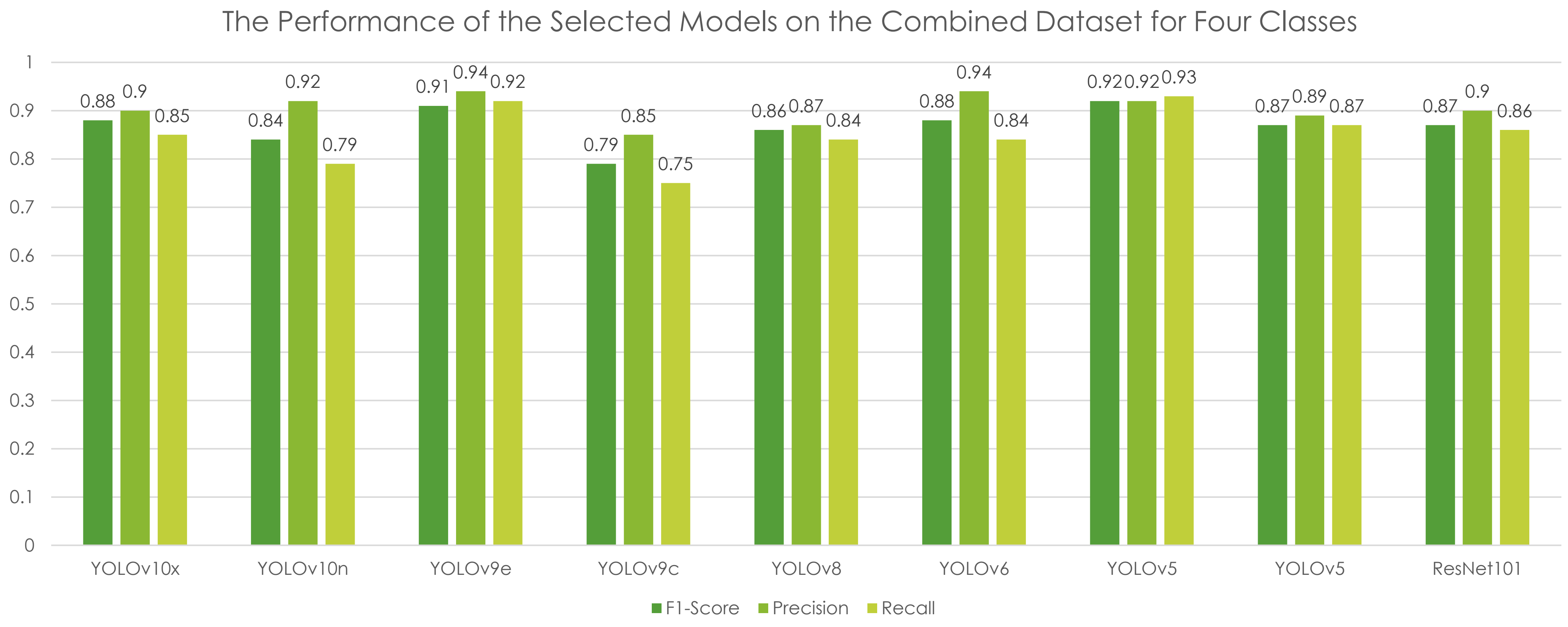
Single Class Results



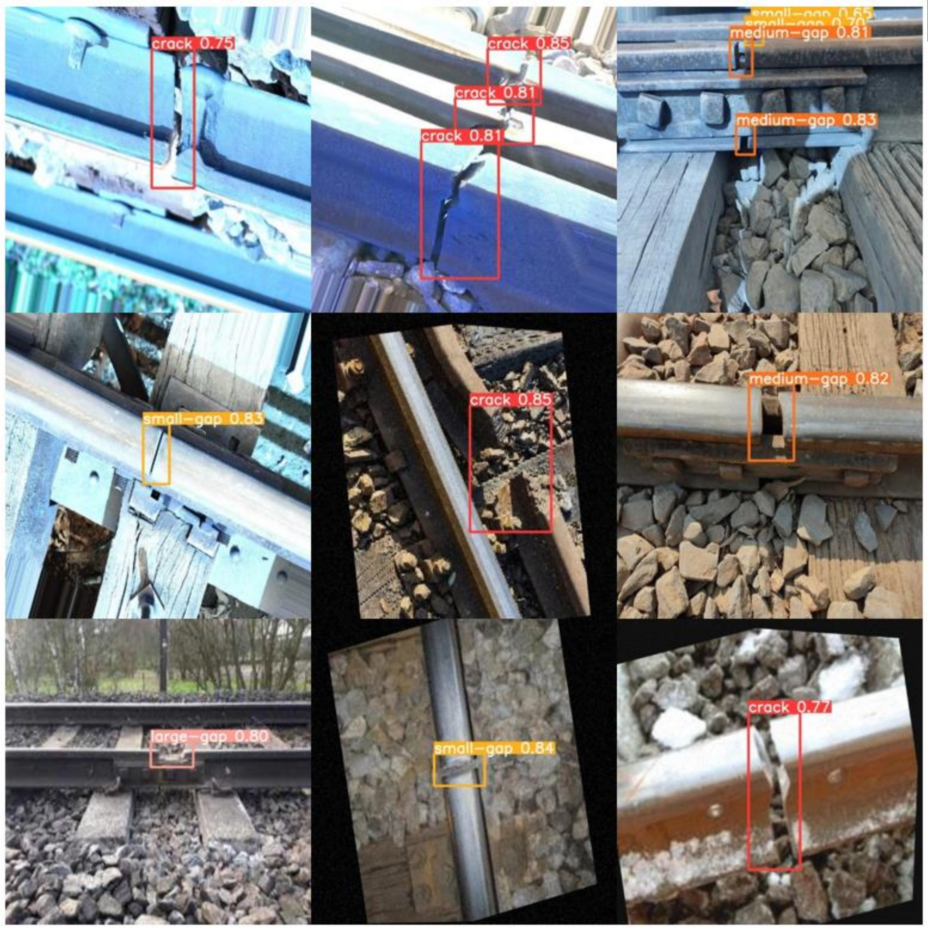
Three-Class Results

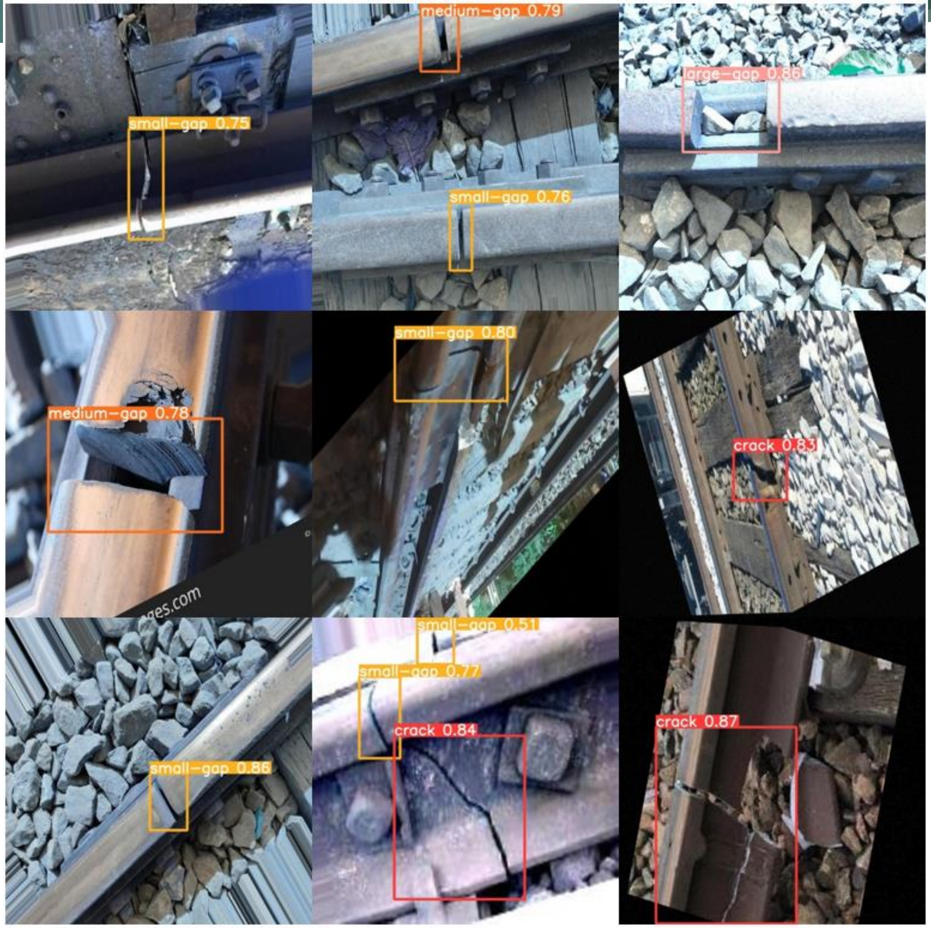


Four-Class Results

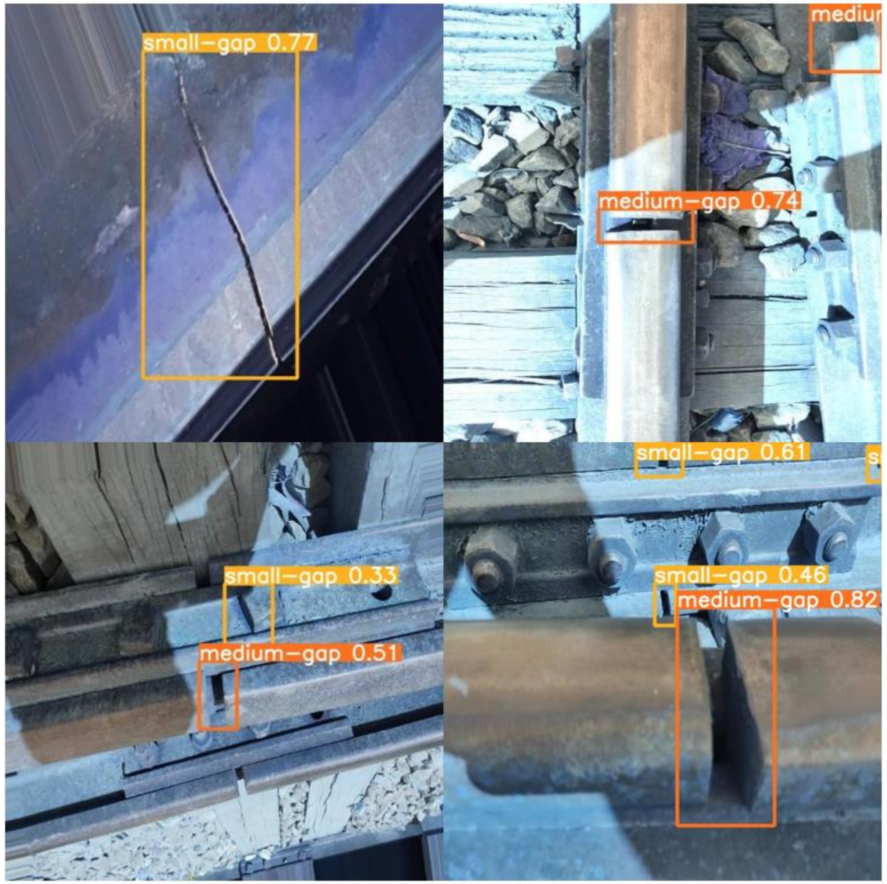


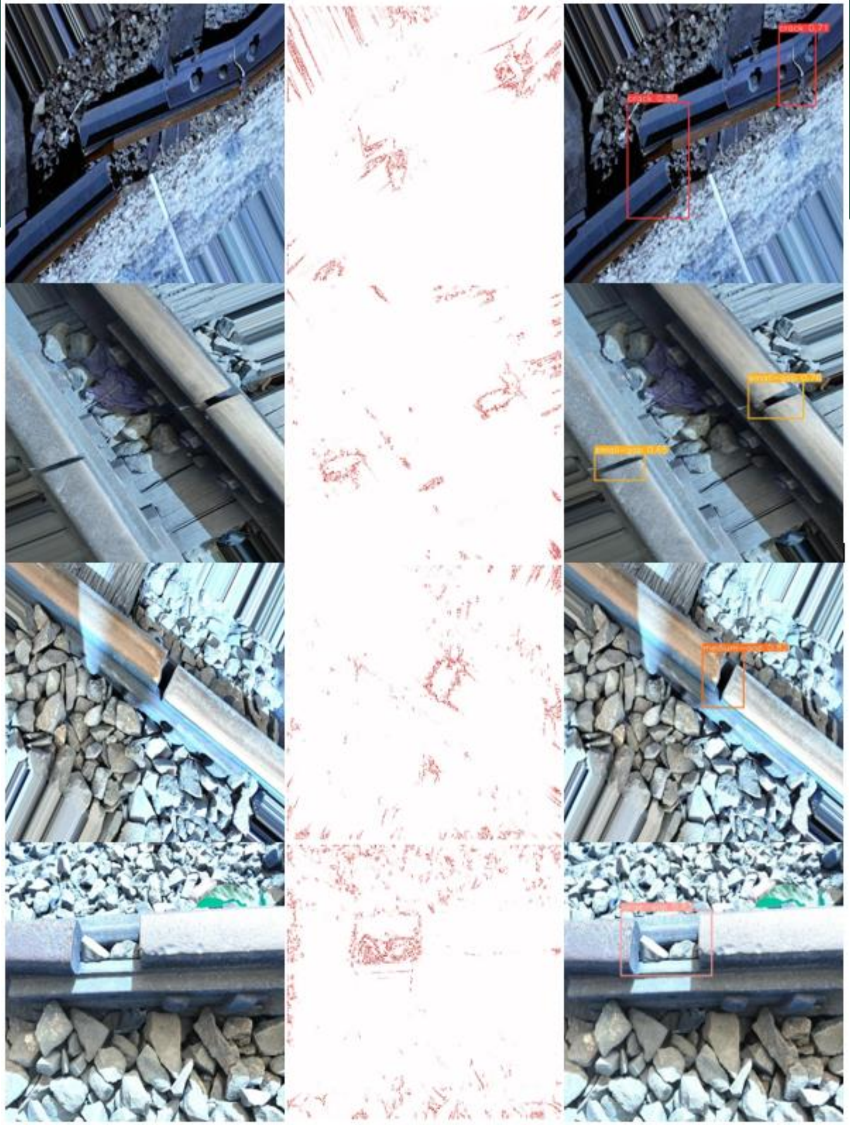
Visual Results





Visual Results





Conclusion

- Successfully able to detect cracks and gaps with haste
- YOLOv5 and YOLOv9 performed the best, with F1 scores of 0.92 and 0.91 respectively
- Datasets have largest impact on accuracy
- Explainable AI allows future research to be more effective

Future Works

- Field testing
- More realistic data
- Newer YOLO models
- Other machine learning models
- Model specialization

Questions?