Construction sites are inherently dynamic and complex, posing numerous safety risks and efficiency challenges. Traditional monitoring methods, primarily manual and error-prone, fail to ensure the safety and efficiency required. Our study proposes a cutting-edge solution using a fine-tuned YOLO v8-based computer vision model designed to accurately and efficiently detect and track construction vehicles like Excavators, Cement trucks, Dump trucks, and Wheel loaders in real-time, thereby enhancing safety protocols and operational workflows on construction sites.

### Introduction

- **Real-Time Detection:** YOLO v8s offers fast processing speeds crucial for the real-time detection of construction vehicles, a key requirement for monitoring dynamic construction sites.
- **Efficiency on Low-Power Devices:** YOLO v8s can perform well on low-power CPUs, deployable in various construction environments without the need for specialized computing resources.

### Why we choose Yolo v8s

<table>
<thead>
<tr>
<th>Model</th>
<th>size</th>
<th>mAP(%)</th>
<th>Speed</th>
<th>params</th>
<th>FLOPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yolov8 N</td>
<td>640</td>
<td>37.3</td>
<td>80.4</td>
<td>3.2</td>
<td>8.7</td>
</tr>
<tr>
<td>Yolov8 S</td>
<td>640</td>
<td>44.9</td>
<td>128.4</td>
<td>11.2</td>
<td>28.6</td>
</tr>
<tr>
<td>Yolov8 M</td>
<td>640</td>
<td>59.2</td>
<td>234.7</td>
<td>25.9</td>
<td>78.9</td>
</tr>
<tr>
<td>Yolov8 L</td>
<td>640</td>
<td>52.9</td>
<td>375.2</td>
<td>43.7</td>
<td>165.2</td>
</tr>
<tr>
<td>Yolov8 X</td>
<td>640</td>
<td>53.9</td>
<td>479.1</td>
<td>68.2</td>
<td>257.8</td>
</tr>
</tbody>
</table>

### The methodology

We have trained our model by creating a new training dataset consisting of 5470 images annotated for different construction equipment then applied different augmentation and optimization processes to expand our dataset by 300% using the following steps:

1. Gather training data for dataset creation
2. Annotate the gathered data to different classes
3. Divide the dataset into training, validation, and testing sets
4. Apply preprocessing algorithms to the dataset
5. Apply augmentation algorithms to the dataset
6. Training the model using the created dataset
7. Optimize the trained model using Openvino
8. Deploy the optimized model into the monitoring system

We optimized our model using Openvino by Intel. Which enabled us to increase the performance of our model by 97% while maintaining an average of 94.3% mAP across all classes. Openvino boosted the model’s processing speed and efficiency, enabling real-time detection of construction vehicles on low-power CPUs without compromising accuracy. This optimization ensures the model can be deployed effectively in diverse construction site environments, making advanced monitoring technology more accessible and practical for enhancing site safety and operational efficiency.

### Monitoring and tracking system

- Video captured using available monitoring cameras
- Start inference process for the coming video feedback
- Get the bounding boxes from the model then apply classification and tracking
- Yolo v8s trained model
- Classification and the tracking layer using ByteTrack within Yolo v8

### Openvino Optimization

![Openvino Optimization](image)

We expanded our dataset by augmentation and optimization processes to include 100 classes with more than 100,000 images which will include construction machinery, workers, tools, structural elements, and construction quality issues like cracks to expand our system monitoring and tracking capabilities to be used as a complete solution for automatic tracking construction sites and create real-time digital twins of any construction site.

### Conclusion

Our research shows the potential of using computer vision models to enhance the efficiency and safety of construction sites offering a highly effective solution to one of the main challenges of the construction industry and enabling the deployment of this solution to any construction site with a minimum cost of setup and with the highest accuracy and inferring performance possible.

**Future Research**

We would work on expanding our training dataset to include 100 classes with more than 100,000 images which will include construction machinery, workers, tools, structural elements, and construction quality issues like cracks to expand our system monitoring and tracking capabilities to be used as a complete solution for automatic tracking construction sites and create real-time digital twins of any construction site.