iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XII **Month of publication:** December 2024

DOI: https://doi.org/10.22214/ijraset.2024.65695

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XII Dec 2024- Available at www.ijraset.com

Automated Object Detection and Tracking for Construction Site Safety

Ms. Mamidi Sirisha
U.G. Student, Guru Nanak Institute of Technology Khanapur, Telangana

Abstract: ***The construction industry is increasingly adopting digital solutions to enhance safety, efficiency, and productivity. This project leverages the YOLOv8 object detection model and ByteTrack algorithm to track and count objects on construction sites. The system enables automated monitoring of personnel, machinery, and safety equipment through video analysis, addressing critical challenges like occlusions and dynamic object interactions. A custom dataset tailored for construction environments ensures high accuracy in detecting safety-critical items, such as personal protective equipment (PPE). This approach demonstrates significant potential for real-time safety assessments and operational optimization, serving as a technological leap in construction practices.***

Keywords: Object Tracking, YOLOv8, Byte Track Algorithm, Construction Site Monitoring, Personal Protective Equipment (PPE) Detection, Real-Time Video Analysis, Machine Learning in Construction, Computer Vision.

I. INTRODUCTION

As construction projects become more complex, traditional monitoring methods often struggle to deliver real-time insights and actionable data. This has created a pressing need for innovative solutions that leverage advanced technologies to address these challenges. The integration of computer vision techniques offers a transformative opportunity to revolutionize construction site management by automating processes and enhancing safety compliance, productivity, and operational efficiency. This project applies advanced object detection and tracking algorithms to address key challenges faced in dynamic construction environments, paving the way for digital transformation in the industry.

Construction sites are inherently dynamic and complex, with the constant movement of personnel, machinery, and equipment. Traditional monitoring methods, such as manual supervision or static CCTV systems, fall short in providing the real-time, actionable insights needed to ensure safety compliance and efficient resource utilization. Critical issues such as delayed response to safety hazards, inefficient workflows, and the inability to track and manage resources in real time hinder productivity. Addressing these limitations requires automated solutions capable of analyzing video feeds with speed, accuracy, and adaptability to handle dynamic and often unpredictable site conditions. To address these challenges, this project leverages advanced computer vision techniques to create a robust, automated monitoring system. The system is built around two key components: YOLOv8, a state-of-the-art object detection model, and Byte Track, a sophisticated tracking algorithm. YOLOv8 enables the system to identify personnel, machinery, and safety equipment in video feeds with exceptional accuracy and speed. Byte Track complements this by maintaining object identities across frames, even in the presence of occlusions or significant appearance changes, ensuring reliable tracking over time. These technologies work together to provide a seamless, real-time solution for monitoring construction sites.

The applications of this project are wide-ranging and directly address critical aspects of construction site management. For safety compliance, the system automates the detection of personal protective equipment (PPE) usage and hazardous situations in real time, enabling quick interventions to prevent accidents. For resource management, it tracks the movement and utilization of machinery and equipment, ensuring optimal allocation and reducing downtime. Additionally, the system provides insights into object counts and movement patterns, enabling workflow optimization and improving overall site efficiency. These applications not only enhance safety but also streamline operations, making construction projects more productive and cost-effective. This project demonstrates the integration of domain-specific knowledge with advanced technological tools, offering significant career advantages. By showcasing expertise in applying computer vision and AI to real-world problems, it positions professionals uniquely in emerging fields such as construction technology, smart infrastructure, and digital transformation. The skills developed, including proficiency in Python, YOLOv8, ByteTrack, and real-time video analysis, align closely with roles in AI development, data analysis, and software engineering. Moreover, the project underscores a problem-solving mindset, emphasizing the ability to address practical challenges through innovative solutions. This interdisciplinary approach not only enhances career prospects but also contributes to advancing construction practices through technology.

©IJRASET: All Rights are Reserved



Fig 1: Object Detection in Construction Sites

II. LITERATURE REVIEW

This chapter reviews existing research on construction site monitoring, highlighting advancements in computer vision, machine learning, and automation. It identifies gaps in real-time tracking and safety assessments, addressing challenges in dynamic construction environments.

Assaad et al. (2023) emphasize that while offsite construction methods offer numerous advantages, existing research primarily focuses on traditional construction techniques, leaving a gap in understanding the productivity factors specific to offsite settings. Recent studies have utilized comprehensive methodologies to identify and prioritize risk factors impacting labor productivity in these projects. Key challenges include inadequate workforce training, logistical issues, errors, and coordination failures. Stakeholders recognize the importance of effective mitigation strategies to address these challenges, with a consensus that improving workforce efficiency is essential for boosting overall productivity in offsite construction projects.

Li et al. (2024) highlight the importance of worker safety in the construction sector, emphasizing the need for improved monitoring and hazard detection methods. Advances in machine vision have led to the development of systems capable of identifying safety protection equipment worn by workers. However, challenges such as computational complexity and hardware limitations hinder the practical deployment of these systems. Research aims to enhance detection accuracy while maintaining efficiency, especially in Internet of Things (IoT) environments. Innovations in computer vision, such as YOLO (You Only Look Once), have enabled real-time detection of construction workers and equipment, improving site activity monitoring. Techniques like Convolutional Neural Networks (CNNs) and deep learning models have demonstrated success in classifying equipment and tracking productivity. These technologies automate manual tasks, reduce human error, and provide real-time insights into site operations, ultimately enhancing both productivity and safety in construction environments.

Upadhyay (2024) demonstrates the use of machine learning for productivity analysis in construction, particularly in equipment monitoring. Machine learning models are employed to detect idle times, analyze equipment usage, and optimize performance. These models automate traditional manual monitoring, improving both accuracy and efficiency. By predicting real-time equipment utilization, they enable better resource allocation and timely interventions to reduce delays. The research highlights the potential of machine learning to enhance productivity on construction sites, supporting improved decision-making and more streamlined operations. Such advancements provide valuable insights for improving operational efficiency and resource management in the construction industry.

©IJRASET: All Rights are Reserved

Nakanishi et al. (2022) discuss the growing integration of construction equipment monitoring into project management through digitalization and automation. Traditional methods are increasingly being replaced by advanced technologies that improve accuracy, safety, and efficiency. Key innovations such as computer vision, GPS, and geospatial tools are now used to monitor both static and dynamic elements on construction sites, including equipment, workers, and materials. A primary focus of their research is on improving safety through real-time equipment tracking while enhancing decision-making by integrating technologies like Building Information Modeling (BIM). These advancements aim to streamline operations and promote safer, more efficient construction site management.

In a related study, Kim et al. (2019) present a vision-based action recognition framework tailored to earthmoving excavators. Their research focuses on automating productivity assessments by analyzing sequential working patterns. The framework involves processes like excavator detection, tracking, and action recognition, highlighting the importance of modeling sequential patterns of visual features and operation cycles. They point out that many existing methods overlook sequential data, which negatively affects recognition performance. Through extensive experimentation, the study demonstrates the framework's ability to accurately recognize excavator actions, facilitating automation in cycle time analysis and productivity monitoring. This approach underscores the critical role of sequential pattern analysis in improving action recognition, offering valuable insights for better equipment management and operational efficiency. Both studies reflect the increasing trend of utilizing advanced technologies to optimize construction site operations and enhance safety and productivity monitoring.

Gap analysis

While existing research has explored the use of computer vision and machine learning in construction site monitoring, a significant gap exists in real-time, accurate tracking and safety assessments tailored to the dynamic nature of construction environments. Most studies focus on detecting objects in static settings or fail to address challenges like occlusions and interactions between objects. This project, utilizing the YOLOv8 model and ByteTrack algorithm, fills this gap by providing robust real-time detection and tracking of personnel, machinery, and PPE, even in complex, cluttered scenes. Furthermore, the creation of a custom dataset for construction sites ensures higher accuracy, directly improving safety monitoring and operational efficiency areas that previous approaches have not fully optimized for construction-specific environments.

III. METHODOLOGY

This chapter outlines the methodology used for real-time object detection and tracking on construction sites. It details the system architecture, including the integration of YOLOv8 for object detection and ByteTrack for tracking, ensuring accurate monitoring.

System Architecture

The system architecture as in Fig 2 integrates with two core components – YOLOv8 for Object Detection and ByteTrack for Object Tracking, to analyze video feeds in real time, enabling accurate detection and tracking of objects on construction sites.

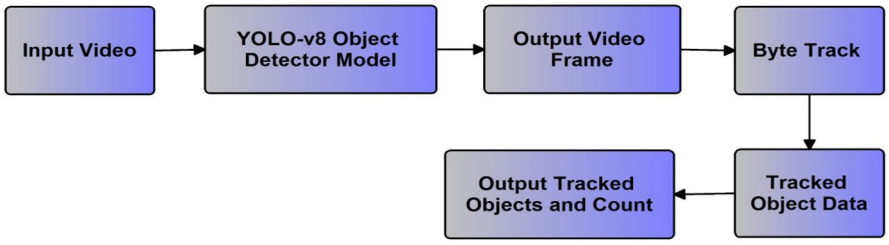


Fig 2: System Architecture

The process flow is as follows:

1) Input Video: A video feed from the construction site serves as the input to the system. This video captures dynamic scenes involving personnel, machinery, and equipment.
2) YOLOv8 Object Detector Model:
    - YOLOv8 processes each frame of the input video to detect multiple objects.
    - It generates bounding boxes and class probabilities for detected objects, enabling the system to identify categories like personnel, machinery, or safety equipment.
    - YOLOv8 uses CSPDarknet53 as its backbone for feature extraction and PANet for feature aggregation, ensuring high detection accuracy and speed suitable for real-time applications.

3) Output Video Frame: The processed video frame from YOLOv8, containing detected objects with bounding boxes and classifications, is passed to the next stage.

4) ByteTrack Tracking Algorithm:
- ByteTrack maintains the identity of detected objects across video frames.
- Its "tracking-by-detection" approach ensures reliable tracking, even under challenging conditions such as occlusions or significant appearance changes.
- This step ensures that objects are consistently tracked throughout the video feed, enabling continuous monitoring of movement and interactions.

5) Tracked Objects: The system outputs tracked objects with unique identities, allowing analysis of their behavior, movement, and interactions.

6) Output Tracked Objects and Count:
- The final output provides detailed information, including the count and movement patterns of tracked objects.
- This data can be used for safety compliance checks (e.g., PPE usage), resource management, and operational optimization.

This architecture is designed to handle the complexity of dynamic construction environments, providing a robust and automated solution for real-time monitoring and analysis.

IV. DESIGN AND ANALYSIS

This chapter focuses on the design and analysis of the object detection and tracking system for construction sites. It covers data collection, preprocessing, model training, hyperparameter tuning, and the integration of YOLOv8 and ByteTrack for real-time object detection, tracking, and safety monitoring.

A. Data Collection and Preprocessing

The dataset used for object detection and tracking on construction sites consists of video footage capturing various construction equipment and personnel. It includes 250 video instances featuring Backhoes, 100 video instances featuring Concrete Mixers, 200 video instances featuring Dump Trucks, 200 video instances featuring Excavators, 500 video instances featuring Man Power (personnel), and 200 video instances featuring Tractors. These videos were recorded under diverse real-world conditions, ensuring variability in lighting, weather, and site layouts. Each video was manually annotated to label the objects, enabling the YOLOv8 model to detect and track these objects accurately. The dataset's diversity in terms of conditions and object behaviors supports the model’s ability to perform robust tracking even in dynamic and challenging construction environments. Data preprocessing involves cleaning and augmenting video data by resizing frames, normalizing pixel values, and applying techniques like rotation and flipping to enhance model robustness against varying conditions on construction sites.

B. Model Training

YOLOv8 Object Detection: YOLO (You Only Look Once) is a real-time object detection system. YOLOv8 is the latest version, designed to offer high accuracy and speed, especially for dynamic environments like construction sites. The model is trained on the annotated dataset to detect objects such as personnel, machinery, and PPE.

Hyperparameter Tuning: Optimization of learning rate, batch size, and other model parameters to achieve optimal performance.

Loss Functions: The model uses multi-task loss functions, including classification loss, bounding box regression loss, and objectness loss, to improve detection accuracy.

C. Tracking with ByteTrack Algorithm

The ByteTrack algorithm is responsible for tracking objects across video frames by associating detections using motion prediction and object features. It effectively maintains the identity of each object over time, even in challenging scenarios such as occlusions or rapid movement. The algorithm utilizes object tracking techniques that predict the motion of objects between frames, ensuring that each object retains a unique identity throughout the video sequence. This capability is crucial in dynamic environments like construction sites, where multiple objects, such as personnel, machinery, and safety equipment, are in constant motion. ByteTrack's ability to handle these complexities ensures accurate and reliable tracking, contributing to real-time monitoring and safety assessment.

©IJRASET: All Rights are Reserved I SJ Impact Factor 7.538 I ISRA Journal Impact Factor 7.894 I

D. Real-Time Video Analysis

The system processes live video streams from construction sites. Each frame is analyzed in real time to detect and track objects, with a focus on ensuring low-latency performance for immediate intervention if safety issues are detected. The system can send alerts when PPE is not detected or when hazardous interactions (e.g., machinery moving too close to personnel) occur.

V. RESULT ANALYSIS

This chapter presents the results of the object detection and tracking system, including a confusion matrix, precision-confidence curve, and performance metrics. It analyzes the model's accuracy in classifying construction equipment and personnel, as well as its ability to improve detection and tracking performance through training and optimization.

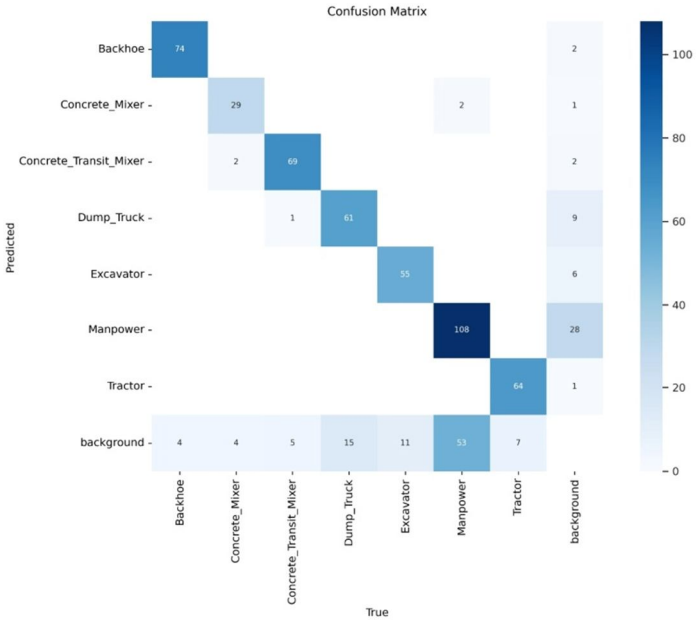


Fig 3: Confusion Matrix

The confusion matrix as shown in Fig 3 indicates that the model correctly classified 74 instances of backhoes, 29 instances of concrete mixers, 69 instances of concrete transit mixers, 61 instances of dump trucks, 55 instances of excavators, 108 instances of manpower, and 64 instances of tractors.

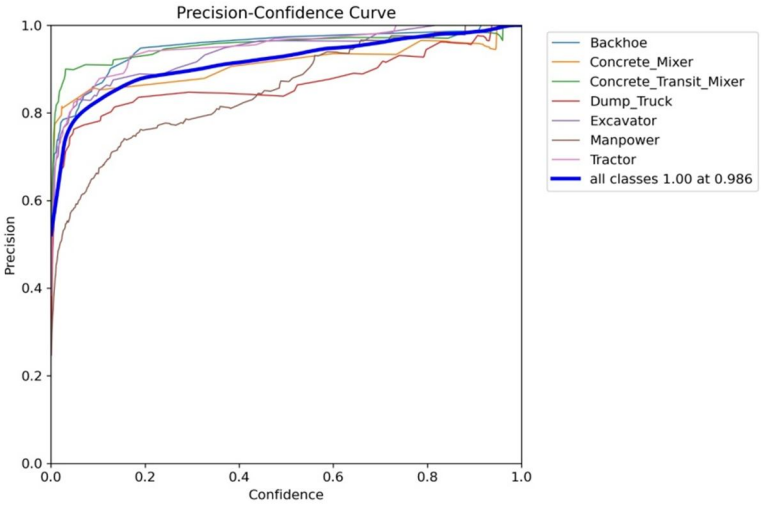


Fig 4: Precision curve

The precision-confidence curve illustrates the performance of the model across different object classes, showing how precision varies with confidence thresholds. The blue line, representing all classes combined, achieves perfect precision (1.00) at a confidence threshold of 0.986, indicating that the model is highly reliable at this threshold. Each individual class, such as backhoes, concrete mixers, dump trucks, and others, shows a similar trend where precision increases as the confidence threshold rises. However, the precision for each class varies slightly, with some classes reaching higher precision earlier than others. Overall, the curve demonstrates the model's strong ability to detect objects with high precision when confident, highlighting its effectiveness in construction site monitoring.

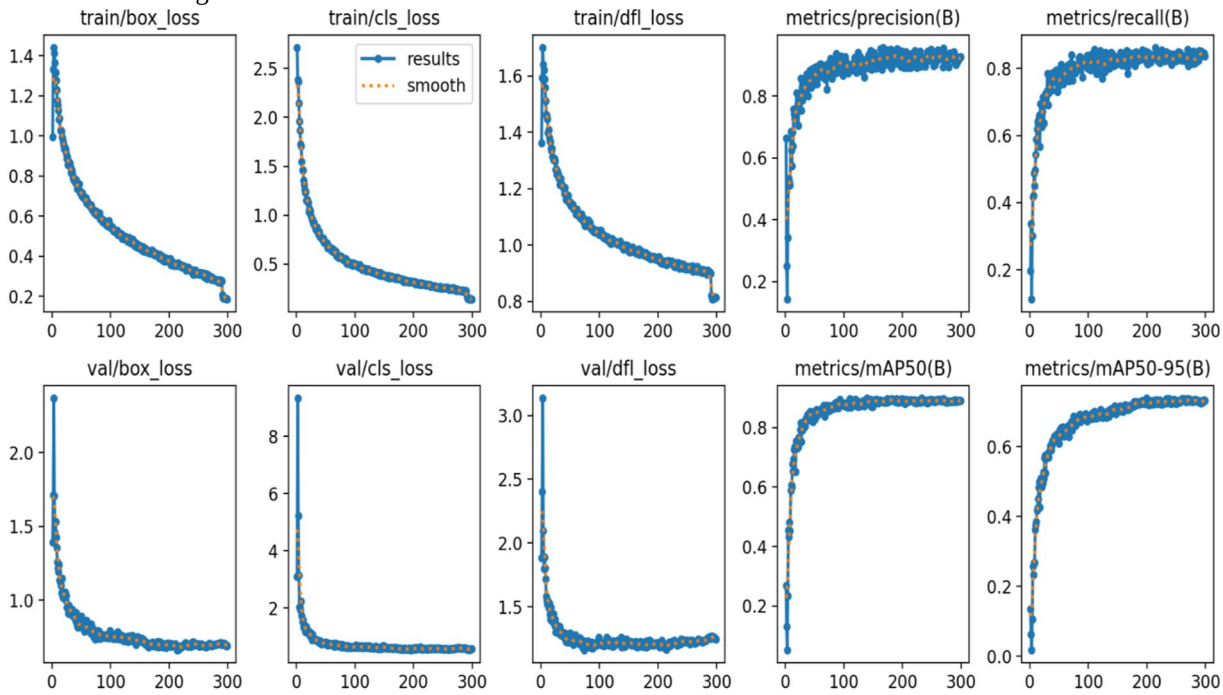


Fig 5: Metrics and Loss Curves

The Fig 5 presents key metrics and loss curves from the training of a deep learning model for object detection. The training loss curves for bounding box regression (train/box_loss), classification (train/cls_loss), and a combined loss function (train/dfl_loss) all show consistent downward trends, indicating the model's improving performance in predicting accurate bounding boxes and classifying objects. Performance metrics such as precision (metrics/precision(B)) and recall (metrics/recall(B)) show steady upward trends, reflecting improvements in the model's ability to correctly identify and detect objects. The mean Average Precision (mAP) at an IoU threshold of 0.5 (metrics/mAP50(B)) shows slight improvement, while the mAP50-95 metric, which considers a broader range of IoU thresholds, exhibits a more pronounced increase, indicating the model's enhanced ability to detect objects with higher overlap to ground truth boxes. Overall, the model is improving in both object detection accuracy and recall over training epochs.

VI. CONCLUSION & FUTURE WORK

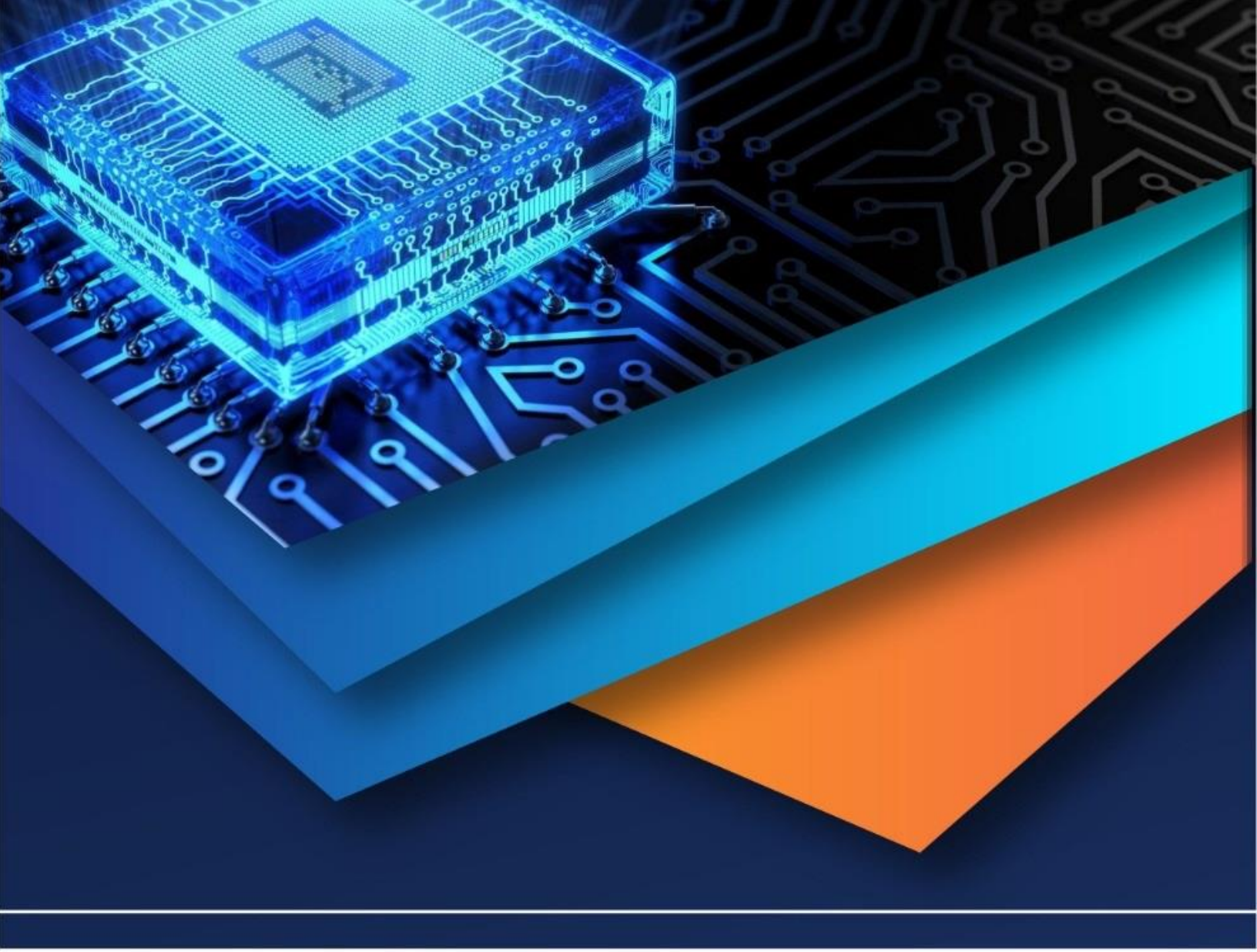
In conclusion, this project demonstrates the successful integration of YOLOv8 for real-time object detection and ByteTrack for object tracking to enhance safety and efficiency on construction sites. By leveraging advanced computer vision techniques, the system accurately identifies and tracks personnel, machinery, and safety equipment, ensuring real-time monitoring of construction environments. The results, including performance metrics like precision, recall, and mAP, highlight the model's effectiveness in detecting and tracking objects with high accuracy, even under challenging conditions such as occlusions and dynamic interactions. This automated system not only improves safety compliance by monitoring PPE usage and hazardous situations but also optimizes resource allocation and workflow efficiency. Future improvements may involve refining detection models, exploring additional tracking algorithms, and utilizing optimized platforms like TensorRT and OpenVINO for enhanced performance across different hardware systems. Overall, this project sets the foundation for advancing construction site management through AI-driven automation, offering significant potential for operational optimization and safety enhancement.

©IJRASET: All Rights are Reserved I SJ Impact Factor 7.538 I ISRA Journal Impact Factor 7.894 I

REFERENCES

[1] Li, Jiaqi, et al. "A Review of Computer Vision-Based Monitoring Approaches for Construction Workers' Work-Related Behaviors." IEEE Access (2024).
[2] Upadhyay, Bhavya, and Ananya Sankrityayan. "YOLO V8: An improved real-time detection of safety equipment in different lighting scenarios on construction sites." (2024).
[3] Assaad, Rayan H., et al. "Key factors affecting labor productivity in offsite construction projects." Journal of Construction Engineering and Management 149.1 (2023): 04022158.
[4] Nakanishi, Yutaro, Takashi Kaneta, and Sayaka Nishino. "A review of monitoring construction equipment in support of construction project management." Frontiers in Built Environment 7 (2022): 632593.
[5] Kim, Jinwoo, and Seokho Chi. "Action recognition of earthmoving excavators based on sequential pattern analysis of visual features and operation cycles." Automation in Construction 104 (2019): 255-264.

©IJRASET: All Rights are Reserved I SJ Impact Factor 7.538 I ISRA Journal Impact Factor 7.894 I



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)