Optimizing Automotive Inventory Management: Harnessing Drones and AI for Precision Solutions

Qian Zhang[1] [a], Dan Johnson[2], Mark Jensen[2], Connor Fitzgerald[2],
Daisy Clavijo Ramirez[2] and Mia Y. Wang[2] [b]

[1]*Department of Engineering, College of Charleston, SC, U.S.A.*
[2]*Department of Computer Science, College of Charleston, SC, U.S.A.*

fi

Keywords: Inventory Management, Deep Learning, UAV Drones, Computer Vision, Artificial Intelligence.

Abstract: Inventory errors within the automotive manufacturing industry pose significant challenges, incurring substantial financial costs and requiring extensive human labor resources. The inherent inaccuracies associated with traditional inventory management practices further exacerbate the issue. To tackle this complex problem, this paper explores the integration of cutting-edge technologies, including UAV (Unmanned Aerial Vehicle) drones, computer vision, and deep learning models, for monitoring inventory in parking lots adjacent to manufacturing plants and harbors before vehicle shipment. These technologies enable real-time, automated inventory tracking and management, offering a more accurate and efficient solution to the problem. Leveraging drones equipped with high-resolution cameras, the system captures real-time imagery of parked vehicles and their components, while deep learning models facilitate precise inventory analysis. This forward-looking approach not only mitigates the costs associated with inventory errors but also equips manufacturers with the agility to optimize their production processes, ensuring competitiveness within the automotive industry.

1 INTRODUCTION

Inventory errors in vehicle manufacturing pose a significant challenge for the industry, with far-reaching implications. Such errors can be highly costly, impacting financial resources, operational efficiency, and human labor. Mismanagement of critical parts and components often leads to production delays, material wastage, and increased operational costs. Given the reliance of vehicle assembly lines on just-in-time (JIT) production systems, inaccuracies in inventory tracking can disrupt entire supply chains, resulting in costly production stoppages and backlogs (Sharma and Gupta, 2020), (Kros et al., 2019). Moreover, manual inventory management, heavily dependent on human labor, is prone to errors, including miscounts and inaccuracies, exacerbating these challenges (Su et al., 2021). These inefficiencies often lead to resource misallocation, such as over-investment in unnecessary parts and delays caused by critical shortages (Khan and Yu, 2022).

Addressing inventory errors in vehicle manufacturing is not merely a financial concern but a strategic imperative. To mitigate these challenges, manufacturers are increasingly adopting modern technologies such as automation, RFID (Radio-Frequency Identification) systems, and advanced inventory management software (Ota et al., 2019). These solutions aim to enhance precision, streamline operations, and reduce costs associated with downtime, excess inventory, and labor inefficiencies. Accurate inventory management is critical for ensuring seamless vehicle production, sustaining profitability, and meeting customer expectations in a highly competitive automotive market.

To tackle the persistent challenges of inventory errors and enhance inventory management efficiency in the automotive industry, this paper proposes a real-time vehicle inventory system leveraging drones, computer vision, and deep learning. The proposed system focuses on monitoring inventory in parking lots outside manufacturing plants or harbors before vehicles are shipped. Drones equipped with high-resolution cameras are deployed to capture real-time aerial imagery of parked vehicles and their associated components. These images are analyzed using com-

[a] https://orcid.org/0000-0003-3166-4291
[b] https://orcid.org/0000-0003-2954-0855

Zhang, Q., Johnson, D., Jensen, M., Fitzgerald, C., Ramirez, D. C. and Wang, M. Y.
Optimizing Automotive Inventory Management: Harnessing Drones and AI for Precision Solutions.
DOI: 10.5220/0013284900003890

In *Proceedings of the 17th International Conference on Agents and Artificial Intelligence (ICAART 2025) - Volume 3*, pages 1140-1145
ISBN: 978-989-758-737-5; ISSN: 2184-433X

puter vision algorithms, enabling precise recognition and tracking of vehicles and parts within the inventory (Ota et al., 2019). The incorporation of deep learning further enhances the system's capabilities, providing real-time communication, data analysis, and actionable insights. This innovative approach not only improves inventory accuracy but also supports better decision-making and resource allocation, paving the way for a more efficient and cost-effective inventory management process.

The remaining sections of the paper are organized as follows: Section 2 presents an overview of current solutions for vehicle detection employing drone and AI technologies. In Section 3, a comprehensive description of the proposed system is provided. Section 4 delves into the details of the experiment's design, execution, and outcomes. Finally, Section 5 brings the paper to a close, discussing its limitations and setting the stage for future work.

2 LITERATURE REVIEW

Efficient inventory management is crucial in modern manufacturing, particularly in the automotive industry, where just-in-time (JIT) production systems demand precise control of parts and components. Inventory errors can result in significant disruptions, such as production delays, resource misallocation, and increased operational costs (Sharma and Gupta, 2020)(Kros et al., 2019). Sharma and Gupta (Sharma and Gupta, 2020) emphasized the impact of manual tracking inefficiencies in JIT systems, advocating for the adoption of automated systems. Similarly, Kros et al. (Kros et al., 2019) conducted an empirical investigation into inventory accuracy in automotive manufacturing, concluding that technological innovations, including automated tools, are necessary to address persistent inaccuracies.

Modern technologies, such as Radio-Frequency Identification (RFID) and AI-based solutions, have emerged as effective strategies for overcoming inventory management challenges. Su et al. (Su et al., 2021) demonstrated the ability of RFID systems to reduce human error rates and provide real-time tracking of inventory, although these systems often require substantial initial investment and infrastructure adaptation. More recently, drones equipped with AI capabilities have gained attention for their potential to revolutionize inventory monitoring. Ota et al. (Ota et al., 2019) highlighted the advantages of using drones for automated vehicle detection and monitoring, emphasizing their ability to conduct real-time aerial surveillance and reduce reliance on manual labor.

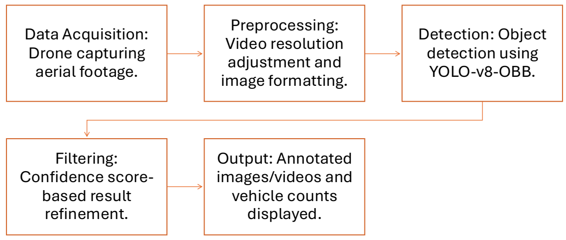

Figure 1: System Overview.

In the realm of vehicle detection and tracking, drones have been successfully employed in various applications. Bisio et al. (Bisio et al., 2021) conducted a comprehensive performance evaluation of leading deep learning (DL)-based object detection techniques, focusing on the RetinaNet framework within the context of the VisDrone-benchmark dataset. Their study provided critical insights into parameter optimization and model selection, setting a foundation for intelligent vehicle detection systems in smart cities. Wang et al. (Wang et al., 2016) introduced a UAV-based vehicle detection and tracking system designed for traffic data collection. Their system leveraged image registration, feature extraction, and tracking across consecutive UAV frames to dynamically detect and track vehicles with high accuracy. Similarly, Xiang et al. (Xiang et al., 2018) proposed a novel framework for vehicle counting using UAVs, integrating techniques like pixel-level foreground detection, image registration, and online-learning tracking to handle both static and dynamic backgrounds. Their results demonstrated over 90% accuracy in vehicle counting for fixed-background videos and 85% accuracy for dynamic ones, showcasing the efficacy of UAV-based solutions in traffic monitoring.

These advancements underscore the potential of drones combined with AI technologies for inventory management and vehicle monitoring. Deep learning models, such as those evaluated by Bisio et al. (Bisio et al., 2021), have proven highly effective in object detection, while UAV-based systems like those described by Wang et al. (Wang et al., 2016) and Xiang et al. (Xiang et al., 2018) demonstrate versatility in diverse scenarios. The proposed system builds on these developments by integrating drones, computer vision, and deep learning to improve inventory accuracy, resource allocation, and operational efficiency in the automotive industry.

3 METHODOLOGY

3.1 System Overview

The proposed system integrates unmanned aerial vehicles (UAVs), computer vision, and deep learning to provide an automated solution for vehicle inventory management in parking lots. Designed for environments such as manufacturing plants and harbors, the system offers real-time, accurate vehicle counting, minimizing errors and enhancing operational efficiency. The overveiew of the system is showing in Figure 1.

A DJI Mavic Pro drone, equipped with a high-resolution camera, captures aerial footage of parking lots. Its extended flight range and stability make it ideal for covering large areas efficiently. The captured footage is preprocessed to a standardized 720p resolution to ensure consistent input quality for subsequent analysis.

Vehicle detection is performed using the YOLO-v8-OBB (You Only Look Once Version 8 with Oriented Bounding Boxes) deep learning model. This advanced object detection approach is particularly effective for densely packed parking lots, as its oriented bounding boxes align with vehicle orientations, improving detection precision and reducing overlaps or false positives. To further enhance detection reliability, the system applies a confidence threshold of 0.8 (80%), processing only high-confidence detections.

The analysis pipeline is implemented using Python libraries, including OpenCV for image processing, NumPy for efficient numerical computations, and the Supervision library for managing detection results. Users interact with the system through a simple HTML interface built with Flask, where video or image files can be uploaded. The system processes the input and outputs annotated frames with bounding boxes, confidence scores, and vehicle counts in real time.

3.2 Dataset

The dataset used in this study consists of 381 images captured using a DJI Mavic Pro drone, which was flown by the researchers over various parking lots, primarily located at grocery stores and apartment complexes. The drone was flown at altitudes ranging from 25 to 35 meters to ensure optimal coverage and image quality. Each image was manually annotated with oriented bounding boxes around vehicle instances, resulting in a total of 2,843 annotations. The annotation process followed specific visibility criteria, requiring at least 60% of a vehicle to be visible, with both windshields clearly discernible. Vehicles that were occluded by trees were excluded from the annotations to prevent misidentification of trees as cars. This ensures that only visible vehicles are included, contributing to the model's accuracy during training.

The dataset is divided into three subsets, following a standard 60/20/20 split: 741 images for training, 73 for validation, and 61 for testing. This partitioning allows for a comprehensive evaluation of the model's performance on unseen data and helps mitigate the risk of overfitting. Before training, the images undergo several preprocessing steps to standardize and enhance their quality. The images are auto-oriented to maintain consistent orientation, then resized to fit within a 640x640 pixel frame, with white borders added as necessary. To further enhance the dataset's diversity and improve the model's robustness, data augmentation techniques are applied. These include horizontal flipping, 90° clockwise and counter-clockwise rotations, cropping with a zoom variation between 0% and 10%, saturation adjustments within a range of -21% to +21%, and the introduction of noise in up to 0.14% of the pixels. Each training image is augmented to generate three variations, thereby expanding the training set and increasing the diversity of vehicle appearances and orientations the model is exposed to.

The main challenges during the creation of the dataset were determining an appropriate threshold for vehicle visibility and handling occlusions caused by trees. These issues were addressed through iterative refinement of annotation guidelines, ensuring that only clear, visible vehicles were included in the dataset. The diversity of parking lots, varying lighting conditions, and different vehicle orientations create a challenging and realistic environment for vehicle detection. These characteristics, along with the preprocessing steps, ensure that the model is well-equipped to generalize to new and unseen parking lot scenarios.

3.3 Neural Network Model

The system utilizes the YOLO-v8-OBB (You Only Look Once Version 8 with Oriented Bounding Boxes) model, which is particularly well-suited for this task due to its ability to detect vehicles with high precision. Unlike traditional models that use axis-aligned bounding boxes, YOLO-v8-OBB employs oriented bounding boxes that align with the orientation of each vehicle. This alignment allows the bounding boxes to more tightly enclose the vehicles, thereby reducing the number of overlapping boxes and minimizing false detections. Furthermore, oriented bound-

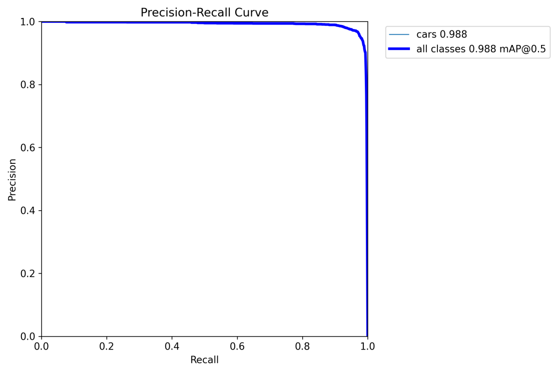

Figure 2: The Final Precision-Recall Curve for the YOLO-v8 OBB Model.

ing boxes help reduce noise during training by excluding background elements and irrelevant objects, which leads to more precise training and improved accuracy during inference. Another advantage of using YOLO-v8-OBB is its ability to efficiently utilize space, which is especially important in densely packed parking lots where vehicles are often parked closely together. This enables the model to accurately distinguish between adjacent vehicles, increasing detection accuracy in these challenging environments.

In addition to the benefits provided by the use of oriented bounding boxes, the YOLO-v8 architecture was selected for its efficiency and accuracy in vehicle detection. The YOLO-v8-OBB model processes video frames in real time, detecting vehicles frame by frame. The supervision library outputs bounding boxes around the detected vehicles with a confidence level greater than 80%. For each frame, the system displays a current count of detected vehicles with a confidence level below 80%. In the case of an image input, the model outputs a total count of detected vehicles. The deployment of the system is facilitated through a local web application that allows users to upload either video or image files for analysis. By leveraging advanced technologies and libraries, the proposed system provides an efficient solution for real-time inventory management, addressing the challenges of inventory errors in the automotive industry.

4 RESULTS AND DISCUSSION

4.1 Results

To evaluate the performance of the YOLO-v8-OBB model, a dataset consisting of aerial drone footage was collected (see Section 3), and various annotation policies were tested. The model was trained using the annotated data, and training metrics were recorded to

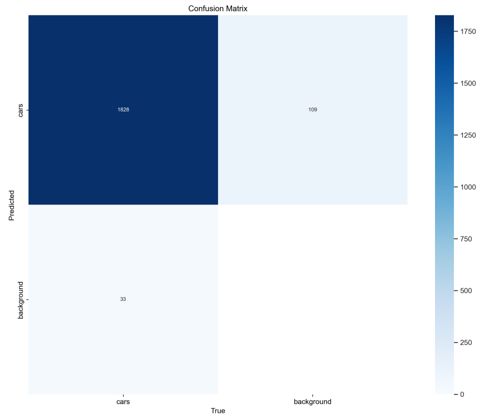

Figure 3: The Confusion Matrix for the final YOLO-v8 OBB model.

monitor its performance. After training, the model was evaluated on previously unseen footage to provide a qualitative and functional comparison.

The performance of the model was assessed using Average Precision (AP), a metric that combines both precision and recall. AP is computed during the evaluation phase and reflects the accuracy of the model in detecting a single class—passenger vehicles. This metric is closely related to Mean Average Precision (mAP), which is used when evaluating models trained on multiple classes, but in this case, only a single class is considered. The AP is derived from the precision and recall values, and is represented by the Area Under the Precision-Recall Curve (AUC), shown in Figure 2. During the validation phase, the model achieved an AP of 0.988, or 98.8%, indicating high detection accuracy.

Intersection over Union (IoU) is used to determine the overlap between predicted and ground truth bounding boxes, where a prediction is considered a true positive if its IoU exceeds a certain threshold. Precision and recall are then calculated, with precision being the ratio of true positives to the total number of predictions, and recall being the ratio of true positives to the total number of actual vehicles in the dataset. The confusion matrix, as shown in Figure 3, provides a more detailed breakdown of the model's classification performance, showing the true positive, false positive, and false negative counts.

4.2 Discussion

During the deployment of the YOLO-v8-OBB model, several challenges were encountered that impacted its performance. One significant issue was occlusion, where vehicles were partially obstructed by objects such as trees or equipment. While the model demonstrated the ability to detect some occluded vehicles,

Figure 4: Occluded vehicles captured by YOLO v8 OBB model.

Figure 5: A classification of a passenger vehicle next to a commercial truck.

occasionally outperforming human observers (Figure 4), it also produced misclassifications in certain scenarios. These results indicate that further refinement of the annotation policies, particularly for partially occluded vehicles, could improve the model's robustness and accuracy in such cases.

Another limitation observed was the edge-of-view problem. Vehicles entering the frame from the periphery were sometimes misclassified as smaller objects, revealing a weakness in the model when handling peripheral areas of the image. This issue could potentially be addressed by incorporating a more diverse set of footage from multiple angles, although this would require additional data collection, which was outside the scope of the current study.

Moreover, the model's performance was found to be sensitive to the altitude of the drone. Variations in drone height led to decreased confidence in predictions, suggesting that the model could benefit from either maintaining a more consistent altitude during data collection or utilizing multi-perspective data to mitigate the effects of height variation. Future research could explore strategies such as incorporating footage from different vantage points or adjusting for varying altitudes to enhance detection accuracy under diverse conditions.

Despite these challenges, the system demonstrated the capability to process video streams and make accurate real-time predictions, validating the feasibility of using the YOLO-v8-OBB model for vehicle detection in practical applications. The model effectively distinguished between passenger vehicles and other vehicle types, as illustrated in Figure 5, further reinforcing its potential for real-world deployment in automotive inventory management.

5 CONCLUSION

This paper presented an innovative approach to automotive inventory management using UAV drones, computer vision, and deep learning models, specifically the YOLO-v8-OBB model. The integration of these technologies offers a significant improvement over traditional inventory methods, addressing issues of inaccuracy while enhancing operational efficiency and reducing costs. The YOLO-v8-OBB model demonstrated high precision in vehicle detection, achieving an average precision (AP) of 98.8% during validation. Its real-time detection capability and effectiveness in complex parking environments make it a promising solution for automating inventory management in the automotive industry.

However, several challenges were identified during the implementation of the system. Issues such as occlusions, where vehicles were partially obstructed by other objects, and the edge-of-view problem, where vehicles entering from the periphery were difficult to classify, posed limitations. These challenges impacted the model's accuracy and highlighted areas for further improvement. Additionally, maintaining a consistent drone altitude was critical for optimal prediction confidence. Future work will focus on addressing these challenges by refining annotation policies, improving data processing methods, and exploring the use of alternative data sources or perspectives. Further research into edge computing solutions and the integration of more sophisticated models could also enhance the system's real-time performance and scalability.

REFERENCES

Bisio, I., Haleem, H., Garibotto, C., Lavagetto, F., and Sciarrone, A. (2021). Performance evaluation and analysis of drone-based vehicle detection techniques from deep learning perspective. *IEEE Internet of Things Journal*, 9(13):10920–10935.

Khan, M. and Yu, S. (2022). Inventory management challenges in the automotive sector. *Journal of Supply Chain Management*, 58(1):90–102.

Kros, J. F. et al. (2019). Inventory accuracy: An empirical investigation in automotive manufacturing. *Production and Operations Management*, 28(7):1736–1752.

Ota, J. et al. (2019). Drones and ai for automated vehicle detection and monitoring. *Robotics and Autonomous Systems*, 118:89–100.

Sharma, S. and Gupta, R. (2020). Impact of inventory errors in just-in-time manufacturing. *International Journal of Production Research*, 58(10):3115–3132.

Su, Q. et al. (2021). Reducing manual inventory errors using rfid technology. *IEEE Transactions on Industrial Informatics*, 17(4):2468–2477.

Wang, L., Chen, F., and Yin, H. (2016). Detecting and tracking vehicles in traffic by unmanned aerial vehicles. *Automation in construction*, 72:294–308.

Xiang, X., Zhai, M., Lv, N., and El Saddik, A. (2018). Vehicle counting based on vehicle detection and tracking from aerial videos. *Sensors*, 18(8):2560.