MANCHESTER
1824
The University of Manchester

Performance Analysis of a drone battery's position and pose estimation algorithm

A dissertation submitted to The University of Manchester for the degree of

Master of Science in Robotics

in the Faculty of Science and Engineering

Year of submission

2024

Student ID

10683124

School of Engineering

Contents

Word count: 3541

List of figures

Declaration of originality

I hereby confirm that this dissertation is my own original work unless referenced clearly to the contrary, and that no portion of the work referred to in the dissertation has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Intellectual property statement

1 Introduction

1.1 Background

Unmanned Aerial Vehicles (UAVs) have seen rapid development and widespread application in recent years, with uses ranging from cargo delivery to agricultural monitoring and pesticide spraying [1], [2]. However, the operational duration and range of UAVs are fundamentally limited by battery capacity [3]. To address this constraint, automated battery hot-swapping technology has emerged as a promising solution to extend the total operation time of UAVs in a single deployment, thereby improving efficiency and reducing manual labor requirements [4], [5].

Considerable research has been conducted on automated battery hot-swapping systems, focusing on both UAV design modifications and ground station configurations. For instance, Toksoz et al. [6] developed an automated battery swap and recharge system to enable persistent UAV missions. Similarly, Lee et al. [7] proposed an autonomous battery swapping system for quadcopters. Kemper et al. [8] explored the design concepts for automated service stations for UAV consumable replenishment. However, these existing approaches typically require UAVs to land precisely at specific locations and orientations, or on specially designed stations. This imposes high demands on the UAV's landing capabilities and makes the systems sensitive to environmental conditions such as strong winds or uneven terrain [9], [10].

1.2 Motivation

The motivation for this project stems from the limitations of existing battery exchange designs, which are largely centered around small drones with specific positioning requirements. These designs typically involve methods like funneling drones into precise positions or using frameworks to secure them for battery replacement [11]. However, such methods are impractical for larger drones due to their size and weight. Moreover, most of these studies have focused on small-scale UAVs, limiting their applicability to larger drones used in cargo transport or agricultural applications [12]. As the use of larger UAVs for commercial applications increases, there is a growing need for battery swapping solutions that can accommodate these larger vehicles [13].

For larger UAVs like the DJI FlyCart 30, a potential approach to automated battery hot-swapping involves using a robotic arm to replace the battery after the drone has landed on a designated pad near the base station. This method requires the robotic arm to accurately locate and determine the position and orientation of the battery on the UAV before attempting to grasp it [14]. Therefore, the primary research question is: How can we develop an accurate and robust algorithm for detecting and determining the position and orientation of batteries in UAV hot-swapping scenarios, particularly for larger drones like the DJI FlyCart 30?

1.3 Aims and Objectives

This study aims to develop and evaluate computer vision algorithms for accurate battery detection and positioning in UAV hot-swapping stations, with a focus on larger drones like the DJI FlyCart 30. The specific objectives are to:

1. Develop computer vision algorithms for battery detection in image feeds.
2. Implement position and orientation estimation techniques for detected batteries.
3. Evaluate algorithm performance using 3D-printed models.
4. Analyze the accuracy and robustness of the developed algorithms through statistical methods.

1.4 General Approach and Methods

The approach will involve creating 3D models of the FlyCart 30 battery top cover, preparing a comprehensive dataset, training YOLOv8-pose models, and conducting experiments using a RealSense depth camera. The performance of different models will be evaluated under various conditions to assess their effectiveness in real-world scenarios. Specifically, the methodology will include:

1. Model creation and dataset preparation using 3D-printed battery models.
2. Training of YOLOv8-pose models with different scales (n, s, m).
3. Experimental setup construction using a RealSense depth camera.
4. Data collection under various test conditions and statistical analysis of results.

1.5 Expected Outcomes

By addressing the challenges associated with battery hot-swapping for larger UAVs, this research aims to contribute to the development of more flexible and robust automated battery replacement systems. The expected outcomes include:

1. A robust computer vision algorithm capable of accurately detecting and determining the pose of UAV batteries.
2. Comparative analysis of different YOLOv8-pose model scales for this specific application.
3. Insights into the performance of the developed algorithms under various environmental conditions.

These outcomes have the potential to extend the operational capabilities of UAVs in diverse applications, particularly for larger drones used in commercial and industrial settings [15].

2 Literature Review

2.1 Evolution of Object Detection Techniques

Object detection has been a fundamental challenge in computer vision for decades, with applications ranging from robotics and autonomous driving to augmented reality. The field has seen significant advancements, particularly with the advent of deep learning techniques [16], [17].

Early approaches relied on hand-crafted features and sliding window techniques. Viola and Jones introduced a seminal work using Haar-like features and AdaBoost for real-time face detection [16]. This was followed by more sophisticated feature descriptors like Histogram of Oriented Gradients (HOG) [17], which proved effective for pedestrian detection.

The introduction of Convolutional Neural Networks (CNNs) revolutionized the field. R-CNN [18], followed by Fast R-CNN [19] and Faster R-CNN [20], significantly improved detection accuracy and speed. However, these two-stage detectors were still computationally intensive for real-time applications.

2.2 YOLO: A Paradigm Shift in Object Detection

The You Only Look Once (YOLO) algorithm, introduced by Redmon et al. [21], marked a paradigm shift by reformulating object detection as a regression problem. This unified, real-time object detection system processed images in a single forward pass through a neural network, achieving real-time detection speeds while maintaining competitive accuracy.

Since its inception, YOLO has undergone several iterations, each addressing limitations of its predecessors. YOLOv2 [22] and YOLOv3 [23] introduced improvements such as batch normalization, anchor boxes, and multi-scale predictions. YOLOv4 [24] further advanced the architecture by integrating state-of-the-art techniques like CSPNet and PANet, significantly boosting performance.

The release of YOLOv5 [25] marked a significant milestone in terms of engineering implementation and deployment flexibility. It offered a range of model sizes to cater to various application needs, from edge devices to high-performance servers. The most recent iteration, YOLOv8 [26], has further refined the architecture, loss functions, and training strategies, pushing the boundaries of detection accuracy and generalization capability.

2.3 Advancements in Pose Estimation

Pose estimation, often intertwined with object detection, has also seen significant progress. PoseCNN [27] introduced an end-to-end solution for 6D pose estimation from RGB images. More recent works like DPOD [28] have further improved accuracy by leveraging dense correspondence prediction.

The integration of multiple sensor modalities, particularly the combination of RGB and depth information, has been a growing trend. Methods like Frustum PointNets [29] effectively combine 2D object detection with 3D point cloud processing for accurate 3D object detection and pose estimation.

2.4 Applications and Real-world Impact

The YOLO family of algorithms has found applications across diverse domains, showcasing its versatility and effectiveness. In industrial settings, Song et al. [30] demonstrated an improved YOLOv5-based object detection method for grasping robots. In agriculture, Yan et al. [31] applied a modified YOLOv5 for real-time apple detection in picking robots.

In the medical field, YOLO variants have been employed for various diagnostic tasks. For instance, Meda et al. [32] applied YOLO-based models for identifying rickets on pediatric wrist radiographs, showcasing the algorithm's potential in assisting medical professionals.

2.5 Current Challenges and Future Directions

Despite these advancements, challenges remain in object detection and pose estimation. Occlusion handling, small object detection, and generalization to novel objects are active areas of research. Recent works like CenterNet [33] and FCOS [34] have proposed anchor-free approaches, aiming to address some of these challenges.

The field is also seeing a shift towards more efficient architectures, with methods like EfficientDet [35] optimizing the balance between accuracy and computational efficiency. Additionally, there's growing interest in self-supervised and weakly-supervised learning approaches to reduce the reliance on large annotated datasets.

2.6 Conclusion

In conclusion, while significant progress has been made in object detection and pose estimation, particularly with the YOLO family of algorithms, there's still room for improvement. Future research directions may focus on more efficient architectures, improved 3D understanding, and better integration of multiple sensor modalities. The continued evolution of these technologies promises to enable more intelligent and responsive computer vision systems across various domains, from autonomous vehicles to medical diagnostics and beyond.

3 Methodology

This study presents an approach for real-time position and pose estimation of drone battery using computer vision and machine learning techniques and evaluates its performance under various conditions. The methodology comprises several key components: experimental setup, data acquisition, pose estimation algorithm, performance metrics, and statistical analysis.

3.1 Experimental Setup

The experiment utilized an Intel RealSense depth camera to capture both color and depth information. The camera was configured to stream color images at a resolution of 640x480 pixels in BGR format and depth images at the same resolution in 16-bit format, both at 30 frames per second. The pyrealsense2 library was employed to interface with the camera and retrieve the image streams.

Two camera angles (45° and 90° between the camera's y-axis and the ground), three battery orientations (Angled, Direct, and Side), and three pre-trained models using YOLOv8n, YOLOv8s, and YOLOv8m, were used as variables in the experiment.

3.2 Data Acquisition

For each combination of conditions (camera angle, battery orientation, and algorithm model), 20 samples were collected, resulting in a total of 360 observations (2 angles × 3 orientations × 3 models × 20 samples). A custom function was developed to capture and process the data, extracting pose information from the color and depth frames.

3.3 Pose Estimation Algorithm

The pose estimation process consists of several steps:

3.3.1 Key-points Detection

A YOLOv8 model, pre-trained on a custom dataset of drone battery's top cover images, was used for key-point detection. The model was designed to identify five specific key-points on the drone battery: the four corners of the main body and the center point of a protruding feature. These key-points are crucial for subsequent pose estimation. The YOLO model processes each color frame and returns the detected key-points along with their confidence scores.

3.3.2 Pose Estimation

The pose estimation process utilizes the detected 2D key-points in conjunction with their corresponding 3D coordinates in the drone battery's body frame. The relationship between 2D image points (u, v) and 3D world points (X, Y, Z) is described by the pinhole camera model:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{1}$$

where (f_x, f_y) are the focal lengths and (c_x, c_y) is the principal point of the camera.

The pose estimation problem is solved using the Perspective-n-Point (PnP) algorithm implemented in OpenCV. This algorithm determines the rotation vector (**r**) and translation vector (**t**) that best align the 3D points with their 2D projections:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2}$$

where $[r_{ij}]$ represents the rotation matrix and $[t_x, t_y, t_z]$ the translation vector.

3.3.3 Data Filtering

To enhance the stability and accuracy of the pose estimates, a Kalman filter is implemented. The filter state vector **x** comprises six elements: three for position (x, y, z) and three for orientation (roll, pitch, yaw). The state transition model **F** and measurement model **H** are defined as identity matrices, assuming a constant velocity model:

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{w}_k \tag{3}$$

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k \tag{4}$$

where $\mathbf{w}_k$ and $\mathbf{v}_k$ represent process and measurement noise, respectively.

The filter predicts the next state based on the previous estimate and updates this prediction using the new measurements. This process helps to smooth out noisy measurements and provide more stable pose estimates.

3.4 Performance Metrics

To assess the system's performance, three primary metrics were utilized:

1. Mean Absolute Error (MAE): This metric measures the average magnitude of errors in a set of predictions, without considering their direction. It is calculated as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i| \tag{5}$$

   where n is the number of observations, y_i is the predicted value, and x_i is the true value.

2. Root Mean Square Error (RMSE): This metric calculates the standard deviation of the residuals (prediction errors). It is more sensitive to large errors than MAE and is computed as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \tag{6}$$

3. Standard Deviation (SD): This metric quantifies the amount of variation or dispersion in a set of values. It is calculated as:

$$\text{SD} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \tag{7}$$

   where μ is the mean of the values.

These metrics were calculated separately for position (X, Y, Z coordinates) and attitude (Roll, Pitch, Yaw angles) estimations.

3.5 Statistical Analysis

To determine the statistical significance of the effects of different factors on system performance, the Kruskal-Wallis H-test was employed. This non-parametric test assesses whether samples originate from the same distribution, making it suitable for comparing two or more independent samples of equal or different sizes.

The Kruskal-Wallis test was performed for each performance metric (MAE, RMSE, SD) across different levels of each factor (camera angle, battery orientation, algorithm model). The test statistic H is calculated as:

$$H = \left[\frac{12}{N(N+1)}\right]\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3(N+1) \tag{8}$$

where N is the total number of observations, R_i is the sum of ranks for group i, and n_i is the number of observations in group i.

The resulting p-value indicates the probability of obtaining the observed results if the null hypothesis (no difference between groups) is true. A p-value less than 0.05 was considered statistically significant, suggesting a genuine effect of the factor on the performance metric.

3.6 Data Visualization and Analysis

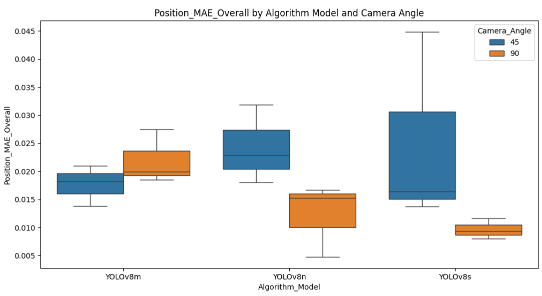
To facilitate the interpretation of results, several types of visualizations were generated:

1. Box plots: To display the distribution of performance metrics across different conditions.
2. Heatmaps: To visualize the performance metrics for various combinations of factors.

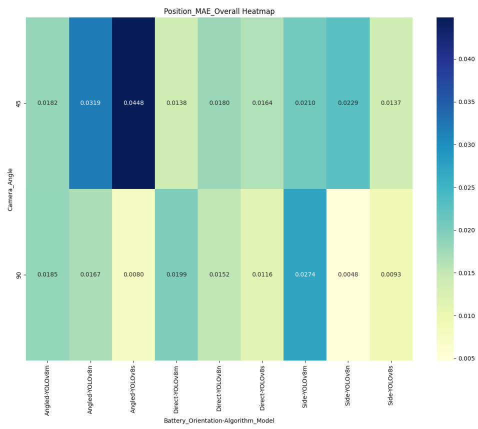
These visualizations were created using appropriate data visualization libraries in Python.

To identify the optimal settings for the pose estimation system, the performance metrics for all combinations of camera angle, battery orientation, and algorithm model were analyzed. The combination yielding the lowest MAE, RMSE, and SD values was considered the best for each specific metric.

This comprehensive methodology allowed for a thorough evaluation of the pose estimation system's performance under various conditions and identification of the most influential factors affecting its accuracy and consistency.



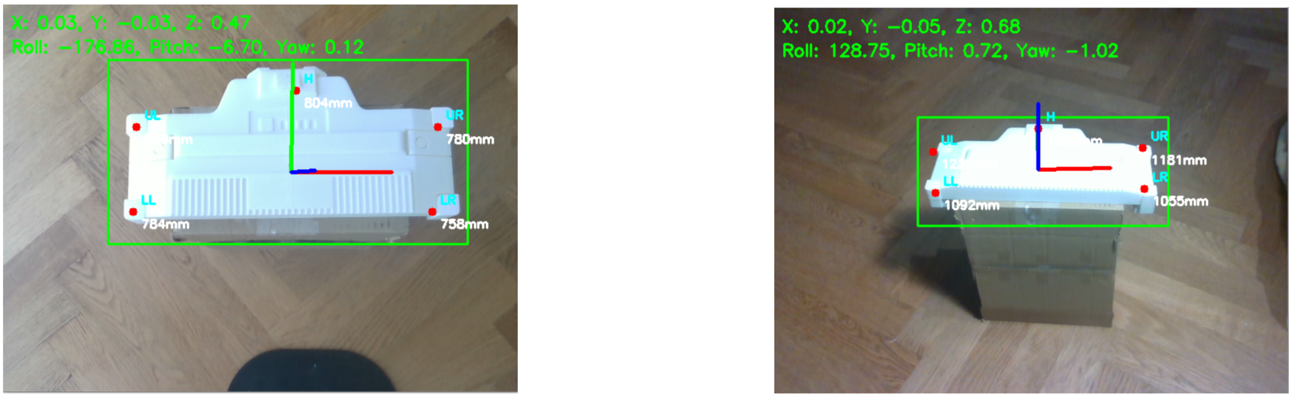
(a) box plot example



(b) heatmap example

Fig. 1. Data Visualization Examples

4 Results



(a) camera 90 degree

(b) camera 45 degree

Fig. 2. code running results

This study investigated the performance of the position and pose estimation system under various conditions, including camera angles, battery orientations, and algorithm models. The key findings are presented below.

4.1 Performance about Position

Position_MAE_Overall (Figures3 and4):

YOLOv8m shows similar performance at 90° and 45° angles, with moderate MAE values. YOLOv8n performs better at 90° than at 45°, with lower and more concentrated MAE values. YOLOv8s performs best at 90°, with the lowest and most stable MAE, but shows poorer and more unstable performance at 45°. The heatmap indicates that YOLOv8s performs worst in the ”Angled” direction at 45° (where MAE is highest), while at 90°, YOLOv8n performs best in the ”Side” direction (where MAE is lowest).

Position_RMSE_Overall (Figures5 and6):

The trend is similar to that of MAE but with more pronounced differences. YOLOv8s has the lowest RMSE at 90°, but the highest and most fluctuating RMSE at 45°. YOLOv8n shows excellent performance at 90°, with low and stable RMSE. The heatmap further confirms that YOLOv8s performs worst in the ”Angled” direction at 45°, while YOLOv8n performs best in the ”Side” direction at 90°.

Position_SD_Overall (Figure7):

YOLOv8m experiences significant SD fluctuation at 45° but is more stable at 90°. YOLOv8n has relatively low SD at both angles, with slightly better stability at 90°. YOLOv8s achieves extremely low and stable SD at 90°, but higher and less stable SD at 45°.

Summary:

Overall, a 90° camera angle generally outperforms 45°, especially in terms of MAE and RMSE. YOLOv8s performs best at 90°, but its performance drops significantly and becomes unstable at

45°. YOLOv8n remains stable across different conditions, particularly excelling at 90°. YOLOv8m shows moderate performance and remains relatively stable across different angles. The "Side" orientation of the battery tends to yield better results at 90°, especially for the YOLOv8n model, while the "Angled" orientation at 45° may lead to poorer performance, particularly for YOLOv8s. Stability of estimations is generally higher at 90° than at 45°, especially for the YOLOv8s model. These results suggest that a 90° camera angle usually provides more accurate and stable position estimation, with YOLOv8n performing well in most scenarios, and YOLOv8s excelling at 90° but not at 45°. The "Side" orientation appears to be more favorable for accurate position estimation, particularly at a 90° camera angle.

4.2 Performance about Pose/Attitude

Attitude_MAE_Overall (Figures8and 9)

YOLOv8m performs poorly at a 90° angle with high and fluctuating MAE values, but it shows stable performance at 45°. Similarly, YOLOv8n is unstable at 90°, with large MAE fluctuations, but it remains stable at 45°. In contrast, YOLOv8s is relatively stable at both angles, though its MAE is slightly higher at 90° compared to 45°. Generally, MAE is lower at 45° than at 90°. The heatmap indicates that, at 90°, YOLOv8m performs the worst (highest MAE) when directly facing the object.

Attitude_RMSE_Overall (Figures 10 and 11)

The RMSE trend is similar to MAE, but the differences are more pronounced. YOLOv8m's RMSE is significantly higher at 90°, especially when directly facing the object. YOLOv8s remains stable and has low RMSE at both angles. In general, RMSE is lower and less varied at 45° across all models.

Attitude_SD_Overall (Figure 12)

YOLOv8n shows the most fluctuation at 90°, indicating poor stability at this angle. YOLOv8s maintains low and stable SD at both angles. YOLOv8m has low SD at 45°, but significant fluctuations at 90°.

Summary

In summary, the 45° angle generally outperforms the 90° angle, particularly in terms of MAE and RMSE. Among the models, YOLOv8s is the most stable, performing well across different angles and metrics. YOLOv8m performs well at 45° but is unstable at 90°, especially when directly facing the object. YOLOv8n shows moderate performance, with poor stability at 90°. Direct facing at 90° tends to result in poorer performance, particularly for the YOLOv8m model. These findings suggest that camera angle, algorithm model, and object orientation are crucial factors in attitude estimation, requiring careful consideration to achieve optimal performance.

4.3 Algorithm Model Performance

While the Kruskal-Wallis test did not reveal statistically significant differences between the YOLO models ($p>0.05$ for all metrics), trends in the raw data suggest that each model has its strengths. YOLOv8n showed the best performance in position MAE and RMSE, particularly at the 90° angle. YOLOv8s excelled in attitude MAE and RMSE, especially at the 45° angle. YOLOv8m demonstrated the most consistent overall performance, having the smallest SD ranges for both position and attitude estimation.

5 Discussion

5.1 Camera Angle Effects

The performance of the 90° angle in position estimation suggests that a top-down view provides more accurate spatial information for determining an object's position. However, the better performance of the 45° angle in attitude estimation might be attributed to the improved visibility of the object's orientation when viewed from an angle, as opposed to directly from above. These findings underscore the importance of camera placement in optimizing system performance, with the choice potentially depending on whether position or attitude estimation is prioritized in a given application.

5.2 Model Performance and Selection

The varying strengths of different YOLO models highlight the importance of model selection based on specific application requirements. Applications prioritizing position accuracy might benefit from using YOLOv8n, while those focusing on attitude estimation could prefer YOLOv8s. In scenarios where consistency across different conditions is crucial, YOLOv8m might be the most suitable choice. These observations suggest that a one-size-fits-all approach may not be optimal for pose estimation tasks.

5.3 Battery Orientation Implications

The consistent superior performance of the Side orientation across different metrics suggests that this orientation might provide a more distinct or stable visual feature for the algorithms to detect and analyze. This finding has important implications for system design and object presentation in pose estimation tasks. It suggests that careful consideration of how objects are presented to the camera can significantly impact estimation accuracy.

5.4 Stability and Consistency Trade-offs

The stability analysis revealed trade-offs between accuracy and consistency among the models. While one model might provide the lowest error on average, another might offer more consistent results across different conditions. This information is crucial for applications where reliability and predictability of the system's performance are as important as raw accuracy. It suggests that the choice of model should consider not just average performance, but also performance consistency across different conditions.

5.5 Limitations

Several limitations of this study should be noted:

- The statistical tests showed few significant differences, which could be due to the limited sample size.
- The study focused on a specific set of conditions and may not generalize to all pose estimation scenarios.
- The trade-offs observed between position and attitude estimation accuracy warrant further investigation.

5.6 Future Work

Based on the findings and limitations of this study, several avenues for future research are proposed:

1. Increase the sample size to potentially reveal more significant effects and provide more robust statistical evidence.
2. Investigate the trade-offs between position and attitude estimation accuracy to inform system design decisions.
3. Explore the possibility of combining models or using ensemble methods to leverage the strengths of each model.
4. Extend the study to include more diverse environmental conditions and object types.
5. Develop adaptive systems that can dynamically adjust camera angles or processing algorithms based on the specific estimation task.
6. Investigate the impact of different lighting conditions and backgrounds on pose estimation performance.
7. Explore the integration of other sensor data (e.g., IMU) to complement vision-based pose estimation.

6 Conclusion

This study provides valuable insights into the performance of pose estimation systems under different conditions. The camera angle significantly influences system performance, with 90° angles favoring position estimation and 45° angles benefiting attitude estimation. Different YOLO models excel in various aspects: YOLOv8n for position accuracy, YOLOv8s for attitude estimation, and YOLOv8m for overall consistency. The battery orientation, particularly the ”Side” orientation, can significantly impact estimation accuracy. Clear trade-offs between accuracy and consistency among different models and conditions were observed. These findings underscore the importance of carefully considering camera angle, algorithm model, and object orientation in system design and deployment. By tailoring these factors to the specific requirements of an application, significant improvements in pose estimation accuracy and consistency can be achieved. Future research addressing the limitations of this study and exploring the proposed avenues could lead to more robust and versatile pose estimation systems capable of adapting to various conditions and applications.

References

[1] H. Shakhatreh, A. H. Sawalmeh, A. Al-Fuqaha, *et al.*, “Unmanned aerial vehicles (uavs): A survey on civil applications and key research challenges,” *IEEE Access*, vol. 7, pp. 48 572–48 634, 2019 (cited on p. 7).

[2] U. M. Mogili and B. Deepak, “Review on application of drone systems in precision agriculture,” *Procedia computer science*, vol. 133, pp. 502–509, 2018 (cited on p. 7).

[3] P. Rajendran and H. Smith, “Review of solar and battery power system development for solar-powered electric unmanned aerial vehicles,” *Advanced Materials Research*, vol. 1125, pp. 641–647, 2015 (cited on p. 7).

[4] N. K. Ure, G. Chowdhary, T. Toksoz, J. P. How, M. A. Vavrina, and J. Vian, “An automated battery management system to enable persistent missions with multiple aerial vehicles,” *IEEE/ASME Transactions on Mechatronics*, vol. 20, no. 1, pp. 275–286, 2015 (cited on p. 7).

[5] K. A. Suzuki, P. K. Filho, and J. R. Morrison, “Automatic battery replacement system for uavs: Analysis and design,” *Journal of Intelligent & Robotic Systems*, vol. 65, no. 1, pp. 563–586, 2012 (cited on p. 7).

[6] T. Toksoz, J. Redding, M. Michini, *et al.*, “Automated battery swap and recharge to enable persistent uav missions,” in *AIAA Infotech@ Aerospace Conference*, 2011, p. 1405 (cited on p. 7).

[7] D. Lee, J. Zhou, and W. T. Lin, “Autonomous battery swapping system for quadcopter,” in *2015 International Conference on Unmanned Aircraft Systems (ICUAS)*, IEEE, 2015, pp. 118–124 (cited on p. 7).

[8] F. P. Kemper, K. A. Suzuki, and J. R. Morrison, “Uav consumable replenishment: Design concepts for automated service stations,” *Journal of Intelligent & Robotic Systems*, vol. 61, no. 1, pp. 369–397, 2011 (cited on p. 7).

[9] H. Herath, H. Herath, S. Sumangala, O. de Silva, D. Chathuranga, and T. D. Lalitharatne, “Design and development of an automated battery swapping and charging station for multi-rotor aerial vehicles,” in *2017 17th International Conference on Control, Automation and Systems (ICCAS)*, IEEE, 2017, pp. 356–361 (cited on p. 7).

[10] K. A. Swieringa, C. B. Hanson, J. R. Richardson, *et al.*, “Autonomous battery swapping system for small-scale helicopters,” in *2010 IEEE International Conference on Robotics and Automation*, IEEE, 2010, pp. 3335–3340 (cited on p. 7).

[11] K. Fujii, K. Higuchi, and J. Rekimoto, “Endless flyer: A continuous flying drone with automatic battery replacement,” in *2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing*, IEEE, 2013, pp. 216–223 (cited on p. 7).

[12] Y. Mulgaonkar and V. Kumar, “Autonomous charging to enable long-endurance missions for small aerial robots,” in *Micro-and Nanotechnology Sensors, Systems, and Applications VI*, International Society for Optics and Photonics, vol. 9083, 2014, 90831S (cited on p. 7).

[13] A. B. Junaid, Y. Lee, and Y. Kim, “Design and implementation of autonomous wireless charging station for rotary-wing uavs,” *Aerospace Science and Technology*, vol. 54, pp. 253–266, 2016 (cited on p. 7).

[14] B. Michini, T. Toksoz, J. Redding, *et al.*, “Automated battery swap and recharge to enable persistent uav missions,” in *AIAA Infotech@ Aerospace Conference*, 2011, p. 1405 (cited on p. 7).

[15] P. R. Palafox, M. Garzón, J. Valente, and A. Barrientos, “Robust visual-aided autonomous takeoff, tracking, and landing of a small uav on a moving landing platform for life-long operation,” *Applied Sciences*, vol. 9, no. 13, p. 2661, 2019 (cited on p. 8).

[16] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, IEEE, vol. 1, 2001, pp. I–I (cited on p. 9).

[17] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR’05)*, IEEE, vol. 1, 2005, pp. 886–893 (cited on p. 9).

[18] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587 (cited on p. 9).

[19] R. Girshick, “Fast r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448 (cited on p. 9).

[20] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” in *Advances in neural information processing systems*, 2015, pp. 91–99 (cited on p. 9).

[21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788 (cited on p. 9).

[22] J. Redmon and A. Farhadi, “Yolo9000: Better, faster, stronger,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271 (cited on p. 9).

[23] J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” *arXiv preprint arXiv:1804.02767*, 2018 (cited on p. 9).

[24] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “Yolov4: Optimal speed and accuracy of object detection,” *arXiv preprint arXiv:2004.10934*, 2020 (cited on p. 9).

[25] G. Jocher *et al.*, *ultralytics/yolov5: v6.0 - YOLOv5n 'Nano' models, Roboflow integration, TensorFlow export, OpenCV DNN support*, version v6.0, Oct. 2021. DOI: `10.5281/zenodo.5563715`. [Online]. Available: `https://doi.org/10.5281/zenodo.5563715` (cited on p. 9).

[26] C. Dong and G. Du, “An enhanced real-time human pose estimation method based on modified yolov8 framework,” *Scientific Reports*, vol. 14, no. 1, p. 8012, 2024 (cited on p. 9).

[27] Y. Xiang, T. Schmidt, V. Narayanan, and D. Fox, “Posecnn: A convolutional neural network for 6d object pose estimation in cluttered scenes,” in *Robotics: Science and Systems*, 2018 (cited on p. 9).

[28] S. Zakharov, I. Shugurov, and S. Ilic, “Dpod: 6d pose object detector and refiner,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1941–1950 (cited on p. 9).

[29] C. R. Qi, W. Liu, C. Wu, H. Su, and L. J. Guibas, “Frustum pointnets for 3d object detection from rgb-d data,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 918–927 (cited on p. 10).

[30] Q. Song, S. Li, Q. Bai, *et al.*, “Object detection method for grasping robot based on improved yolov5,” *Micromachines*, vol. 12, no. 11, p. 1273, 2021 (cited on p. 10).

[31] B. Yan, P. Fan, X. Lei, Z. Liu, and F. Yang, “A real-time apple targets detection method for picking robot based on improved yolov5,” *Remote Sensing*, vol. 13, no. 9, p. 1619, 2021 (cited on p. 10).

[32] K. C. Meda, S. S. Milla, and B. S. Rostad, “Artificial intelligence research within reach: An object detection model to identify rickets on pediatric wrist radiographs,” *Pediatric Radiology*, vol. 51, no. 5, pp. 782–791, 2021 (cited on p. 10).

[33] X. Zhou, D. Wang, and P. Krähenbühl, “Objects as points,” *arXiv preprint arXiv:1904.07850*, 2019 (cited on p. 10).

[34] Z. Tian, C. Shen, H. Chen, and T. He, “Fcos: Fully convolutional one-stage object detection,” in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 9627–9636 (cited on p. 10).

[35] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10 781–10 790 (cited on p. 10).

Appendices

A Github link for codes

https://github.com/Jiahao-Geng/106833124.git

B List of Result charts

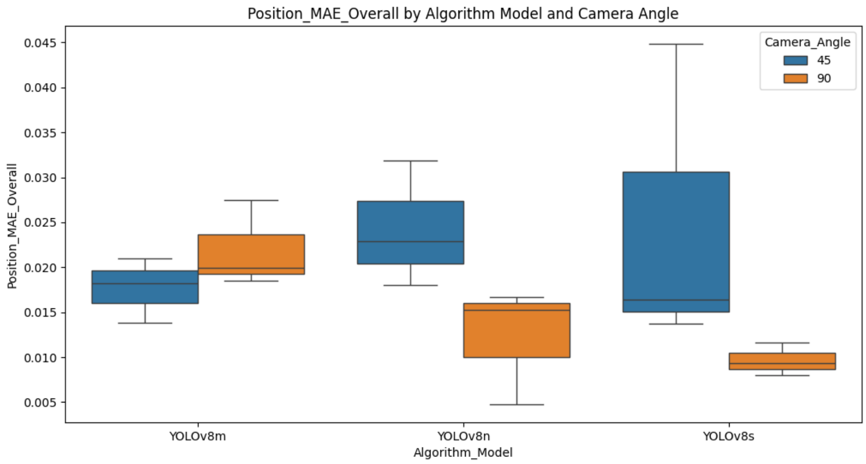


Fig. 3. position MAE overall boxplot

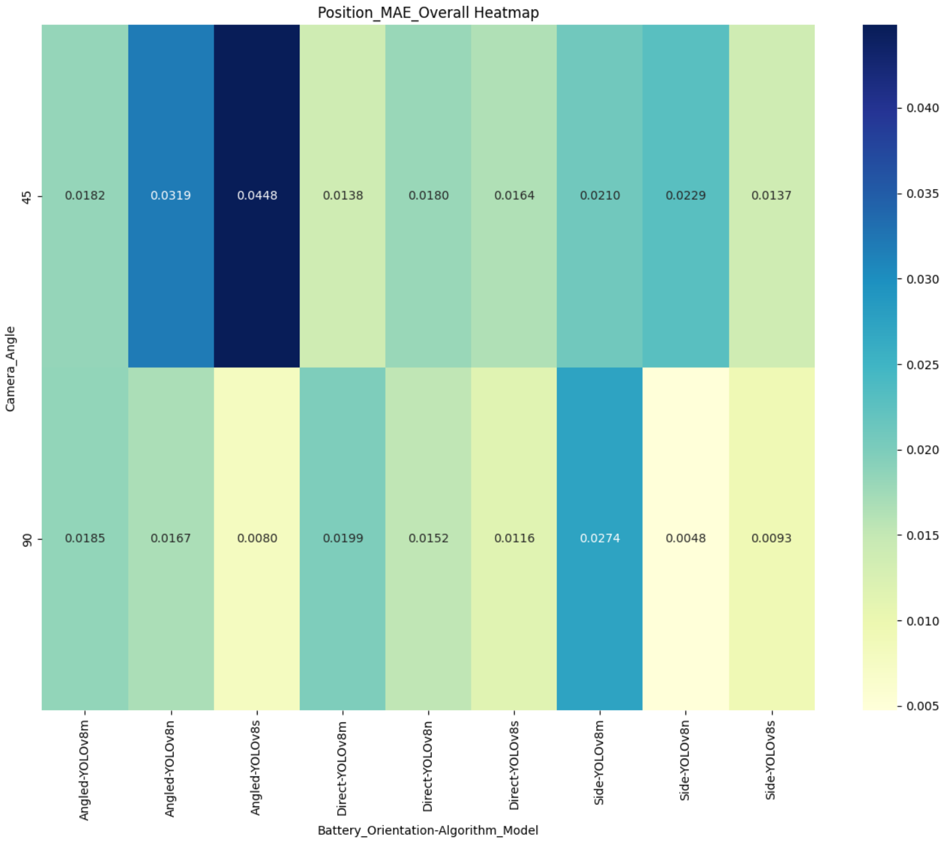


Fig. 4. position MAE overall heatmap

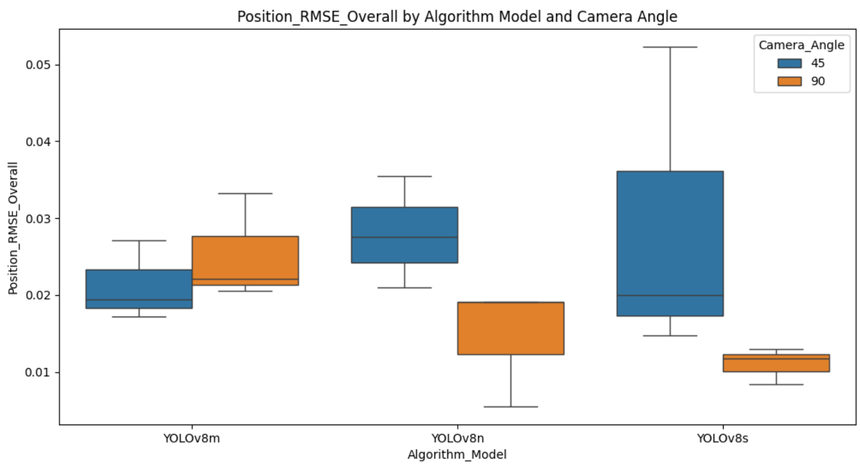


Fig. 5. position RMSE overall boxplot

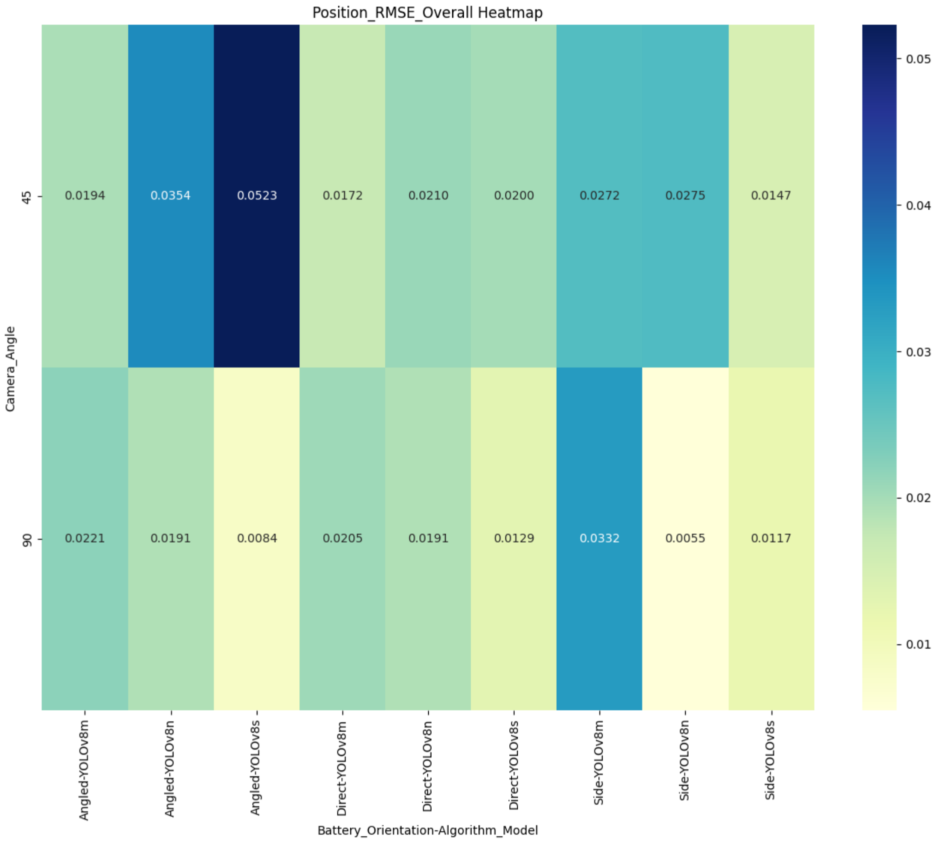


Fig. 6. position RMSE overall heatmap

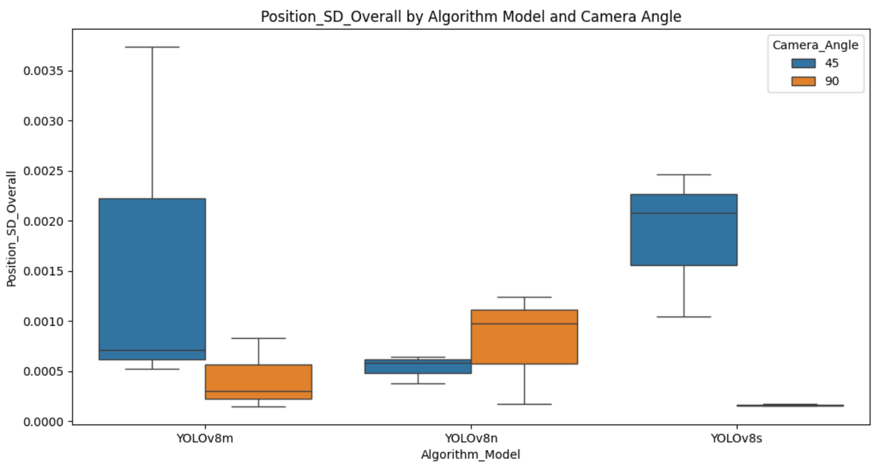


Fig. 7. position SD overall boxplot

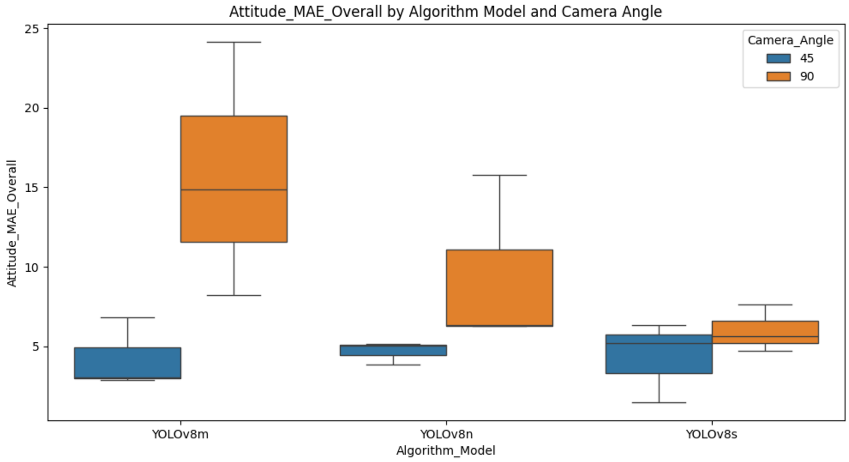


Fig. 8. attitude MAE overall boxplot

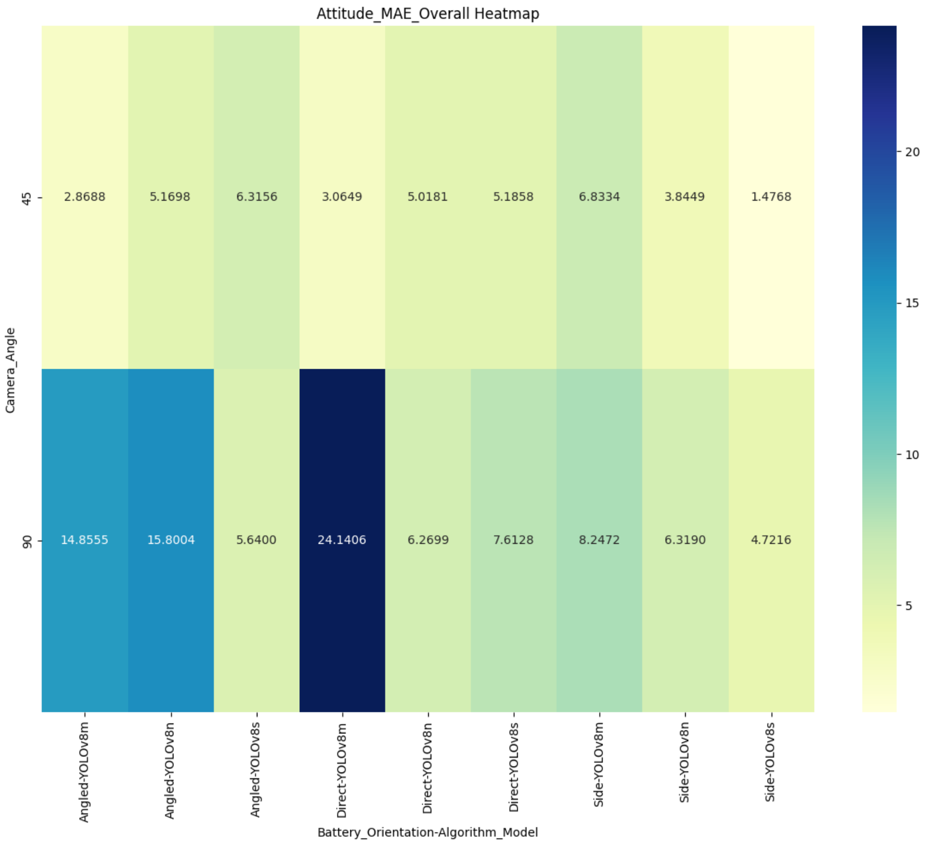


Fig. 9. attitude MAE overall heatmap

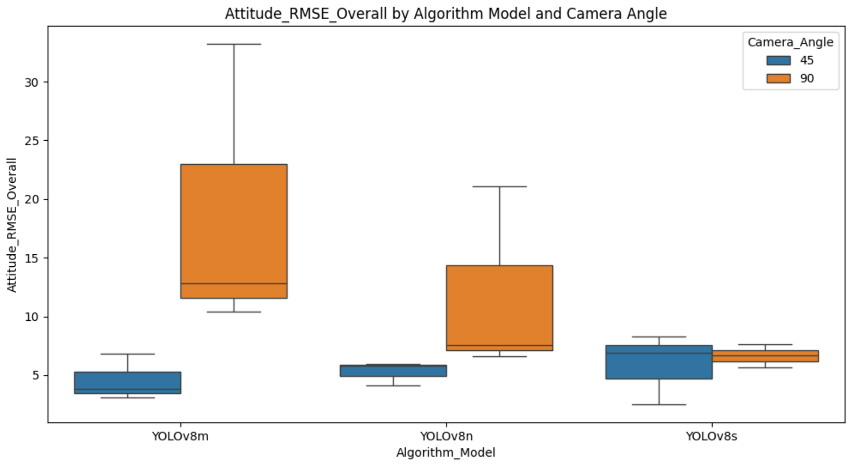


Fig. 10. attitude RMSE overall boxplot

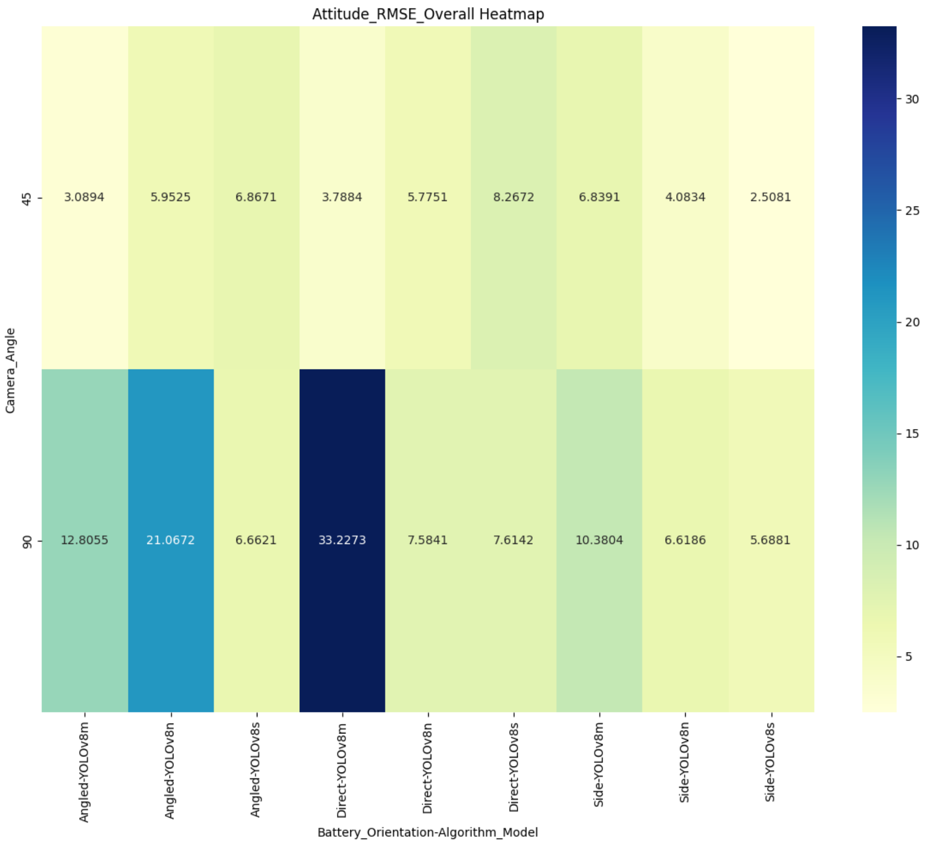


Fig. 11. attitude RMSE overall heatmap

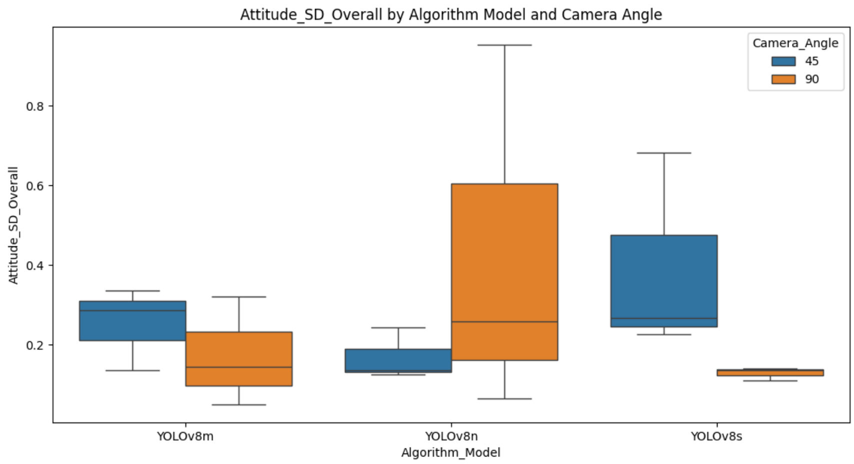


Fig. 12. attitude SD overall boxplot

C Project outline

Project Outline - Jiahao Geng 10683124

Scope

This project focuses on developing and evaluating computer vision algorithms for battery detection and positioning in UAV hot-swapping stations. The research will be primarily software-based, utilizing computer vision and machine learning techniques to identify batteries and determine their position and pose. The project will not involve hardware implementation of a robotic arm or actual battery swapping mechanisms. The study will be limited to the analysis of the position and pose estimation algorithm, with an emphasis on the DJI FlyCart 30 drone battery model.
The primary research question is: How does computer vision algorithms (yolov8-pose and PnP) perform under different conditions for detecting and determining the position and orientation of batteries in UAV hot-swapping scenarios, particularly for larger drones like the DJI FlyCart 30?

Motivation

The motivation behind this project stems from the limitations of existing battery exchange designs, which are largely centered around small drones with specific positioning requirements. These designs typically involve methods like funneling drones into precise positions or using frameworks to secure them for battery replacement. However, these methods are impractical for larger drones due to their size and weight. For example, large agricultural drones and delivery drones cannot easily be funneled or precisely positioned using current methods. Therefore, there is a clear need for a more flexible and robust solution that can handle the unique challenges posed by larger drones. Improving battery swapping stations to efficiently manage these larger UAVs will enhance their operational efficiency and broaden their applicability across various commercial sectors.

Aim

To develop and evaluate computer vision algorithms for accurate battery detection and positioning in UAV hot-swapping stations, with a focus on larger drones like the DJI FlyCart 30.

Objectives

1. Develop computer vision algorithms for battery detection in images feeds.
2. Implement position and orientation estimation techniques for detected batteries.
3. Evaluate algorithm performance using 3D-printed models.
4. Analyze the accuracy and robustness of the developed algorithms through statistical methods.

D Risk assessment

The University of Manchester

MANCHESTER 1824

General Risk Assessment Form

Date: 2024/07/01	Assessed by: Jiahao Geng	Checked / Validated* by: (3)	Location: Home	Assessment ref no: GRAF_10683124	Review date: 2024/07/01
Task / premises: Doing research for MSc project and dissertation at home					

Activity (8)	Hazard (9)	Who might be harmed and how (10)	Existing measures to control risk (11)	Risk rating (12)	Result (13)
Working from home	Deliveries to home addresses and GDPR	home working Student	1. When having items delivered to home addresses to enable working staff should be aware of implications of sharing home address with any suppliers.	Low	A
Working from home	Poor posture, repetitive movements, long periods looking at DSE (display screen equipment)	home working Student Back strain (due to poor posture). Repetitive Strain Injury (RSI) to upper limbs. Eye strain.	1. Set up workstation to a comfortable position with good lighting and natural light where possible 2. Take regular breaks away from the screen, at least some activity at your workstation every 20mins and a 5 minute break from workstation every hour. 3. Regularly stretch your arms, back, neck, wrists and hands to avoid repetitive strain injuries. Refer to workstation exercises here 4. Set up a desktop working space where possible and try to avoid working on a laptop without a docking station	Low	A

Result : T = trivial, A = adequately controlled, N = not adequately controlled, action required, U = unknown risk

Activity (8)	Hazard (9)	Who might be harmed and how (10)	Existing measures to control risk (11)	Risk rating (12)	Result (13)
Working from home	Stress / Wellbeing	home working Student Psychosocial effects, Work / Life imbalance, Anxiety	1. Please refer to Seven rules of home working published by AMBS 2. Regular contact meetings with manager and peers, Skype, Zoom, Phone 3. Define working hours, set a start & close daily routine, get dressed and prioritise your tasks.	Low	A
Working from home	Misuse of electrical appliance, faulted electrical appliance.	home working Student Electric shock, burns and fire	1. All office equipment used in accordance with the manufacturer's instructions 2. Visual checks before use to make sure equipment, cables and free from defects 3. Avoid daisy chaining and do not overload extension leads 4. University IT equipment brought home should already be PAT tested 5. The domestic electrical supply and equipment owned by the employee is the responsibility of the employee to maintain 6. Liquid spills cleaned up immediately 7. Defective plugs, cables and equipment should be taken out of use	Med	A
Working from home	Obstructions and trip hazards	home working Student Slips, trips and falls causing physical injury	1. Floors and walkways kept clear of items, e.g. boxes, packaging, equipment etc 2. Furniture is arranged such that movement of people and equipment are not restricted 3. Make sure all areas have good level of lighting 4. Reasonable standards of housekeeping maintained 5. Trailing cables positioned neatly away from walkways 6. Cabinet drawers and doors kept closed when not in use	Med	A

Result : T = trivial, A = adequately controlled, N = not adequately controlled, action required, U = unknown risk

Activity (8)	Hazard (9)	Who might be harmed and how (10)	Existing measures to control risk (11)	Risk rating (12)	Result (13)
Working from home	Fire	home working Student Risk of burns, Smoke inhalation, asphyxiation	1. In the event of a fire evacuate out of the building and call the fire brigade 2. All waste, including combustible waste, removed regularly. 3. Heaters located away from combustible materials and switched off when office is left unattended 4. Avoid daisy chaining and do not overload extension leads 5. Test smoke alarm routinely and replace batteries every 6-12 months 6. Please refer to fire brigade Home Fire Safety and Smoke Alarms	Med	A
Working from home	High risk activities	home working Student Personal injuries / accidents	1. Home working is restricted to the use of laptops, computers and low power equipment which complies to < 42 Volts operation & < 3 Amps total current consumption and which cannot exceed > 40degC operational temperature 2. No practical hardware work must be undertaken which requires tools, power-tools, soldering or any other sources of physical or chemical hazard	Low	A
Working from home	Accident / Incidents	home working Student Injuries from home working activities	1. If you suffer an accident / incident whilst working at home in relation to your workstation, please report the event to your line manager and the School Safety Advisor to complete an accident / incident form	Low	A

Result : T = trivial, A = adequately controlled, N = not adequately controlled, action required, U = unknown risk

Activity (8)	Hazard (9)	Who might be harmed and how (10)	Existing measures to control risk (11)	Risk rating (12)	Result (13)
Working from home	Manual handling of items delivered to a home address Interaction with the delivery driver	home working Student Personal injuries / accidents / infection	1. Use kinetic lifting techniques e.g. feet apart, load held close and in front of the body. If lifting off the floor, bend knees and keep the spine neutral. 2. Ensure there is a firm grip on the item whilst moving 3. Ensure trip hazards are removed on route from the front door to where the item is to be located. 4. Do not store large, heavy, fragile or cumbersome items at height (eg on high shelves or on top of cabinets/bookcases etc)	Low	A
Working from home	Performing craftwork at home Injuries from using tools	home working Student Cuts, punctures	1. Refer to craft tools usage guidelines. 2. Use appropriate personal protective equipment (such as gloves). 3. Keep the workbench clean and organized to avoid scattered tools. 4. Properly store sharp tools to prevent access by children and pets.	Med	A
Working from home	Cooking at home	home working Student Burns from hot surfaces, cuts from sharp utensils, and foodborne illnesses	1. Follow kitchen safety guidelines and use appropriate cooking equipment. 2. Keep a first aid kit nearby for minor injuries. 3. Ensure proper food handling and storage to prevent contamination. 4. Regularly clean and maintain kitchen appliances.	Med	A

Result : T = trivial, A = adequately controlled, N = not adequately controlled, action required, U = unknown risk

Action plan (14)				
Ref No	Further action required	Action by whom	Action by when	Done

Result : T = trivial, A = adequately controlled, N = not adequately controlled, action required, U = unknown risk

Notes to accompany General Risk Assessment Form

This form is the one recommended by Health & Safety Services, and used on the University's risk assessment training courses. It is strongly suggested that you use it for all new assessments, and when existing assessments are being substantially revised. However, its use is not compulsory. Providing the assessor addresses the same issues; alternative layouts may be used.

(1) **Date** : Insert date that assessment form is completed. The assessment must be valid on that day, and subsequent days, unless circumstances change and amendments are necessary.

(2) **Assessed by** : Insert the name and signature of the assessor. For assessments other than very simple ones, the assessor should have attended the University course on risk assessments (link to STDU)

(3) **Checked / Validated* by** : delete one.

Checked by : Insert the name and signature of someone in a position to check that the assessment has been carried out by a competent person who can identify hazards and assess risk, and that the control measures are reasonable and in place. The checker will normally be a line manager, supervisor, principal investigator, etc. Checking will be appropriate for most risk assessments.

Validated by : Use this for higher risk scenarios, eg where complex calculations have to be validated by another "independent" person who is competent to do so, or where the control measure is a strict permit-to-work procedure requiring thorough preparation of a workplace. The validator should also have attended the University's risk assessment course or equivalent, and will probably be a chartered engineer or professional with expertise in the task being considered. Examples of where validation is required include designs for pressure vessels, load-bearing equipment, lifting equipment carrying personnel or items over populated areas, and similar situations.

(4) **Location** : insert details of the exact location, ie building, floor, room or laboratory etc

(5) **Assessment ref no** : use this to insert any local tracking references used by the school or administrative directorate

(6) **Review date** : insert details of when the assessment will be reviewed as a matter of routine. This might be in 1 year's time, at the end of a short programme of work, or longer period if risks are known to be stable. Note that any assessment must be reviewed if there are any significant changes – to the work activity, the vicinity, the people exposed to the risk, etc

(7) **Task / premises** : insert a brief summary of the task, eg typical office activities such as filing, DSE work, lifting and moving small objects, use of misc electrical equipment. Or, research project [title] involving the use of typical laboratory hardware, including fume cupboards, hot plates, ovens, analysis equipment, flammable solvents, etc.

(8) **Activity** : use the column to describe each separate activity covered by the assessment. The number of rows is unlimited, although how many are used for one assessment will depend on how the task / premises is sub-divided. For laboratory work, activities in one particular lab or for one particular project might include; use of gas cylinders, use of fume cupboard, use of computer or other electrical equipment, use of lab ovens, hot plates or heaters, use of substances hazardous to health, etc

(9) **Hazard** : for each activity, list the hazards. Remember to look at hazards that are not immediately obvious. For example, use of a lathe will require identification of the machine hazards, but also identification of hazards associated with the use of cutting oils (dermatitis), poor lighting, slipping on oil leaks, etc. The same activity might well have several hazards associated with it. Assessment of simple chemical risks (eg use of cleaning chemicals in accordance with the instructions on the bottle) may be recorded here. More complex COSHH

assessments eg for laboratory processes, should be recorded on the specific COSHH forms (link).

(10) **Who might be harmed and how** : insert everyone who might be affected by the activity and specify groups particularly at risk. Remember those who are not immediately involved in the work, including cleaners, young persons on work experience, maintenance contractors, Estates personnel carrying out routine maintenance and other work. Remember also that the risks for different groups will vary. Eg someone who needs to repair a laser may need to expose the beam path more than users of the laser would do. Vulnerable groups could include children on organised visits, someone who is pregnant, or employees and students with known disabilities or health conditions (this is not a definitive list).

For each group, describe how harm might come about, eg an obstruction or wet patch on an exit route is a hazard that might cause a trip and fall; use of electrical equipment might give rise to a risk of electric shock; use of a ultraviolet light source could burn eyes or skin.

(11) **Existing measures to control the risk** : list all measures that already mitigate the risk. Many of these will have been implemented for other reasons, but should nevertheless be recognised as means of controlling risk. For example, restricting access to laboratories or machine rooms for security reasons also controls the risk of unauthorised and unskilled access to dangerous equipment. A standard operating procedure or local rules (eg for work with ionising radiation, lasers or biological hazards) will often address risks. Some specific hazards may require detailed assessments in accordance with specific legislation (eg COSHH, DSEAR, manual handling, DSE work). Where this is the case, and a detailed assessment has already been done in another format, the master risk assessment can simply cross-reference to other documentation. For example, the activity might be use of a carcinogen, the hazard might be exposure to hazardous substances, the existing control measures might all be listed in a COSHH assessment. Controls might also include use of qualified and/or experienced staff who are competent to carry out certain tasks; an action plan might include training requirements for other people who will be carrying out those tasks.

(12) **Risk Rating** : the simplest form of risk assessment is to rate the remaining risk as high, medium or low, depending on how likely the activity is to cause harm and how serious that harm might be.

The risk is **LOW** - if it is most unlikely that harm would arise under the controlled conditions listed, and even if exposure occurred, the injury would be relatively slight.
The risk is **MEDIUM** - if it is more likely that harm might actually occur and the outcome could be more serious (eg some time off work, or a minor physical injury.
The risk is **HIGH** - if injury is likely to arise (eg there have been previous incidents, the situation looks like an accident waiting to happen) and that injury might be serious (broken bones, trip to the hospital, loss of consciousness), or even a fatality.

Schools or administrative directorates may choose to use other rating systems. Typical amongst these are matrices (of 3x3, 4x4, 5x5 or even more complex) which require the assessor to select a numerical rating for both "likelihood that harm will arise" and "severity of that harm". These may give a spurious sense of accuracy and reliability – none are based on quantitative methods. There are methods of estimating risk quantitatively, and these may be appropriate for complex design of load bearing structures and the like. Advice on methods of risk assessment is available from HSS. Whatever system of assessment is adopted, it is **essential** that the assessor has received suitable training and is familiar with the meaning of the terms (or numbers) used.

(13) **Result** : this stage of assessment is often overlooked, but is probably the most important. Assigning a number or rating to a risk does not mean that the risk is necessarily adequately controlled. The options for this column are:

T = trivial risk. Use for very low risk activities to show that you have correctly identified a hazard, but that in the particular circumstances, the risk is insignificant.

A = adequately controlled, no further action necessary. If your control measures lead you to conclude that the risk is low, and that all legislative requirements have been met (and University policies complied with), then insert A in this column.

N = not adequately controlled, actions required. Sometimes, particularly when setting up new procedures or adapting existing processes, the risk assessment might identify that the risk is high or medium when it is capable of being reduced by methods that are reasonably practicable. In these cases, an action plan is required. The plan should list the actions necessary, who they are to be carried out by, a date for completing the actions, and a signature box for the assessor to sign off that the action(s) has been satisfactorily completed. Some action plans will be complex documents; others may be one or two actions that can be completed with a short timescale.

U = unable to decide. Further information required. Use this designation if the assessor is unable to complete any of the boxes, for any reason. Sometimes, additional information can be obtained readily (eg from equipment or chemicals suppliers, specialist University advisors) but sometimes detailed and prolonged enquiries might be required. Eg is someone is moving a research programme from a research establishment overseas where health and safety legislation is very different from that in the UK.

For T and A results, the assessment is complete.
For N or U results, more work is required before the assessment can be signed off.

(14) **Action Plan**. Include details of any actions necessary in order to meet the requirements of the information in Section 11 'Existing measures to control the risk'. Identify someone who will be responsible for ensuring the action is taken and the date by which this should be completed. Put the date when the action has been completed in the final column.