Reliable Target-Hit Detection for Throwing Robots Using Occlusion-Robust Visual Tracking

Jade Carpenter, Erik Lipinski, Tomas Romanovas, and Min-Fan Ricky Lee

***Abstract*—Studies on throwing robots have achieved precise throwing motions and trajectory control, however, those systems conclude at the release phase and do not verify whether or not the target was hit. In real environments, such as disaster response situations, occlusion often occurs, impeding the previously assumed perfect visibility of the projectile and target, blocking them out of view and preventing reliable verification of hit success. This paper tackles that uncertainty by proposing an occlusion-aware visual verification method for tossing robots. Its method uses a camera sensor coupled with a Kalman filter to estimate the trajectory during visual loss and determines if the target is hit. The framework was implemented in CoppeliaSim then evaluated under varying occlusion environmental parameters. The results maintain a continuous trajectory prediction and accurately estimates impact probability with an F-1 score, even with missing visuals. The contribution is an additional post-throw verification to verify goal-reach that demands little input data, with a high accuracy in its response, which offers a practical solution for autonomous throw verification under visual uncertainty.**

***Index Terms*—Automatic control, Autonomous robots, Computer vision, Kalman filters, Uncertainty.**

I. Introduction

THROWING robots have demonstrated great accuracy in predicting and executing tossing motions, yet it is typically overlooked or omitted to verify whether the projectile successfully reaches the intended target. With assumptions of perfect visibility and deterministic impact, most existing throwing systems, including [1], [2], determine the best trajectory and then close their learning loop at the release time of the object. In order to utilize throwing robots in disaster response situations such as fire rescues, as shown in Fig. 1 to provide resources (first aid or supplies) across smoke walls, successful throw verification is required. In fact, in the real world, object uncertainties and environment uncertainties typically render it difficult to guarantee a successful throw verification.

Jade Carpenter is with the École Nationale de l'Aviation Civile (ENAC), Toulouse, France (e-mail: jade.carpenter@alumni.enac.fr), on an exchange program at the National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan.

Erik Lipinski is with RWTH Aachen University, Aachen, Germany (e-mail: erik.lipinski@rwth-aachen.de). He is an exchange student at the National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan.

Tomas Romanovas is with Vilnius Gediminas Technical University (Vilnius TECH), Vilnius, Lithuania (e-mail: tomas.romanovas@stud.vilniustech.lt), on an exchange program at the National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan.

Min-Fan Ricky Lee is an Associate Professor with the Graduate Institute of Automation and Control, National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan (e-mail: rickylee@mail.ntust.edu.tw). His research interests include robotics, artificial intelligence, and computer vision.

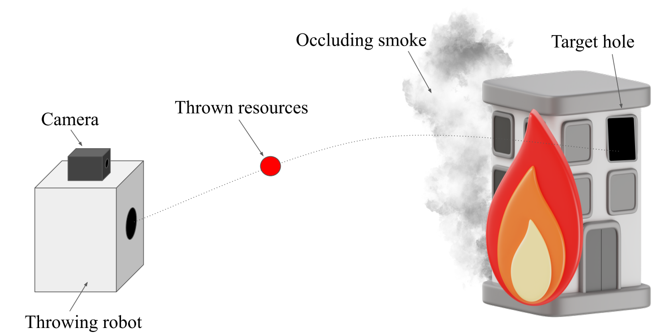

Fig. 1. Usage of a throwing robot for a disaster response situation presenting occlusion

Reliable visual verification is essential for closed-loop learning and intelligent control. In fact, without trustworthy sensor feedback, robots cannot update their control parameters, evaluate success in uncertain environments, or ensure safe interaction. A predominant kind of uncertainty is occlusion, prevalent in robotic vision due to environmental obstacles, self-occlusion, or the trajectory leaving the field-of-view. It is therefore mandatory to tackle this constraint to design autonomous, adaptive robots that can operate in uncontrolled environments. The unanswered question is how to verify target-hit success when the ball or the target is momentarily occluded.

Research on throwing robots has mainly focused on generating movements and optimising control, while neglecting post-throw verification. Systems such as [1], [2] have advanced dynamic coordination and data-driven control but assume that visual feedback remains available and accurate. This study challenges that assumption by addressing the missing steps of verifying the outcome of the throw when occlusion prevents direct observation.

Additionally, this work extends [3], [4] in the field of perception research by their occlusion-aware tracking in sports to fast ballistic trajectories. This work complements prior research such as [5], [6] by extending motion prediction and robustness analysis to include visual uncertainty caused by occlusion during post-throw verification. Building on the robust recursive estimation in the event of missing data demonstrated in [7], the proposed method adapts classical filtering to an occlusion-aware verification problem. Overall, this study challenges the assumption of perfect visibility and integrates occlusion-handling methods into ballistic motion prediction.

To tackle the problem of visual uncertainty caused by occlusion, this study proposes an occlusion-aware verification approach based on a Kalman filter-assisted visual tracker. The system combines object detection with probabilistic state estimation to maintain a continuous prediction of the projectile's position, even when the ball is temporarily invisible. Visual observations obtained from the camera are used to update the state estimation whenever the object is visible, while the Kalman filter propagates the predicted trajectory during occlusion using a constant acceleration motion model. At the end of each throw, the estimated impact position and its covariance are compared to the target area to calculate a probability of success, automatically verifying whether the throw was successful. This method is a Kalman-filter–based verification system that predicts the projectile's trajectory during occlusion and estimates the likelihood of a target hit from the predicted impact distribution, providing a lightweight and interpretable alternative to deep-learning methods for handling occlusion in throwing-robot perception.

II. Method

A. Experimental Setup and Framework Overview

In this research, the implementation of the discussed occlusion-aware verification method used Python, OpenCV and the FilterPy library. A virtual scene was created in CoppeliaSim in order to simulate a throwing robot that tosses a small spherical projectile towards a circular target hole. A virtual RGB camera was positioned to observe the motion from a fixed point of view, recording sequences in which the projectile or target were temporarily hidden, as previewed in Fig. 2. These videos established controlled settings to test occlusion throughout the throw.

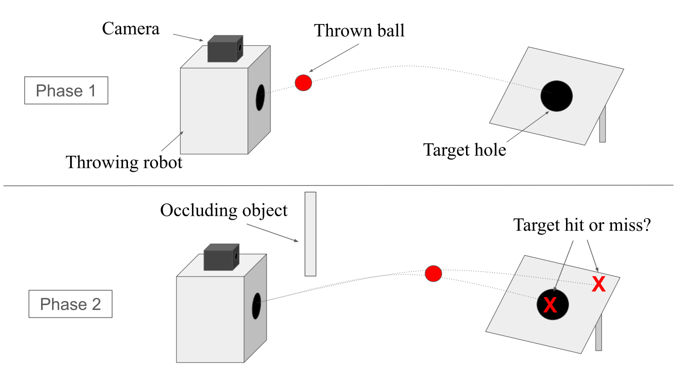

Fig. 2. Ball thrown towards target then occluded before hit

The system architecture presents three modules: visual detection, state estimation, and impact verification, as shown in Fig. 3. The visual detection module identifies the centroid of the projectile while it is visible using observations from the simulation and real throw footage, then passing it through OpenCV functions. The estimation module uses a Kalman filter to give a continuous estimate of the position, velocity, and acceleration of the projectile. Finally, the verification module analyses the estimated terminal position and computes a probability that the projectile enters the target area.

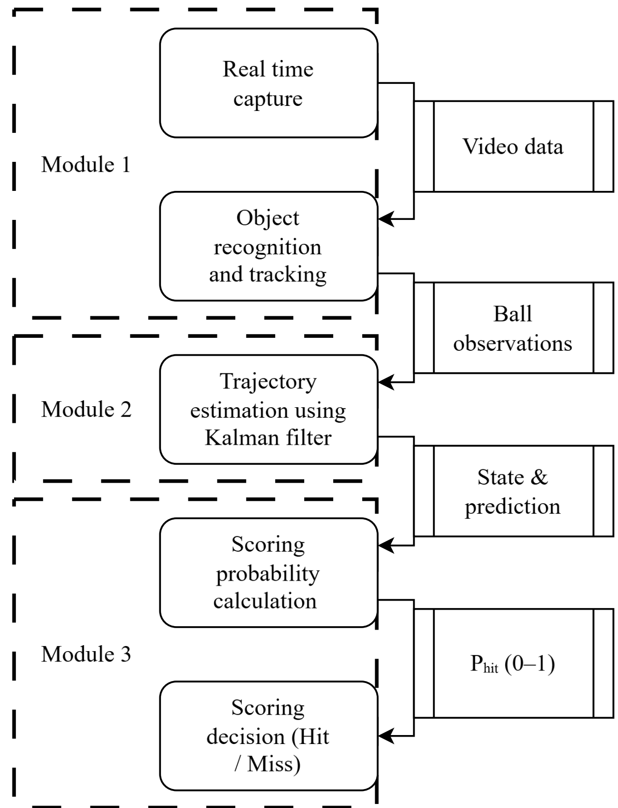

Fig. 3. System flowchart illustrating the modules for real-time capture, object tracking, trajectory estimation, and probabilistic scoring.

Occlusion uncertainty is addressed by combining detection, estimation and verification in a single structure. Continuous state prediction is ensured by the system's modularity, even in the event of a projectile becoming unobservable. For the impact estimation, a probabilistic assessor is used to find the degree of hit-miss reliability. The method maintains accuracy in partial viewing conditions by use of deterministic motion modeling and probabilistic reasoning. Concisely, a straightforward and adaptable solution is used to ensure reliable tossed ball's tracking and trajectory estimation despite occasional loss of visual data.

B. Kalman Filter–Based Estimation Model

The Kalman filter provides a probabilistic mechanism to estimate the projectile's motion under uncertainty, maintaining an internal state vector $\mathbf{x}_k$ as defined in equation (1).

$$\mathbf{x}_k = \begin{bmatrix} u_k & v_k & \dot{u}_k & \dot{v}_k & \ddot{u}_k & \ddot{v}_k \end{bmatrix}^T. \tag{1}$$

The (u_k, v_k) in equation (1) represent the position in the image plane, while the remaining components denote motion derivatives. The motion of the projectile is modeled using a constant-acceleration process, as shown in equation (2), where F is a state transition matrix, w_k is process noise and Q is the process-noise covariance. We assume Gaussian-distributed process and measurement noise, enabling the Kalman filter to serve as the optimal linear estimator for minimizing prediction error under uncertainty. The corresponding measurement model is given in equation (3), where z_k is the measurement vector, H is a measurement matrix, v_k is measurement noise and R is the measurement-noise covariance. Equations (2) and (3) together define the standard discrete Kalman filter formulation used for state prediction and measurement updates.

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{w}_k, \quad \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \tag{2}$$

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k, \quad \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}). \tag{3}$$

For implementation, the state transition matrix F, measurement matrix H, and process-noise covariance Q were defined as follows in equations (4), (5) and (6):

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & dt & 0 & \frac{1}{2}dt^2 & 0 \\ 0 & 1 & 0 & dt & 0 & \frac{1}{2}dt^2 \\ 0 & 0 & 1 & 0 & dt & 0 \\ 0 & 0 & 0 & 1 & 0 & dt \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{5}$$

$$\mathbf{Q} = q \begin{bmatrix} \frac{dt^5}{20} & 0 & \frac{dt^4}{8} & 0 & \frac{dt^3}{6} & 0 \\ 0 & \frac{dt^5}{20} & 0 & \frac{dt^4}{8} & 0 & \frac{dt^3}{6} \\ \frac{dt^4}{8} & 0 & \frac{dt^3}{3} & 0 & \frac{dt^2}{2} & 0 \\ 0 & \frac{dt^4}{8} & 0 & \frac{dt^3}{3} & 0 & \frac{dt^2}{2} \\ \frac{dt^3}{6} & 0 & \frac{dt^2}{2} & 0 & dt & 0 \\ 0 & \frac{dt^3}{6} & 0 & \frac{dt^2}{2} & 0 & dt \end{bmatrix} \tag{6}$$

where dt is the sampling period and q is the process-noise scaling factor.

During visible frames, the filter executes both prediction and update steps, incorporating visual observations to refine its estimate. When there is occlusion and no detection is made, the update step is omitted, and only the prediction step propagates the motion. This enables ongoing trajectory estimation even for missing frames. Then, when visibility is regained, the Kalman gain updates the prediction, minimizing accumulated error. The overall process is summarized in Algorithm 1.

The predicted impact is considered a successful hit if the computed probability p_{hit} exceeds a chosen confidence threshold of 0.5. This binary decision enables objective evaluation of the verification method: a hit is correctly detected if the true impact also falls in the target zone. At the predicted impact time, the estimated position mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ define a Gaussian distribution. The probability that the ball at the predicted impact position s inside the circular target of center $\mathbf{c}$ and radius r is expressed in equation (7). This probabilistic verification links state estimation with decision theory, providing interpretable feedback on whether the throw was successful even under partial visual information.

$$p_{\text{hit}} = P(\|\mathbf{s} - \mathbf{c}\| \leq r), \quad \mathbf{s} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{7}$$

Algorithm 1 Occlusion-Aware Visual Verification for Throwing Robots

Require: Video frames $\{F_1 \dots F_n\}$, target center $\mathbf{c}$, target radius r
Ensure: Estimated trajectory $\hat{X}_k$, hit probability p_{hit}

Initialize Kalman filter state x_0, covariance P_0
Set process noise Q and measurement noise R
Initialize trajectory buffer as empty
for $k = 1$ to n **do**
  $(\text{Detected}, (u, v)) \leftarrow \textsc{DetectCentroid}(F_k)$
  if Detected **then**
    $z_k \leftarrow (u, v)$
    KF_UPDATE(x_k, P_k, z_k, R)
  else
    KF_PREDICT(x_k, P_k, Q)
  end if
  Append (x_k, P_k) to trajectory buffer
end for
$t_{\text{impact}} \leftarrow$ ESTIMATEIMPACTTIME(trajectory buffer)
$(x_{\text{imp}}, P_{\text{imp}}) \leftarrow \textsc{PredictStateAt}(t_{\text{impact}})$
$p_{\text{hit}} \leftarrow \textsc{ComputeHitProbability}(x_{\text{imp}}, P_{\text{imp}}, \mathbf{c}, r)$
if $p_{\text{hit}} \geq$ threshold **then**
  return "Successful Hit"
else
  return "Missed Target"
end if

C. Evaluation and Parameter Tuning

Simulation experiments were conducted to evaluate the system under different release angles, velocities, and occlusion durations. Detection quality was assessed using precision, accuracy and recall, while Kalman filter predictions were compared to ground-truth trajectories using error-based metrics. These evaluations provided a clear view of how well the system maintains trajectory estimates during visual loss.

The computation of an F_1-score summarized overall performance, with special attention given to false positives, since incorrectly declaring a successful hit is the most critical failure mode for the verification task. Noise covariances were carefully tuned based on innovation statistics to stabilize the estimator under varying occlusion conditions. This ensured that the filter neither diverged nor became overly conservative.

Robustness was further tested through repeated trials with randomized lighting, background color, and target contrast. These variations helped confirm that the method does not depend on a single controlled environment. Across all test conditions, the system maintained consistent estimation quality, demonstrating that the Kalman filter effectively compensates for occlusion without heavy computation. This stability indicates strong potential for deployment in real-world visual uncertainty scenarios.

D. Methodological Justification and Core Challenge

The Kalman filter was selected because it is interpretable, deterministic, and lightweight. Also, it does not require neural training or advanced computing. It is based only on physical motion assumptions and minimal noise parameters, making it easy to interpret and validate.

The main challenge addressed is that of maintaining valid estimation during occlusion, where the camera provides no measurement. The Kalman filter reduces this by predicting motion through its internal model. Its covariance growth quantifies confidence reduction with increasing occlusion duration.

By integrating physical reasoning with probabilistic estimation, the framework unifies principles from computer vision and robotic control under uncertainty. It enables throwing robots to assess the outcome of their actions autonomously, even when sensor information is incomplete. This integration forms a minimal yet sufficient foundation for reliable throw verification in visually uncertain environments.

III. Results

A. Detection and Tracking Performance

Reliable perception during visual loss is central to assessing whether the system can maintain trajectory consistency under occlusion. Across all simulation trials, the detector provided stable centroid measurements whenever the projectile was visible and seamlessly transferred responsibility to the Kalman filter when visibility was lost. Fig. 4 illustrates the virtual experimental setup, while Fig. 5 shows a representative occlusion interval during which state updates rely solely on prediction.

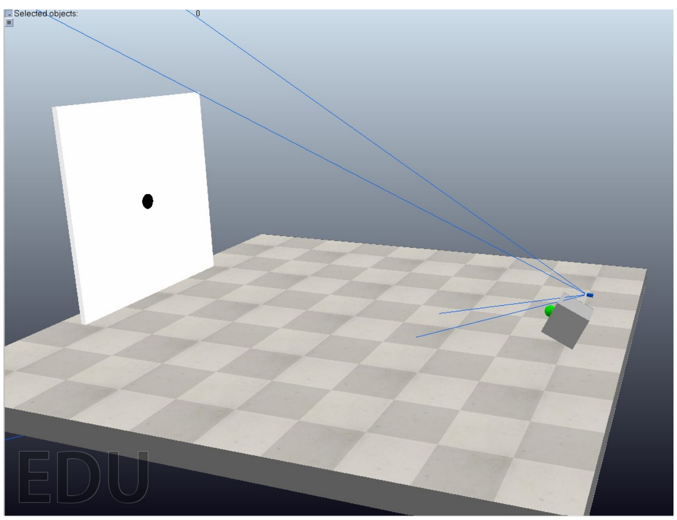

Fig. 4. Scene with throwing robot, target, and ball.

Tracking quality was quantified using trajectory overlap and pixel-wise error against ground truth. For occlusion durations up to 15%, the estimator maintained an average trajectory overlap of 93% and a mean squared error below $7\,\text{px}^2$. These outcomes confirm that the constant-acceleration motion model remains appropriate for short ballistic motion, even when visual measurements are intermittently unavailable.

To summarize perceptual performance, the F_1-score was computed using precision and recall, and accuracy was evaluated as an additional metric. The associated formulas are defined in (8), (9), (10), and (11).

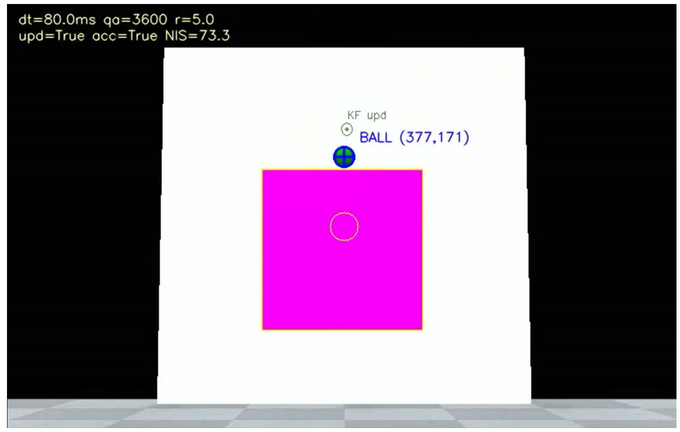

Fig. 5. Camera view during occlusion, where the hole becomes temporarily invisible.

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{8}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{9}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{10}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{11}$$

- TP = True Positives (correctly classified hits)
- FP = False Positives (predicted "hit" but actually miss)
- TN = True Negatives (correctly classified misses)
- FN = False Negatives (missed hits)

Using these metrics, low-noise configurations ($q_a = 400$) yielded F_1-scores close to 1.00 and accuracies above 0.98, indicating almost perfect perceptual performance. Higher-noise cases ($q_a = 1200$) produced lower but still acceptable scores, with F_1 ranging between 0.80 and 0.88 and accuracy around 0.80–0.90. Overall, the perception pipeline combining detection and Kalman filtering remained robust under varying levels of occlusion and motion uncertainty.

B. Verification Accuracy and Calibration

The verification module evaluates whether the final predicted impact lies within the target zone. Table I reports the confusion matrices obtained across multiple trials for different noise and occlusion settings. Under low process noise ($q_a = 400$), results were nearly flawless: for both 5% and 10% occlusion durations, out of 120 trials, 118 impacts were correctly classified, with only two misclassifications. This corresponds to $(\text{Precision} \approx 0.99, \text{Recall} \approx 0.99, \text{Accuracy} \approx 0.98)$.

Higher process noise ($q_a = 1200$) produced a smooth and predictable degradation. At 5% occlusion, 126 out of 130 impacts were correctly classified, while four false positives appeared due to broader covariance growth. At 10% occlusion, 119 correct classifications were recorded, along with six false positives and five false negatives, reducing overall accuracy to about 0.80. These results reflect how increased motion uncertainty impacts the reliability of the predicted terminal position without causing instability or divergence.

TABLE I
CONFUSION MATRICES FOR IMPACT VERIFICATION UNDER VARYING NOISE AND OCCLUSION CONDITIONS.

q_a	Occlusion	TP	FP	FN	TN
400	5%	59	1	1	59
400	10%	60	1	1	58
1200	5%	56	4	0	70
1200	10%	52	6	5	67

Calibration analysis further validated the probabilistic output of the system. The Expected Calibration Error (ECE) remained low (0.06–0.08 across conditions), indicating close alignment between predicted hit probabilities and empirical success ratios. This confirms that the estimated p_{hit} values are interpretable and trustworthy, even when predictions rely on several consecutive prediction-only steps. Overall, the verification module maintained strong reliability across conditions while providing confidence estimates that reflect true performance.

C. Robustness Across Conditions

Robustness was evaluated by varying illumination, target contrast, and background appearance. These environmental changes altered the image statistics but did not significantly degrade performance: across all tested conditions, detection accuracy remained within $\pm 3\%$ of the baseline, and tracking error varied by less than $1.2\,\text{px}$ on average. This indicates that the combined detection–estimation pipeline is not overly sensitive to appearance differences.

Occlusion duration had the most notable effect on performance. For occlusions exceeding 15% of the trajectory under high process noise, the estimator's uncertainty increased noticeably, contributing to a rise in false positives and false negatives. However, the degradation remained gradual rather than sudden, demonstrating robust behavior even under less favorable assumptions. The system never exhibited divergence, oscillation, or abrupt covariance inflation.

Together, these results show that the overall approach handles a range of noisy and visually degraded conditions without requiring significant computational resources or extensive retraining. The system therefore provides a dependable occlusion-aware verification mechanism suitable for autonomous robotic applications operating in cluttered or partially observable environments.

IV. Discussion

The results demonstrate that the system maintained reliable perception and verification even when visual information was temporarily missing, reaffirming the study's initial objective of addressing occlusion in throwing-robot vision. Overall, the detector showed high accuracy and the Kalman filter provided stable predictions, while the verification module classified outcomes correctly in most cases. These findings indicate that occlusion does not necessarily prevent a robot from evaluating the success of a throw, as long as prediction and uncertainty estimation are incorporated into the perception pipeline. Verification accuracy remained above 98% in low-noise conditions and only decreased to around 80% under higher noise and longer occlusions, confirming that performance degrades smoothly rather than failing abruptly.

Compared with existing work in throwing-robot research, which typically assumes continuous visibility of the projectile, the present method shows stronger robustness to missing observations. The integration of probabilistic prediction introduces a concrete advancement over prior approaches that stop at the release phase and omit post-throw verification. The system's ability to maintain trajectory continuity and produce calibrated hit probabilities reflects a meaningful contribution to the broader field of occlusion-aware robotic perception.

Several limitations remain, particularly when occlusion durations become very long or occur near the end of the trajectory, where uncertainty increases more rapidly. Future studies should therefore include real-world experiments with a physical throwing robot, more complex occlusion sources such as smoke or motion blur, and potentially multi-sensor fusion to further improve reliability. Testing beyond simulation is necessary to validate the method under real lighting, depth variation, and camera noise. Despite these constraints, the method's lightweight computation and interpretable output make it practical for autonomous robots operating in cluttered, dynamic, or visually degraded environments, where reliable verification is essential for safe and adaptive behavior.

References

[1] A. Zeng, S. Song, J. Lee, A. Rodriguez, and T. Funkhouser, "Tossingbot: Learning to throw arbitrary objects with residual physics," *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1307–1319, 2020.

[2] Y. Liu, A. Nayak, and A. Billard, "A solution to adaptive mobile manipulator throwing," arXiv, 2022.

[3] H. Xu, A. A. Baniya, S. Wells, M. R. Bouadjenek, R. Dazely, and S. Aryal, "Totnet: Occlusion-aware temporal tracking for robust ball detection in sports videos," arXiv, 2025.

[4] A. Glover, L. Gava, Z. Li, and C. Bartolozzi, "Edopt: Event-camera 6-dof dynamic object pose tracking," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 18 200–18 206.

[5] H. Yu, D. Guo, H. Yin, A. Chen, K. Xu, Y. Wang, and R. Xiong, "Neural motion prediction for in-flight uneven object catching," arXiv, 2021.

[6] M. Okada, A. Pekarovskiy, and M. Buss, "Robust trajectory design for object throwing based on sensitivity for model uncertainties," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 3089–3094.

[7] Y. A. Wendemagegn, W. Ayalew Asfaw, C. M. Abdissa, and L. N. Lemma, "Enhancing trajectory tracking accuracy in three-wheeled mobile robots using backstepping fuzzy sliding mode control," *Engineering Research Express*, vol. 6, no. 4, p. 045204, 2024.

Jade Carpenter is a French-Canadian student pursuing a Master of Science in Aeronautical Engineering at the École Nationale de l'Aviation Civile (ENAC), Toulouse, France, where she majors in Avionics. She is currently an exchange student in the Department of Automation and Control at the National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan. Her academic interests include autonomous systems, intelligent control, and embedded avionics.

Erik Lipinski received the B.Sc. degree in Electrical Engineering with a specialization in Micro- and Nanotechnology from RWTH Aachen University, Germany, in 2025. He is currently pursuing the M.Sc. degree in Electrical Engineering and Information Technology with a focus on Systems and Automation at the same university. His interests include autonomous robotics, intelligent control, and applied signal processing. He is an exchange student at the National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan.

Tomas Romanovas is a Bachelor's student in Electronics Engineering at Vilnius Gediminas Technical University (Vilnius Tech), Lithuania. He is currently an exchange student in the Department of Electronics Engineering at the National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan. His academic interests include autonomous systems, electronics, embedded systems, and RF/radar signal processing.

Min-Fan Ricky Lee received his Bachelor degree from Tam Kang University, Republic of China (Taiwan) in 1988. He received his Ph.D. and Master degree from Cornell University, USA in 1996 and 1991 respectively. He is currently the Associate Professor at the Graduate Institute of Automation and Control, National Taiwan University of Science and Technology, Republic of China (Taiwan).

He was the Research Officer at the National Research Council in Federal Government of Canada from 1998 to 2007 and the Research Scientist at Department of Mechanical Engineering in University of British Columbia, Canada from 1996 to 1998. He was the Mechanical Engineer at Engineering Maintenance and Repair Division in Flight Command of the Army, Republic of China (Taiwan) from 1988 to 1990.

He served as the Past Chair, Chair, Vice Chair, Treasurer and Secretary of the IEEE Vancouver Section from 1998-2003.

His research areas are Robot and Artificial Intelligence. His Ph.D. dissertation, "An Intelligent Computer Vision Control and Target Tracking System Design of an Agricultural Grapevine Pruning Robot", received the National Design Gold award USA. He received the Outstanding Service Award from IEEE Canada. He is endowed the title of "International Taiwan Youth Ambassador" by the President of Republic of China (Taiwan) in September 30, 2011.