Hierarchical Reinforcement Learning for Deer Detection and Deterrence

Ebasa Temesgen [1], Greta Brown[1], Haifeng Huang[1], Jack Swanberg[1], Sree Ganesh Lalitaditya Divakarla[1]
Maria Gini[1]

Abstract— **Wildlife crop damage, particularly done by deer, poses a risk to modern agriculture, threatening efforts to increase sustainability and productivity. This paper proposes a hierarchical multi-agent reinforcement learning (MARL) framework to detect and deter deer incursions using unmanned aerial vehicles (UAVs). At the high level, a centralized mission planner coordinates and allocates tasks among multiple UAVs, each tasked with monitoring a designated field sector. Each UAV employs Multi-Agent Proximal Policy Optimization (MAPPO) for local control, constructing flight paths, and generating pursuit maneuvers. We reduce the computational complexity and improve scalability by dividing the system into hierarchical high- and low-level policies. To address the scarcity of annotated real-world deer imagery, a photorealistic simulation environment built on Unreal Engine provides a safe and customized testbed for training detection and control policies. Although this work is in its early stages, the results so far show promising progress toward solving the problem of deterring deer.**

I. INTRODUCTION

Precision agriculture has become increasingly crucial in the threat of climate change and the growing need to optimize crop yields while minimizing environmental impacts [1]. Despite continuous improvement in mechanized processes, wildlife incursions, particularly by deer, remain a persistent challenge for farmers, resulting in substantial crop losses and economic losses [2]. This becomes a difficult challenge, especially in the North and Midwest regions, where the growing season is short. Conventional methods for monitoring and repelling deer, such as physical fencing or manual patrols, often lack adequate coverage and real-time responsiveness on large farms. In addition, the high cost of conventional methods like fencing often makes them infeasible for small-scale farms.

Unmanned aerial vehicles (UAVs), or drones, have proven to be an invaluable tool for the detection and deterrence of wildlife in agriculture [3]. UAVs offer a mobile and adaptable presence that stationary deterrents lack. When equipped with cameras (including thermal infrared), UAVs can autonomously scout large fields to detect and respond to intruding deer in real-time. Research has shown that drone-based thermal imaging can reliably spot deer hidden in croplands [4]. Once detected, UAVs can also act as a haze or herd animals away in a non-lethal manner [5]. Field studies and trials indicate that drones can successfully steer wildlife away from areas where they are not wanted. An example of this ability was shown for elephants, where they were driven away from farms in conflict areas [6]. Compared to manual patrolling or fixed scare devices, UAVs can respond immediately to incursions and have a much larger effective range, making them a compelling component of a precision agriculture toolkit. They can also effectively handle both detection and deterrence tasks simultaneously. However, relying on a single monolithic model to handle both detection and deterrence can limit scalability and adaptability in multi-agent scenarios.

In response to these limitations, we propose a hierarchical reinforcement learning (RL) framework for the efficient detection and deterrence of deer. Using quadcopter UAVs with onboard sensors and real-time processing, our framework seeks to detect deer and perform selective deterrence autonomously. The use of UAVs creates a system that can easily scale to any farm size, as well as be able to quickly adapt based on changes in the environment or landowner's goals. Recent advances in agent-based decision-making and generative models for visual understanding [7] show promise in providing robust detection and planning in complex environments. This hierarchical framework allows for effectively distributing tasks and reacting to events with the correct behavior.

A. Multi-Agent and Hierarchical Approaches

In multi-agent settings, coordination between autonomous UAVs can significantly improve coverage, reduce response time, and improve robustness when deployed on large or fragmented farmlands [8]. We explore a hierarchical RL framework in which a top-level agent provides macro-level decisions, such as task allocation and flight planning, and bottom-level agents are responsible for local navigation and deterrence actions. This approach allows for flexibility, whereby the system can accommodate varied conditions, such as unexpected deer movement patterns or varying environments.

At the higher level, an agent coordinates and distributes tasks among N UAVs, each to patrol and frighten deer in different parts of the field. After receiving a path plan, each UAV agent performs deterrence actions at a local level using Multi-Agent Proximal Policy Optimization (MAPPO) [9]. This includes whether to deviate from a planned path when detecting a target or to continue patrol. Logistic operations, such as returning to a charging station if the battery is running low, are also managed by an auxiliary subroutine.

*Work supported in part by the AI-CLIMATE Institute

[1] The authors are with the Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, 55455, USA {temes021, brow6802, huan2802, swanb045, divak014, gini}@umn.edu

B. Simulation and Synthetic Data Generation

As annotated datasets of deer in farmland settings are scarce, bridging the sim-to-real gap is essential for effective deep learning training [10]. To alleviate this data bottleneck, we use sophisticated simulation environments, including Unreal Engine, to generate synthetic scenarios that faithfully mimic real-world farmland conditions and deer behavior. The UAVs operating in these simulations collect large-scale data for both detection (e.g., using YOLO-based object recognition frameworks [11]) and policy learning in RL, accelerating model development.

C. Contributions

In this work, we focus on early stage results that highlight the viability of a hierarchical RL strategy for the detection and deterrence of deer through cooperative UAVs. Specifically, we:

1) Develop a hierarchical RL framework that separates high-level task allocation from low-level route execution, enabling dynamic and efficient multi-agent coordination.
2) Demonstrate how synthetic data from game engines can mitigate the lack of labeled deer imagery, improving both detection and deterrence performance.
3) Integration of YOLOv5 model, fine tuned for our usecase and measured its performance in simulation, supporting deer detection task and downstream UAV coordination.

II. Related Work

A. Precision Agriculture and Wildlife Management

Wildlife incursion is a persistent problem in precision agriculture, particularly for fields in proximity to natural habitats. Several methods have been proposed for the detection of animals and the deterrence of wildlife that causes crop damage. Bapat et al. [12] demonstrated a wireless sensor network (WSN) system utilizing passive infrared sensors and in situ deterrents such as flashing lights and noise. It is a modular and low-cost system. However, WSN-based systems can be limited in coverage flexibility, remoteness feasibility of farms, and responsiveness when the behavior of animals is unpredictable.

Unmanned aerial vehicles (UAVs) have become more accessible for aerial surveillance and wildlife monitoring. Axford et al. [13] surveyed deep learning techniques for wildlife detection in aerial imagery, highlighting key challenges such as occlusion, small object detection, and real-time inference with small targets, dense vegetation, and small training data. The authors identify key research gaps, such as the need for better domain adaptation techniques and more annotated datasets. This challenge is reflected in the work of Crossling et al. [14], who explored deep learning image classification pipelines using VGG16 and ResNet50 CNN architectures for the detection of five wildlife species in farming contexts. The study highlights the importance of training on data that reflect deployment conditions. The very fact that clean stock photos were used led to poor generalization in the domain.

Lyu et al. [15] developed a thermal drone-based detection system using an enhanced Faster R-CNN model with ResNet and pyramidal feature networks. Their method achieved high accuracy even in partially obscured settings, demonstrating the viability of thermal imaging for the monitoring of deer. However, their work does not extend to active deterrence. Bhusal et al. [16] presented a UAV-based pest bird deterrence system that autonomously deploys drones to chase birds upon detection. The approach shows how UAVs can act as mobile deterrents, but is not designed for continuous monitoring or coordination between multiple drones.

Some studies explore ground-based deterrence strategies. Geerthik et al. [17] proposed an IoT-based detection and humane deterrence system that uses context-specific responses (e.g., alarms and water sprays). Laguna et al. [18] evaluated the use of portable light and ultrasonic devices to repel red deer. Their field results indicate a measurable reduction in deer activity; however, the effectiveness of deterrence declined over time as the animals habituated to the stimulus over several weeks. The study suggests that, while such devices may be effective initially, their impact diminishes with repeated use in a fixed location.

B. Multi-Agent Reinforcement Learning

Recent MARL algorithms have made it feasible for multiple agents, including ground robots and UAVs, to learn coordinated policies. QMIX [19] is a value-based MARL algorithm that factorizes joint value functions and significantly outperforms previous methods in cooperative tasks. Multi-Agent Deep Deterministic Policy Gradient (MADDPG) [20] extends DDPG to multi-agent settings with a centralized critic for each agent. This framework handles both cooperative and competitive interactions by allowing each agent to learn a continuous policy while considering the actions of other agents.

Multi-agent PPO (MAPPO) [9] demonstrates that a simple PPO-based algorithm achieves strong performance on various cooperative MARL benchmarks, including Particle World, StarCraft Multi-Agent Challenge, Google Research Football, and Hanabi. MAPPO matched or exceeded the performance of more complex off-policy methods in final rewards and sample efficiency while requiring minimal hyperparameter tuning and no special architectural modifications.

These MARL advances are not purely theoretical; they have also been successfully applied to real robots. For example, MARL-based UAV swarms have been trained to cover agricultural fields cooperatively, achieving robust coverage with minimal overlap. Boubin et al. [21] implemented a MARL policy on physical drones for crop scouting, which yielded a 31% increase in profits over non-learning baselines in field tests.

C. Hierarchical Reinforcement Learning

Hierarchical frameworks are being explored to handle long-horizon decision making in robotics by decomposing tasks into subtasks. A recent survey [22] outlines the evolution of Hierarchical Reinforcement Learning (HRL), from early handcrafted hierarchies to the Options framework and goal-conditioned policies, highlighting how hierarchy can improve sample efficiency through temporal abstraction. In practice, HRL has enabled for more scalable robot behaviors. For example, Wang et al. [23] design a hierarchical policy for multi-UAV air combat, where a high-level planner selects maneuvers and lower-level controllers execute them. This yielded superior win rates against non-hierarchical approaches in simulation. Likewise, hierarchical policies have improved the navigation safety of ground robots. Zhu and Hayashibe [24] use a two-tier RL agent consisting of a waypoint planner and a motion controller to successfully navigate cluttered environments with fewer collisions.

D. Synthetic Data and Sim-to-Real Transfer

Closing the reality gap is necessary for the deployment of learning-based systems in field robots and UAV scenarios. One such approach is to train perception models on synthetic images produced from simulators. Guo et al. [25] follow this path using Unreal Engine / AirSim to produce a labeled dataset of UAV images and then test the performance of detectors learned in such synthetic images when actually deployed on real aerial images. These studies guide improvements in object detection in precision agriculture through more realistic and diverse simulations. On the control side, researchers employ high-fidelity simulation and domain randomization to facilitate the transfer of policies to real robots.

Kaufmann et al. [26] demonstrated a giant leap in sim-to-real transfer by training a drone racing policy in simulation to outperform human champions in real-world races. Their success was based on a photorealistic simulator and the stepwise increment of real-world effects, highlighting the power of contemporary simulators for high-speed UAV flight. Another research direction tackles the challenge of sim-to-real tuning automation. Du et al. [27] developed an auto-tuning approach that incrementally optimizes the simulator parameters so that they better represent reality, essentially enhancing policy transfer without a lengthy trial and error.

III. Problem Formulation and Environment Setup

A. Problem Formulation

We model the proposed UAV-based deer detection and deterrence system as a multi-agent Markov Decision Process (MMDP), formalized by the tuple $(\mathcal{S}, \mathcal{A}_1, \ldots, \mathcal{A}_N, P, R, \gamma)$, where:

- $\mathcal{S}$ denotes the global state space, which includes UAV positions, battery levels, deer detections (from onboard vision), time of day, and current deterrence status.
- $\mathcal{A}_i$ is the action space of the agent i, including decisions such as continuing along the pre-assigned path, deviating to track a deer, or activating deterrence mechanisms (e.g., sound or lights).
- $P : \mathcal{S} \times \mathcal{A}_1 \times \cdots \times \mathcal{A}_N \to \mathcal{S}$ is the state transition function determined by the joint actions of all agents and the dynamics of the environment.
- $R : \mathcal{S} \times \mathcal{A}_1 \times \cdots \times \mathcal{A}_N \to \mathbb{R}$ is the global reward function that encourages effective deer deterrence, area coverage, and energy efficiency.
- $\gamma \in [0, 1)$ is the discount factor.

B. Agent Observations, Actions, and Rewards

Each UAV operates as an autonomous agent with partial observability, receiving local sensory input and onboard detections. The observation space o_i for agent i at time t includes the UAV's own GPS position and velocity, relative positions and velocities of nearby detected deer (from YOLO-based detection and tracking), the battery status as a normalized scalar (0 to 1), whether a deterrence action was recently triggered (boolean), the distance to an assigned patrol waypoint, and a local occupancy map representing other nearby UAVs or obstacles (if within communication range).

The action space $\mathcal{A}_i$ consists of discrete high-level decisions:

- `Patrol`: continue following the assigned path.
- `Pursue`: deviate to follow and monitor a detected deer.
- `Deter`: activate the deterrence mechanism (e.g., sound or light).
- `Return`: trigger return-to-home (RTH) behavior when the battery is low.

The reward function r_t is shaped to encourage effective behavior and cooperation:

$$R_t = \alpha_1 \cdot D_t - \alpha_2 \cdot E_t - \alpha_3 \cdot R_t + \alpha_4 \cdot C_t \tag{1}$$

where:

- D_t: Deter success (e.g., distance a deer moves away after deterrence),
- E_t: Energy consumption (flight + deterrence),
- R_t: Redundancy penalty (e.g., multiple UAVs engaging the same deer),
- C_t: Coverage gain (novel area patrolled).

The reward coefficients $(\alpha, \beta, \delta, \gamma)$ are tuned to balance deterrence performance with energy efficiency and coverage.

IV. Hierarchical Multi-Agent UAV Architecture

In this paper, we propose a hierarchical multi-agent drone system optimized for efficient wildlife tracking and active animal deterrence, in this instance, for deer populations to minimize agricultural losses. The system includes three distinct, yet interconnected levels: a High-Level Planner, a Mid-Level Reinforcement Learning (RL) Controller, and a Low-Level Detection and Precision Landing Controller.

A. High-Level Planning Layer

The top layer computes exhaustive area-coverage trajectories algorithmically using grid-coverage methods or path-planning coverage algorithms like Rapidly Exploring Random Trees (RRT*) [28]. The layer gives each drone a named

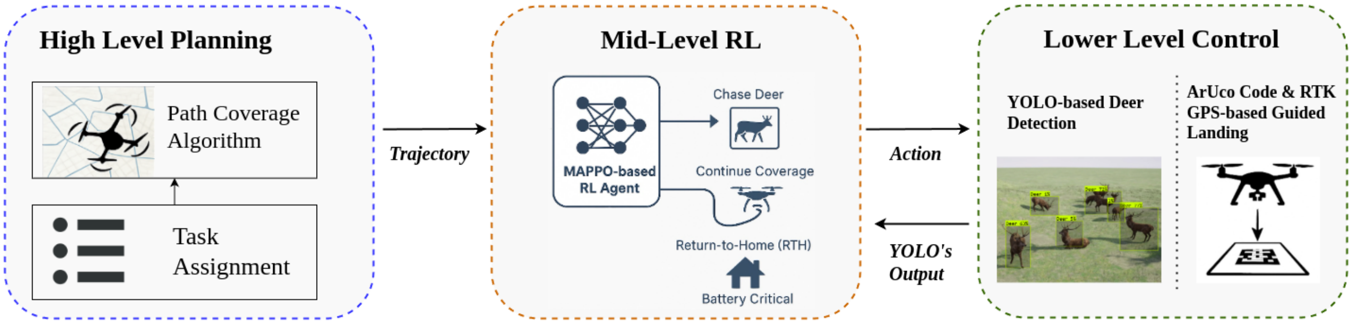


Fig. 1: Hierarchical drone architecture illustrating the interactions among High-Level Planning, Mid-Level Reinforcement Learning Controller, and Low-Level Detection and Landing functionalities.

trajectory or waypoint and optimizes spatial coverage according to mission objectives and environmental restrictions. It guarantees high-level mission consistency while lowering computational overhead in the lower layers by abstracting the global mission to a formal set of paths.

The high-level controller includes a task allocation mechanism that can assign specific subtasks (e.g., surveying a subregion) to the individual drones, which can be done manually with Human-in-the-loop or Algorithmically (e.g., Auction-based [29]) where an auctioning algorithm (or another decentralized coordination scheme) automatically distributes tasks by optimizing metrics such as fuel consumption, time-to-completion, or priority level.

B. Mid-Level Reinforcement Learning Controller

We provide a decision-making module that uses Multi-Agent Proximal Policy Optimization (MAPPO) at the intermediate level. The lower-level controller provides real-time deer detection feedback to the RL controller, while the high-level layer provides trajectory waypoints. These inputs and drone conditions, such battery levels, are used to dynamically choose drone activities. In particular, the RL controller can:

- Chasing observed deer: When deer are spotted, chase behavior is activated to aggressively keep animals away from important places.
- Continuing coverage: Keeping an eye out and maximizing coverage efficiency while adhering to recommended routes when no deer are found.
- Return-to-Home (RTH): When drone battery levels get close to critical thresholds, precision landing maneuvers are initiated for recharging.

C. Low-Level Detection and Precision Landing Layer

The lowest layer incorporates two critical functions: animal detection and precision landing. YOLO (You Only Look Once), a high-performance real-time object detection model, is used to accurately identify deer from drone-acquired imagery. Detection outcomes are relayed in real time to the mid-level RL controller to select the preferred deterrence method.

Additionally, this layer handles drone landing and recharging with centimeter-level precision. We achieve precise landings using Real-Time Kinematic (RTK) GPS, augmented with visual guidance through an ArUco marker-based landing system. When the battery reaches a critically low level, an RTH and autonomous landing mode are initiated from the mid-level controller. The system utilizes RTK to position itself above the charging station with centimeter-level precision so that the charging station is visible and the need to search the environment for the station is avoided. The charging station is equipped with an ArUco marker, which assists in the system's ability to precisely locate the desired landing point. A key benefit of using the ArUco markers is that they allow us to robustly determine the position and orientation of the drone relative to the charging station, helping to ensure that the drone does not land in a position that leaves it unable to charge. As the drone descends towards the desired landing position, it updates its estimated position based off the camera feed of the ArUco marker, enabling it to make minuscule adjustments to perturbations that would not have been detectable using GPS alone. This design also helps improve performance in adverse conditions, such as in the presence of gusty winds or temporary loss of GPS.

D. Integrated System Operation

The precision and dependability of the most recent detection and landing techniques, the adaptive decision-making power of MAPPO-based RL, and the resilience of traditional path planning approaches are all combined in this hierarchical system that divides tasks across layers. The architecture facilitates smooth layer integration, guaranteeing mission efficiency and dependability while enabling real-time adaptation to changing environmental circumstances.

V. Synthetic Data Generation and Training

A. Simulation Environment

To enable high-fidelity environmental behaviors and physics, we configured our simulation framework using Unreal Engine 4.27.2 in conjunction with the AirSim plugin. To enable different simulation situations, our replicated agricultural setting has a variety of topography, vegetation, and lighting conditions. Real drone action is simulated by sensor models such as RGB cameras for vision-based perception. The behaviors of deer, which follow a leader-follower model, were accurately imitated. In order to ensure realistic herd

Fig. 2: A snapshot of the simulated deer herd exhibiting diverse behaviors such as grazing, standing, and resting, mimicking realistic wildlife dynamics for training.

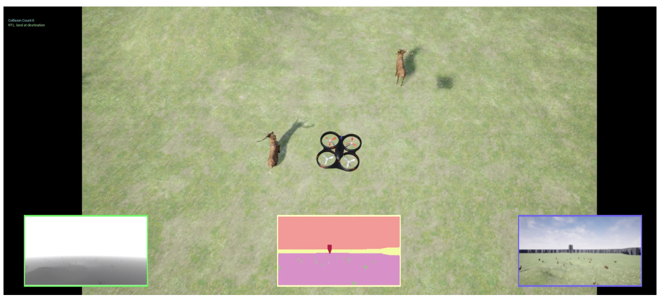


Fig. 3: Top-down view of a UAV detecting and responding to deer in the simulated farmland environment. Images show depth, semantic, and RGB visual streams used for detection and navigation.

movement, the lead deer moves randomly across the environment, while the following deer engage in random behaviors like grazing and resting but always stay close to the leader.

B. Synthetic Data Generation Approach

We used a modular approach to generate synthetic farmland and wildlife scenarios. A master deer character was defined with parameters that can be modified to determine the size of the herd. Upon simulation startup, follower deer are procedurally created from the master character, with randomized positions, orientations, and scales to simulate real-world variations in a herd. Herd behaviors included roving, resting, and feeding, plus responses to simulated threats such as the presence of UAVs or other intrusions. Visual realism was enhanced by attention to animation design, realistic texture, lighting, and environmental settings such as dynamic weather and plant variation, thus offering robust synthetic data sets for visual detection model learning and RL policy learning.

C. Sim-to-Real Transfer Strategies

To bridge the gap between sim-to-real, we will use domain randomization techniques such as variations in lighting, deer spawning locations, vegetation density, and wind speed and direction. which were systematically randomized to improve generalizability. In addition, sensor noise and camera parameters were randomized to replicate real-world imperfections. Limited real-world data collected from field trials will be used to validate the trained model on synthetic data.

VI. Experimental Setup and Preliminary Findings

A. Experimental Setup

There are three UAVs operating in designated patrol zones on a simulated farming area of around 100 hectares were used in the initial trials, which were conducted utilizing the simulation system provided. In order to replicate actual deployment settings, UAV agents were first setup with randomized battery levels and beginning locations. The tests were designed to assess UAV performance in a variety of situations, such as:

- Environmental conditions: Environmental factors include the cycles of day and night as well as different weather conditions like wind speed and illumination.
- Deer herd variability: At the beginning of each episode, the herd sizes were randomly assigned and ranged from 5 to 15.
- Operational constraints: Specified battery levels that initiate return-to-home (RTH) procedures for UAVs.

A camera-based precision landing system was used using ArUco markers. The landing system was implemented using a Proportional-Integral controller (PID) that interfaced with the PX4 Autopilot software running on the Pixhawk flight controller. The camera feed then guided the quadcopter's position relative to the ArUco marker and sent an updated goal location to the PX4 Autopilot. The landing experiments were done in Gazebo with ROS2.

To simulate the real environment, we took some ideas about how to insert the deer into the environment and make it to react when the drone gets close. Therefore, when the environment starts, the deer will start at different levels at random, including resting, feeding, and shifting. We also set the alert threat system to the deer character that led them to escape when they got an alert from the drone. The extension system helps the simulation become more real since the animal can run away from any unknown item; however, it also enhances the challenge of detection since it requires long-range detection from the camera, so the drone would not need to reach the deer closely and detect them.

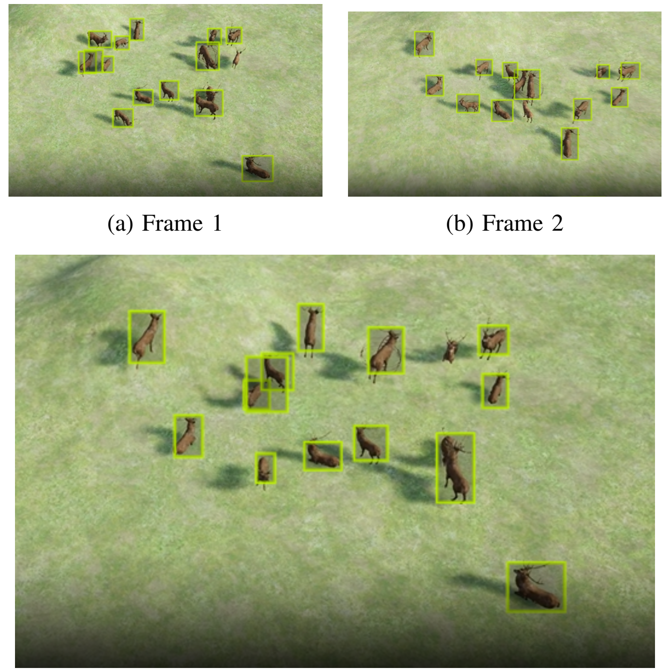
In the simulation, the AI Engine provided the deer model with functionality to control the deer movement (with Inverse Kinematics setting), resting, and gazing behaviors. Unreal Engine provides us with Non-player character (NPC) behaviors, which randomize the movements or standing and activate fast movement when the setting interface becomes close to the deer. Otherwise, the deer will remain resting and will not move fast.

B. Preliminary Findings

The preliminary findings thus far show the effective detection capabilities of the trained UAV agents. In our custom simulation scenarios, UAVs were able to identify deer in real-time. The next step will be to coordinate responses to deer incursions as proposed in the above sections.

As shown in Figure 5, the UAV was able to identify multiple deer in our custom-built Unreal World simulation setup.

Fig. 4: An onboard perspective capturing the UAV's precision landing on the gazebo.



(a) Frame 1

(b) Frame 2

(c) Frame 3

Fig. 5: Simulation snapshots illustrating deer detection across multiple frames using the YOLOv5-based UAV agent. Consistent detection across perspectives supports robust multi-agent coordination and real-time response in virtual testbeds.

During live simulation, the model outputted bounding boxes with respective confidence scores. This is then further passed through a temporal-spatial filtering. The model gave a mean average precision (mAP @ 0.5) of 90.5% with a precision score of 93.9% and recall of 86.9%, which confirms a true detection within the simulation domain.

Using this detection pipeline, the UAV agents were able to identify deer incursions in near real time. This detection will then be used to trigger coordinated multi-agent responses.

VII. Conclusions and Future Work

In conclusion, this work develops the use of artificial data generation and suggests a systematic hierarchical reinforcement learning approach for effective deer deterrent using UAVs. Even though this study just started, the findings thus far point to a potentially novel approach to achieve autonomy in the deer detection and deterrent challenge. The near-term next stages include iteratively improving the model's performance and its reliability. This is a challenging problem that needs to be tested in the actual world. We believe that our proposed approach is a scalable strategy to lessen environmental damage and agricultural loss caused by deer.

ACKNOWLEDGMENT

We would like to thank Chris James, a farmer in Minnesota who collaborated with us. We also thank Mario Jerez, Resha Tejpaul, and Dr. Saad Bedros for their invaluable assistance during deployments. We also want to thank Code Like Me, who is a YouTuber, and sharing his instruction on how to set up and input animal characters into the Unreal Engine.

References

[1] H. C. J. Godfray, J. R. Beddington, I. R. Crute, L. Haddad, D. Lawrence, J. F. Muir, J. Pretty, S. Robinson, S. M. Thomas, and C. Toulmin, "Food security: the challenge of feeding 9 billion people," *science*, vol. 327, no. 5967, pp. 812–818, 2010.

[2] C. N. Boyer, L. Chen, G. Perez-Quesada, and S. A. Smith, "Unwelcomed guests: Impact of deer harvest on corn and soybean wildlife damage," *Crop Protection*, vol. 183, p. 106753, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0261219424001819

[3] T. R. Obermoller, A. S. Norton, E. S. Michel, and B. S. Haroldson, "Use of drones with thermal infrared to locate white-tailed deer neonates for capture," *Wildlife Society Bulletin*, vol. 45, no. 4, pp. 682–689, 2021.

[4] J. Cukor, J. Bartoška, J. Rohla, J. Sova, and A. Machálek, "Use of aerial thermography to reduce mortality of roe deer fawns before harvest," *PeerJ*, vol. 7, p. e6923, 2019.

[5] X. Li, H. Huang, A. V. Savkin, and J. Zhang, "Robotic herding of farm animals using a network of barking aerial drones," *Drones*, vol. 6, no. 2, 2022. [Online]. Available: https://www.mdpi.com/2504-446X/6/2/29

[6] N. Hahn, A. Mwakatobe, J. Konuche, N. de Souza, J. Keyyu, M. Goss, A. Chang'a, S. Palminteri, E. Dinerstein, and D. Olson, "Unmanned aerial vehicles mitigate human–elephant conflict on the borders of tanzanian parks: a case study," *Oryx*, vol. 51, no. 3, pp. 513–516, 2017.

[7] K. Liu, Z. Tang, D. Wang, Z. Wang, B. Zhao, and X. Li, "Coherent: Collaboration of heterogeneous multi-robot system with large language models," *arXiv preprint arXiv:2409.15146*, 2024.

[8] E. Bondi, D. Dey, A. Kapoor, J. Piavis, S. Shah, F. Fang, B. Dilkina, R. Hannaford, A. Iyer, L. Joppa *et al.*, "Airsim-w: A simulation environment for wildlife conservation with uavs," in *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, 2018, pp. 1–12.

[9] C. Yu, A. Velu, E. Vinitsky, J. Gao, Y. Wang, A. Bayen, and Y. Wu, "The surprising effectiveness of ppo in cooperative multi-agent games," *Advances in neural information processing systems*, vol. 35, pp. 24 611–24 624, 2022.

[10] S. James, P. Wohlhart, M. Kalakrishnan, D. Kalashnikov, A. Irpan, J. Ibarz, S. Levine, R. Hadsell, and K. Bousmalis, "Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks," in *Proceedings of the IEEE/CVF conference on*

computer vision and pattern recognition, 2019, pp. 12 627–12 637.

[11] R. Varghese and M. Sambath, "Yolov8: A novel object detection algorithm with enhanced performance and robustness," in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*. IEEE, 2024, pp. 1–6.

[12] V. Bapat, P. Kale, V. Shinde, N. Deshpande, and A. Shaligram, "Wsn application for crop protection to divert animal intrusions in the agricultural land," *Computers and electronics in agriculture*, vol. 133, pp. 88–96, 2017.

[13] D. Axford, F. Sohel, M. Vanderklift, and A. Hodgson, "Collectively advancing deep learning for animal detection in drone imagery: Successes, challenges, and research gaps," *Ecological informatics*, p. 102842, 2024.

[14] C. J. Crossling, H. Vadapalli, and D. van der Haar, "Automated pest control: Computer vision for wildlife surveillance," in *Proceedings of the 2024 7th International Conference on Computational Intelligence and Intelligent Systems*, 2024, pp. 38–43.

[15] H. Lyu, F. Qiu, L. An, D. Stow, R. Lewison, and E. Bohnett, "Deer survey from drone thermal imagery using enhanced faster r-cnn based on resnets and fpn," *Ecological Informatics*, vol. 79, p. 102383, 2024.

[16] S. Bhusal, M. Karkee, U. Bhattarai, Y. Majeed, and Q. Zhang, "Automated execution of a pest bird deterrence system using a programmable unmanned aerial vehicle (uav)," *Computers and Electronics in Agriculture*, vol. 198, p. 106972, 2022.

[17] S. Geerthik, B. Vishal *et al.*, "Iot-based system for humane animal deterrence and sustainable crop management," in *2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*. IEEE, 2025, pp. 770–775.

[18] E. Laguna, P. Palencia, A. J. Carpio, J. Mateos-Aparicio, C. Herraiz, C. Notario, J. Vicente, V. Montoro, and P. Acevedo, "Evaluation of a combined and portable light-ultrasound device with which to deter red deer," *European Journal of Wildlife Research*, vol. 68, no. 4, p. 50, 2022.

[19] T. Rashid, M. Samvelyan, C. S. De Witt, G. Farquhar, J. Foerster, and S. Whiteson, "Monotonic value function factorisation for deep multi-agent reinforcement learning," *Journal of Machine Learning Research*, vol. 21, no. 178, pp. 1–51, 2020.

[20] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," *Advances in neural information processing systems*, vol. 30, 2017.

[21] J. Boubin, C. Burley, P. Han, B. Li, B. Porter, and C. Stewart, "Marble: Multi-agent reinforcement learning at the edge for digital agriculture," in *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)*. IEEE, 2022, pp. 68–81.

[22] M. Hutsebaut-Buysse, K. Mets, and S. Latré, "Hierarchical reinforcement learning: A survey and open research challenges," *Machine Learning and Knowledge Extraction*, vol. 4, no. 1, pp. 172–221, 2022.

[23] H. Wang and J. Wang, "Enhancing multi-uav air combat decision making via hierarchical reinforcement learning," *Scientific Reports*, vol. 14, no. 1, p. 4458, 2024.

[24] W. Zhu and M. Hayashibe, "A hierarchical deep reinforcement learning framework with high efficiency and generalization for fast and safe navigation," *IEEE Transactions on industrial Electronics*, vol. 70, no. 5, pp. 4962–4971, 2022.

[25] C. Guo, Y. Yu, and J. Huang, "Learning from synthetic data for object detection on aerial images," in *International Conference on Image, Signal Processing, and Pattern Recognition (ISPP 2024)*, vol. 13180. SPIE, 2024, pp. 832–838.

[26] E. Kaufmann, L. Bauersfeld, A. Loquercio, M. Müller, V. Koltun, and D. Scaramuzza, "Champion-level drone racing using deep reinforcement learning," *Nature*, vol. 620, no. 7976, pp. 982–987, 2023.

[27] Y. Du, O. Watkins, T. Darrell, P. Abbeel, and D. Pathak, "Auto-tuned sim-to-real transfer," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 1290–1296.

[28] E. Galceran and M. Carreras, "A survey on coverage path planning for robotics," *Robotics and Autonomous systems*, vol. 61, no. 12, pp. 1258–1276, 2013.

[29] E. Nunes and M. Gini, "Multi-robot auctions for allocation of tasks with temporal constraints," in *Proc. of the AAAI Conference on Artificial Intelligence*, 2015, pp. 2110–2116.