# Deep Reinforcement Learning Approach for Prioritized Heterogeneous Object Tracking in UAV Swarms

1st B. Nithya
*Dept. of Computer Science and Engg.*
*National Institute of Technology*
Tiruchirappalli, India
nithya@nitt.edu

2nd Neethu Subash
*Dept. of Computer Science and Engg.*
*National Institute of Technology*
Tiruchirappalli, India
neethu.subash@gmail.com

3rd Hari Vignesh
*Department of Data Science*
*St.Joseph's College*
Tiruchirappalli, India
harivignesh3090@gmail.com

***Abstract*—Deep Reinforcement learning(DRL) has recently showcased its boundless potential in Multi-agent collaboration. There exist DRL-based models that strive to unravel the Unmanned aerial vehicle (UAV) Multi-target tracking predicaments, however, the conundrum of tracing desired targets persists as most research endeavors solely focus on the agents' policies while neglecting the dynamic environment, which harbors elusive and unpredictable targets. The proposed model executes Prioritized Heterogeneous Object Tracking (PHOT) using DRL, to furnish a solution for preferred Multi-target tracking by upholding the policy of UAV swarms. Tracking preferred targets assumes paramount importance as these targets necessitate meticulous monitoring within the ever-changing environmental backdrop, particularly during surveillance missions employing UAV swarms, which can harness the full potential of the UAV swarms in real-time. A YOLO v4-based algorithm is put forth, which can trail prioritized targets in a tracking by detection paradigm. To maintain the policy of UAV swarms, a dueling double deep Q-learning algorithm is deployed, complemented by a reward function that stimulates UAVs to track the preferred targets instantaneously. The UAVDT dataset is employed for the simulations, and the results evince that PHOT exhibits robustness, suitable for deployment in a real-time environment, amassing a highly commendable average cumulative reward over time, and enabling actions predicated on unidentical prioritized targets.**

***Keywords*—DRL, UAV, Multi-Target Tracking, YOLO**

## I. Introduction

In the recent years, the realm of Multi-agent collaboration employing DRL has achieved remarkable strides, particularly in the realm of UAVs which is harnessed in various domains for Traffic Surveillance and Monitoring, Rescue Expeditions, Agriculture, Photography and Product Conveyance. The technical progressions in UAVs are skyrocketing, paving the way for extensive employment across diverse domains. UAVs can be extensively deployed for monitoring missions, where tracking multiple targets in a dynamic milieu assumes great significance. Deep reinforcement learning approaches [1] render them more apt for UAVs than deep learning-based approaches, as they can adeptly learn from the inherently unpredictable environment. Although existing DRL-based investigations have tackled the UAV Multi-target tracking conundrum, the issue of effectively tracking desired targets in the face of dissimilar and uncertain targets persists. Most prior research has primarily focused on optimizing the agents' policies, disregarding the environment's dynamic nature and the imperative to track prioritized targets.

Tracking prioritized targets is essential for real-time surveillance missions, as it allows the UAV swarm to leverage its full potential by focusing on specific targets in the given dynamic context. This research proposes a YOLO v4 [2]based algorithm to address the preferred Multi-target tracking problem using UAV swarms which solves the object detection and tracking. By adopting the tracking-by-detection approach of YOLO v4, proposed PHOT can effectively track prioritized targets considering the presence of unidentical and uncertain targets. [3]unidentical targets refers to that the targets can look homogenous by it's appperance but it may not. Through providing ID for each of the detected objects identity of objects can be used for specific tracking purposes. To maintain the policy of UAV swarms, the dueling double [4] Deep Q-learning algorithm is adapted, which allows the agents to learn and update their policies efficiently. Markov decision process is adapted to make agent take actions according to the movement of specified targets. One critical element of the algorithm is developing a reward system that motivates UAVs to prioritize tracking targets in real-time [5]. This approach encourages the swarm to adjust its actions according to the importance of the targets leading to decision-making capabilities during surveillance operations.

Intricate threads of related endeavors unravel in Section 2 of the paper. Section 3 unveils the PHOT model, vividly portraying the prioritization and UAV policies. Section 4 unleashes a scorching beam of light on the analysis of results, illuminating the path ahead. Lastly, Section 5 summarizes the findings and tantalizes hints about future work that may take shape.

## II. Related Work

Through this literature review, the aim is to highlight the various strategies employed by researchers in UAV-based multi-target inspection. The review will highlight the strengths and limitations of existing methods, highlight the need for future work to address these challenges and extend the scope of the review to include real-world data and much more difficult. Furthermore, real-time monitoring is required in UAV applications, where timely decision-making is paramount. For this reason, the PHOT model contributes to developing prioritized planning to pursue many different objectives in real-time. This approach holds great promise to increase UAV-based multi-target tracking systems' overall performance and accuracy.

In paper [6] a method, for tracking targets in swarms of UAVs using deep reinforcement learning. This approach handles information through the Cartogram FR. While it achieves a good tracking ratio and scalability it assumes a controlled environment with targets, which limits its practical usefulness.The paper [7] introduces an algorithm for searching and tracking targets with UAVs while considering energy refueling. They use a quantum probability model and the upper confidence tree algorithm. With a 75% tracking ratio there is potential for difficulty with moving targets and larger search areas. Addressing path finding and target tracking in crowds the paper [8] utilizes an A* planner. Distributed deep neural network for multi-UAV scenarios. It reduces target search time successfully, although further work is required to handle targets. The paper [9] presents the Meta TD3 algorithm designed to track UAV targets in environments using deep reinforcement learning. It outperforms algorithms but acknowledges challenges when deploying in real-world situations. In paper [10], they propose a Context-aware Multi-task Siamese Network specifically for object tracking in UAV videos. This method surpasses existing approaches but lacks exploration of generalizability and analysis of complexity. The paper puts forward RTD Net—a real-time object detection network tailored for UAV vision employing CNN and Transformer models [11]. It can achieve high mAP (mean precision) and perform in real time but it still requires evaluation, in more complex scenarios. The work in [12] presents a control system for UAVs that uses DRL to track targets. The system demonstrates improved accuracy and efficiency compared to methods. It is sensitive to the range of communication. The paper [13] it proposes an approach for real-time detection and recognition of objects using a VTOL UAV with a fixed-wing design. This approach achieves accuracy and speed in detection. Is limited to a single UAV platform and requires testing under different conditions. The algorithm in [14] introduces a methodology that combines learning with UAV data for object detection and tracking. It achieves precision and recall; however, its focus is primarily on tracking objects lacking extensive real-world testing. Finally, the researchers propose a framework [15] specifically designed for accurate object tracking in UAV videos. The framework outperforms existing methods in managing ID switching; however, further work is needed to enhance its performance in tracking objects and crowded scenarios.

**Inferences**

- Existing methods mostly fail to adapt to complex, crowded scenarios.
- Tracking accuracy has been very less in existing models.
- Dynamic scenarios are challenging to the current models.

**Motivation**

- The proposed PHOT is capable of detection and tracking in complex environments.
- Prioritization algorithm is proposed, which could overcome the existing methods which assume homogenous targets in the environment.
- DRL empowers the efficient distribution of resources, guaranteeing that UAVs channel their endeavors towards objects that crave greater attention or possess superior significance, thus maximizing the exploitation of UAV resources.
- DRL models possess the remarkable ability to assimilate and acquire wisdom from novel information and encounters. This empowers the UAV swarm to navigate through ever-evolving situations, encompassing mobile targets, fluctuating environmental circumstances, and unforeseen impediments.

## III. The Proposed PHOT Model

The PHOT model, as envisioned, seeks to exalt and trail distinct targets that are dynamically ascertained while the UAV soars, while upholding the UAVs' policy to galvanize the pursuit of the prioritized targets. As the Workflow is shown in Fig. 1 firstly, the Yolo v4 model is trained using the UAVDT dataset for object detection and tracking. Prioritisation tracking algorithm 1 to track the specified object using YOLO v4. The policy of UAV is framed using a double-deep Q-learning network. According to the context of the targets, reward shaping and the Markov decision process are framed to encourage the agent to track the preferred targets. The DDQN agent is then made to train with UAVDT to learn the policy. The results of the PHOT model are compared with existing methods, which show a high average cumulative reward and high accuracy in tracking prioritised targets.

### A. Prioritization

*1) Training the Model:* The YOLOv4 algorithm, a renowned and potent deep learning-based technique for real-time object detection, hones the object detection model. Throughout the training phase, the model acquires the ability to discern and pinpoint objects in images or video frames by forecasting bounding boxes and their respective class probabilities.

*2) Object Detection:* The detection phase entails employing a YOLOv4 model to unearth objects in real-time video frames captured by the UAV's camera. As the UAV soars and seizes video frames, the YOLOv4 model scrutinizes each frame, revealing sundry objects that grace the scene. The model

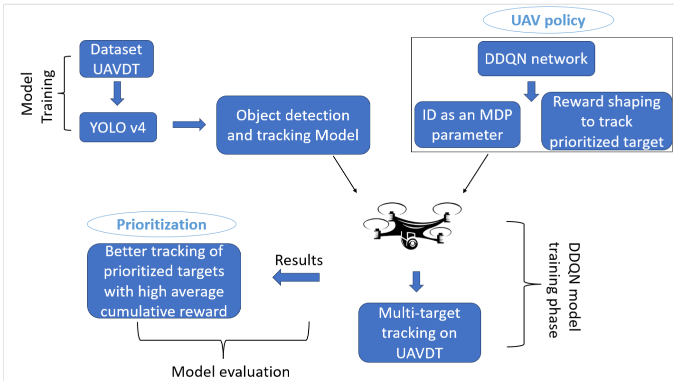


Fig. 1. Procedure of the proposed work.

excels at spotting cars, trucks, or other predetermined classes that compose the training data.

YOLOv4, a distinguished algorithm for real-time object detection, boasts celerity and precision in equal measure. It cleaves the input image into a grid, prophesying bounding boxes and class probabilities for every grid cell. This endows it with efficiency in real-time applications like UAV-rooted object detection. During the detection phase, the YOLOv4 model successively processes each video frame, pinpointing the whereabouts and categories of objects that permeate the frame. The customary output of the detection process encompasses the coordinates of bounding boxes encircling the detected objects and their corresponding class labels and confidence scores. Armed with this information, the UAV can make informed decisions or undertake actions predicated on the objects that pervade its milieu. For instance, it can trace the course of vehicles, surveil traffic, or detect and respond to potential hazards or anomalies in the scene. It is crucial to underscore that the veracity and efficacy of the YOLOv4 model during real-time object detection hinge on factors such as the caliber of the training data, the intricacy of the objects under scrutiny, and the hardware resources procurable for inference on the UAV.

*3) Workflow of the PHOT Model:* The proposed PHOT Algorithm 1 cascades forth to unravel the enigma of object detection and tracking within the vivid tapestry of video frames. It diligently scrutinizes each frame of the video sequence and instigates the ignition of variables to track these elusive entities. The algorithm embarks on a relentless odyssey through the vast expanse of frames, deftly discerning objects through the prism of bounding boxes(bboxes) and meticulously configuring these boxes to encapsulate their respective targets. It deftly employs the measure of the distance between the center point (c) of the current frame and the center point (c') of the preceding frame to ascertain the uninterrupted continuity of the objects. If this distance be found to be less than 20 pixels, and the count be no greater than 2, it dutifully bestows upon the object a distinctive identifier (id) and promptly advances the numerical value of id. For subsequent frames, if the distance proves to be less than

**Algorithm 1:** Proposed PHOT Algorithm

$V \leftarrow$ Videoframes;
Initialize $id$ as 0;
Initialize $count$ as 0;
Initialize an empty array $arr$;
**while** *true* **do**
$bboxes \leftarrow$ Detect objects in $V$;
**for** *each $bboxes$* **do**
Configure bounding box for object $k$;
**if** *$count \leq 2$* **then**
Compare distance $d$ between center point $c'$ of previous frame and $c$ of current frame;
**if** *$d < 20$ pixels* **then**
Configure $id$;
Increment $id$ value by 1;
**end**
**end**
**else**
Compare distance $d$ between center point $c'$ of previous frame and $c$ of current frame;
**if** *$d < 20$ pixels* **then**
Retain the previously configured $id$;
Increment $id$ value by 1;
**end**
**else**
New $id$;
**end**
**end**
$Flag \leftarrow 0$;
**if** *key pressed* **then**
$ids \leftarrow$ {user input};
$Flag \leftarrow 1$;
**end**
**if** *$Flag = 1$* **then**
**for** *$id$ in $ids$* **do**
Configure center point and bbox for the $id$;
**end**
$arr \leftarrow$ [save $ids$];
**end**
**end**
**end**

20 pixels, it maintains the same id while incrementing its value for next object; conversely, if the distance exceeds this threshold, a fresh id is promptly assigned. Suppose a key is pressed, thus signifying the user's intervention. In that case, the algorithm graciously allows custom ids to be bestowed upon specific objects, which are then preserved within the cherished confines of the array. In a ceaseless display of alacrity, the algorithm tirelessly processes frames, compares the center points of objects, and accordingly bestows upon them their rightful ids. The ids are then saved so that it can be used by the DDQN agent to make decision. The algorithm

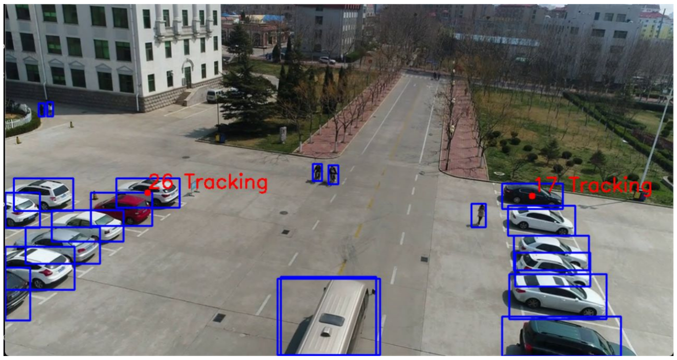


Fig. 2. Prioritization using Prioritization algorithm.

works until the it doesn't get a videoframe as input.

*4) Object Tracking with Prioritization:* Rather than simultaneously monitoring all identified objects, PHOT can selectively focus on specific targets, allowing users to handpick their desired subject for tracking. Upon pressing the spacebar at any given moment during runtime, the UAV's tracking algorithm shall patiently await the entrance of the ID. As if on cue, an input prompt shall manifest on the screen, beseeching the individual to input the target's unique ID, thus initiating the object's pursuit.

1) User Input and Target Selection: By giving the target ID as input to the algorithm 1, specifies the choice of object/objects that is to be prioritized for tracking. The target ID and the detection ID given during the detection by the YOLOv4 model are the same. Enter "2" when requested if the YOLOv4 model identifies three objects with IDs 1, 2, and 3, and ID 2 is the one that is explicitly intended to monitor. Similarly, you can prompt with "2, 3" if several targets are detected.
2) Tracking the Preferred Target: After receiving the user's input, the algorithm executes and track the object, thus it will focus its attention on tracking the specified target as shown in Fig. 2 (e.g., object of ID 2). The tracking algorithm will continuously follow the selected object, with real-time updates on its position.
3) Further Detection and Tracking: As the UAV continues its flight, the YOLOv4 model will keep detecting objects in each frame, and the tracking algorithm will update the position of the preferred target as it moves. The process continues iteratively until another object ID is chosen.

### B. UAV Policy

*1) Dueling Double Deep Q-Network:* The Dueling Double (DQN) as shown in Fig. 3 emerges as a modified rendition of the Deep Q Network (DQN) algorithm. Its purpose is to enrich the efficacy and steadiness of value-based reinforcement learning. Instead of merely utilizing a neural network, this approach bifurcates it into two distinctive fragments: the value stream and the advantage stream. The value stream prophesies the anticipated value of existence within a state, while the advantage stream assesses the comparison between each action

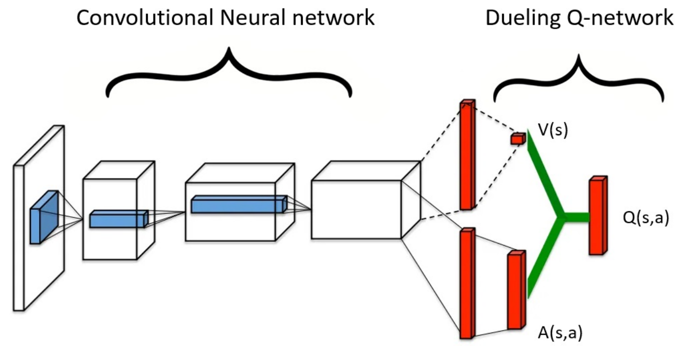


Fig. 3. Architecture of Dueling Double deep q-learning network.

and the action value. These two streams amalgamate within a stratum, which computes the Q values by compounding both flows. This entitles the algorithm to assimilate knowledge regarding both state values and action advantages simultaneously. Dueling Double DQN integrates Q learning to magnify stability by utilizing two neural networks to approximate Q values. This mitigates the commonly encountered overestimation bias in DQN and Q learning methodologies. By amalgamating dueling and double Q learning, Dueling DQN amplifies learning efficiency, diminishes overestimation bias, and establishes itself as a dependable algorithm for tackling intricate reinforcement learning tasks, such as training UAV swarms from the ground.

*2) Markov decision process(MDP):* States, actions, transition probabilities, rewards, and a discount factor characterize MDP. The core tenet of MDP lies in the Markov property, where the future state hinges solely on the present state and the chosen action, detached from past states and actions. When it comes to UAV object detection and tracking with deep reinforcement learning, MDP takes center stage in modeling the sequential decision-making process of the UAV agent. The states embody the UAV's observations, be it images or sensor data, while actions mirror the array of movements the UAV can execute. By harnessing the power of deep reinforcement learning with MDP, the UAV agent can discern optimal choices for object detection and tracking, deftly navigating the environment while enhancing its tracking prowess based on the rewards it receives. Thus, MDP empowers the UAV to deftly handle partial observations, learn from its experiences, and adapt its tracking strategies in response to the ever-evolving environment. The UAVs rely solely on localized interaction garnered from their neighboring entities, and the parameters of the MDP model unfurl as follows: (N, S, G, Z, L, T) where:

- $N$ is the set of $n$ agents,
- $S$ is the global state space and defined by $s \in S$,
- $G$ is the communication graph for the UAV swarm,
- $L$ denotes the interaction set that collects all UAVs' local interactions $l \in L$,
- $T$ denotes the state possibility transition model from state $s$ to the next state $s' \in S$.

For any local interaction by the UAV, it should have the observation set which is defined by

$$o(i,k) = (x_T^{(k)}, y_T^{(k)}, v_{xT}^{(k)}, v_{yT}^{(k)}, \mathrm{id}_T^{(k)}) \tag{1}$$

In (1), $v_x$ and $v_y$ denote the velocity of target in $x$ and $y$ direction respectively. A new parameter id of the target $k$ is added where it can be obtained during runtime of UAV.

*3) Reward function:* The Dueling Double Deep Q-Learning algorithm relies heavily on the reward function, imparting vital guidance that steers learning. This function assigns a numerical value to each state-action pair, indicating the immediate rewards or penalties the agent incurs when executing a specific action in a given state. The Dueling architecture deftly separates the value function and advantage streams. Within this framework, the reward function assumes a pivotal role in ascertaining the significance of different actions' advantage values in relation to their state values. This, in turn, empowers the algorithm to deftly estimate Q-values for myriad actions and make well-informed decisions throughout the learning process.

Each UAV must go beyond proficient tracking of previously observed targets to effectively pursue multiple targets. It must exhibit dynamic tracking capabilities to cater to the ever-evolving prioritization needs. Additionally, the UAV should adhere to a policy that aligns with the context of multi-target tracking, ensuring seamless cohesiveness in its operations. Reward shaping as follows :

$$r_{\text{tar}}^{(i,k)} = \begin{cases} 0, & \text{else} \\ 1, & \text{id(k)} \end{cases} \tag{2}$$

Hence, the reward function (2) sculpts the goals of the agent to trail the coveted targets. The condition id(k) stems from the observation parameter of MDP, where k denotes the observed target. If the agent mislay the target, it receives nil regarding reward.

## IV. Result Analysis

The UAVDT dataset stands as a vast and intricate collection utilized to educate and scrutinize object detection (DET), solitary object tracking (SOT), and multi-object tracking (MOT) algorithms. It encompasses 80,000 frames, meticulously extracted from 10 hours of unprocessed videos captured via a UAV platform.

- Dataset Magnitude: The dataset's dimensions encompass a staggering 80,000 frames, substantiating its adequacy in effectively molding deep learning models.
- Video Sequences: The dataset comprises 100 video sequences. These sequences are judiciously cherry-picked from the 10 hours of UAV footage, procured from diverse locations. This curation ensures an extensive array of environments and scenarios, thereby amplifying the dataset's comprehensiveness.
- Data Typology: The dataset encompasses annotations for object detection, solitary object tracking, and multi-object tracking missions. Ergo, it offers ground truth labels for both the spatial coordinates and categorical classification of the objects, alongside the trajectories of mobile entities for comprehensive tracking evaluations.
- Formidable Predicaments: Due to the presence of numerous confounding variables, including but not limited to, erratic weather patterns, mercurial illuminative circumstances, intricate backdrops, and occlusions, this dataset is regarded as a demanding and formidable adversary. These trials and tribulations facilitate meticulously assessing the detection and tracking algorithms' resilience and adaptability.

In computer vision, particularly within the bounds of UAV applications, the UAVDT dataset serves as a valuable resource for researchers and developers. It empowers them to educate and appraise their object detection and tracking algorithms within the parameters of real-world UAV scenarios. The effective utilization of this dataset allows for the measurement of performance and the comparison of the efficacy of various methodologies. It is paramount to acknowledge that the efficacy and diversity of the dataset are pivotal in the triumph and generalization of machine learning models. With its colossal size and a plethora of frames extracted from a diverse range of environments, the UAVDT dataset is an invaluable asset for advancing research in UAV-based computer vision tasks.

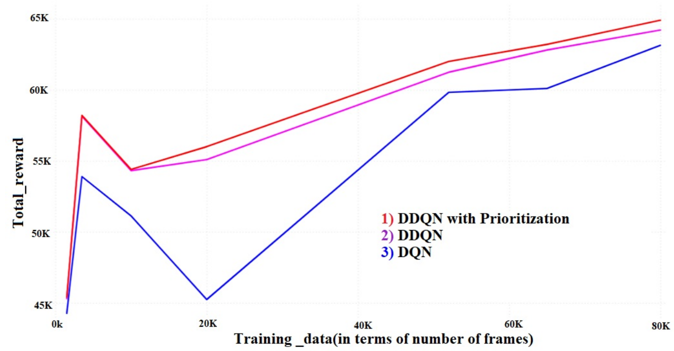


Fig. 4. Comparision between Total reward obtained by DQN and DDQN with and without prioritization.

### A. Deep Q-learning vs Dueling Double Q-learning with and without prioritization

Opting for the Dueling Double Deep Q-Learning (Dueling DDQN) network, with prioritization of targets, for the training of the UAVDT dataset is a cunning choice, rooted in its superiority over other algorithms like DQN and DDQN. The training process unveils remarkable cumulative rewards Fig. 4, showcasing the efficacy of this approach in tackling your specialized mission: detecting and tracking objects within UAV-based scenarios.

### B. Cumulative reward Comparision

In this proposed PHOT model, cumulative reward flourishes as it expands shown in Table I, for the reward function cunningly shapes the UAVs' progress. They are bestowed a

TABLE I. CUMULATIVE REWARD COMPARISON

| No of UAVs/Targets | Reward [1] | Reward (Proposed) |
|---|---|---|
| 5/5 | -2.5387 | -1.50 |
| 10/10 | 47.3673 | 63.01 |
| 20/20 | -210.3694 | 20.53 |


Fig. 5. Preferred targets vs Tracked targets.

positive reward of 1 when they diligently pursue the prioritized target. Moreover, this very process facilitates the scalability of UAV swarms. Also, the assumption of the homogenous target is not considered, thus allowing us to track Multiple targets at the convenience by providing the specific target's or it's id.

### C. Accuracy of Tracking Prioritized targets

The agent's tracking prowess, as revealed by the 90% accuracy in capturing the desired or designated targets Fig. 5. Throughout 500 episodes, it never fails to snatch them, showcasing its remarkable ability to learn and apply its tracking skills with utmost proficiency.

- Accuracy: Achieving an accuracy of 90% is a momentous triumph, particularly in endeavors like object tracking, where the surroundings can be intricate and demanding. It demonstrates the agent's ability to accurately discern and pursue its desired targets.
- Training Episodes: The agent's attainment of this level of precision in just 500 episodes astounds, and signifies an optimal and potent training process. It insinuates that the reinforcement learning algorithm swiftly assimilates from the bestowed rewards and experiences, rapidly mastering its craft.
- Preferred/Specified Targets: The agent's knack for tracing favored or designated targets is crucial for practical applications. Ranking particular targets according to pre-established criteria enables more astute and deliberate decision-making in real-world scenarios.

## V. Conclusion

The potential of UAVs to detect and track objects can be leveraged undoubtedly, but potential real-world outcomes and difficulties are challenging. The proposed PHOT model prioritizes the given/preferred target through the Prioritized Tracking algorithm and works with good accuracy in tracking the preferred targets by maintaining the policy of UAVs with Identity as a parameter in Markov Decision Process. Reward shaping is done to encourage the agent to track the preferred target. The real-time difficulties of deploying UAVs in traffic and surveillance missions can be reduced. The most important aspect of deploying a UAV swarm in a dynamic environment is that the accuracy of the model varies unexpectedly as there are uncertain scenarios may happen. The future work will be focussed on this ground where actual deployment and performance analysis of the proposed model in an environment with complex real-world scenarios.

## Acknowledgment


The authors acknowledge Science and Engineering Research Board (SERB), a statutory body of the Department of Science and Technology (DST), Government of India who provided financial support for this research.


## References


[1] [3] D. Han, H. J. Kim, and W.-J. Chung, "Deep reinforcement learning with double Q-learning and prioritization for aggressive driving in autonomous vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 1, pp. 37

[2] Y. Chen, C. Li, and J. Wang, "Deep reinforcement learning for robot navigation in unknown environments," IEEE Transactions on Robotics, vol. 37, no. 4, pp. 1210–1223, 2021.

[3] Z. Zhang, X. Wang, W. Li, and R. Wang, "YOLOv3-based multi-object tracking with CNN feature fusion and data association"," IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 2, pp. 787–797, 2019.

[4] Y. Hen, C. Li, and J. Wang, "Deep reinforcement learning for robot navigation in unknown environments," IEEE Transactions on Robotics, vol. 37, no. 4, pp. 1210–1223, 2021. Transactions on Robotics, 37(4), 1210–1223.

[5] Z. Zhang, X. Wang, W. Li, and R. Wang, "Real-time multi-object tracking with YOLOv3"," IEEE Transactions on Intelligent Transporta-Tion Systems, vol. 21, no. 11, pp. 5045–5054, 2020.

[6] W. Zhou, Z. Liu, J. Li, X. Xu, and L. Shen, "Multi-target tracking for unmanned aerial vehicle swarms using deep reinforcement learning," Neurocomputing, vol. 466, pp. 285–297, 2021.

[7] T. Wang, R. Qin, Y. Chen, H. Snoussi, and C. Choi, "A reinforcement learning approach for UAV target searching and tracking," Multimed. Tools Appl., vol. 78, no. 4, pp. 4347–4364, 2019.

[8] S. H. Silva, P. Rad, N. Beebe, K.-K. R. Choo, and M. Umapathy, "Cooperative unmanned aerial vehicles with privacy preserving deep vision for real-time object identification and tracking," J. Parallel Distrib. Comput., vol. 131, pp. 147–160, 2019.

[9] B. Li, Z. Gan, D. Chen, and D. Sergey Aleksandrovich, "UAV maneuvering target tracking in uncertain environments based on deep reinforcement learning and meta-learning," Remote Sens. (Basel), vol. 12, no. 22, p. 3789, 2020.

[10] H. Yu et al., "The unmanned aerial vehicle benchmark: Object detection, tracking and baseline," Int. J. Comput. Vis., vol. 128, no. 5, pp. 1141–1159, 2020.

[11] T. Ye, W. Qin, Z. Zhao, X. Gao, X. Deng, and Y. Ouyang, "Real-time object detection network in UAV-vision based on CNN and transformer," IEEE Trans. Instrum. Meas., vol. 72, pp. 1–13, 2023.

[12] J. Moon, S. Papaioannou, C. Laoudias, P. Kolios, and S. Kim, "Deep reinforcement learning multi-UAV trajectory control for target tracking," IEEE Internet Things J., vol. 8, no. 20, pp. 15441–15455, 2021.

[13] S. Sonkar, P. Kumar, R. C. George, T. P. Yuvaraj, D. Philip, and A. K. Ghosh, "Real-time object detection and recognition using fixed-wing LALE VTOL UAV," IEEE Sens. J., vol. 22, no. 21, pp. 20738–20747, 2022.

[14] A. A. Micheal, K. Vani, S. Sanjeevi, and C.-H. Lin, "Object detection and tracking with UAV data using deep learning," J. Ind. Soc. Remote Sens., vol. 49, no. 3, pp. 463–469, 2021.

[15] A. A. Micheal and K. Vani, "Deep learning-based multi-class multiple object tracking in UAV video," J. Ind. Soc. Remote Sens., vol. 50, no. 12, pp. 2543–2552, 2022.