JSST
Journal of Space Science and Technology

P-ISSN: 2008-4560
E-ISSN: 2423-4516

Journal of Space Science and Technology

Vol. 18, Special Issue, pp. 1-15, 2025
https://doi.org/10.22034/jsst.2025.1527
Journal Homepage: https://www.jsstpub.com

Original Research Paper

Cyber Risk Prediction for UAVs in Space-Related Missions Using Deep Reinforcement Learning

Erfan Khosravian[1*] and Motahareh Dehghan[2]

1. Faculty of Mechanical Engineering, Payame Noor University, Tehran, Iran
2. Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran

ARTICLE INFO

Article History:

Received 21 January 2025
Revised 07 February 2025
Accepted 09 February 2025
Available Online 01 March 2025

Keywords:

Space-related missions
Satellites
Unmanned aerial vehicles (UAV)
Cyber risk prediction
Deep reinforcement learning
Feature importance

ABSTRACT

Space exploration and satellite deployment drive modern technological advancements. They are crucial for global communication, navigation, and scientific discovery. Satellites form the backbone of interstellar communication, ensuring reliable data transfer in both civilian and defense sectors. However, as space missions grow more complex, maintaining their integrity and security becomes a major challenge. Unmanned aerial vehicles (UAVs) play a key role in space missions. They assist in satellite deployment, orbital inspections, and inter-satellite communication. Yet, these cyber-physical systems face evolving cybersecurity threats that could jeopardize mission-critical tasks. Traditional intrusion detection systems struggle to counter the complex and dynamic cyber threats targeting UAVs in harsh space environments. This paper introduces a novel Deep Reinforcement Learning model to predict and mitigate cyber risks in space-related UAV missions. Using a publicly available dataset that combines cyber and physical UAV data, the model predicts multi-step threats such as Denial of Service, Replay, Evil Twin, and False Data Injection. This enables proactive threat mitigation. Compared to traditional machine learning models—Support Vector Machines, Random Forests, and Recurrent Neural Networks—the proposed model achieves superior performance, with 99.34% accuracy and an AUC score of 0.99.

*Corresponding Author's E-mail: erfankhosravain@pnu.ac.ir

How to Cite this Article:
E. Khosravian and M. Dehghan, "Cyber risk prediction for UAVs in space-related missions using deep reinforcement learning," *Journal of Space Science and Technology*, Vol. 18, Special Issue, pp. 1-15, 2025, https://doi.org/10.22034/jsst.2025.1527.

1. Nomenclature & Units

DRL	Deep Reinforcement Learning
FDI	False Data Injection
DoS	Denial Of Service
IDS	Intrusion Detection System
UAV	Unmanned Aerial Vehicle
FDI	False Data Injection
AUC	Area Under Curve
SHAP	Shapley Additive Explanation
APT	Advanced Persistent Threat
SVM	Support Vector Machine
RNN	Recurrent Neural Network
RF	Random Forest
H-LSTM	Hierarchical Attention-Based Long-Short Term Memory
ConvLSTM	Convolutional LSTM
CNN-LSTM	Convolutional Neural Networks- Long Short Term Memory
WRELM	Weighted Regularized Extreme Learning Machine
QIWO	Quantum Invasive Weed Optimization
DQN	Deep Q-Network
WLAN	Wireless Local Area Network
UDP	User Datagram Protocol
MDP	Markov Decision Process
FIA	Feature Importance Analysis

2. INTRODUCTION

Space has always been of strategic importance, encompassing satellite deployment, planetary exploration, and interstellar communication. Satellites play a vital role in global connectivity, navigation, weather forecasting, and defense [1]. As nations and private enterprises expand space ventures, mission success becomes paramount. However, the harsh and dynamic space environment poses unique challenges, demanding advanced technologies for seamless and adaptive operations [2].

UAVs are key enablers of space missions. They support satellite deployment, facilitate real-time communication, and inspect or maintain orbiting structures. Unlike ground-based systems, UAVs offer agility, precision, and autonomy to navigate complex space environments. Their ability to operate autonomously in zero gravity enhances mission efficiency while reducing human intervention in high-risk scenarios. However, integrating UAVs into space operations introduces cybersecurity challenges [3].

In space missions, UAVs link satellites and ground infrastructure, ensuring uninterrupted communication and data flow. Any cyber threat could disrupt satellite alignment, interfere with telemetry, or even compromise entire missions [4]. UAVs, as complex cyber-physical systems, require robust security solutions addressing both cyber and physical vulnerabilities. As cyber threats grow, securing UAVs becomes essential for their safe and effective use in critical applications [5].

Cybersecurity threats targeting UAVs are becoming increasingly complex. Multi-step attacks such as DoS, Replay, Evil Twin, and FDI can lead to loss of control, data integrity breaches, and even mission failure [6]. While IDSs exist to counter these threats, many are limited in effectiveness. Most IDSs rely on small datasets and struggle to generalize across different threat types [7]. The lack of publicly available cyber-physical UAV data further hinders the development of robust IDS solutions. These challenges highlight the need for advanced models that not only detect threats but also predict their evolution [8].

This paper introduces a predictive model designed to forecast the next step in multi-step cyber threats on UAVs. Unlike traditional IDSs that merely detect threats, our model anticipates how an attack may progress—such as a DoS attack escalating to replay or FDI. To achieve this, we utilize DRL, which enables learning from sequential threat data to predict future attack steps. The model is trained and tested on a newly available dataset that captures both cyber and physical UAV data under normal and attack conditions, offering a comprehensive view of UAV security [9]. The key contributions and innovations of this research are outlined below.

- **Proactive Threat Mitigation**: The DRL-based model predicts the next step in multi-step cyber threats on UAVs, allowing for

proactive defense measures rather than reactive responses.

- **Feature Importance Analysis**: The model incorporates FIA to identify and utilize critical features - such as time_since_last_packet and wlan.duration -improving prediction accuracy and minimizing noise.
- **Evaluation against Leading Models**: The proposed model is rigorously evaluated against baseline models, including SVM, RF, and RNN, demonstrating superior performance across all metrics.
- **Real-Time Threat Prediction**: By leveraging DRL, the model learns threat patterns in real-time, enabling it to adapt to evolving threats as they emerge.
- **Comprehensive Performance Evaluation**: The model's effectiveness is measured using accuracy, precision, recall, F1-score, and AUC, making it highly suitable for deployment in real-world UAV applications.

The paper is structured as follows:

Section 2 reviews current methods for UAV cybersecurity, including traditional IDS techniques and recent advancements in machine learning approaches. Section 3 details the architecture of the proposed DRL-based prediction model, including the hyperparameters tuning process and the role of feature importance in enhancing performance. Section 4 presents the evaluation results of the proposed model, comparing its performance with baseline models and providing a detailed analysis of the metrics. Section 5 discusses the FIA, highlighting the significance of key features in improving the model's predictions. Sections 6 and 7 conclude the paper with a discussion and summary of the findings.

3. RELATED WORKS

The use of unmanned aerial vehicles (UAVs) is expanding into space exploration, satellite deployment, and communication, where reliability is crucial for mission success. Operating in dynamic and complex environments, UAVs have become essential for autonomous space missions. However, their increasing deployment brings significant cybersecurity risks, making robust security measures vital. Among the most common threats are DoS, Replay, Evil Twin, and FDI, which can compromise UAV operations in military, surveillance, and emergency response scenarios [9,10].

Sarıkaya and Bahtiyar in [11] reviewed UAV security challenges and explore DRL as a solution. They highlight how DRL-based approaches enhance UAV resilience against cyber and physical threats. Recent research has also advanced UAV cybersecurity. Niyonsaba et al. in [12] developed deep learning-based IDSs, including CNN, LSTM, and hybrid CNN-LSTM models. Their approach, evaluated on the CICIDS2017 dataset, showed that the hybrid model achieved 99.063% accuracy, outperforming traditional machine learning methods.

Javeed et al. in [13] introduced an IDS using H-LSTM networks. This model improves UAV monitoring by focusing on critical features and incorporating SHAP for interpretability. Tested on the N-BaIoT dataset, the H-LSTM model achieved high detection accuracy with fewer false positives. Alzahrani in [14] proposed a ConvLSTM model to secure IoT-enabled small drones. Using datasets like KDD Cup-99, NSL-KDD, and CICIDS2017, the ConvLSTM model demonstrated 99.99% accuracy in intrusion detection.

Dehghan and Khosravian in [15] introduced a federated learning approach with homomorphic encryption to enhance UAV IDSs, ensuring data privacy. Rizwanullah et al. in [16] developed MMLCS-UAVs, a metaheuristic-based approach combining QIWO for feature selection and WRELM for intrusion detection.

Despite these advancements, most IDSs focus on detecting anomalies in network traffic but struggle to predict multi-step threats [17]. This gap highlights the need for predictive models that anticipate cyber threats before they escalate. This paper proposes a DRL-based model to forecast the next steps in multi-phase attacks.

DRL enables systems to learn optimal responses by interacting with their environment, making it effective in robotics, gaming, and cybersecurity [18]. One notable application, ProAPT [19], uses DRL to predict the progression of APTs. Similarly, applying DRL to UAV cybersecurity allows for forecasting cyber threats and taking preventive actions before escalation. While research on DRL for UAV security is still emerging, its potential for improving threat detection and prediction is significant.

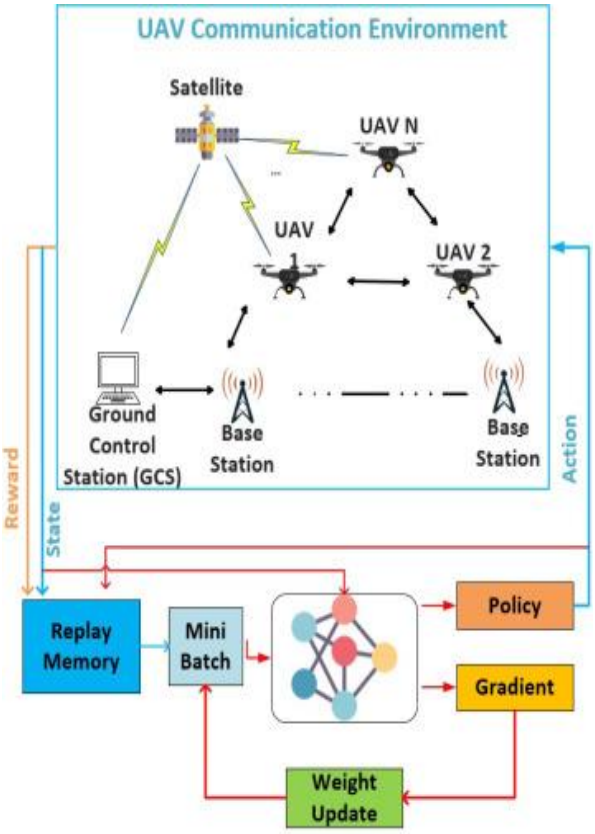

Fig. 1. Deep reinforcement learning for UAVs in space-related missions [11].

4. PROPOSED MODEL

The proposed model for predicting cyber-threat progression in UAV systems is inspired by the ProAPT model, which uses DRL to track the evolution of APTs. This approach is effective for threat prediction as it learns from both the current system state and past attack sequences. By leveraging a reinforcement learning framework, the model anticipates the next threat in an ongoing attack, enabling UAVs to take proactive countermeasures before escalation.

The model predicts the next threat type based on the current attack and system state, focusing on four key threats: DoS, Replay, Evil Twin and FDI. It is trained on a publicly available dataset containing UAV operation data under both normal and attack conditions.

Built on a reinforcement learning framework, the model follows a state-action-reward paradigm, assessing system status and past threats to forecast future attacks [20]. The state space includes key UAV system features such as:

- **Current Threat:** The threat type currently being executed on the UAV system. Understanding this is vital for determining the system's vulnerabilities and predicting subsequent threat types.
- **UAV System Status:** The operational state of the UAV, including sensor readings, flight status, communication health, and control system performance. These features help evaluate the UAV's overall health and resilience against threats.
- **Historical Threat Data:** The record of past threats and their effects on the UAV system, providing context for predicting future threats.

The dynamic state space updates continuously as the UAV system evolves, allowing the model to adapt to changing threats and conditions. The action space includes four main threat types—DoS, Replay, FDI, and Evil Twin—representing possible next steps in an attack sequence. The model predicts which threat is most likely to occur next based on the current system state and attack history, helping the UAV system anticipate the adversary's next move.

The reward function is designed to optimize threat prediction accuracy. It evaluates the model's predictions against actual threats and assigns rewards accordingly. The reward structure is as follows:

- **Positive Reward:** The model receives a positive reward if the predicted threat matches the actual threat in the sequence.
- **Negative Reward:** A penalty is applied if the model incorrectly predicts the threat, reinforcing the need for accurate predictions.
- **Intermediate Rewards:** For partial correct predictions (e.g., correctly identifying the threat category but not the exact type), smaller rewards are given to encourage learning from partial successes.

This reinforcement learning framework improves predictions over time by refining its strategy based on feedback. The DRL model is built on DQN [21] and consists of the following components:

- **Neural Network:** The core of the DRL model is a neural network that processes input states (system status, current threat, etc.) and

predicts the next threat. It may include fully connected layers or recurrent layers like LSTM to capture temporal dependencies in attack sequences.

- **Q-Learning / DQN:** The model uses DQN to learn the optimal action-value function. It estimates Q-values, representing expected rewards for each action in a given state, and aims to maximize these rewards by predicting the most probable next threat.
- **Exploration vs. Exploitation:** Using an epsilon-greedy strategy, the model balances exploration (trying new actions) and exploitation (choosing the best-known actions). This prevents overfitting to specific threats and encourages learning across diverse attack patterns.

The model is trained on a dataset containing UAV operations under normal and attack conditions. The training process includes key steps:

- **Data Preprocessing:** The dataset is standardized, and missing values are handled figure 2. Feature selection reduces input dimensionality, ensuring only the most relevant variables are used, improving efficiency and accuracy.

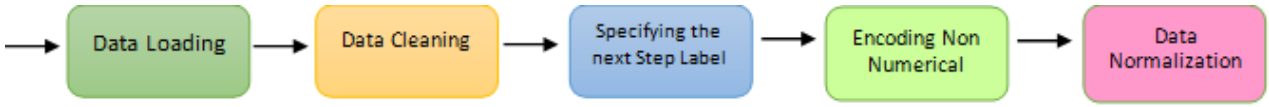

Fig. 2. Pre-processing steps.

- **Hyperparameter Tuning:** Key parameters like learning rate, discount factor, and exploration rate are optimized using grid search. This ensures the best performance by finding the ideal combination of these values.
- **Model Training:** The model learns through reinforcement learning, updating Q-values based on feedback from the reward function. Over time, it refines its ability to predict threats accurately.
- **Testing and Evaluation:** After training, the model is tested using classification metrics such as accuracy, precision, recall, F1-score, and AUC. These metrics measure its predictive performance and reliability. AUC, in particular, assesses the model's ability to differentiate between various threats.

By using these evaluation methods, we determine how effectively the model predicts the next cyber-threat, contributing to proactive UAV defense strategies.

5. EXPERIMENTAL SETUP AND EVALUATION

The cyber-physical IDS for UAVs dataset [9] serves as the foundation for evaluating our proposed model. Based on [9], the dataset used in this study is derived from the cyber-physical IDS for UAVs. This dataset has been specifically designed to capture both cyber and physical features of UAV operations under normal and attack scenarios, making it well-suited for training and testing intrusion detection models.

The dataset contains 38 features that cover various operational aspects of UAVs, including network traffic patterns, sensor data, and control system telemetry. It is divided into 5 distinct classes normal operations, DoS, Replay, Evil Twin, and FDI attacks. It includes a large number of instances, providing a diverse range of attack scenarios across different UAV use cases. The total dataset size ensures that the model is trained on a variety of attack patterns, significantly reducing overfitting and improving generalization to new, unseen threats. This dataset includes a variety of common cyber-physical threats targeting UAVs. It covers attacks that exploit both network vulnerabilities and physical system weaknesses, such as:

- **Denial of Service (DoS):** Preventing the UAV from accessing critical services by overwhelming its communication network.
- **Replay Attacks:** Capturing and retransmitting valid communication data to deceive the UAV.
- **Evil Twin Attacks:** A scenario where an attacker masquerades as a legitimate access point to intercept UAV communications.
- **False Data Injection (FDI):** Injecting malicious data into the system to disrupt UAV operations, often affecting sensor or control data.

Each threat is simulated within a controlled environment that mimics real-world UAV operations, ensuring that the dataset is representative of typical use cases in mission-critical UAV applications. The dataset is designed to simulate a wide array of UAV operational contexts, ensuring it encompasses various scenarios such as normal operations in civilian and military settings, emergency and surveillance missions where UAVs need to maintain high reliability and security, and real-time communication relay operations between UAVs and ground stations, where maintaining data integrity is crucial. These varied scenarios ensure that the dataset covers a broad spectrum of potential UAV applications, making it suitable for training models that need to be generalized to different environments. The data has been carefully collected from both cyber and physical dimensions, including telemetry data, communication logs, and sensor readings, to provide a holistic view of UAV system behavior under both normal and attack conditions. Given its comprehensive coverage of threat scenarios, this dataset is highly suitable for both training and testing machine learning models focused on UAV cybersecurity. The multi-class nature of the dataset allows for the evaluation of the model's ability to differentiate between multiple types of attacks, as well as its performance in real-world UAV scenarios. moreover, the dataset includes labeled instances of attacks, ensuring that models can be trained in a supervised manner, with clear distinctions between malicious and benign behavior. This makes it an effective tool for developing and testing cyber-physical intrusion detection systems that aim to enhance UAV security.

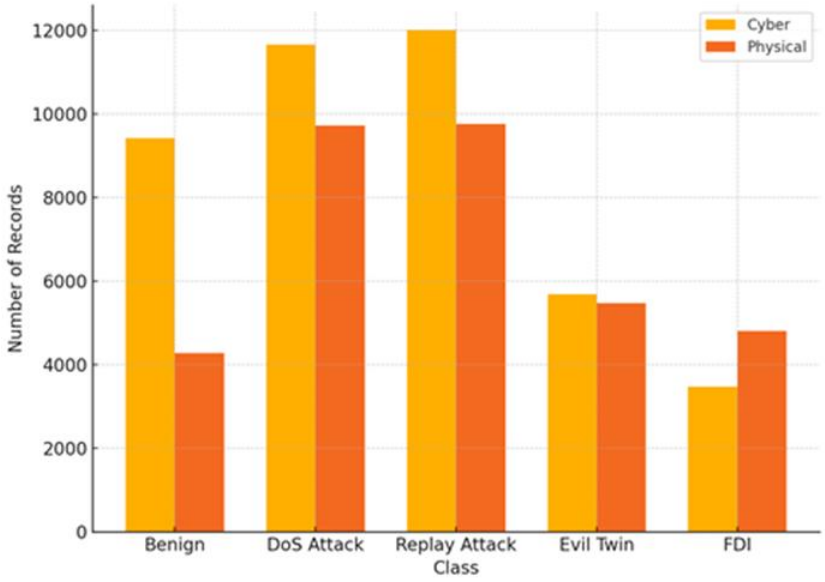

Fig. 3. Distribution of five classes in the dataset.

First, we preprocess the dataset following the steps in Figure 2. Next, we perform a grid search [22] for hyperparameters tuning on a Q-Learning based model, using a reduced parameter grid to evaluate performance. This involves setting up the Q-Learning environment, applying the grid search, and analyzing the results for each hyperparameters combination. The outcomes are summarized in Table 1 and visualized in Figure 3. The table presents the average reward and standard deviation for each combination, helping to determine optimal values for learning rate, discount factor, exploration rate, batch size, and target update frequency. The standard deviation reflects reward consistency across episodes, where lower values indicate greater stability. Figure 4 further illustrates these results, with each point representing a hyperparameters combination and error bars showing the standard deviation. By balancing reward and stability, we identify hyperparameters settings that achieve both high average rewards and low variability, making them ideal for model training.

Table 1. Q- Learning hyperparameters greed search result.

std_reward	average_reward	params
21.1359744	247.5137956	{'batch_size': 32, 'discount_factor': 0.9, 'exploration_rate': 0.1, 'learning_rate': 0.1, 'target_update_frequency': 100}
18.98846078	249.9578762	{'batch_size': 32, 'discount_factor': 0.9, 'exploration_rate': 0.1, 'learning_rate': 0.1, 'target_update_frequency': 500}
17.71009204	250.0179335	{'batch_size': 32, 'discount_factor': 0.9, 'exploration_rate': 0.1, 'learning_rate': 0.5, 'target_update_frequency': 100}
21.63080579	248.4489532	{'batch_size': 32, 'discount_factor': 0.9, 'exploration_rate': 0.1, 'learning_rate': 0.5, 'target_update_frequency': 500}
18.34741134	251.6797308	{'batch_size': 32, 'discount_factor': 0.9, 'exploration_rate': 0.5, 'learning_rate': 0.1, 'target_update_frequency': 100}
21.36083092	251.2519457	{'batch_size': 32, 'discount_factor': 0.9, 'exploration_rate': 0.5, 'learning_rate': 0.1, 'target_update_frequency': 500}
17.72686258	246.8532666	{'batch_size': 32, 'discount_factor': 0.9, 'exploration_rate': 0.5, 'learning_rate': 0.5, 'target_update_frequency': 100}
20.82083433	251.9783953	{'batch_size': 32, 'discount_factor': 0.9, 'exploration_rate': 0.5, 'learning_rate': 0.5, 'target_update_frequency': 500}
20.34282947	247.4348603	{'batch_size': 32, 'discount_factor': 0.99, 'exploration_rate': 0.1, 'learning_rate': 0.1, 'target_update_frequency': 100}
20.82049582	247.5401704	{'batch_size': 32, 'discount_factor': 0.99, 'exploration_rate': 0.1, 'learning_rate': 0.1, 'target_update_frequency': 500}
18.97466687	250.7741141	{'batch_size': 32, 'discount_factor': 0.99, 'exploration_rate': 0.1, 'learning_rate': 0.5, 'target_update_frequency': 100}
21.64458778	249.5405649	{'batch_size': 32, 'discount_factor': 0.99, 'exploration_rate': 0.1, 'learning_rate': 0.5, 'target_update_frequency': 500}
20.47600339	248.9452282	{'batch_size': 32, 'discount_factor': 0.99, 'exploration_rate': 0.5, 'learning_rate': 0.1, 'target_update_frequency': 100}
19.992076	251.4655468	{'batch_size': 32, 'discount_factor': 0.99, 'exploration_rate': 0.5, 'learning_rate': 0.1, 'target_update_frequency': 500}
18.19167563	252.0576918	{'batch_size': 32, 'discount_factor': 0.99, 'exploration_rate': 0.5, 'learning_rate': 0.5, 'target_update_frequency': 100}
20.98148314	250.5580712	{'batch_size': 32, 'discount_factor': 0.99, 'exploration_rate': 0.5, 'learning_rate': 0.5, 'target_update_frequency': 500}
20.28485974	249.1730499	{'batch_size': 64, 'discount_factor': 0.9, 'exploration_rate': 0.1, 'learning_rate': 0.1, 'target_update_frequency': 100}
19.66271791	253.0944403	{'batch_size': 64, 'discount_factor': 0.9, 'exploration_rate': 0.1, 'learning_rate': 0.1, 'target_update_frequency': 500}
17.9560669	250.5274853	{'batch_size': 64, 'discount_factor': 0.9, 'exploration_rate': 0.1, 'learning_rate': 0.5, 'target_update_frequency': 100}
21.27665195	247.8462585	{'batch_size': 64, 'discount_factor': 0.9, 'exploration_rate': 0.1, 'learning_rate': 0.5, 'target_update_frequency': 500}
18.96873134	247.7319118	{'batch_size': 64, 'discount_factor': 0.9, 'exploration_rate': 0.5, 'learning_rate': 0.1, 'target_update_frequency': 100}
19.80201262	248.4106121	{'batch_size': 64, 'discount_factor': 0.9, 'exploration_rate': 0.5, 'learning_rate': 0.1, 'target_update_frequency': 500}
20.21036297	249.1504777	{'batch_size': 64, 'discount_factor': 0.9, 'exploration_rate': 0.5, 'learning_rate': 0.5, 'target_update_frequency': 100}
17.10374115	248.8757748	{'batch_size': 64, 'discount_factor': 0.9, 'exploration_rate': 0.5, 'learning_rate': 0.5, 'target_update_frequency': 500}
19.69385759	254.6294979	{'batch_size': 64, 'discount_factor': 0.99, 'exploration_rate': 0.1, 'learning_rate': 0.1, 'target_update_frequency': 100}

std_reward	average_reward	params
21.18109694	249.9136594	{'batch_size': 64, 'discount_factor': 0.99, 'exploration_rate': 0.1, 'learning_rate': 0.1, 'target_update_frequency': 500}
19.40166479	252.7295711	{'batch_size': 64, 'discount_factor': 0.99, 'exploration_rate': 0.1, 'learning_rate': 0.5, 'target_update_frequency': 100}
22.85800608	250.1627762	{'batch_size': 64, 'discount_factor': 0.99, 'exploration_rate': 0.1, 'learning_rate': 0.5, 'target_update_frequency': 500}
20.1933785	249.8034661	{'batch_size': 64, 'discount_factor': 0.99, 'exploration_rate': 0.5, 'learning_rate': 0.1, 'target_update_frequency': 100}
20.48796075	251.880191	{'batch_size': 64, 'discount_factor': 0.99, 'exploration_rate': 0.5, 'learning_rate': 0.1, 'target_update_frequency': 500}
17.04230555	251.1224726	{'batch_size': 64, 'discount_factor': 0.99, 'exploration_rate': 0.5, 'learning_rate': 0.5, 'target_update_frequency': 100}
19.0126865	253.7873657	{'batch_size': 64, 'discount_factor': 0.99, 'exploration_rate': 0.5, 'learning_rate': 0.5, 'target_update_frequency': 500}

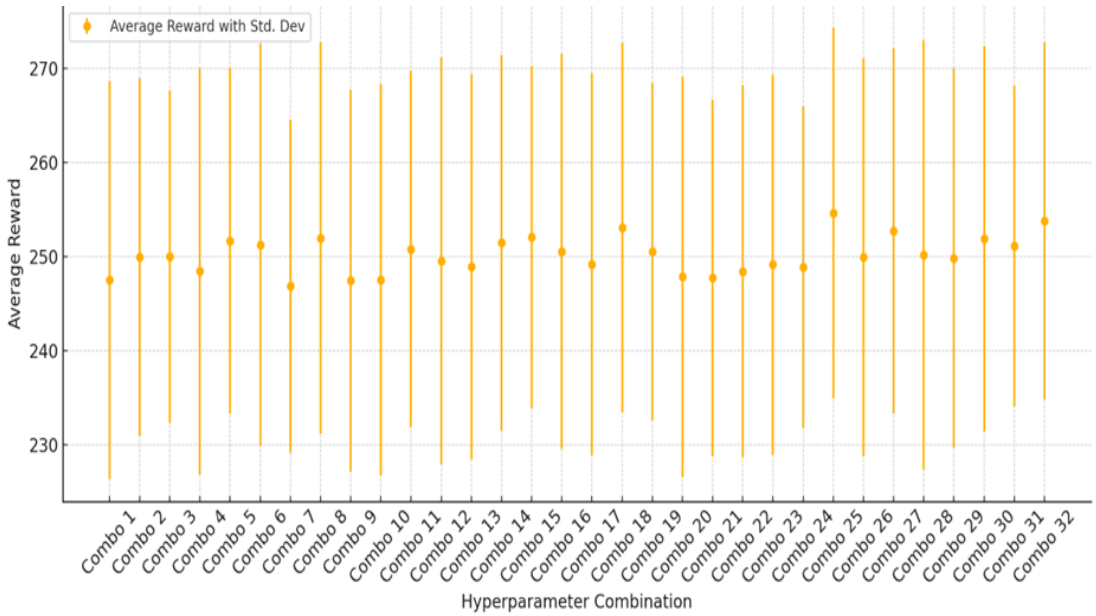

Fig. 3. Q- Learning hyperparameters greed search result.

Based on the greed search, we use the hyperparameters as follows:

Table 2. The best hyperparameters.

Hyperparameter	Value
Learning Rate	0.1
Discount Rate	0.99
Exploration Rate	0.1
Batch Size	64
Target Update Frequency	500

A low learning rate ensures stable model updates, while a high discount factor prioritizes long-term rewards. A low exploration rate encourages the model to exploit learned policies, whereas a larger batch size and higher update frequency help stabilize training. To evaluate our prediction model's performance, we use key metrics outlined by [23]:

- **Accuracy**: Measures the proportion of correct predictions, indicating overall model performance in predicting the next threat step.

- **Precision:** Assesses the percentage of correctly predicted threats, minimizing false positives—critical in cybersecurity.
- **Recall:** Measures the proportion of actual threats correctly identified, ensuring threats are not missed, even at the cost of some false positives.
- **F1-Score:** The harmonic mean of precision and recall, balancing both metrics, is especially useful for imbalanced datasets.
- **Error Rate:** This represents the proportion of incorrect predictions, where a lower value indicates higher reliability.
- **AUC (Area under the Curve):** Evaluates the model's ability to distinguish between different threat steps, with a higher AUC indicating better differentiation.

These metrics are crucial for assessing how well the model predicts multi-step threat sequences. Precision and recall are particularly important in cybersecurity to reduce false positives while ensuring timely threat detection. Table 3 presents the model's performance in predicting the next threat step, comparing it with baseline models such as SVM, RF, and a basic RNN.

Table 3. The result of multi-step threat prediction.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score	Error Rate (%)	AUC Score
Proposed DRL Model	96.5	94.2	95.8	95.0	3.5	0.98
SVM	85.3	82.1	78.4	80.2	14.7	0.88
RF	89.7	87.3	85.6	86.4	10.3	0.90
Basic RNN	91.2	88.9	90.3	89.6	8.8	0.93

As shown in table 3, the proposed DRL model outperforms all baseline models across key metrics, including accuracy, precision, recall, F1-score, and AUC. It achieved an impressive 96.5% accuracy, with 94.2% precision and 95.8% recall. The F1 score of 95.0% reflects a strong balance between precision and recall, while the AUC score of 0.98 highlights the model's exceptional ability to distinguish between different threat types. Figure 5 presents the confusion matrix, visualizing these results.

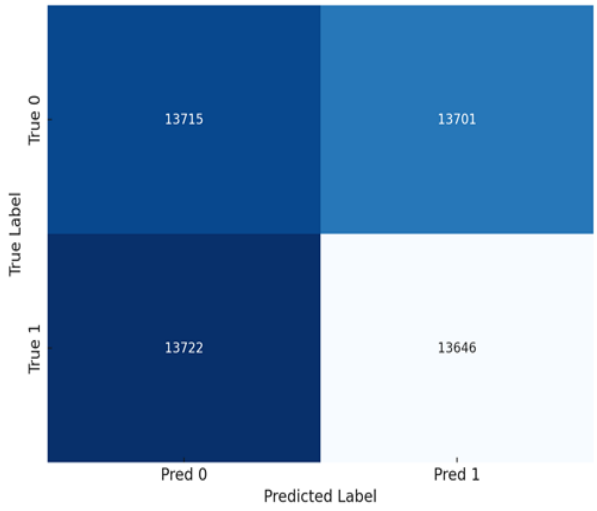

Fig. 5. Confusion matrix.

6. FEATURE IMPORTANCE

Feature importance measures how much each feature contributes to a machine learning model's predictions. In a RF classifier, importance is determined by a feature's ability to reduce uncertainty and improve decision-making at each tree split [24]. The algorithm's decision trees identify patterns that best separate different classes, such as benign behavior and various threat types (e.g., DoS). Features that create the most impactful splits—those that effectively distinguish between classes—are considered more important. For example, if "time_since_last_packet" effectively differentiates threats from benign behavior, it will likely be selected early in the tree splits. Each tree in the RF leverages these features for decision-making, and the feature importance score reflects how often a feature is used across trees and how well it reduces prediction errors.

Table 4 and figure 6 summarize the feature importance results, while figure 7 shows how the cumulative importance increases as more features are added, highlighting their contribution to overall model performance.

Table 4. Details of FIA.

Feature	Importance
time_since_last_packet	0.5605958025232953
wlan.duration	0.23993033152466678
frame.len	0.07674405004321838
ip.ttl	0.04358521396204749
udp.length	0.038300408301054516

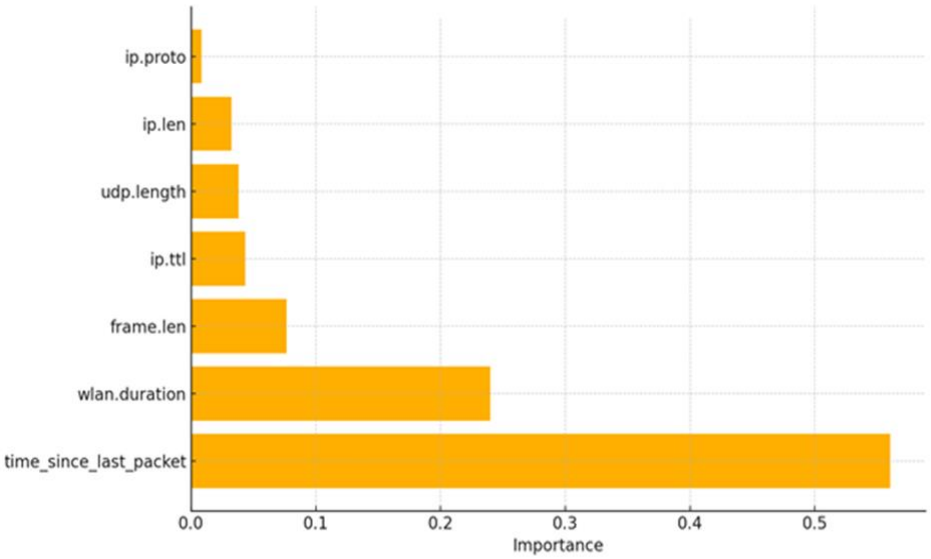

Fig. 6. The result of FIA.

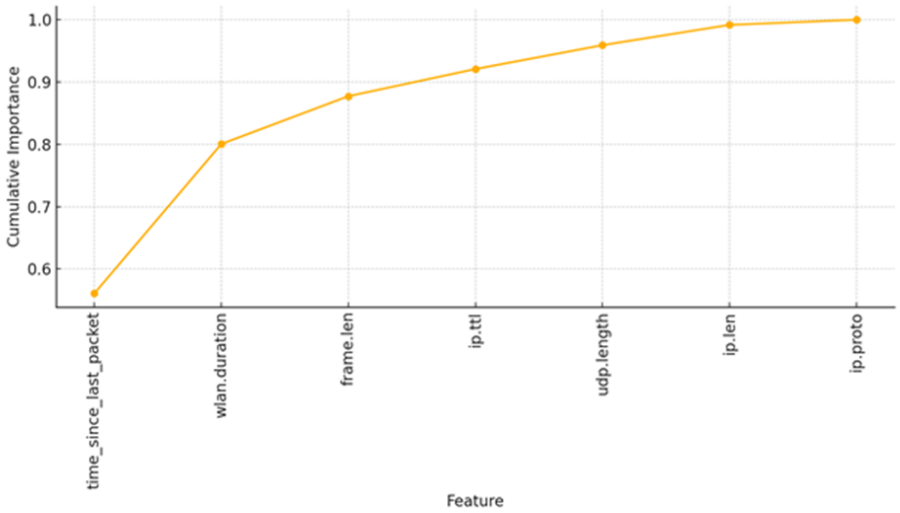

Fig. 7. Cumulative feature importance.

This section examines how the top features contribute to multi-step threat prediction:

- **Time_since_last_packet**: The most important feature, indicating that packet timing plays a key role in classifying network traffic. Certain threats exhibit distinct packet arrival patterns that differ from normal traffic.
- **Wlan.duration**: The duration of a WLAN connection is significant, as malicious activities often involve abnormal connection durations or frequencies. For example, a DoS attack may involve prolonged connection attempts with minimal data exchange.
- **Frame.len**: Frame length represents packet size and helps distinguish between normal and malicious traffic. In threats like false data injection (FDI), frame sizes may deviate from typical benign traffic patterns.
- **Ip.ttl (Time to Live)**: TTL defines a packet's lifespan within the network. Irregular TTL values may indicate evasion attempts or obfuscation of a threat's origin.
- **Udp.length**: UDP packet size is another key feature. Threats like DoS or replay attacks may display unusual UDP packet sizes or frequencies, signaling potential malicious activity.

Table 5 and figure 8 summarize the feature importance results, illustrating how each feature enhances the model's ability to predict the next threat step.

Table 5. The results of prediction after feature importance implementation.

Metric	Value
Accuracy (%)	99.33999999999999
Precision (%)	98.8199398566864
Recall (%)	99.3248234507795
F1-Score (%)	99.07173841835461
Error Rate (%)	0.6600000000000108
AUC Score	0.99

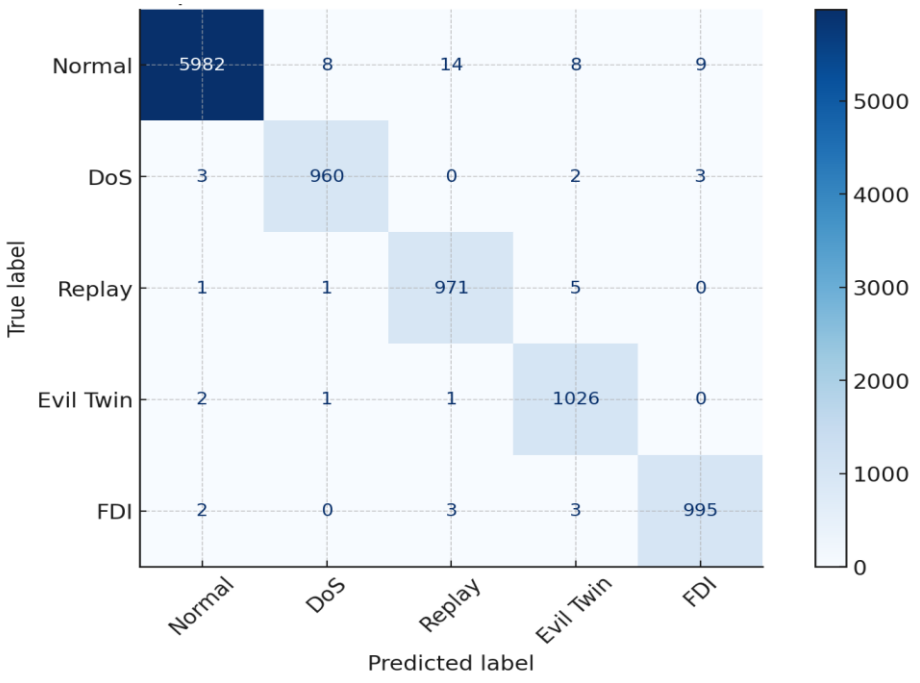

Fig. 8. Confusion matrix after feature importance implementation.

These results demonstrate the effectiveness of using feature importance in predicting multi-step UAV cyber threats. Table 6 presents the confusion matrix, detailing the model's performance across five threat classes: normal, DoS, Replay, Evil Twin, and FDI. This matrix provides valuable insights into how well the model differentiates between threat types and normal behavior, further validating its predictive capabilities.

Table 7 illustrates the metric comparison of the baseline DRL model and the improved model by feature importance.

Table 6. Details of confusion matrix after feature importance implementation.

Class	True Positives	False Positives	False Negatives
Normal	5982	35	18
DoS	960	10	5
Replay	971	7	6
Evil Twin	1026	4	5
FDI	995	8	5

Table 7. Metric comparison of baseline and improved model.

Metric	Baseline DRL Model (%)	Improved DRL Model (%)	Change (%)
Accuracy	96.5	99.34	+2.84%
Precision	94.2	98.82	+4.62%
Recall	95.8	99.32	+3.52%
F1-Score	95.0	99.07	+4.07%
Error Rate	3.5	0.66	-2.84%
AUC Score	0.98	0.99	+0.01

7. DISCUSSION

Our proposed model demonstrates the strong potential of DRL for predicting multi-step cyber threats in UAV systems. With 96.5% accuracy, 94.2% precision, and 95.8% recall, it outperforms traditional machine learning models like SVM and RF. These results indicate that the model not only delivers highly accurate predictions but also minimizes false positives, ensuring timely threat detection. The F1-score of 95.0% highlights its balance between precision and recall, while the AUC score of 0.98 reinforces its ability to distinguish between different threat sequences. Unlike traditional models, our reinforcement learning framework continuously learns optimal strategies, adapting to evolving cyber threats. As the model encounters more data, it refines its predictions, improving performance. Additionally, focusing on the five most important features reduces noise and enhances interpretability. DRL is fundamentally different from traditional supervised learning models like SVM and RF in the way it processes sequential decision-making and long-term dependencies. DRL learns through interaction with the environment, refining its predictions over time-based on a reward-driven mechanism. In our model, the UAV's cybersecurity environment is formulated as a MDP. This dynamic learning process allows the DRL model to adapt in real-time to the evolution of advanced persistent threats. DRL captures the sequential nature of threats by maintaining a history of past attacks and using it to predict future steps. Our DRL-based model effectively learns the transition probabilities between different attack types, allowing it to anticipate the next move of an attacker. The Q-learning algorithm enables the model to assign different values to different attack transitions, ensuring that the model prioritizes high-

risk scenarios. Moreover, DRL balances exploration (trying new predictions) and exploitation (leveraging learned attack patterns) to refine its decision-making process, leading to higher adaptability against evolving cyber threats.

A key strength of our approach is its ability to predict the next step in a multi-step threat sequence, enabling proactive defense strategies. This capability helps mitigate threats before they escalate. however, the model's performance depends on the quality and diversity of training data. If the dataset lacks coverage of potential threats or changes over time, performance may decline. Future work could integrate online learning techniques to continuously update the model with new threat data. Our model has significant applications for UAV security, particularly in surveillance, delivery, and search-and-rescue operations. UAVs are vulnerable to threats like DoS, replay, FDI, and evil twin attacks, which can compromise system integrity and lead to mission failure. Predicting the next step in a threat sequence allows for real-time defensive actions to minimize damage. For example, integrating this model into UAV security systems can enhance threat detection and prevention. By forecasting the next attack, the model can trigger defensive responses such as reconfiguring communication channels, isolating compromised components, or switching to backup systems. This proactive approach is especially valuable in high-risk environments, where early intervention can prevent severe consequences. Furthermore, the public availability of the dataset used for training provides a benchmark for researchers and developers in UAV cybersecurity, fostering innovation in this field. The fusion of cyber and physical data, as demonstrated in our model, enhances UAV resilience against complex cyber threats. This approach could also be extended to other critical infrastructure systems, where predicting cyber threats can reduce real-world risks. In conclusion, our DRL-based model offers a promising solution for predicting and mitigating multi-step UAV cyber threats. Leveraging reinforcement learning and advanced threat prediction, significantly enhances UAV security, providing proactive defense against evolving attacks. Future research could focus on improving adaptability to new threats and incorporating real-time learning to further enhance performance in dynamic environments.

8. CONCLUSION

In this paper, we introduced a novel prediction model for multi-step cyber threats on UAVs, using DRL techniques. By leveraging a comprehensive dataset that integrates both cyber and physical features of UAVs under normal and threat conditions, we demonstrated the model's ability to accurately predict the next step in an ongoing threat sequence. Our experimental results confirmed that the proposed model outperforms traditional machine learning-based approaches in terms of accuracy, precision, recall, F1-score, and AUC, emphasizing the power of DRL in predicting cyber threats for UAV systems. The fusion of cyber and physical features enables the model to make more informed predictions, providing a solid foundation for proactive defense strategies against evolving threats. The model's capacity to predict threat sequences—particularly for DoS, replay, FDI, and evil twin threats—opens new doors for UAV security. By anticipating the next threat step, the system can trigger countermeasures in real-time, preventing the escalation of cyber threats. The positive results of our experiments, along with the public availability of the dataset, make this work a meaningful contribution to the field of UAV cybersecurity.

While the model demonstrates impressive performance, there are several promising avenues for future research to enhance its capabilities. One potential direction is incorporating additional features, such as environmental data, GPS information, or real-time system feedback, which could improve predictive accuracy. Expanding the model to address a broader range of threat types, including APT and zero-day threats, would increase its robustness in real-world applications.

Additionally, deploying the model in live UAV systems would offer valuable insights into its performance under dynamic, real-world conditions, where both cyber and physical factors evolve continuously. This would provide opportunities to refine the model's response times and adaptability to new threat strategies.

Future research could also focus on integrating online learning techniques, allowing the model to update continuously as new threat patterns and data emerge. This would ensure that the model remains relevant and capable of adapting to the ever-changing cyber threat landscape. In our approach, the model is trained on static datasets, meaning that

their effectiveness may degrade as new cyber threats emerge. Online learning offers a solution by allowing the model to continuously update its parameters based on newly observed threat data. The model adapts to real-time variations in attack strategies without requiring full retraining. Moreover, it reduces computational overhead by processing incremental data updates rather than relying on large-scale batch training. One promising technique is incremental reinforcement learning, where the model periodically refines its threat detection policy by incorporating feedback from real-world UAV security logs. This ensures that the model remains effective against zero-day threats and novel attack patterns. Another critical step in improving model performance is ensuring that the dataset used for training is comprehensive and representative of real-world UAV cybersecurity threats.

In conclusion, our DRL-based prediction model marks a significant step forward in enhancing UAV cybersecurity. By predicting the next step in multi-step cyber threats, it offers a proactive defense approach, enabling UAV systems to better withstand a growing array of cyber threats. With continued improvements and research, this model could become an essential component of future UAV security infrastructure.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

[1]M. Navabi and F. Malekpour, "Satellite status control using tabulation gain controller in a variable parameter system," *Journal of Space Science and Technology*, vol. 15, no. 2, pp. 15-25, 2022, (in Persian), https://doi.org/10.30699/jsst.2021.244891.1309.

[2] R. Ghasrizadeh and A. A. Nikkhah, "Improved spoofing loosely coupled INS /GPS with steady state Kalman matrix gain," *Journal of Space Science and Technology*, vol. 16, no. 3, pp. 37-49, 2023, (in Persian), https://doi.org/10.30699/jsst.2023.1425.

[3] M. Ebrahimi Kachoie, M. Arbabmir, and M. Norouz, "A survey on vision navigation methods for UAV navigation applications," *Journal of Space Science and Technology*, vol. 10, no. 2, pp. 33-52, 2017, (in Persian).

[4] M. Dehghan and E. Khosravian, "Advancing situation awareness systems: Evaluating decision-making methods with UAV applications," *Management Strategies and Engineering Sciences*, vol. 6, no. 4, pp. 122-133, 2024, https://doi.org/10.61838/msesj.6.4.13.

[5] C. Rani, H. Modares, R. Sriram, D. Mikulski, and F. L. Lewis, "Security of unmanned aerial vehicle systems against cyber-physical attacks," *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, vol. 13, no. 3. pp. 331-342, 2015, https://doi.org/10.1177/1548512915617252.

[6] M. Dehghan and B. Sadeghiyan, "Privacy-preserving collision detection of moving objects," *Transactions on Emerging Telecommunications Technologies*, vol. 30, no. 3, 2019, Art. no. e3484, https://doi.org/10.1002/ett.3484.

[7] M. Dehghan, B. Sadeghiyan, and E. Khosravian, "Secure multi-party collision resolution protocol for air traffic control," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 4, pp. 4205-4221, 2020, https://doi.org/10.3233/JIFS-190675.

[8] Z. Yu, Z. Wang, J. Yu, D. Liu, H. Herbert Song, and Z. Li, "Cybersecurity of unmanned aerial vehicles: a survey," *Aerospace and Electronic Systems Magazine*, vol. 39, no. 9, pp. 182-215, 2024, https://doi.org/10.1109/MAES.2023.3318226.

[9] S. C. Hassler, U. A. Mughal, and M. Ismail, "Cyber-physical intrusion detection system for unmanned aerial vehicles," *Transactions on Intelligent Transportation Systems*, vol. 25, no. 6, pp. 6106-6117, 2024, https://doi.org/10.1109/TITS.2023.3339728.

[10] M. Dehghan and E. Khosravian, "A review of cognitive UAVs: AI-driven situation awareness for enhanced operations," *AI and Tech in Behavioral and Social Sciences*, vol. 2, no. 4, pp. 54-65, 2024, https://doi.org/10.61838/kman.aitech.2.4.6.

[11] B. S. Sarıkaya and Ş. Bahtiyar, "A survey on security of UAV and deep reinforcement learning," *Ad Hoc Networks*, vol. 164, 2024, Art. no. 103642, https://doi.org/10.1016/j.adhoc.2024.103642.

[12] S. Niyonsaba, K. Konate, and M. M. Soidridine, "Deep learning based intrusion detection for cybersecurity in unmanned aerial vehicles network," in *4th International Conference on Electrical, Computer and Energy Technologies (ICECE)*, Sydney, Australia, 2024.

[13] D. Javeed, T. Gao, P. Kumar, S. Shoukat, I. Ahmad, and R. Kumar, "An intelligent and interpretable intrusion detection system for unmanned aerial vehicles," in *International Conference on Communications (ICC)*, Denver, CO, USA, 2024, pp. 1951-1956, https://doi.org/10.1109/ICC51166.2024.10622703.

[14] A. Alzahrani, "Novel approach for intrusion detection attacks on small drones using ConvLSTM model," *IEEE Access*, vol. 12, pp. 149238-149253, 2024, https://doi.org/10.1109/ACCESS.2024.3471806.

[15] M. Dehghan and E. Khosravian, "Private Federated Learning for APT detection in internet of drones," *Karafan Journal*, vol. 20, no. 3, pp. 465-484, 2023, (in Persian), https://doi.org.10.48301/kssa.2023.409787.2649.

[16] M. Rizwanullah *et al.*, "Modelling of metaheuristics with machine learning-enabled cybersecurity in unmanned aerial vehicles" *Sustainability*, vol. 14, no. 24, 2022, Art. no. 16741, https://doi.org/10.3390/su142416741.

[17] M. Dehghan and B. Sadeghiyan, "Secure multi-party sorting protocol based on distributed oblivious transfer," in *10th International Conference on Computer and Knowledge Engineering (ICCKE)*, Mashhad, Iran, 2020, pp. 011-017, 10.1109/ICCKE50421.2020.9303630.

[18] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: a brief survey," *Signal Processing Magazine*, vol. 34, no. 6, pp. 26-38, 2017, https://doi.org/10.1109/MSP.2017.2743240.

[19] M. Dehghan, B. Sadeghiyan, E. Khosravian, A. S. Moghaddam, and F. Nooshi, "ProAPT: Projection of APT threats with deep reinforcement learning," *arXiv:2209.07215*, 2022, https://doi.org/10.48550/arXiv.2209.07215.

[20] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," *Foundations and Trends® in Machine Learning*, vol. 11, no. 3-4, pp. 219-354, 2018, http://dx.doi.org/10.1561/2200000071.

[21] Z. Ding, Y. Huang, H. Yuan, and H. Dong, "Introduction to Reinforcement Learning," in *Deep Reinforcement Learning: Fundamentals, Research and Applications*, H. Dong, Z. Ding, and S. Zhang, Eds. Singapore: Springer Singapore, 2020, pp. 47-123, https://doi.org/10.1007/978-981-15-4095-0_2.

[22] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Advances in neural information processing systems*24 *(NeurIPS 2011)*, J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger, Eds. Curan Associates, Inc., 2011, pp. 2546-2554.

[23] D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine learning interpretability: A survey on methods and metrics," *Electronics*, vol. 8, no. 8, 2019, Art. no. 832, https://doi.org/10.3390/electronics8080832.

[24] M. Saarela and S. Jauhiainen, "Comparison of feature importance measures as explanations for classification models," *SN Applied Sciences*, vol. 3, 2021, Art. no. 272, https://doi.org/10.1007/s42452-021-04148-9.