Reinforcement Learning for GNSS Spoofing Detection: A Multi-Class DQN Approach with TEXBAT

REINFORCEMENT LEARNING FOR GNSS SPOOFING DETECTION: A MULTI-CLASS DQN APPROACH WITH TEXBAT

Abdullah Al Noman Chowdhury[1], *Elham Ahmadi*[1], *Mohammed Elmusrati*[1], *Heidi Kuusniemi*[2], *Jani Boutellier*[1]

[1]School of Technology and Innovations, University of Vaasa, Finland
[2]Faculty of Information Technology and Communication Sciences, Tampere University, Finland

ABSTRACT

Global Navigation Satellite Systems (GNSS) are critical for positioning and timing, but the weak and unencrypted nature of civil GNSS signals makes them highly vulnerable to spoofing attacks. Existing detection methods, often based on handcrafted metrics or supervised learning, lack robustness across diverse scenarios. In this paper, we formulate GNSS spoofing detection as a multi-class reinforcement learning problem and propose a Deep Q-Network (DQN) trained on GNSS tracking features from the Texas Spoofing Test Battery (TEXBAT) dataset. The proposed approach adopts a tracking-level, non-sequential formulation evaluated on static and replay-like spoofing scenarios. Using nine spoof-sensitive features across multiple Pseudo-Random Noise (PRN) codes, the agent classifies clean and spoofed signals into seven classes, achieving 88.1% multi-class accuracy and a macro F1-score of 0.884. The proposed method outperforms SVM, random forest, and autoencoder baselines, demonstrating the effectiveness of reinforcement learning for GNSS spoofing detection. By learning decision policies through interaction and reward feedback, reward shaping improves sensitivity to stealth and replay spoofing without increasing false alarms.

Index Terms— GNSS, spoofing, deep reinforcement learning, resilient positioning, interference

1. INTRODUCTION

Global Navigation Satellite Systems (GNSS) are critical for positioning, navigation, and timing (PNT) services in aviation, maritime operations, telecommunications, and critical infrastructure [1]. Civil GNSS signals are weak at the receiver front end and openly accessible, making them susceptible to intentional interference [2]. Spoofing, where counterfeit signals mislead the receiver, poses serious risks including aviation safety incidents and timing failures [3].

Recent evidence confirms that spoofing is an operational reality. Between 2022 and 2024, widespread GNSS disruptions were reported in the Baltic and Nordic regions, prompting safety advisories from the European Union Aviation Safety Agency (EASA) [4]. Similar disruptions were observed in the Black Sea and Eastern Mediterranean, affecting both aviation and maritime sectors [5]. These incidents highlight the need for robust spoofing detection methods.

Conventional detection methods rely on signal quality monitoring, correlation distortion, angle-of-arrival discrimination, and consistency checks between pseudoranges and navigation solutions [6, 7]. While foundational, these approaches require specialized hardware or careful threshold tuning and degrade under sophisticated spoofing. More recently, signal processing and ML/DL techniques have been applied to spoofing detection. Babić *et al.* [8] employed discrete wavelet transforms with support vector machines (SVMs), while Borhani-Darian *et al.* [9] analyzed I/Q samples using convolutional and Cross Ambiguity Function (CAF) based techniques. Semanjski *et al.* [10] studied supervised classification under static and dynamic GNSS conditions. However, these studies remain supervised, binary focused, or data intensive, and none evaluate a multi-class reinforcement learning (RL) framework across multiple Texas Spoofing Test Battery (TEXBAT) scenarios.

On the TEXBAT dataset, Mahroof *et al.* reported F1-scores near 0.90 using kNN and SVM, but their work was limited to binary classification [11]. An autoencoder-based anomaly detector [12] achieved high alert rates for overpowered and replay attacks, but failed entirely on stealth attack types. While these methods improve robustness over classical monitoring, they remain constrained by binary formulations, reliance on labeled data, and limited generalization under advanced spoofing.

In this work, we propose an RL framework for multi-class GNSS spoofing detection, a dimension not addressed in prior studies. Unlike supervised approaches that depend on static labels, RL learns adaptive decision policies through interaction and reward feedback. Although RL has shown promise in communication security and spectrum monitoring [13, 14, 15], its application to GNSS spoofing remains largely unexplored [16]. A notable exception is Dasgupta *et al.*, who applied RL using in-vehicle sensors rather than

This research is funded by the EU–Interreg Aurora project "TRUST".

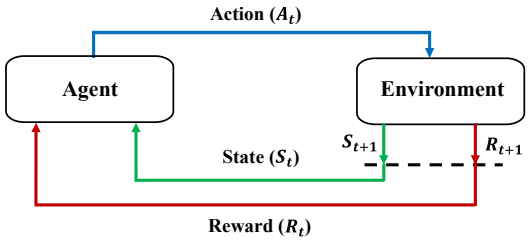

Fig. 1. RL framework, illustrating agent–environment interaction through actions, states, and rewards.

GNSS tracking features [17]. To bridge this gap, we formulate spoofing detection as a classification task and develop a Deep Q-Network (DQN) trained directly on raw tracking features from the TEXBAT dataset.

The main contributions of this paper are threefold: first, we formulate GNSS spoofing detection as a multi-class reinforcement learning problem, requiring discrimination among seven classes (clean plus six spoofing scenarios); second, we employ a DQN in which the agent learns decision policies through reward feedback, improving adaptability across spoofing types compared to supervised ML; and third, we present a comprehensive evaluation demonstrating superior performance over classical supervised models (SVM, RF) and an autoencoder baseline.

2. REINFORCEMENT LEARNING BACKGROUND

Reinforcement learning models the interaction between an agent and an environment, typically formalized as a Markov Decision Process (MDP) $(\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P}, \gamma)$ [18]. At each time step t, the agent observes a state $s_t \in \mathcal{S}$, selects an action $a_t \in \mathcal{A}$, according to a policy $\pi(a|s)$, receives a reward $r_t = \mathcal{R}(s_t, a_t)$, and transitions to a new state s_{t+1} with probability $\mathcal{P}(s'|s_t, a_t)$. $\gamma \in [0, 1]$ denotes the discount factor which determines the importance of future rewards. The objective of the agent is to maximize the expected discounted return as:

$$G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k}, \quad J(\pi) = \mathbb{E}_\pi[G_t], \tag{1}$$

where the state-value function under policy π is $V^\pi(s) = \mathbb{E}_\pi\left[G_t \mid s_t = s\right]$, and the action-value function is $Q^\pi(s, a) = \mathbb{E}_\pi\left[G_t \mid s_t = s, a_t = a\right]$. These functions satisfy the Bellman expectation equations [18]:

$$V^\pi(s) = \sum_a \pi(a|s) \sum_{s'} \mathcal{P}(s'|s, a)\left[\mathcal{R}(s, a) + \gamma V^\pi(s')\right]. \tag{2}$$

RL methods are broadly categorized into value-based methods, which estimate action value functions, policy-based methods, which directly optimize the policy and actor-critic methods, which combine both [18]. A canonical value-based

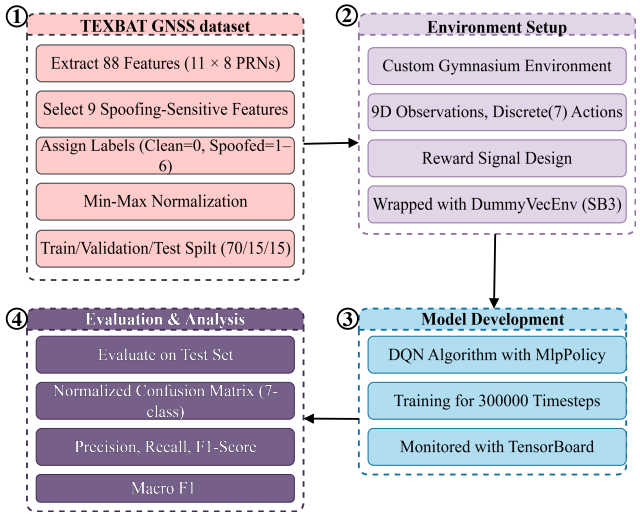

Fig. 2. Overall Methodology pipeline for GNSS spoofing detection using RL and DQN.

algorithm is Q-learning, whose update rule is:

$$Q(s, a) \leftarrow Q(s, a) + \alpha\left[r + \gamma \max_{a'} Q(s', a') - Q(s, a)\right]. \tag{3}$$

3. PROBLEM FORMULATION

In this work, GNSS spoofing detection is formulated as a single step stateless MDP, see Fig. 1. The state s is a concatenated GNSS tracking feature vector of nine spoofing-sensitive parameters across eight PRNs. The action a corresponds to predicting one of seven classes (clean or spoofed scenarios), and the reward r is $+1$ for a correct classification, and -1 otherwise. Figure 2 presents a high-level overview of the proposed RL-based spoofing detection pipeline, including data preprocessing, feature engineering, environment formulation, agent training, and evaluation, with each component detailed in the following sections.

3.1. Dataset Description

This study utilizes the TEXBAT dataset from the University of Texas Radionavigation Laboratory [19]. TEXBAT consists of high-fidelity GPS L1 C/A signal recordings containing both authentic (unspoofed) and spoofed scenarios. The recordings encompass a variety of realistic attack conditions in static (fixed receiver antenna) and dynamic (moving receiver) settings. In this work, the focus is placed exclusively on static and replay-like scenarios for training and evaluation, as summarized in Table 1. This choice enables a tracking-level formulation where signal features remain stable and well-defined, while avoiding additional complexities introduced by receiver motion. The tracking data were provided in CSV format, and meaningful column headers were added based on the repository documentation [12].

Table 1. TEXBAT scenarios and their inclusion in this study.

ID	Scenario Type	Used
cleanStatic (cs)	Authentic (no spoofing)	Yes
ds1	Abrupt takeover (static)	Yes
ds2	Time-push, overpowered	Yes
ds3	Time-push, matched-power	Yes
ds4	Position-push, matched-power	Yes
ds5	Time-push (dynamic)	No
ds6	Position-push (dynamic)	No
ds7	Stealth sparse spoofing	Yes
ds8	Replay/phase-synchronized	Yes

3.2. Feature Selection and Preprocessing

From the tracking logs, a set of nine spoofing-sensitive features was retained for training and evaluation: *acq_doppler_hz*, *acq_doppler_step*, *fs*, *prompt_i*, *prompt_q*, *cn0_db_hz*, *carrier_doppler_hz*, *pseudorange_m*, and *rx_time*. Using up to eight PRN channels simultaneously, the observation vector can reach 72 dimensions per epoch. All features are normalized to the range $[0, 1]$ using Min–Max scaling to ensure a balanced contribution across input dimensions:

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}. \tag{4}$$

The combined data set was split into training, validation and test sets using a stratified split of 70/15/15 to preserve the class distributions. The split was performed at the tracking-feature sample level to evaluate discriminative performance under a tracking-level formulation, rather than transfer across temporally disjoint recordings.

3.3. Feature Analysis

Exploratory analysis confirmed that the chosen features capture distinguishing signatures of spoofing. For instance, C/N_0 remains stable under authentic conditions but exhibits irregular fluctuations and drops during spoofing. Carrier Doppler evolves smoothly for authentic signals but shows abrupt deviations under spoofing, while pseudorange measurements reveal discontinuous shifts in spoofed cases.

3.4. Reinforcement Learning Environment

The spoofing detection problem is formulated as a stateless, single-step RL task. The observation $\mathbf{s}_t \in [0, 1]^9$ is the 9-dimensional normalized feature vector for one PRN. The action space is $\mathcal{A} = \{0, 1, 2, 3, 4, 5, 6\}$, corresponding to one clean and six spoofed scenarios. The reward is defined as:

$$r_t = \begin{cases} +1, & \text{if the prediction is correct,} \\ -1, & \text{otherwise.} \end{cases} \tag{5}$$

To emphasize the difficulty of stealth (ds7) and replay (ds8) attacks, a shaped reward scheme was applied, correct classifications of ds7/ds8 yield $+2$, while incorrect classifications yield -2. This biases the agent toward learning more discriminative policies for subtle spoofing scenarios.

3.5. DQN Architecture and Training

The spoofing classification task was solved using a DQN implemented with the Stable-Baselines3 (SB3) framework [20]. The custom RL environment was implemented as a subclass of `gym.Env` and wrapped with `DummyVecEnv` to enable efficient vectorized training. Each episode corresponds to a single GNSS signal snapshot, ensuring independence between samples.

The Q-network follows the default `MlpPolicy` architecture provided by SB3, consisting of a two-layer feed-forward MLP with 64 hidden units per layer and ReLU activations. The input is the 9-dimensional normalized feature vector, and the output layer produces 7 Q-values corresponding to the clean and six spoofed classes. The online and target networks share the same architecture.

DQNs extend Q-learning by approximating $Q(s, a)$ with a neural network parameterized by θ [18, 21]. The parameters are optimized by minimizing the temporal difference loss:

$$\mathcal{L}(\theta) = \mathbb{E}_{(s,a,r,s')}\Big[\big(y - Q_\theta(s, a)\big)^2\Big], \tag{6}$$

where the target is:

$$y = r + \gamma \max_{a'} Q_{\theta^-}(s', a'). \tag{7}$$

4. EXPERIMENTAL RESULTS

The model was trained with a learning rate of 1×10^{-3}, discount factor $\gamma = 0.99$, a replay buffer of 10,000 transitions, and a mini batch size of 64. The target network was updated every 500 steps. The exploration followed a ϵ-greedy strategy, linearly decaying from 1.0 to 0.05 during the first 50% of the training. Training was run for 300,000 timesteps in total.

Model convergence and generalization were monitored using a validation environment, with training stability assessed via standard learning metrics.

4.1. Evaluation Setup

The DQN model was evaluated on a strictly held-out test set comprising 28,391 samples, covering clean and spoofed GNSS signal observations. Performance was quantified using accuracy, precision, recall, and F1-score, both per class and macro averaged. Visual diagnostics were generated to interpret the decision boundaries of the classifier. In addition, an ablation experiment was conducted by training the same DQN architecture with and without reward shaping to assess its impact on detection performance.

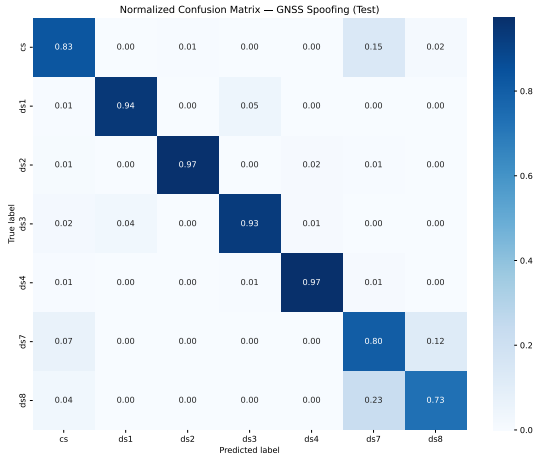

Fig. 3. Normalized confusion matrix of the DQN predictions on the TEXBAT test set.

4.2. Confusion Matrix Analysis

Figure 3 shows the normalized confusion matrix for the test set. The clean class (cs) was correctly classified in 83% of cases, though about 15% of clean signals were misclassified as stealth (ds7). The four classical spoofing scenarios (ds1–ds4) achieved high classification accuracy (94–97%). Stealth (ds7) and replay (ds8) attacks remained more challenging, with accuracies of 80% and 73%, respectively.

Without reward shaping, the F1-scores of ds7 and ds8 decrease to 0.68 and 0.69, compared to 0.74 and 0.78 obtained with the shaped reward. These results demonstrate that the DQN robustly captures observable spoofing distortions but struggles with stealthy manipulations designed to mimic authentic signals. Clean-signal performance remains largely unchanged (F1 $\approx$ 0.80 without shaping versus 0.83 with shaping), indicating that reward shaping improves sensitivity to stealth and replay spoofing without increasing false positives.

4.3. Per-Class Precision, Recall, and F1-Score

To further quantify per-class performance, precision, recall, and F1-scores of the proposed method are reported in Table 2. Classical spoofing scenarios (ds1–ds4) achieved consistently high values, with ds2 obtaining the best results (precision 0.987, recall 0.967, F1 = 0.977). The clean class achieved precision 0.839 and recall 0.831, reflecting false negatives caused by misclassification as ds7. Stealth (ds7) and replay (ds8) showed weaker performance, with F1-scores of 0.735 and 0.780, respectively. Overall, the model achieved a macro F1-score of 0.884 and an average test accuracy of 88.1%.

4.4. Comparison with Prior Work

Most prior TEXBAT studies focused on binary spoofing detection under selected scenarios. For example, kNN and SVM

Table 2. Performance of the DQN model on the TEXBAT test set, showing per class precision, recall, and F1-scores.

Class	Precision	Recall	F1-score
cs	0.839	0.831	0.835
ds1	0.964	0.941	0.953
ds2	0.987	0.967	0.977
ds3	0.938	0.929	0.934
ds4	0.970	0.973	0.971
ds7	0.677	0.803	0.735
ds8	0.838	0.730	0.780
Macro avg.	**0.888**	**0.882**	**0.884**

Table 3. Performance comparison between the DQN model and the Autoencoder on the TEXBAT dataset.

Method	cs	ds1	ds2	ds3	ds4	ds7	ds8
Autoencoder	–	87.5%	100%	75%	62.5%	**0%**	100%
DQN	**83%**	**94%**	**97%**	**93%**	**97%**	**80%**	**73%**

reported accuracies near 0.92–0.94 on ds3/ds8 [11]; however, such binary formulations are not directly comparable to the multi-class setting considered here. As a broader benchmark, we evaluate the open source autoencoder baseline [12], which applies per-channel mean squared error (MSE) thresholds for anomaly detection across TEXBAT scenarios. While effective for overpowered and replay attacks, it fails entirely on stealth spoofing (ds7). In contrast, the proposed DQN jointly classifies seven classes (clean + six spoofing scenarios) and consistently outperforms the autoencoder. Unlike static threshold-based methods, the RL agent learns decision policies through interaction and feedback, enabling more adaptive behavior.

These results confirm the effectiveness of RL for multi-class GNSS spoofing detection. Limitations remain in stealth attacks (ds7, ds8) and false positives on clean signals, motivating future work on temporal modeling and sequential RL frameworks to capture spoofing dynamics beyond single-snapshot classification.

5. CONCLUSION

We presented a reinforcement learning framework for multi-class GNSS spoofing detection using a DQN on TEXBAT dataset features. By framing spoofing detection as a multi-class task, RL can distinguish multiple spoofing scenarios beyond simple clean/spoofed states. Unlike supervised or anomaly-based methods, RL adapts policies through interaction and feedback, generalizing to unseen attacks and reducing reliance on large labeled datasets. These results underscore its potential for resilient GNSS spoofing detection. Future work focuses on evaluating dynamic TEXBAT scenarios.

6. REFERENCES

[1] Scott Madry et al., *Global navigation satellite systems and their applications*, Springer, 2015.

[2] Fabio Dovis, *GNSS interference threats and countermeasures*, Artech House, 2015.

[3] Zachary Clements, James E Yoder, and Todd E Humphreys, "Carrier-phase and imu based gnss spoofing detection for ground vehicles," in *Proceedings of the ION International Technical Meeting, Long Beach, CA*, 2022, pp. 83–95.

[4] European Union Aviation Safety Agency, "Easa safety information bulletin 2024-02: Gnss jamming and spoofing in europe," 2024, Available at: `https://www.easa.europa.eu/`.

[5] Radionavigation Laboratory, University of Texas at Austin, "Spoofing in the mediterranean disrupting everyday activities," 2024, Available at: https://radionavlab.ae.utexas.edu/spoofing-in-the-mediterranean-disrupting-everyday-activities/.

[6] Michael Turner, Stephen Wimbush, Christoph Enneking, and Andriy Konovaltsev, "Spoofing detection by distortion of the correlation function," in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, 2020, pp. 566–574.

[7] Xiaomin Wei, Cong Sun, Xinghua Li, and Jianfeng Ma, "Gnss spoofing detection for uavs using doppler frequency and carrier-to-noise density ratio," *Journal of Systems Architecture*, vol. 153, pp. 103212, 2024.

[8] Katarina Babić, Marta Balić, and Dinko Begušić, "Gnss spoofing detection based on wavelets and machine learning," *Electronics*, vol. 14, no. 12, pp. 2391, 2025.

[9] Parisa Borhani-Darian, Haoqing Li, Peng Wu, and Pau Closas, "Detecting gnss spoofing using deep learning," *EURASIP Journal on advances in signal processing*, vol. 2024, no. 1, pp. 14, 2024.

[10] Silvio Semanjski, Ivana Semanjski, Wim De Wilde, and Alain Muls, "Use of supervised machine learning for gnss signal spoofing detection with validation on real-world meaconing and spoofing data—part i," *Sensors*, vol. 20, no. 4, pp. 1171, 2020.

[11] Asra Mahroof, Imtiaz Nabi, Salma Zainab Farooq, and Najam Abbas Naqvi, "Machine learning-based detection of spoofing attacks in gnss: a study using texbat dataset," in *2024 14th international conference on electrical engineering (ICEENG)*. IEEE, 2024, pp. 90–95.

[12] T. Seki, "Gnss spoof detector," `https://github.com/seki5405/gnss_spoof_detector`, 2020, Open-source autoencoder baseline on TEXBAT.

[13] Nguyen Cong Luong, Dinh Thai Hoang, Shimin Gong, Dusit Niyato, Ping Wang, Ying-Chang Liang, and Dong In Kim, "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE communications surveys & tutorials*, vol. 21, no. 4, pp. 3133–3174, 2019.

[14] Yonghua Wang, Zifeng Ye, Pin Wan, and Jiajun Zhao, "A survey of dynamic spectrum allocation based on reinforcement learning algorithms in cognitive radio networks," *Artificial intelligence review*, vol. 51, no. 3, pp. 493–506, 2019.

[15] Felix Obite, Aliyu D Usman, and Emmanuel Okafor, "An overview of deep reinforcement learning for spectrum sensing in cognitive radio networks," *Digital Signal Processing*, vol. 113, pp. 103014, 2021.

[16] Abdullah Al Noman Chowdhury, "Machine learning and statistical methods for gnss spoofing detection: A systematic review and deep reinforcement learning simulation," M.S. thesis, University of Vaasa, 2025.

[17] Sagar Dasgupta, Tonmoy Ghosh, and Mizanur Rahman, "A reinforcement learning approach for global navigation satellite system spoofing attack detection in autonomous vehicles," *Transportation research record*, vol. 2676, no. 12, pp. 318–330, 2022.

[18] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 2018.

[19] Todd E. Humphreys, Jahshan A. Bhatti, Daniel P. Shepard, and Kyle D. Wesson, "The texas spoofing test battery: Toward a standard for evaluating gps signal authentication techniques," in *Proc. ION GNSS*, Nashville, TN, 2012, Institute of Navigation, pp. 3569–3583.

[20] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann, "Stable-baselines3: Reliable reinforcement learning implementations," *Journal of machine learning research*, vol. 22, no. 268, pp. 1–8, 2021.

[21] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.