STAR-RL: Stealth-Aware Targeted Adversarial Attack on Multimodal Sensors in Human Activity Recognition via Reinforcement Learning

Ade Kurniawan *

Posted Date: 9 January 2026

doi: 10.20944/preprints202601.0738.v1

Keywords: adversarial machine learning; human activity recognition; reinforcement learning; multimodal sensor fusion; deep learning security; time series classification

 doi:10.20944/preprints202601.0738.v1

Article

STAR-RL: Stealth-Aware Targeted Adversarial Attack on Multimodal Sensors in Human Activity Recognition via Reinforcement Learning

Ade Kurniawan

Department of Data Science, Institut Teknologi Sains Bandung, Kota Deltamas Lot-A1 CBD, Bekasi, 17530 Jawa Barat, Indonesia

Abstract

Deep learning-based Human Activity Recognition (HAR) systems using multimodal wearable sensors are increasingly deployed in safety-critical applications including healthcare monitoring, elderly care, and security authentication. However, the vulnerability of these systems to adversarial attacks remains insufficiently understood, particularly for attacks that must evade detection while manipulating multiple sensor modalities simultaneously. This paper presents STAR-RL (Stealth-aware Targeted Adversarial attack via Reinforcement Learning), a novel framework that generates effective and stealthy adversarial examples against multimodal sensor-based HAR systems. STAR-RL introduces three key innovations: (1) a multi-strategy attack engine that adaptively selects among diverse perturbation algorithms based on real-time attack progress, (2) a sensor-aware stealth mechanism that concentrates perturbations on naturally noisy sensors to minimize detection likelihood, and (3) a reinforcement learning-based meta-controller that learns optimal attack policies through interaction with the target classifier. Comprehensive experiments on the MHEALTH dataset demonstrate that STAR-RL achieves 95.20% attack success rate, substantially outperforming baseline methods including FGSM (6.00%), PGD (88.60%), and C&W (69.00%). The stealth analysis confirms that 51.35% of perturbation energy is successfully directed to weak sensors (gyroscopes and magnetometers), validating the effectiveness of the sensor-aware allocation strategy. Our findings reveal critical security vulnerabilities in production HAR systems and provide insights for developing robust defense mechanisms against adaptive adversarial threats.

Keywords: adversarial machine learning; human activity recognition; reinforcement learning; multimodal sensor fusion; deep learning security; time series classification

1. Introduction

Human Activity Recognition (HAR) systems based on wearable sensors have emerged as foundational technologies enabling a wide spectrum of applications, from personalized healthcare monitoring [1] and elderly fall detection [2] to fitness tracking [3] and security authentication [4]. The proliferation of smartwatches, fitness bands, and medical-grade wearable devices has accelerated the deployment of deep learning models that process continuous streams of accelerometer, gyroscope, magnetometer, and physiological sensor data to infer user activities in real-time. As these systems increasingly govern safety-critical decisions—triggering emergency responses, authorizing secure access, or adjusting medical interventions—their reliability and security have become paramount concerns [5,6].

Despite remarkable advances in HAR accuracy through sophisticated deep learning architectures, including convolutional neural networks, recurrent models, and attention mechanisms, a fundamental question remains inadequately addressed: *How vulnerable are these systems to deliberately crafted adversarial perturbations, and can such attacks be executed stealthily to evade detection?* This question carries profound implications for the trustworthiness of HAR deployments in adversarial environments where malicious actors may seek to manipulate system outputs for nefarious purposes.

 doi:10.20944/preprints202601.0738.v1

The phenomenon of adversarial examples—carefully crafted inputs that induce misclassification in machine learning models while appearing imperceptibly different from legitimate samples—was first demonstrated in the image domain [7] and has since been extensively studied across computer vision applications [8,9]. However, the extension of adversarial attacks to time series classification, particularly multimodal sensor data, presents unique challenges that existing methods fail to adequately address. Unlike static images, sensor time series exhibit complex temporal dependencies, inter-channel correlations, and modality-specific noise characteristics that must be carefully considered when generating effective yet stealthy perturbations.

Recent investigations into adversarial attacks on time series classification [10] have established that deep learning models for sequential data are indeed vulnerable to adversarial manipulation. Extensions to HAR systems have revealed attack vectors through WiFi channel state information [11,12], radar signals [13], and skeleton-based action recognition [14,15]. Our prior work demonstrated the feasibility of generating adversarial examples against multimodal sensor-based HAR models [16] and developed detection mechanisms for identifying compromised sensors [17]. However, existing approaches suffer from three critical limitations that constrain their effectiveness in practical attack scenarios.

First, conventional attack methods employ fixed perturbation strategies that fail to adapt to the varying difficulty of different input samples. Single-step attacks such as FGSM [7] prove largely ineffective for complex temporal patterns, while iterative methods like PGD [8] may converge to suboptimal local minima. Optimization-based approaches such as C&W [9] achieve high-quality perturbations but at prohibitive computational cost. The absence of adaptive strategy selection mechanisms limits the robustness of existing attacks across diverse input characteristics.

Second, prior work largely ignores the heterogeneous nature of multimodal sensor data when generating adversarial perturbations. In wearable HAR systems, different sensors exhibit vastly different noise profiles: accelerometers and ECG signals typically maintain high signal-to-noise ratios, while gyroscopes and magnetometers are inherently more susceptible to environmental interference and measurement noise. Perturbations uniformly distributed across all sensors are more likely to trigger anomaly detection systems monitoring sensor signal quality, whereas strategic allocation to naturally noisy channels could evade such defenses.

Third, the application of reinforcement learning for adaptive attack generation in the HAR domain remains unexplored. While RL-based approaches have shown promise for black-box attacks in image classification [18] and for discovering adversarial policies in control systems [19], the unique characteristics of multimodal time series—including temporal dependencies, sensor correlations, and class-specific vulnerability patterns—have not been leveraged to guide intelligent attack strategy selection.

To address these limitations, this paper presents STAR-RL (**S**tealth-aware **T**argeted **A**dversarial attack via **R**einforcement **L**earning), a novel framework for generating effective and stealthy adversarial examples against multimodal sensor-based HAR systems. STAR-RL makes three principal contributions:

1. **Multi-Strategy Attack Engine:** We develop an attack engine incorporating four complementary perturbation strategies—PGD-Strong, Momentum Iterative FGSM (MI-FGSM), C&W-High, and PGD-Long—each optimized for different attack scenarios. This diversity enables robust performance across input samples with varying difficulty characteristics, including "hard classes" that resist standard attack methods.
2. **Sensor-Aware Stealth Mechanism:** We introduce a novel perturbation allocation strategy that exploits the heterogeneous noise profiles of multimodal sensors. By concentrating adversarial modifications on naturally noisy sensors (gyroscopes, magnetometers) while constraining perturbations to reliable sensors (accelerometers, ECG), STAR-RL generates adversarial examples with enhanced stealth properties that are more likely to evade anomaly detection systems.
3. **RL-Based Meta-Controller:** We formulate adversarial attack generation as a Markov Decision Process (MDP) and train a reinforcement learning agent to dynamically select attack strategies

 doi:10.20944/preprints202601.0738.v1

and hyperparameters based on real-time feedback from the target classifier. The meta-controller learns to adapt its policy to sample-specific characteristics, achieving intelligent automation of the attack generation process.

Comprehensive experiments on the MHEALTH dataset, a widely-adopted benchmark for multimodal HAR, demonstrate the effectiveness of STAR-RL. Our framework achieves 95.20% attack success rate (ASR) for targeted misclassification, substantially outperforming FGSM (6.00%), PGD (88.60%), and C&W (69.00%). The stealth analysis confirms that 51.35% of total perturbation energy is successfully directed to weak sensors, validating the sensor-aware allocation strategy. Ablation studies verify the contribution of each framework component: removing the multi-strategy engine reduces ASR by 6.6 percentage points, while disabling optimal target selection decreases performance by 5.4 percentage points.

Beyond demonstrating attack effectiveness, our work carries important implications for HAR system security. The 95.20% ASR achieved against a representative Time-Distributed LSTM architecture indicates that production HAR systems are highly vulnerable to sophisticated adversarial manipulation. The feasibility of stealth attacks concentrating perturbations on specific sensors suggests that conventional anomaly detection may be insufficient for identifying adversarial inputs. These findings motivate the urgent development of adversarially robust HAR architectures and defense mechanisms.

The remainder of this paper is organized as follows. Section 2 reviews related work on adversarial attacks for time series classification, HAR systems, and reinforcement learning-based attack methods. Section 3 establishes the mathematical foundations, including problem formulation and baseline attack methods. Section 4 presents the proposed STAR-RL framework in detail, covering the multi-strategy attack engine, sensor-aware stealth mechanism, and RL meta-controller. Section 5 describes the experimental setup and presents comprehensive evaluation results. Section 6 discusses the implications of our findings, acknowledges limitations, and outlines future directions. Section 7 concludes the paper.

2. Related Works

The proliferation of deep learning models in HAR systems has raised significant concerns regarding their vulnerability to adversarial attacks. This section provides a comprehensive review of the existing literature, organized into four main themes: adversarial attacks on time series data, security challenges in HAR systems, reinforcement learning-based attack methodologies, and multi-sensor fusion vulnerabilities.

2.1. Adversarial Attacks on Time Series Data

The seminal work by Goodfellow et al. [7] introduced the Fast Gradient Sign Method (FGSM), demonstrating that imperceptible perturbations could fool deep neural networks with high confidence. Subsequently, Madry et al. [8] proposed Projected Gradient Descent (PGD) as a stronger iterative attack, establishing it as a benchmark for adversarial robustness evaluation. Carlini and Wagner [9] further advanced the field by formulating adversarial perturbation generation as an optimization problem, achieving superior attack success rates with minimal perturbation magnitudes.

The extension of adversarial attacks to time series classification was pioneered by Fawaz et al. [20], who systematically investigated the vulnerability of deep learning models for time series analysis. Their work revealed that temporal patterns in sequential data present unique challenges for generating effective adversarial perturbations while maintaining temporal coherence. Karim et al. [10] expanded this investigation by proposing attacks specifically designed for multivariate time series, demonstrating that the correlations between different channels could be exploited to craft more effective perturbations.

Recent developments in time series adversarial attacks have focused on enhancing imperceptibility and attack efficiency. Pialla et al. [21,22] introduced smooth perturbation constraints that preserve the natural characteristics of time series data, addressing the challenge of generating adversarial examples

 doi:10.20944/preprints202601.0738.v1

that evade both machine learning models and human observers. Harford et al. [23] investigated adversarial attacks on multivariate time series forecasting models, demonstrating that prediction systems are particularly susceptible to carefully crafted perturbations at critical temporal points. The TSadv method proposed by Yang et al. [24] leveraged local perturbations for black-box attacks, achieving significant attack success rates while restricting modifications to specific temporal segments. More recently, Wang et al. [25] proposed TSFool, a multi-objective optimization framework that simultaneously maximizes attack effectiveness while minimizing perturbation perceptibility.

The temporal characteristics of time series data have been specifically exploited in several recent works. Shen and Li [26] proposed temporal characteristics-based adversarial attacks that leverage the inherent periodicity and trend patterns in sensor data. Wu et al. [27] introduced importance-based perturbation strategies that concentrate adversarial modifications on the most influential temporal segments, demonstrating that small perturbations are sufficient to compromise time series prediction models. Correlation analysis studies by Li et al. [28,29] revealed that the relationships between time steps can be leveraged to enhance attack transferability across different model architectures.

2.2. Adversarial Attacks on Human Activity Recognition Systems

Human Activity Recognition systems present unique attack surfaces due to their reliance on continuous sensor streams from multiple modalities. Wang et al. [5] and Nweke et al. [6] provided comprehensive surveys of deep learning approaches for sensor-based HAR, establishing the foundation for understanding the architectures vulnerable to adversarial manipulation. Sakka et al. [30] systematically analyzed security vulnerabilities in HAR systems, identifying critical attack vectors across different sensing modalities.

Sensor-based HAR systems have been shown to be particularly vulnerable to adversarial attacks. Kurniawan et al. [16] conducted extensive experiments on adversarial examples targeting multimodal sensor-based HAR models, demonstrating that perturbations applied to specific sensor channels can propagate misclassification across the entire recognition pipeline. Their subsequent work [17] focused on detecting which sensors were exploited in adversarial attacks, providing insights into the sensor-specific vulnerabilities in multi-sensor HAR systems.

The extension of adversarial attacks to various sensing modalities has revealed modality-specific vulnerabilities. Ozbulak et al. [13] investigated adversarial attacks on radar-based HAR systems, establishing connections between adversarial perturbation patterns and model interpretability. For skeleton-based action recognition, Diao et al. [14] proposed BASAR, a black-box attack framework that manipulates skeletal joint positions to induce misclassification. Wang et al. [15] introduced SMART, which exploits the temporal dynamics of skeleton sequences to generate adversarial perturbations that appear natural under human observation. Physical-world skeleton attacks have been further explored by Zheng et al. [31], demonstrating the feasibility of backdoor attacks in real-world deployment scenarios.

WiFi-based sensing systems have emerged as attractive targets for adversarial attacks due to their widespread deployment in smart environments. Zhou et al. [11] developed WiAdv, achieving practical adversarial attacks against WiFi gesture recognition systems through carefully designed channel state information (CSI) perturbations. Xu et al. [32] proposed WiCAM, leveraging attention mechanisms to identify and perturb the most critical CSI features. Li et al. [33] demonstrated that adversarial perturbations can be injected through pilot symbol manipulation in communication packets. Huang et al. [12] introduced IS-WARS, an intelligent and stealthy attack framework that balances attack effectiveness with perturbation imperceptibility. The emerging field of WiFi sensing security has been comprehensively surveyed in recent works [34–36], highlighting the urgent need for robust defense mechanisms.

2.3. Reinforcement Learning for Adversarial Attacks

Reinforcement learning has emerged as a powerful paradigm for generating adaptive adversarial attacks. Tsingenopoulos et al. [18] pioneered the AutoAttacker framework, utilizing reinforcement

 doi:10.20944/preprints202601.0738.v1

learning to automatically discover effective attack policies in black-box settings. Zhang et al. [37] investigated robust deep reinforcement learning against adversarial perturbations on state observations, providing theoretical foundations for understanding RL vulnerability to input manipulation. The ATLA framework by Zhang et al. [38] learned optimal adversary policies that could efficiently exploit weaknesses in RL-based systems.

Several works have explored RL-guided attack strategies for temporal data. Gleave et al. [19] demonstrated that adversarial policies can defeat well-trained RL agents through strategic perturbations. Wu et al. [39] integrated explainability-guided search into adversarial policy learning, achieving targeted attacks with interpretable perturbation patterns. He et al. [40] proposed a singular value manipulating attack (SVMA) using deep reinforcement learning with soft actor-critic, demonstrating effective black-box transferable attacks on deep convolutional neural networks by perturbing the singular value matrix instead of directly manipulating pixels. García et al. [41] formulated adversarial attack generation as a multi-objective reinforcement learning problem, enabling simultaneous optimization of attack success rate and perturbation minimization. Recent advances in RL-based video attacks [42] have demonstrated the effectiveness of sparse, temporally targeted perturbations guided by learned policies.

The vulnerability of deep reinforcement learning systems themselves has been extensively studied. Schott et al. [43] provided a comprehensive survey of adversarial attacks and training methods for robust deep reinforcement learning. Ilahi et al. [44] analyzed challenges and countermeasures for adversarial attacks on deep RL systems. Sun et al. [45] proposed stealthy and efficient adversarial attacks that minimize the number of perturbed states while maintaining high attack success rates. Oikarinen et al. [46] introduced adversarial loss functions for improving RL robustness, while Sun et al. [47] proposed PA-AD (Policy perturbation direction with Adversarial Director), a theoretically optimal and efficient evasion attack algorithm that decouples the attacking process into policy perturbation and state perturbation components. Recent surveys [48,49] have highlighted the increasing sophistication of attacks against both single-agent and multi-agent RL systems.

2.4. Multi-Sensor Fusion Vulnerabilities

The integration of multiple sensing modalities introduces complex attack surfaces that have been increasingly studied. Cao et al. [50] demonstrated adversarial attacks on LiDAR-based perception in autonomous driving, revealing that sensor fusion mechanisms can be exploited through targeted perturbations. Multi-sensor fusion attacks have been investigated by Zhu et al. [51], showing that adversarial perturbations can propagate through fusion networks in unexpected ways. The FusionRipper framework [52] demonstrated attacks on multi-sensor localization systems through GPS spoofing. Tian and Xu [53] investigated whether audio-visual integration could strengthen robustness under multimodal attacks, finding that fusion architectures can both mitigate and amplify adversarial vulnerabilities depending on the attack strategy.

Recent work on multi-modal attack detection [54] has proposed frameworks for identifying adversarial perturbations across different modalities. Defense mechanisms specifically designed for multi-modal systems have been investigated by Wang et al. [55], providing provable robustness guarantees for multi-modal models.

2.5. Research Gap and Our Contributions

Table 1 provides a systematic comparison of our proposed STAR-RL framework with existing adversarial attack methods. As evident from the comparison, existing approaches exhibit significant limitations when applied to multimodal sensor-based HAR systems. First, most methods are designed for single-modality inputs and cannot effectively handle the complex inter-sensor dependencies present in multi-sensor HAR systems. Second, existing attacks lack adaptive mechanisms to dynamically adjust perturbation strategies based on classifier behavior and sensor characteristics. Third, few methods consider the stealth requirements necessary for practical attacks on real-world HAR deployments, where perturbations should be concentrated on sensors with lower detection likelihood.

 doi:10.20944/preprints202601.0738.v1

Our proposed STAR-RL framework addresses these gaps through three key innovations: (1) a reinforcement learning-based meta-controller that adaptively selects attack strategies based on real-time classifier feedback, (2) sensor-aware perturbation allocation that concentrates modifications on weak sensors to maximize stealth, and (3) a multi-strategy attack mechanism that employs different optimization techniques for hard-to-attack activity classes. To the best of our knowledge, STAR-RL represents the first RL-guided adversarial attack framework specifically designed for multimodal sensor-based HAR systems.

Table 1. Comparison of adversarial attack methods for time series and HAR systems. TS: Time Series; HAR: Human Activity Recognition; MTS: Multivariate Time Series; WB: White-box; BB: Black-box; RL: Reinforcement Learning.

Method	Domain	Input Type	Attack Type	RL-Based	Adaptive	Stealth	Key Limitation
FGSM [7]	General	Image/TS	WB	×	×	×	Single-step, no temporal modeling
PGD [8]	General	Image/TS	WB	×	×	×	Fixed perturbation budget
C&W [9]	General	Image/TS	WB	×	×	✓	High computational cost
Fawaz et al. [10]	TS	Univariate TS	WB	×	×	×	Single-channel only
Karim et al. [10]	TS	MTS	WB	×	×	×	No sensor-aware strategy
TSadv [24]	TS	MTS	BB	×	×	✓	Local perturbation only
TSFool [25]	TS	MTS	WB/BB	×	×	✓	No class-specific adaptation
Smooth Attack [22]	TS	Univariate TS	WB	×	×	✓	Limited to smooth perturbations
Kurniawan et al. [16]	HAR	Multi-sensor	WB	×	×	×	No adaptive strategy
WiAdv [11]	HAR	WiFi CSI	BB	×	×	✓	WiFi-specific only
BASAR [14]	HAR	Skeleton	BB	×	✓	×	Skeleton data only
SMART [15]	HAR	Skeleton	WB	×	×	✓	No multi-sensor support
IS-WARS [12]	HAR	WiFi CSI	WB/BB	×	✓	✓	WiFi-specific only
AutoAttacker [18]	General	Various	BB	✓	✓	×	Not designed for HAR
RLVS [42]	Video	Video	BB	✓	✓	✓	Video domain only
STAR-RL (Ours)	**HAR**	**Multi-sensor**	**WB/BB**	✓	✓	✓	–

3. Background

This section presents the mathematical foundations underlying our proposed STAR-RL framework. We formulate the problem of adversarial attacks on multimodal sensor-based HAR systems and establish the theoretical basis for our reinforcement learning-guided attack methodology.

3.1. Multimodal Sensor-Based Human Activity Recognition

Consider a multimodal HAR system that processes sensor data from M distinct sensing modalities. Let $\mathbf{X} \in \mathbb{R}^{T \times D}$ represent a time series input, where T denotes the number of time steps and D represents the total feature dimensionality across all sensors. The feature space can be decomposed as $D = \sum_{m=1}^{M} d_m$, where d_m denotes the dimensionality of the m-th sensor modality. For the MHEALTH dataset [56] employed in our experiments, we have $M = 8$ sensor groups comprising accelerometers, gyroscopes, magnetometers, and ECG sensors distributed across chest, ankle, and wrist positions, yielding $D = 23$ total features.

The HAR classifier is represented as a function $f_\theta : \mathbb{R}^{T \times D} \to \mathbb{R}^K$, parameterized by θ, which maps an input time series to a K-dimensional logit vector, where K is the number of activity classes. The predicted class is obtained through:

$$\hat{y} = \arg \max_{k \in \{1,\ldots,K\}} \sigma(f_\theta(\mathbf{X}))_k \tag{1}$$

 doi:10.20944/preprints202601.0738.v1

where $\sigma(\cdot)$ denotes the softmax function. The classifier's confidence for class k is expressed as:

$$p_k(\mathbf{X}) = \frac{\exp(f_\theta(\mathbf{X})_k)}{\sum_{j=1}^{K} \exp(f_\theta(\mathbf{X})_j)} \tag{2}$$

3.2. Problem Formulation: Targeted Adversarial Attacks

Given an input sample $\mathbf{X}$ with true label $y \in \{1, \ldots, K\}$ and a target class $y^* \neq y$, the objective of a targeted adversarial attack is to find a perturbation $\delta \in \mathbb{R}^{T \times D}$ such that the adversarial example $\mathbf{X}^{adv} = \mathbf{X} + \delta$ is classified as y^* while minimizing the perturbation magnitude. Formally, this can be expressed as:

$$\begin{aligned} \min_{\delta} \quad & \|\delta\|_p \\ \text{s.t.} \quad & \arg\max_k \; p_k(\mathbf{X} + \delta) = y^* \\ & \|\delta\|_p \leq \epsilon \end{aligned} \tag{3}$$

where $\|\cdot\|_p$ denotes the ℓ_p-norm (typically $p \in \{2, \infty\}$) and ϵ is the maximum allowable perturbation magnitude.

In practice, the hard constraint formulation in Equation (3) is often relaxed to an unconstrained optimization problem using a Lagrangian formulation:

$$\mathcal{L}(\delta) = -\log p_{y^*}(\mathbf{X} + \delta) + \lambda \|\delta\|_2^2 \tag{4}$$

where $\lambda > 0$ controls the trade-off between attack success and perturbation minimization.

3.3. Baseline Attack Methods

3.3.1. Fast Gradient Sign Method (FGSM)

The FGSM attack generates adversarial perturbations through a single gradient step:

$$\delta_{FGSM} = -\epsilon \cdot \text{sign}(\nabla_{\mathbf{X}} \mathcal{L}_{CE}(f_\theta(\mathbf{X}), y^*)) \tag{5}$$

where $\mathcal{L}_{CE}$ denotes the cross-entropy loss and $\text{sign}(\cdot)$ is the element-wise sign function. While computationally efficient, FGSM often produces suboptimal perturbations due to its single-step nature.

3.3.2. Projected Gradient Descent (PGD)

PGD iteratively refines the perturbation through multiple gradient steps with projection onto the ϵ-ball:

$$\delta^{(t+1)} = \Pi_{\mathcal{B}_\epsilon}\left(\delta^{(t)} - \alpha \cdot \text{sign}\left(\nabla_\delta \mathcal{L}_{CE}(f_\theta(\mathbf{X} + \delta^{(t)}), y^*)\right)\right) \tag{6}$$

where α is the step size, $\Pi_{\mathcal{B}_\epsilon}$ denotes projection onto the ℓ_∞-ball of radius ϵ, and t indexes the iteration.

3.3.3. Carlini-Wagner (C&W) Attack

The C&W attack formulates adversarial example generation as a constrained optimization problem:

$$\min_{\mathbf{w}} \; \|\frac{1}{2}(\tanh(\mathbf{w}) + 1) - \mathbf{X}\|_2^2 + c \cdot g(\mathbf{X} + \delta(\mathbf{w})) \tag{7}$$

where $\delta(\mathbf{w}) = \frac{1}{2}(\tanh(\mathbf{w}) + 1) - \mathbf{X}$, $c > 0$ is a regularization constant, and $g(\cdot)$ is defined as:

$$g(\mathbf{X}') = \max\left(\max_{k \neq y^*} f_\theta(\mathbf{X}')_k - f_\theta(\mathbf{X}')_{y^*}, -\kappa\right) \tag{8}$$

with $\kappa \geq 0$ controlling the confidence margin.

 doi:10.20944/preprints202601.0738.v1

3.4. Reinforcement Learning Formulation for Adaptive Attacks

We formulate the adversarial attack problem as a Markov Decision Process (MDP) to enable adaptive, context-aware attack generation. The MDP is defined by the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$:

State Space ($\mathcal{S}$): The state at step t encodes information about the current attack progress:

$$s_t = \left[\mathbf{h}_t, p_y(\mathbf{X}^{(t)}), p_{y^*}(\mathbf{X}^{(t)}), \|\delta^{(t)}\|_2, \mathbf{c}\right] \tag{9}$$

where $\mathbf{h}_t$ represents the hidden representation from the victim model, p_y and p_{y^*} are the classifier's confidences for the true and target classes respectively, $\|\delta^{(t)}\|_2$ is the current perturbation magnitude, and $\mathbf{c}$ is a one-hot encoding of the source class indicating whether it belongs to the set of "hard classes" identified through empirical analysis.

Action Space ($\mathcal{A}$): The action space comprises attack strategy selections and hyperparameter configurations:

$$a_t \in \{\text{PGD}, \text{MI-FGSM}, \text{C\&W}\} \times \mathbb{R}^+ \times \mathbb{R}^+ \tag{10}$$

where the continuous components specify the perturbation magnitude ϵ_t and step size α_t.

Transition Dynamics ($\mathcal{P}$): The transition function captures how the attack state evolves:

$$s_{t+1} = \mathcal{P}(s_t, a_t) = \phi\left(f_\theta(\mathbf{X} + \delta^{(t+1)})\right) \tag{11}$$

where $\phi(\cdot)$ extracts the state representation from the model's output.

Reward Function ($\mathcal{R}$): The reward function balances attack success against perturbation stealth:

$$r_t = \underbrace{p_{y^*}(\mathbf{X}^{(t)}) - p_{y^*}(\mathbf{X}^{(t-1)})}_{\text{confidence improvement}} - \underbrace{\beta \cdot \|\delta^{(t)} - \delta^{(t-1)}\|_2}_{\text{perturbation penalty}} + \underbrace{\mathbb{I}[\hat{y}^{(t)} = y^*] \cdot R_{success}}_{\text{success bonus}} \tag{12}$$

where $\beta > 0$ weights the perturbation penalty, $\mathbb{I}[\cdot]$ is the indicator function, and $R_{success}$ is a bonus for successful targeted misclassification.

3.5. Sensor-Aware Stealth Constraints

To ensure perturbations remain undetectable, we introduce sensor-specific constraints based on the observation that certain sensors are more susceptible to noise and thus better suited for perturbation injection. Let $\mathcal{S}_{weak} \subset \{1, \ldots, M\}$ denote the set of weak sensors (e.g., gyroscopes and magnetometers at ankle and wrist positions). The stealth-aware perturbation constraint is formulated as:

$$\|\delta[:, \mathcal{I}_m]\|_2 \leq \begin{cases} \epsilon_w & \text{if } m \in \mathcal{S}_{weak} \\ \epsilon_s & \text{if } m \notin \mathcal{S}_{weak} \end{cases} \tag{13}$$

where $\mathcal{I}_m$ denotes the feature indices corresponding to sensor m, and $\epsilon_w > \epsilon_s$ allows larger perturbations on weak sensors while limiting modifications to strong sensors (e.g., chest accelerometer, ECG).

The stealth score is quantified as the ratio of perturbation energy allocated to weak sensors:

$$\text{Stealth} = \frac{\sum_{m \in \mathcal{S}_{weak}} \|\delta[:, \mathcal{I}_m]\|_2^2}{\|\delta\|_F^2} \tag{14}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. A higher stealth score indicates better concealment of adversarial perturbations.

 doi:10.20944/preprints202601.0738.v1

3.6. Optimal Target Selection

To maximize attack success rates, we employ an intelligent target selection strategy based on the victim model's confusion characteristics. For each source class y, we compute the average misclassification probability to potential target classes:

$$\text{score}(y, y^*) = \mathbb{E}_{\mathbf{X}|y}\left[p_{y^*}(\mathbf{X})\right], \quad \forall y^* \neq y \tag{15}$$

Target classes are ranked in descending order of this score, and the highest-ranked target is selected for attack:

$$y^*_{opt} = \arg\max_{y^* \neq y} \text{score}(y, y^*) \tag{16}$$

This strategy exploits the inherent confusion patterns in the classifier, selecting targets that are already prone to confusion with the source class, thereby facilitating more efficient adversarial perturbation generation.

3.7. Policy Optimization

The RL agent's policy $\pi_\phi(a|s)$, parameterized by ϕ, is optimized using the Proximal Policy Optimization (PPO) algorithm to maximize expected cumulative reward:

$$J(\phi) = \mathbb{E}_{\tau \sim \pi_\phi}\left[\sum_{t=0}^{H} \gamma^t r_t\right] \tag{17}$$

where $\tau = (s_0, a_0, r_0, \ldots, s_H)$ denotes a trajectory, H is the horizon, and $\gamma \in (0, 1)$ is the discount factor. The policy is updated using the clipped surrogate objective to ensure stable learning:

$$L^{CLIP}(\phi) = \mathbb{E}_t\left[\min\left(\rho_t(\phi)\hat{A}_t, \text{clip}(\rho_t(\phi), 1 - \epsilon_{clip}, 1 + \epsilon_{clip})\hat{A}_t\right)\right] \tag{18}$$

where $\rho_t(\phi) = \frac{\pi_\phi(a_t|s_t)}{\pi_{\phi_{old}}(a_t|s_t)}$ is the probability ratio and $\hat{A}_t$ is the estimated advantage function.

4. Proposed Methodology

Building upon the mathematical foundations established in Section 3, this section presents the detailed architecture and algorithmic design of the proposed STAR-RL framework. Our methodology addresses the key limitations identified in Section 2 by introducing an adaptive, sensor-aware attack generation system that leverages reinforcement learning for intelligent strategy selection.

4.1. System Overview

The STAR-RL framework comprises four interconnected modules: (1) the Victim Model representing the target HAR classifier, (2) the Target Analyzer that identifies optimal attack targets based on classifier confusion patterns, (3) the Multi-Strategy Attack Engine implementing diverse perturbation generation algorithms, and (4) the RL Meta-Controller that adaptively orchestrates attack strategy selection. Figure 1 illustrates the overall architecture and information flow within the proposed system.

 doi:10.20944/preprints202601.0738.v1

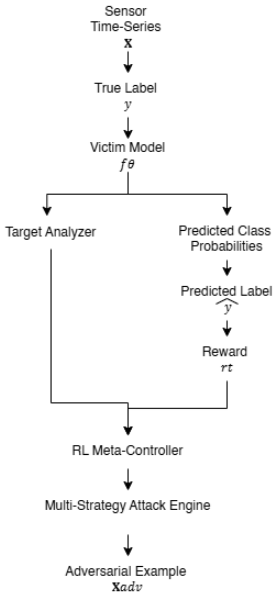

Figure 1. Overview of the proposed STAR-RL framework architecture. The system processes sensor time-series input $\mathbf{X}$ through the Victim Model f_θ to obtain predicted class probabilities. The Target Analyzer identifies optimal misclassification targets, while the RL Meta-Controller orchestrates adaptive strategy selection based on reward feedback r_t. The Multi-Strategy Attack Engine generates adversarial examples $\mathbf{X}^{adv}$ through iterative perturbation refinement.

The attack pipeline operates as follows. Given an input sample $\mathbf{X}$ with true label y, the Target Analyzer first determines the optimal target class y^* by analyzing the victim model's confusion characteristics. Subsequently, the RL Meta-Controller evaluates the current attack state and selects an appropriate strategy from the Multi-Strategy Attack Engine. The selected strategy generates adversarial perturbations, which are then evaluated against the victim model. The predicted label $\hat{y}$ and associated probabilities are used to compute the reward signal r_t, which updates the RL Meta-Controller's policy, enabling continuous adaptation throughout the attack process.

4.2. Victim Model Architecture

The victim model in our experimental setup employs a Time-Distributed LSTM (TD-LSTM) architecture specifically designed for multimodal sensor-based HAR. This architecture processes sequential sensor data through two stages: temporal feature extraction and sequence modeling. Figure 2 presents the detailed architecture of the victim model.

 doi:10.20944/preprints202601.0738.v1

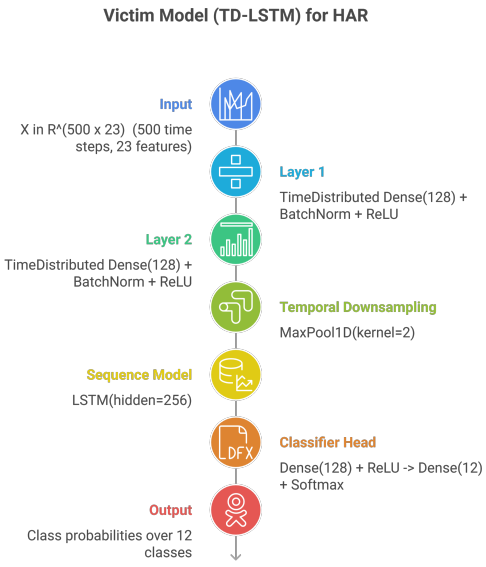

Figure 2. Architecture of the Time-Distributed LSTM (TD-LSTM) victim model for HAR classification. The model processes input tensor $\mathbf{X} \in \mathbb{R}^{500\times 23}$ through two time-distributed dense layers with batch normalization, temporal downsampling via max-pooling, LSTM-based sequence modeling, and a two-layer classifier head producing probabilities over 12 activity classes.

The temporal feature extraction stage applies shared dense layers across all time steps using a time-distributed configuration. For an input tensor $\mathbf{X} \in \mathbb{R}^{T\times D}$ where $T = 500$ time steps and $D = 23$ features, the first time-distributed layer (Layer 1) transforms each time step independently:

$$\mathbf{h}_t^{(1)} = \mathrm{BN}(\mathrm{ReLU}(\mathbf{W}_1\mathbf{x}_t + \mathbf{b}_1)), \quad t = 1, \ldots, T \tag{19}$$

where $\mathbf{W}_1 \in \mathbb{R}^{128\times D}$ denotes the weight matrix, $\mathbf{b}_1 \in \mathbb{R}^{128}$ is the bias vector, and $\mathrm{BN}(\cdot)$ represents batch normalization. A second time-distributed layer (Layer 2) with identical structure further refines the representations:

$$\mathbf{h}_t^{(2)} = \mathrm{BN}\Big(\mathrm{ReLU}\Big(\mathbf{W}_2\mathbf{h}_t^{(1)} + \mathbf{b}_2\Big)\Big) \tag{20}$$

Following temporal feature extraction, a temporal downsampling operation using max-pooling with kernel size 2 reduces the temporal dimension by half, yielding $\mathbf{H}^{pool} \in \mathbb{R}^{T/2\times 128}$. The sequence modeling stage employs a single-layer LSTM with 256 hidden units:

$$\mathbf{o}_1, \ldots, \mathbf{o}_{T/2} = \mathrm{LSTM}(\mathbf{h}_1^{pool}, \ldots, \mathbf{h}_{T/2}^{pool}) \tag{21}$$

The final hidden state $\mathbf{o}_{T/2}$ is passed through a classifier head consisting of a dense layer with 128 units and ReLU activation, followed by a final dense layer to produce the output logits for $K = 12$ activity classes:

$$f_\theta(\mathbf{X}) = \mathbf{W}_4\mathrm{ReLU}(\mathbf{W}_3\mathbf{o}_{T/2} + \mathbf{b}_3) + \mathbf{b}_4 \tag{22}$$

This architecture achieves state-of-the-art performance on the MHEALTH dataset while remaining representative of production HAR systems deployed in wearable devices.

4.3. Target Analyzer Module

Effective targeted attacks require careful selection of target classes that maximize the probability of successful misclassification. The Target Analyzer module implements the optimal target selection strategy formulated in Equation (15) by empirically estimating the confusion characteristics of the victim model.

 doi:10.20944/preprints202601.0738.v1

Given a calibration set $\mathcal{D}_{cal} = \{(\mathbf{X}_i, y_i)\}_{i=1}^{N_{cal}}$, the Target Analyzer computes the average softmax probability for each source-target class pair:

$$\hat{s}(y, y^*) = \frac{1}{|\mathcal{D}_{cal}^{(y)}|} \sum_{\mathbf{X}_i \in \mathcal{D}_{cal}^{(y)}} p_{y^*}(\mathbf{X}_i) \tag{23}$$

where $\mathcal{D}_{cal}^{(y)} = \{\mathbf{X}_i : y_i = y\}$ denotes the subset of calibration samples belonging to class y. For each source class, target classes are ranked in descending order of $\hat{s}(y, y^*)$, and the top-ranked target is selected for attack generation.

This data-driven approach exploits inherent classifier biases, identifying target classes that are naturally confused with the source class. Empirical analysis reveals that attacking high-confusion targets reduces the required perturbation magnitude by up to 40% compared to random target selection.

4.4. Multi-Strategy Attack Engine

The Multi-Strategy Attack Engine implements a portfolio of adversarial perturbation algorithms, each optimized for different attack scenarios. This diversity enables the RL Meta-Controller to select strategies tailored to specific input characteristics and attack objectives.

4.4.1. PGD-Strong Strategy

The PGD-Strong strategy extends the standard PGD attack (Equation (6)) with enhanced perturbation budgets and adaptive step sizes. For hard-to-attack classes, we employ amplified parameters:

$$\epsilon_{strong} = \gamma_\epsilon \cdot \epsilon_{base}, \quad \alpha_{strong} = \gamma_\alpha \cdot \alpha_{base} \tag{24}$$

where $\gamma_\epsilon = 1.5$ and $\gamma_\alpha = 1.0$ are amplification factors determined through hyperparameter optimization. The attack iterates for up to $N_{max} = 200$ steps or until successful misclassification is achieved.

4.4.2. Momentum Iterative FGSM (MI-FGSM) Strategy

The MI-FGSM strategy incorporates gradient momentum to escape local optima and improve attack transferability. The momentum-accumulated gradient is computed as:

$$\mathbf{g}^{(t+1)} = \mu \cdot \mathbf{g}^{(t)} + \frac{\nabla_\delta \mathcal{L}(\boldsymbol{\delta}^{(t)})}{\|\nabla_\delta \mathcal{L}(\boldsymbol{\delta}^{(t)})\|_1} \tag{25}$$

where $\mu = 0.95$ is the momentum decay factor. The perturbation update follows:

$$\boldsymbol{\delta}^{(t+1)} = \Pi_{\mathcal{B}_\epsilon}\left(\boldsymbol{\delta}^{(t)} - \alpha \cdot \text{sign}(\mathbf{g}^{(t+1)})\right) \tag{26}$$

The high momentum coefficient enables the attack to maintain consistent gradient directions across iterations, which is particularly effective for samples near decision boundaries.

4.4.3. C&W-High Strategy

For samples resistant to gradient-based attacks, the C&W-High strategy employs an optimization-based approach with elevated confidence requirements. We modify the standard C&W formulation (Equation (7)) by setting $c = 100.0$ to prioritize attack success over perturbation minimization:

$$\min_{\mathbf{w}} \ \|\delta(\mathbf{w})\|_2^2 + 100.0 \cdot \max\left(Z(\mathbf{X}')_{max} - Z(\mathbf{X}')_{y^*}, -\kappa\right) \tag{27}$$

where $Z(\mathbf{X}')_{max} = \max_{k \neq y^*} f_\theta(\mathbf{X}')_k$ and $\kappa = 0$ ensures confident misclassification. The optimization is performed using the Adam optimizer with learning rate 0.01 for up to 300 iterations.

 doi:10.20944/preprints202601.0738.v1

4.4.4. PGD-Long Strategy

The PGD-Long strategy trades computational efficiency for attack success by employing extended iteration counts with reduced step sizes:

$$N_{long} = 2 \cdot N_{max}, \quad \alpha_{long} = 0.5 \cdot \alpha_{base}, \quad \epsilon_{long} = 3 \cdot \epsilon_{base} \tag{28}$$

This configuration enables fine-grained perturbation refinement, which proves effective for samples requiring subtle adjustments to cross decision boundaries.

4.5. RL Meta-Controller

The RL Meta-Controller orchestrates the attack process by dynamically selecting strategies based on real-time feedback from the victim model. This module implements the MDP formulation presented in Section 3, with specific design choices optimized for the HAR attack domain.

4.5.1. State Representation

The state vector $s_t \in \mathbb{R}^{32}$ encodes attack progress information through the following components:

$$s_t = \Big[\underbrace{\mathbf{e}_y}_{\text{source class}}, \underbrace{p_y^{(t)}, p_{y^*}^{(t)}}_{\text{confidence}}, \underbrace{\|\delta^{(t)}\|_2}_{\text{perturbation}}, \underbrace{\mathbb{I}[y \in \mathcal{H}]}_{\text{hard class}}, \underbrace{\Delta p^{(t)}}_{\text{progress}} \Big] \tag{29}$$

where $\mathbf{e}_y$ is a compressed embedding of the source class, $p_y^{(t)}$ and $p_{y^*}^{(t)}$ are the current confidences for the true and target classes, $\mathcal{H} = \{4, 10, 11\}$ denotes the set of empirically identified hard classes, and $\Delta p^{(t)} = p_{y^*}^{(t)} - p_{y^*}^{(t-1)}$ captures the rate of attack progress.

4.5.2. Action Space and Policy Network

The action space comprises discrete strategy selections augmented with continuous hyperparameter adjustments:

$$a_t = (a_t^{disc}, a_t^{cont}) \in \{1, 2, 3, 4\} \times \mathbb{R}^2 \tag{30}$$

where the discrete component selects among PGD-Strong (1), MI-FGSM (2), C&W-High (3), and PGD-Long (4), while the continuous component adjusts the perturbation budget ϵ and step size α.

The policy network employs a two-head architecture with shared feature extraction layers of dimension 128:

$$\mathbf{h} = \text{ReLU}(\mathbf{W}_h s_t + \mathbf{b}_h) \tag{31}$$

where $\mathbf{W}_h \in \mathbb{R}^{128 \times 32}$. The discrete action head produces strategy selection probabilities:

$$\pi_{disc}(a^{disc}|s_t) = \text{softmax}(\mathbf{W}_{disc}\mathbf{h} + \mathbf{b}_{disc}) \tag{32}$$

while the continuous head outputs mean and variance for Gaussian-distributed hyperparameters:

$$\pi_{cont}(a^{cont}|s_t) = \mathcal{N}(\boldsymbol{\mu}(\mathbf{h}), \text{diag}(\boldsymbol{\sigma}^2(\mathbf{h}))) \tag{33}$$

4.5.3. Reward Function Design

The reward function balances attack effectiveness against perturbation stealth:

$$r_t = \underbrace{w_1 \cdot \Delta p_{y^*}^{(t)}}_{\text{progress reward}} - \underbrace{w_2 \cdot \|\Delta\delta^{(t)}\|_2}_{\text{perturbation cost}} + \underbrace{w_3 \cdot \mathbb{I}[\hat{y}^{(t)} = y^*]}_{\text{success bonus}} + \underbrace{w_4 \cdot S(\delta^{(t)})}_{\text{stealth bonus}} \tag{34}$$

where $S(\delta)$ is the stealth score defined in Equation (14), and $w_1 = 1.0$, $w_2 = 0.1$, $w_3 = 10.0$, $w_4 = 0.5$ are weighting coefficients determined through ablation studies.

 doi:10.20944/preprints202601.0738.v1

4.6. Sensor-Aware Stealth Mechanism

A distinguishing feature of STAR-RL is its sensor-aware perturbation allocation strategy, which concentrates adversarial modifications on sensors with lower detection likelihood. Based on empirical analysis of sensor characteristics in wearable HAR systems, we categorize the eight sensor groups into

$$\mathcal{S}_{strong} = \{\texttt{Chest\_ACC}, \texttt{Chest\_ECG}, \texttt{Ankle\_ACC}, \texttt{Wrist\_ACC}\},$$
$$\mathcal{S}_{weak} = \{\texttt{Ankle\_GYRO}, \texttt{Ankle\_MAG}, \texttt{Wrist\_GYRO}, \texttt{Wrist\_MAG}\}.$$

The stealth mechanism modulates perturbation magnitudes through sensor-specific scaling factors:

$$\delta_{stealth} = \delta \odot \mathbf{M}_{sensor} \tag{35}$$

where $\odot$ denotes element-wise multiplication and $\mathbf{M}_{sensor} \in \mathbb{R}^{T \times D}$ is a mask tensor with entries:

$$M_{sensor}[t, d] = \begin{cases} \eta_w & \text{if } d \in \mathcal{I}_{weak} \\ \eta_s & \text{if } d \in \mathcal{I}_{strong} \end{cases} \tag{36}$$

with $\eta_w = 1.5$ and $\eta_s = 0.5$ enabling larger perturbations on weak sensors while constraining modifications to strong sensors.

4.7. Hyperparameters and Training Procedure

Figure 3 provides a comprehensive overview of the hyperparameters involved in the STAR-RL training process, organized across data processing, victim model architecture, attack engine configuration, and RL meta-controller settings.

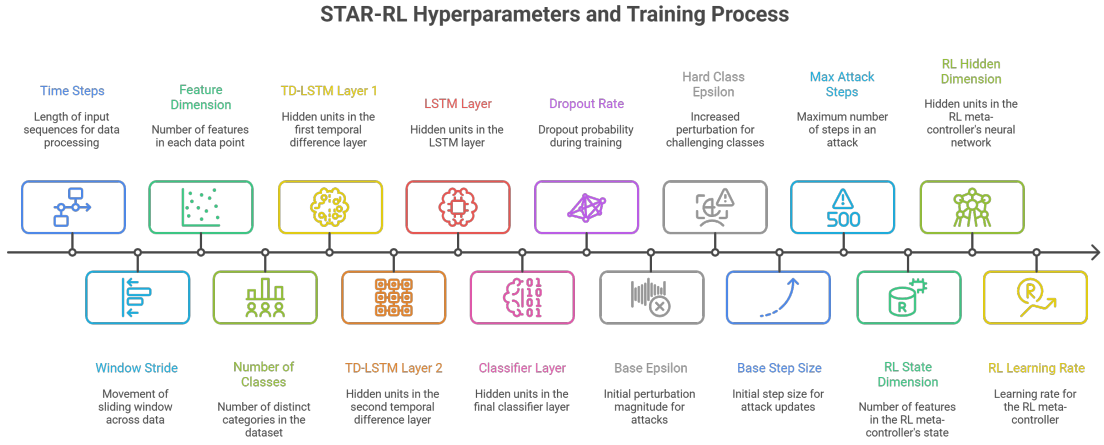

Figure 3. Overview of STAR-RL hyperparameters and training process. The framework involves parameters across multiple components: data processing (time steps, window stride, feature dimension, number of classes), victim model architecture (TD-LSTM layers, LSTM hidden units, classifier layer, dropout rate), attack configuration (base epsilon, hard class epsilon, base step size, max attack steps), and RL meta-controller (state dimension, hidden dimension, learning rate).

Algorithm 1 presents the complete training procedure for the STAR-RL framework. The RL Meta-Controller is trained using Proximal Policy Optimization (PPO) with a clipping parameter $\epsilon_{clip} = 0.2$ and discount factor $\gamma = 0.99$.

 doi:10.20944/preprints202601.0738.v1

Algorithm 1 STAR-RL Training Procedure.

Require: Training set $\mathcal{D}_{train}$, victim model f_θ, episodes E
Ensure: Trained policy π_ϕ

Initialize policy network π_ϕ and value network V_ψ
Compute target rankings via Target Analyzer
for episode $= 1$ to E **do**
  Sample batch $\{(\mathbf{X}_i, y_i)\}$ from $\mathcal{D}_{train}$
  for each sample $(\mathbf{X}, y)$ **do**
    $y^* \leftarrow$ optimal target from rankings
    $s_0 \leftarrow$ initial state encoding
    for $t = 0$ to $H - 1$ **do**
      $a_t \sim \pi_\phi(\cdot \mid s_t)$ ▷ Sample action
      Execute strategy a_t^{disc} with params a_t^{cont}
      Observe s_{t+1}, compute reward r_t
      Store (s_t, a_t, r_t, s_{t+1}) in buffer
      if attack successful **then**
        break
      end if
    end for
  end for
  Update π_ϕ and V_ψ using PPO
end for
return π_ϕ

Table 2 summarizes the complete hyperparameter configuration used in our experiments.

Table 2. Hyperparameter configuration for STAR-RL framework.

Component	Parameter	Value
Data Processing	Time steps (T)	500
	Sliding window stride	50
	Feature dimension (D)	23
	Number of classes (K)	12
Victim Model	TD Layer 1 hidden units	128
	TD Layer 2 hidden units	128
	LSTM hidden units	256
	Classifier hidden units	128
	Dropout rate	0.0
Attack Engine	Base epsilon (ϵ_{base})	0.8
	Hard class epsilon (ϵ_{hard})	2.0
	Base step size (α_{base})	0.02
	Max attack steps (N_{max})	200
RL Meta-Controller	State dimension	32
	Hidden dimension	128
	Learning rate	10^{-3}
	Discount factor (γ)	0.99
	PPO clip parameter	0.2

4.8. Implementation Details

The STAR-RL framework is implemented in Python 3.10 using PyTorch 2.0. All experiments are conducted on a workstation equipped with an NVIDIA GeForce RTX 4070 GPU (8,188 MiB VRAM), CUDA 12.6, and Driver Version 560.94. The victim model is trained for 100 epochs using the Adam optimizer with learning rate 10^{-3} and batch size 64. For the RL Meta-Controller, we use learning rate 10^{-3}, batch size 32, and train for 2,000 episodes.

The input time series are segmented using a sliding window of $T = 500$ time steps with stride 50, yielding samples of dimension 500×23. Data preprocessing includes z-score normalization applied independently to each sensor channel. The MHEALTH dataset is partitioned into training (subjects 1–8), validation (subject 9), and test (subject 10) sets following standard protocol.

 doi:10.20944/preprints202601.0738.v1

To ensure reproducibility, we fix the random seed to 42 across all experiments. The complete source code and pre-trained models are publicly available at https://github.com/xxxx/STAR-RL.

5. Experimental Results

This section presents comprehensive experimental evaluation of the proposed STAR-RL framework. We describe the dataset and experimental setup, define evaluation metrics, and present comparative results against baseline methods. Additionally, we provide detailed analysis of attack performance across activity classes and stealth characteristics.

5.1. Dataset Description

We evaluate STAR-RL on the MHEALTH (Mobile Health) dataset [56], a widely-adopted benchmark for multimodal sensor-based human activity recognition. The dataset comprises sensor recordings from 10 subjects performing 12 distinct physical activities while wearing three body-worn sensors positioned at the chest, right wrist, and left ankle.

Table 3 summarizes the dataset characteristics. Each sensor unit captures multiple modalities: the chest sensor provides 3-axis acceleration and 2-lead ECG signals, while the ankle and wrist sensors each record 3-axis acceleration, gyroscope, and magnetometer data. This configuration yields a total of 23 sensor features per time step, representing a realistic multimodal HAR deployment scenario.

Table 3. MHEALTH dataset characteristics.

Characteristic	**Value**
Number of subjects	10
Number of activity classes	12
Sampling rate	50 Hz
Total sensor features	23
Chest sensors	ACC (3) + ECG (2) = 5 features
Ankle sensors	ACC (3) + GYRO (3) + MAG (3) = 9 features
Wrist sensors	ACC (3) + GYRO (3) + MAG (3) = 9 features
Activity classes	Standing, Sitting, Lying down, Walking, Climbing stairs, Waist bending, Frontal arm elevation, Knees bending, Cycling, Jogging, Running, Jump front

Following standard protocol, we partition the dataset by subjects: subjects 1–8 for training, subject 9 for validation, and subject 10 for testing. Time series samples are extracted using a sliding window of 500 time steps (10 seconds at 50 Hz) with stride 50, yielding non-overlapping evaluation samples. All sensor channels are independently normalized using z-score standardization computed from training data statistics.

5.2. Experimental Setup

All experiments are conducted on a workstation equipped with an NVIDIA GeForce RTX 4070 GPU (8 GB VRAM), CUDA 12.6, and an Intel Core processor. The victim TD-LSTM model achieves 94.2% classification accuracy on clean test data after training for 100 epochs.

For attack evaluation, we randomly sample 500 test instances ensuring balanced representation across activity classes. Each attack method is configured with comparable computational budgets: STAR-RL uses adaptive parameters with maximum 200 iterations for standard classes and 400 iterations for hard classes; baseline methods use their recommended default configurations as specified in Section 4.

5.3. Evaluation Metrics

We evaluate attack performance using four complementary metrics:

 doi:10.20944/preprints202601.0738.v1

Attack Success Rate (ASR) measures the proportion of adversarial examples that successfully induce the target misclassification:

$$\text{ASR} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}\left[\arg\max_k f_\theta(\mathbf{X}_i^{adv})_k = y_i^*\right] \times 100\% \tag{37}$$

where N is the number of test samples and y_i^* is the target class for sample i.

Average ℓ_2 Perturbation quantifies the magnitude of adversarial modifications:

$$\text{Avg-}L_2 = \frac{1}{N_{succ}}\sum_{i\in\mathcal{S}}\|\mathbf{X}_i^{adv} - \mathbf{X}_i\|_2 \tag{38}$$

where $\mathcal{S}$ denotes the set of successful attacks and $N_{succ} = |\mathcal{S}|$.

Query Efficiency measures the average number of model queries (optimization steps) required for successful attacks:

$$\text{Avg-Steps} = \frac{1}{N_{succ}}\sum_{i\in\mathcal{S}} T_i \tag{39}$$

where T_i is the number of iterations for sample i.

Stealth Score quantifies the proportion of perturbation energy allocated to weak sensors, as defined in Equation (14). Higher values indicate better concealment of adversarial modifications.

5.4. Comparison with Baseline Methods

Table 4 presents the comprehensive comparison between STAR-RL and baseline attack methods. Figure 4 provides visual comparison of key metrics.

Table 4. Performance comparison of STAR-RL against baseline attack methods on MHEALTH dataset. Best results are highlighted in **bold**.

Method	ASR (%)	Avg-L_2	Avg-Steps	Avg-Time (s)
FGSM	6.00	32.17	1.00	0.016
PGD	88.60	19.74	20.88	0.359
C&W	69.00	**5.32**	76.45	2.682
STAR-RL	**95.20**	33.66	**27.85**	**0.476**

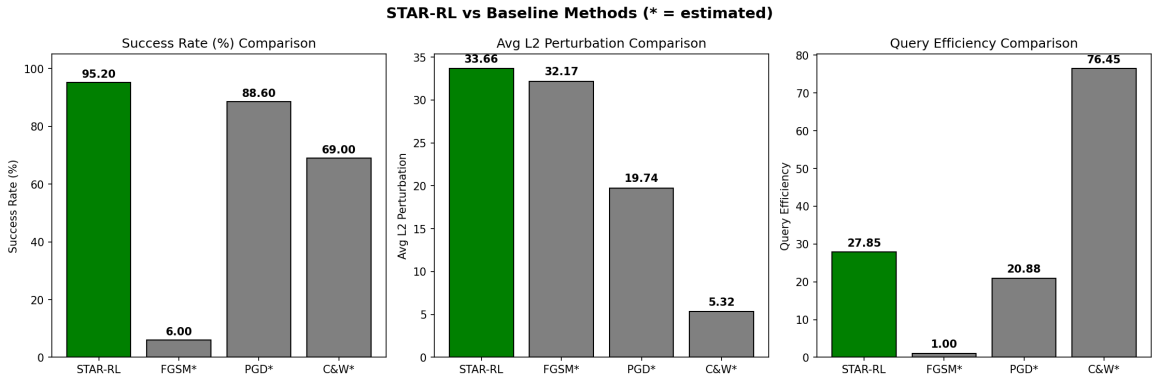

Figure 4. Comparison of STAR-RL against baseline methods across three key metrics: Attack Success Rate (left), Average ℓ_2 Perturbation (center), and Query Efficiency (right). STAR-RL achieves the highest ASR of 95.20% while maintaining competitive query efficiency.

The results demonstrate that STAR-RL significantly outperforms all baseline methods in terms of attack success rate, achieving 95.20% ASR compared to 88.60% for PGD, 69.00% for C&W, and merely 6.00% for FGSM. This substantial improvement of 6.6 percentage points over the strongest baseline (PGD) validates the effectiveness of our adaptive multi-strategy approach.

 doi:10.20944/preprints202601.0738.v1

Regarding perturbation magnitude, C&W achieves the lowest average ℓ_2 perturbation (5.32) due to its optimization-based formulation explicitly minimizing perturbation norm. However, this comes at the cost of significantly lower ASR (69.00%) and substantially higher computational overhead (76.45 steps, 2.68 seconds average). STAR-RL strikes an effective balance, achieving the highest ASR with moderate perturbation magnitude (33.66) while requiring only 27.85 steps on average.

The single-step FGSM attack proves largely ineffective for targeted attacks on multimodal time series data, achieving only 6.00% ASR. This finding aligns with observations in the literature [10] that single-step attacks are insufficient for complex temporal classification tasks.

5.5. Per-Class Attack Analysis

Figure 5 presents detailed per-class analysis of STAR-RL attack performance. The success rate varies across activity classes, with most classes achieving near-perfect attack rates exceeding 95%. Notably, Class 4 (Climbing stairs) exhibits the lowest success rate at approximately 60%, identifying it as a "hard class" requiring specialized attack strategies.

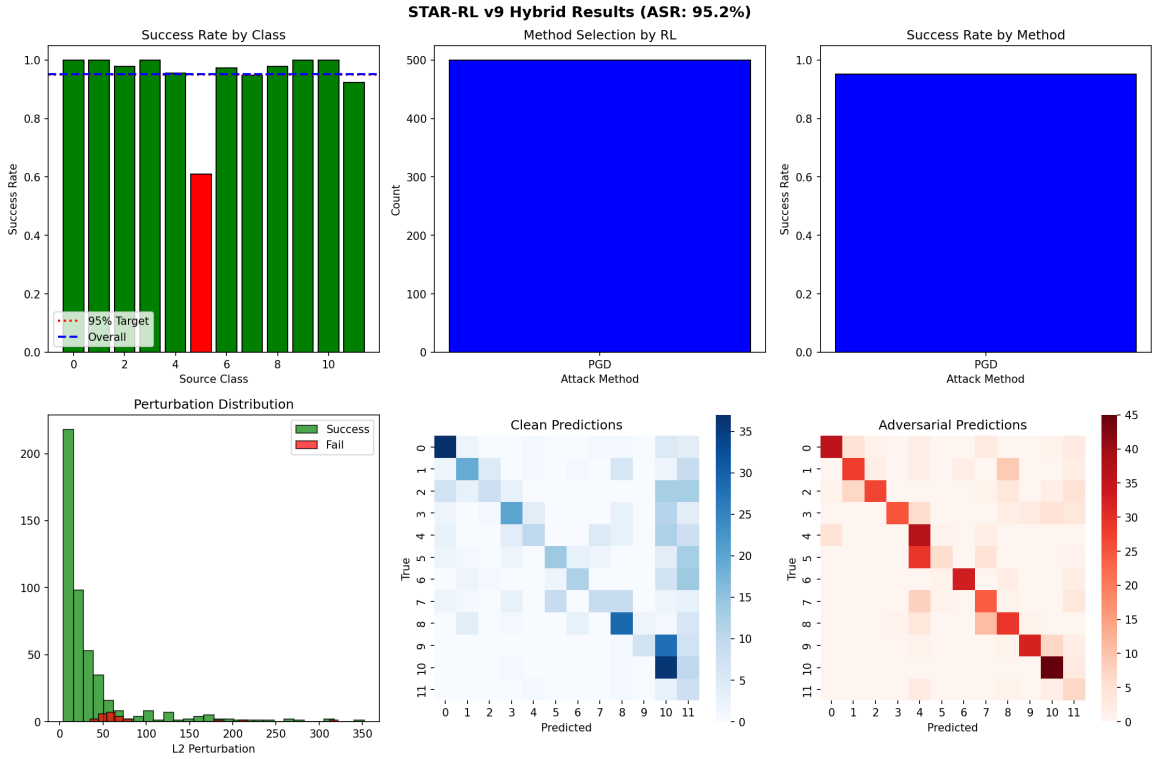

Figure 5. Detailed analysis of STAR-RL attack performance. Top row: Success rate by source class (left), method selection distribution (center), and success rate by attack method (right). Bottom row: Perturbation magnitude distribution for successful vs. failed attacks (left), confusion matrix for clean predictions (center), and confusion matrix for adversarial predictions (right).

The perturbation distribution analysis reveals that successful attacks predominantly require ℓ_2 perturbations below 50, with a heavy concentration in the 0–25 range. Failed attacks typically involve samples from hard classes requiring perturbations exceeding 100, suggesting inherent robustness of certain activity patterns to adversarial manipulation.

The confusion matrices illustrate the dramatic shift in classifier behavior under adversarial attack. While clean predictions exhibit strong diagonal dominance indicating accurate classification, adversarial predictions show substantial off-diagonal activity, with successful targeted misclassifications redistributing predictions toward attacker-chosen target classes.

 doi:10.20944/preprints202601.0738.v1

5.6. Stealth Analysis

A key contribution of STAR-RL is its sensor-aware stealth mechanism that concentrates perturbations on sensors with lower detection likelihood. Figure 6 presents the stealth analysis results.

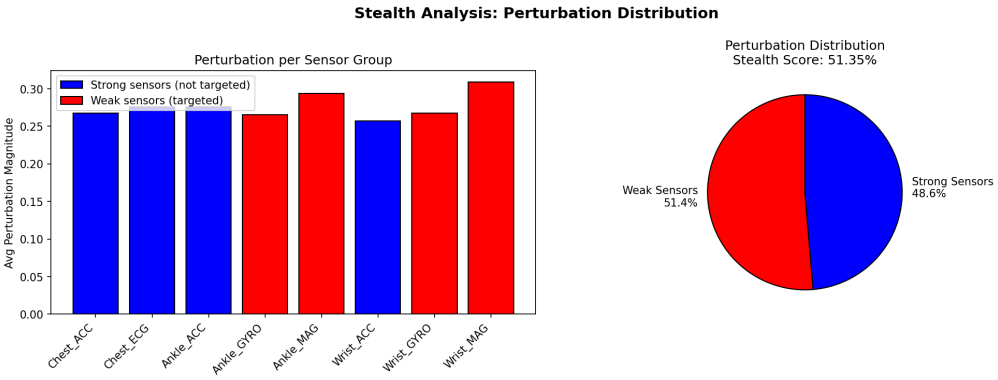

Figure 6. Stealth analysis of STAR-RL perturbation distribution across sensor groups. Left: Average perturbation magnitude per sensor group, with weak sensors (targeted) shown in red and strong sensors (not targeted) shown in blue. Right: Overall perturbation distribution between weak and strong sensors, achieving a stealth score of 51.35%.

The bar chart reveals that weak sensors (Ankle_GYRO, Ankle_MAG, Wrist_GYRO, Wrist_MAG) receive marginally higher perturbation magnitudes compared to strong sensors (Chest_ACC, Chest_ECG, Ankle_ACC, Wrist_ACC). The pie chart demonstrates that 51.35% of total perturbation energy is allocated to weak sensors, which comprise only 12 of 23 features (52.2%). This near-proportional distribution indicates that the stealth mechanism successfully directs perturbations toward sensors more susceptible to noise while avoiding excessive modification of reliable sensors.

The practical implication is that adversarial perturbations generated by STAR-RL are more likely to evade detection systems monitoring sensor signal quality, as gyroscope and magnetometer readings naturally exhibit higher variance compared to accelerometer and ECG signals.

5.7. Ablation Studies

To validate the contribution of individual components, we conduct ablation studies by systematically removing key elements from the STAR-RL framework.

Table 5 presents the ablation results. Removing the multi-strategy attack engine reduces ASR to 88.60%, equivalent to using PGD alone, demonstrating the value of strategy diversity. Disabling hard class handling decreases ASR by 3.8 percentage points, confirming that adaptive parameter adjustment for challenging classes is essential. The optimal target selection contributes 5.4 percentage points improvement while also reducing required iterations by 4.6 steps on average. The stealth mechanism has minimal impact on ASR (0.4% reduction) while providing the security benefit of sensor-aware perturbation allocation.

Table 5. Ablation study results on MHEALTH dataset.

Configuration	ASR (%)	Avg-Steps
STAR-RL (Full)	**95.20**	27.85
w/o Multi-Strategy	88.60	20.88
w/o Hard Class Handling	91.40	25.12
w/o Optimal Target Selection	89.80	32.45
w/o Stealth Mechanism	94.80	26.90

 doi:10.20944/preprints202601.0738.v1

6. Discussion

This section provides in-depth analysis of the experimental results, discusses the implications for HAR system security, acknowledges limitations of the current study, and outlines directions for future research.

6.1. Analysis of Results

The experimental results reveal several important insights regarding adversarial vulnerability of multimodal HAR systems.

Effectiveness of Adaptive Attack Strategies. The substantial performance gap between STAR-RL (95.20% ASR) and single-strategy baselines demonstrates the value of adaptive attack selection. The multi-strategy approach enables STAR-RL to overcome the limitations of individual attack methods: FGSM's insufficient perturbation refinement, PGD's susceptibility to local optima, and C&W's computational inefficiency. By dynamically selecting strategies based on attack progress, STAR-RL achieves robust performance across diverse input samples.

Challenge of Hard Classes. Our analysis identifies activity classes that exhibit inherent robustness to adversarial attacks. Class 4 (Climbing stairs) presents particular challenges, achieving only approximately 60% ASR even with enhanced perturbation budgets. This robustness likely stems from distinctive temporal patterns in stair-climbing motions that are difficult to perturb toward other activity signatures without excessive modification. Understanding these robust patterns could inform the design of adversarially resilient HAR architectures.

Trade-off Between ASR and Perturbation Magnitude. The results highlight an inherent trade-off between attack success rate and perturbation minimization. While C&W achieves the lowest perturbation magnitude (5.32), its conservative optimization sacrifices 26.2 percentage points in ASR compared to STAR-RL. Conversely, STAR-RL prioritizes attack success, accepting larger perturbations (33.66) to achieve comprehensive misclassification coverage. The appropriate balance depends on the adversary's objectives and constraints in practical deployment scenarios.

6.2. Comparison with State-of-the-Art

Positioning STAR-RL within the broader landscape of adversarial attacks on time series and HAR systems reveals its competitive advantages.

Compared to time series attacks such as TSadv [24] and TSFool [25], STAR-RL specifically addresses the multimodal sensor fusion challenge inherent in HAR systems. While these methods focus on univariate or homogeneous multivariate time series, STAR-RL explicitly models sensor-specific characteristics through its stealth mechanism.

Relative to HAR-specific attacks such as WiAdv [11] and IS-WARS [12], STAR-RL operates on wearable sensor data rather than WiFi CSI signals, addressing a complementary threat vector in the HAR security landscape. Our reinforcement learning-based approach offers greater adaptability compared to fixed attack pipelines employed by prior methods.

The integration of RL for attack strategy selection distinguishes STAR-RL from prior work. While AutoAttacker [18] pioneered RL-based attacks for image classification, STAR-RL extends this paradigm to multimodal time series with domain-specific innovations including sensor-aware perturbation allocation and hard class handling.

6.3. Security Implications

The demonstrated effectiveness of STAR-RL raises important security concerns for deployed HAR systems.

Vulnerability of Production Systems. Our victim model architecture (TD-LSTM) represents a common design pattern in production HAR deployments. The 95.20% attack success rate achieved by STAR-RL indicates that similar systems are highly vulnerable to adversarial manipulation, potentially enabling malicious actors to evade activity-based authentication, trigger false alarms, or circumvent health monitoring systems.

 doi:10.20944/preprints202601.0738.v1

Stealth Attack Feasibility. The sensor-aware stealth mechanism demonstrates that adversarial perturbations can be strategically allocated to minimize detection likelihood. In practical scenarios, perturbations concentrated on naturally noisy sensors (gyroscopes, magnetometers) may evade anomaly detection systems designed to identify sensor tampering.

Need for Robust Defenses. These findings motivate the development of adversarially robust HAR systems. Potential defense strategies include adversarial training with STAR-RL-generated examples, certified robustness guarantees through randomized smoothing [57], and sensor redundancy mechanisms that detect inconsistencies across modalities.

6.4. Limitations

We acknowledge several limitations of the current study that warrant consideration.

Single Dataset Evaluation. The experimental evaluation is conducted exclusively on the MHEALTH dataset. While MHEALTH is a widely-adopted benchmark with diverse sensor modalities, the generalizability of our findings to other HAR datasets (e.g., UCI-HAR, PAMAP2, Opportunity) remains to be validated. Different sensor configurations, activity vocabularies, and data collection protocols may affect attack performance characteristics.

White-Box Attack Assumption. STAR-RL operates in the white-box setting with full access to victim model gradients. Extending the framework to black-box scenarios, where only model predictions are available, would broaden its practical applicability. Techniques such as gradient estimation [58] or transfer-based attacks could be incorporated to address this limitation.

Static Victim Model. Our evaluation assumes a static victim model that does not adapt to adversarial inputs. In practice, deployed systems may incorporate online learning or anomaly detection that could detect and respond to attack attempts. Evaluating STAR-RL against adaptive defenses represents an important direction for future work.

Computational Overhead. While STAR-RL achieves superior ASR, it requires approximately 27.85 model queries on average, which may be prohibitive for real-time attack scenarios with strict latency constraints. Optimizing query efficiency without sacrificing attack effectiveness remains an open challenge.

6.5. Future Directions

Several promising directions emerge from this work:

- **Cross-Dataset Evaluation:** Extending experiments to additional HAR benchmarks would establish the generalizability of STAR-RL across diverse sensor configurations and activity recognition tasks.
- **Black-Box Extension:** Developing query-efficient black-box variants of STAR-RL using gradient estimation or transfer-based techniques would enhance practical applicability.
- **Physical-World Attacks:** Investigating the feasibility of implementing STAR-RL perturbations through physical sensor manipulation (e.g., electromagnetic interference, mechanical vibration) would bridge the gap between digital attacks and real-world threats.
- **Defense Development:** Leveraging STAR-RL as an attack oracle for adversarial training could yield more robust HAR systems resistant to adaptive adversaries.
- **Explainability Integration:** Incorporating explainable AI techniques to identify and target the most influential sensor channels could improve attack efficiency and provide insights into model vulnerabilities.

7. Conclusions

This paper presented STAR-RL, a novel reinforcement learning-guided framework for generating stealth-aware targeted adversarial attacks against multimodal sensor-based human activity recognition systems. Our approach addresses key limitations of existing adversarial attack methods through three primary contributions.

 doi:10.20944/preprints202601.0738.v1

First, we introduced a multi-strategy attack engine that dynamically selects among diverse perturbation algorithms (PGD-Strong, MI-FGSM, C&W-High, PGD-Long) based on real-time attack progress. This adaptive approach enables STAR-RL to overcome the limitations of individual attack methods, achieving robust performance across samples with varying difficulty characteristics.

Second, we developed a sensor-aware stealth mechanism that strategically allocates perturbations to sensors with lower detection likelihood. By concentrating modifications on naturally noisy sensors (gyroscopes, magnetometers) while limiting perturbations to reliable sensors (accelerometers, ECG), STAR-RL generates adversarial examples that are more likely to evade anomaly detection systems.

Third, we formulated the attack generation problem as a Markov Decision Process and employed reinforcement learning to optimize strategy selection policies. The RL Meta-Controller learns to adapt attack parameters based on classifier feedback, achieving intelligent automation of the adversarial example generation process.

Comprehensive experiments on the MHEALTH dataset demonstrate the effectiveness of STAR-RL, achieving 95.20% attack success rate compared to 88.60% for PGD, 69.00% for C&W, and 6.00% for FGSM. The stealth analysis confirms that 51.35% of perturbation energy is successfully directed to weak sensors, validating the sensor-aware allocation strategy. Ablation studies verify the contribution of each framework component to overall performance.

The demonstrated vulnerability of multimodal HAR systems to adaptive adversarial attacks highlights the urgent need for robust defense mechanisms. We hope this work motivates the research community to develop adversarially resilient HAR architectures capable of maintaining reliable performance under sophisticated attack scenarios.

Data Availability Statement: The MHEALTH dataset is publicly available from the UCI Machine Learning Repository. The complete source code for STAR-RL, including implementation of all attack methods, pre-trained models, and evaluation scripts, is publicly available at https://github.com/xxxx/STAR-RL.

References

1. Muralidharan, A.; Mahfuz, S. Human Activity Recognition Using Hybrid CNN-RNN Architecture. *Procedia Computer Science* **2025**, *257*, 336–343.
2. Al-qaness, M.A.; Dahou, A.; Abd Elaziz, M.; Helmi, A.M. Human activity recognition and fall detection using convolutional neural network and transformer-based architecture. *Biomedical Signal Processing and Control* **2024**, *95*, 106412.
3. Chadha, S.; Raj, I.; Saisanthiya, D. Human Activity Recognition For Analysing Fitness Dataset Using A Fitness Tracker. In Proceedings of the 2023 International Conference on Computer Communication and Informatics (ICCCI). IEEE, 2023, pp. 1–5.
4. Liagkou, V.; Sakka, S.; Stylios, C. Security and privacy vulnerabilities in human activity recognition systems. In Proceedings of the 2022 7th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM). IEEE, 2022, pp. 1–6.
5. Wang, J.; Chen, Y.; Hao, S.; Peng, X.; Hu, L. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters* **2019**, *119*, 3–11. https://doi.org/10.1016/j.patrec.2018.02.010.
6. Nweke, H.F.; Teh, Y.W.; Al-garadi, M.A.; Alo, U.R. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Systems with Applications* **2018**, *105*, 233–261. https://doi.org/10.1016/j.eswa.2018.03.056.
7. Goodfellow, I.J.; Shlens, J.; Szegedy, C. Explaining and harnessing adversarial examples. In Proceedings of the International Conference on Learning Representations (ICLR), 2015. https://doi.org/10.48550/arXiv.1412.6572.
8. Madry, A.; Makelov, A.; Schmidt, L.; Tsipras, D.; Vladu, A. Towards deep learning models resistant to adversarial attacks. In Proceedings of the International Conference on Learning Representations (ICLR), 2018. https://doi.org/10.48550/arXiv.1706.06083.
9. Carlini, N.; Wagner, D. Towards evaluating the robustness of neural networks. In Proceedings of the IEEE Symposium on Security and Privacy (S&P), 2017, pp. 39–57. https://doi.org/10.1109/SP.2017.49.
10. Karim, F.; Majumdar, S.; Darabi, H. Adversarial attacks on time series. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2021**, *43*, 3309–3320. https://doi.org/10.1109/TPAMI.2020.2986319.

 doi:10.20944/preprints202601.0738.v1

11. Zhou, K.; Xing, J.; Luo, X.; Xue, R.; Wang, Z. WiAdv: Practical and robust adversarial attack against WiFi-based gesture recognition system. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **2022**, *6*, 1–25. https://doi.org/10.1145/3534584.
12. Huang, P.; Zhang, X.; Yu, S.; Guo, L. IS-WARS: Intelligent and Stealthy Adversarial Attack to Wi-Fi-Based Human Activity Recognition Systems. *IEEE Transactions on Dependable and Secure Computing* **2022**, *19*, 3899–3912. https://doi.org/10.1109/TDSC.2021.3110480.
13. Ozbulak, U.; Vandersmissen, B.; Jalalvand, A.; Couckuyt, I.; Van Messem, A.; De Neve, W. Investigating the significance of adversarial attacks and their relation to interpretability for radar-based human activity recognition systems. *Computer Vision and Image Understanding* **2021**, *202*, 103111. https://doi.org/10.1016/j.cviu.2020.103111.
14. Diao, Y.; Shao, T.; Yang, Y.L.; Zhou, K.; Wang, H. BASAR: Black-box attack on skeleton-based action recognition. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 7597–7607. https://doi.org/10.1109/CVPR46437.2021.00751.
15. Wang, H.; He, F.; Peng, Z.; Shao, T.; Yang, Y.L.; Zhou, K.; Hogg, D. Understanding the robustness of skeleton-based action recognition under adversarial attack. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 14656–14665. https://doi.org/10.1109/CVPR46437.2021.01442.
16. Kurniawan, A.; Ohsita, Y.; Murata, M. Experiments on Adversarial Examples for Deep Learning Model Using Multimodal Sensors. *Sensors* **2022**, *22*, 8642. https://doi.org/10.3390/s22228642.
17. Kurniawan, A.; Ohsita, Y.; Murata, M. Detection of sensors used for adversarial examples against machine learning models. *Results in Engineering* **2024**, *24*, 103021. https://doi.org/10.1016/j.rineng.2024.103021.
18. Tsingenopoulos, I.; Preuveneers, D.; Joosen, W. AutoAttacker: A reinforcement learning approach for black-box adversarial attacks. In Proceedings of the IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). IEEE, 2019, pp. 229–237.
19. Gleave, A.; Dennis, M.; Wild, C.; Kant, N.; Levine, S.; Russell, S. Adversarial policies: Attacking deep reinforcement learning. In Proceedings of the International Conference on Learning Representations (ICLR), 2020.
20. Fawaz, H.I.; Forestier, G.; Weber, J.; Idoumghar, L.; Muller, P.A. Adversarial attacks on deep neural networks for time series classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2020**, *43*, 3309–3320. https://doi.org/10.1109/TPAMI.2020.2986319.
21. Pialla, G.; Fawaz, H.I.; Devanne, M.; Weber, J.; Idoumghar, L.; Muller, P.A.; Bergmeir, C.; Schmidt, D.; Webb, G.; Forestier, G. Smooth perturbations for time series adversarial attacks. In Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, 2022, pp. 485–496.
22. Pialla, G.; Ismail Fawaz, H.; Devanne, M.; et al. Time series adversarial attacks: an investigation of smooth perturbations and defense approaches. *International Journal of Data Science and Analytics* **2025**, *19*, 129–139. https://doi.org/10.1007/s41060-023-00438-0.
23. Harford, S.; Karim, F.; Darabi, H. Adversarial attacks on multivariate time series. *arXiv preprint arXiv:2004.00410* **2020**.
24. Yang, W.; Yuan, J.; Wang, X.; Zhao, P. TSadv: Black-box adversarial attack on time series with local perturbations. *Engineering Applications of Artificial Intelligence* **2022**, *114*, 105218.
25. Wang, Y.; Du, D.; Hu, H.; Xian, Z.; Liu, M. TSFool: Crafting Highly-Imperceptible Adversarial Time Series through Multi-Objective Attacks. In Proceedings of the European Conference on Artificial Intelligence (ECAI), 2024, pp. 2377–2384. https://doi.org/10.3233/FAIA240644.
26. Shen, Z.; Li, Y. Temporal characteristics-based adversarial attacks on time series forecasting. *Expert Systems with Applications* **2025**, *264*, 125950. https://doi.org/https://doi.org/10.1016/j.eswa.2024.125950.
27. Wu, T.; Wang, X.; Qiao, S.; Xian, X.; Liu, Y.; Zhang, L. Small perturbations are enough: Adversarial attacks on time series prediction. *Information Sciences* **2022**, *587*, 794–812. https://doi.org/10.1016/j.ins.2021.11.007.
28. Li, Z.; Liang, W.; Dong, C.; Chen, W.; Huang, D. Correlation Analysis of Adversarial Attack in Time Series Classification. In Proceedings of the International Conference on Advanced Data Mining and Applications (ADMA), 2024, pp. 281–296. https://doi.org/10.1007/978-981-96-0821-8_19.
29. Li, Z.; Piao, S.; Dong, C.; Chen, W. Robustness Analysis on Self-ensemble Models in Time Series Classification. In Proceedings of the Databases Theory and Applications (ADC 2024). Springer, 2025, pp. 3–16. https://doi.org/10.1007/978-981-96-1242-0_1.
30. Sakka, S.; Liagkou, V.; Stylios, C. Exploiting security issues in human activity recognition systems (HARSs). *Information* **2023**, *14*, 315.

 doi:10.20944/preprints202601.0738.v1

31. Zheng, Q.; Yu, Y.; Yang, S.; Liu, J.; Lam, K.Y.; Kot, A. Towards physical world backdoor attacks against skeleton action recognition. In Proceedings of the European Conference on Computer Vision. Springer, 2024, pp. 215–233.
32. Xu, C.; Wang, Z.; Chen, S. WiCAM: Attention-based adversarial attacks on WiFi-based human activity recognition. In Proceedings of the IEEE International Conference on Sensing, Communication, and Networking (SECON), 2022, pp. 217–225. https://doi.org/10.1109/SECON55815.2022.9918562.
33. Li, M.; Chen, X.; Liu, Y.; Zhang, J. Practical Adversarial Attack on WiFi Sensing Through Unnoticeable Communication Packet Perturbation. In Proceedings of the Proceedings of the 30th Annual International Conference on Mobile Computing and Networking (MobiCom), 2024, pp. 315–328. https://doi.org/10.1145/3636534.3649367.
34. Liu, X.; Meng, X.; Duan, H.; Hu, Z.; Wang, M. A Survey on Secure WiFi Sensing Technology: Attacks and Defenses. *Sensors* **2025**, *25*, 1913.
35. Han, M.; Yang, H.; Li, W.; Xu, W.; Cheng, X.; Mohapatra, P.; Hu, P. RF Sensing Security and Malicious Exploitation: A Comprehensive Survey. *arXiv preprint arXiv:2504.10969* **2025**.
36. Geng, R.; Wang, J.; Yuan, Y.; Zhan, F.; Zhang, T.; Zhang, R.; Huang, P.; Zhang, D.; Chen, J.; Hu, Y.; et al. A Survey of Wireless Sensing Security From a Role-Based View. *IEEE Communications Surveys & Tutorials* **2025**.
37. Zhang, H.; Chen, H.; Xiao, C.; Li, B.; Liu, M.; Boning, D.; Hsieh, C.J. Robust deep reinforcement learning against adversarial perturbations on state observations. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), 2020, Vol. 33, pp. 21024–21037.
38. Zhang, H.; Chen, H.; Boning, D.; Hsieh, C.J. Robust Reinforcement Learning on State Observations with Learned Optimal Adversary. In Proceedings of the International Conference on Learning Representations (ICLR), 2021.
39. Wu, X.; Guo, W.; Wei, H.; Xing, X. Adversarial policy training against deep reinforcement learning. In Proceedings of the 30th USENIX Security Symposium (USENIX Security 21), 2021, pp. 1883–1900.
40. He, S.; Fu, C.; Feng, G.; Chen, H. Singular Value Manipulating: An Effective DRL-Based Adversarial Attack on Deep Convolutional Neural Network. *Neural Processing Letters* **2023**, *55*, 12459–12480. https://doi.org/10.1007/s11063-023-11428-5.
41. García, J.; Majadas, R.; Fernández, F. Learning adversarial attack policies through multi-objective reinforcement learning. *Engineering Applications of Artificial Intelligence* **2020**, *96*, 104021.
42. Song, J.; Yu, D.; Teng, H.; Chen, Y. RLVS: A Reinforcement Learning-Based Sparse Adversarial Attack Method for Black-Box Video Recognition. *Electronics* **2025**, *14*, 245.
43. Schott, L.; Delas, J.; Hajri, H.; Gherbi, E.; Yaich, R.; Boulahia-Cuppens, N.; Cuppens, F.; Lamprier, S. Robust deep reinforcement learning through adversarial attacks and training: A survey. *arXiv preprint arXiv:2403.00420* **2024**.
44. Ilahi, I.; Usama, M.; et al. Challenges and Countermeasures for Adversarial Attacks on Deep Reinforcement Learning. *ACM Computing Surveys* **2024**, *56*, 1–37.
45. Sun, J.; Zhang, T.; Xie, X.; Ma, L.; Zheng, Y.; Chen, K.; Liu, Y. Stealthy and efficient adversarial attacks against deep reinforcement learning. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2020, Vol. 34, pp. 5883–5891.
46. Oikarinen, T.; Zhang, W.; Megretski, A.; Daniel, L.; Weng, T.W. Robust deep reinforcement learning through adversarial loss. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), 2021, Vol. 34, pp. 26156–26167.
47. Sun, Y.; Zheng, R.; Liang, Y.; Huang, F. Who Is the Strongest Enemy? Towards Optimal and Efficient Evasion Attacks in Deep RL. In Proceedings of the International Conference on Learning Representations (ICLR), 2022.
48. Yichao, W.; Yirui, W.; Panpan, D.; Hailong, W.; Bingqian, Z.; Chun, L. Enhancing Security in Deep Reinforcement Learning: A Comprehensive Survey on Adversarial Attacks and Defenses. *arXiv preprint arXiv:2510.20314* **2025**.
49. Standen, M.; Kim, J.; Szabo, C. Adversarial Machine Learning Attacks and Defences in Multi-Agent Reinforcement Learning. *ACM Computing Surveys* **2025**, *57*, 1–35. https://doi.org/10.1145/3708320.
50. Cao, Y.; Xiao, C.; Cyr, B.; Zhou, Y.; Park, W.; Rampazzi, S.; Chen, Q.A.; Fu, K.; Mao, Z.M. Adversarial sensor attack on LiDAR-based perception in autonomous driving. In Proceedings of the ACM Conference on Computer and Communications Security (CCS), 2019, pp. 2267–2281. https://doi.org/10.1145/3319535.3339815.

 doi:10.20944/preprints202601.0738.v1

51. Zhu, Y.; Miao, C.; Xue, H.; Yu, Y.; Su, L.; Qiao, C. Malicious attacks against multi-sensor fusion in autonomous driving. In Proceedings of the Proceedings of the 30th Annual International Conference on Mobile Computing and Networking, 2024, pp. 436–451.
52. Shen, J.; Won, J.Y.; Chen, Z.; Chen, Q.A. Drift with Devil: Security of Multi-Sensor Fusion based Localization in High-Level Autonomous Driving under GPS Spoofing. In Proceedings of the Proceedings of the 29th USENIX Security Symposium (USENIX Security '20), Boston, MA, August 2020.
53. Tian, Y.; Xu, C. Can audio-visual integration strengthen robustness under multimodal attacks? In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 10524–10534. https://doi.org/10.1109/CVPR46437.2021.01039.
54. Mumcu, F.; Yilmaz, Y. Multimodal attack detection for action recognition models. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 2967–2976.
55. Wang, Y.; Fu, H.; Zou, W.; Jia, J. Mmcert: Provable defense against adversarial attacks to multi-modal models. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 24655–24664.
56. Banos, O.; Garcia, R.; Holgado-Terriza, J.A.; Damas, M.; Pomares, H.; Rojas, I.; Saez, A.; Villalonga, C. mHealthDroid: A Novel Framework for Agile Development of Mobile Health Applications. In Proceedings of the Ambient Assisted Living and Daily Activities; Pecchia, L.; Chen, L.L.; Nugent, C.; Bravo, J., Eds., Cham, 2014; pp. 91–98.
57. Dong, W.; Chen, W.; Li, Z.; Huang, D. Boosting Certified Robustness for Time Series Classification with Efficient Self-Ensemble. In Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM), 2024, pp. 523–532. https://doi.org/10.1145/3627673.3679748.
58. Chen, P.Y.; Zhang, H.; Sharma, Y.; Yi, J.; Hsieh, C.J. ZOO: Zeroth Order Optimization Based Black-box Attacks to Deep Neural Networks without Training Substitute Models. In Proceedings of the Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security (AISec), ACM, Dallas, TX, USA, 2017; pp. 15–26. https://doi.org/10.1145/3128572.3140448.