Anomaly Detection in Road-Tanker Fuel Transport: A Deep Learning and Simulation Study

Naemeh Mohammadpour
Amirkabir University of Technology

ABSTRACT

Fuel theft in road-tanker fleets is challenging to detect due to scarce labels, sensor noise, and distribution shifts. We present a reproducible, simulation-driven framework that combines a configurable trip simulator (geofence, speed, tank level, and outflow with realistic noise) with two detection strategies. An unsupervised LSTM Autoencoder (AE) is trained on normal windows to produce anomaly-sensitive auxiliary features. On top of these, we build: (i) an end-to-end BiLSTM+Attention model (Seq-Attn) that fuses temporal context with auxiliary features, and (ii) a hybrid LSTM+RF approach that encodes each window with an LSTM and classifies the embedding via a Random Forest. We evaluate on three scenarios---Baseline, Class Imbalance, and Distribution Shift---using trip-wise splits and theft-centric metrics (Precision, Recall, F1, PR-AUC, ROC-AUC). Results show both models achieve near-perfect detection in the baseline. Under imbalance and shift, Seq-Attn achieves consistently higher recall with almost perfect precision, while LSTM+RF remains competitive but misses slightly more thefts. These findings suggest using Seq-Attn as the primary detector, with LSTM+RF as a supplementary model for interpretability and ensemble robustness. All code and artifacts are released for end-to-end reproducibility and future validation.

Keywords: Fuel theft detection, BiLSTM, Attention mechanism, LSTM Autoencoder, Random Forest, Simulation, Anomaly detection, Time series classification

1. INTRODUCTION

Machine learning methods are generally categorized into three main paradigms: supervised learning, where models are trained on labeled data (for example, predicting equipment failures using annotated sensor logs); unsupervised learning, which aims to discover hidden patterns in unlabeled data (for example, clustering routes to identify abnormal fuel-consumption behaviors); and reinforcement learning, which optimizes sequential decision-making through interaction with the environment (for example, robotic control systems) [1,2,3,4].

The methods applied in this paper combine unsupervised learning, through an LSTM Autoencoder used for extracting anomaly-sensitive features, with supervised learning, through Seq-Attn and LSTM+RF models designed for three-class classification. This hybrid approach enables us to exploit the ability of unsupervised techniques to uncover hidden patterns in unlabeled data, while simultaneously leveraging the accuracy of supervised models in detecting fuel theft incidents.

Fuel theft in road-tanker fleets poses a significant operational and financial challenge for the logistics and energy sectors. Unauthorized fuel discharge during transit not only leads to direct economic losses but also creates safety, compliance, and sustainability concerns. Traditional monitoring systems—which typically combine GPS tracking, fuel-level sensors, and geofence rules—can detect abrupt tank-level drops or out-of-depot discharge events. However, these rule-based methods are highly sensitive to sensor noise, calibration drift, and operational heterogeneity, often producing excessive false alarms or missing subtle theft incidents.

This motivates the development of robust, data-driven approaches that can generalize across routes, vehicles, and environmental conditions.

In recent years, deep learning models for time-series anomaly detection have emerged as powerful alternatives. Unsupervised autoencoder-based approaches, including LSTM Autoencoders and Variational Autoencoders (VAE), learn to reconstruct "normal" operational patterns and flag deviations as anomalies. Sequential architectures such as LSTMs capture long-range temporal dependencies in sensor data, while attention mechanisms further enhance end-to-end models by focusing on the most informative temporal regions. In parallel, hybrid pipelines that combine deep sequence embeddings with classical ensemble classifiers, such as Random Forests, have demonstrated strong performance in related anomaly-detection domains, balancing representational power with interpretability.

Despite these advances, fuel-theft detection in moving tanker fleets remains underexplored. Real-world deployments face three main challenges:
(i) **Data scarcity**, since verified theft labels are rare and expensive to obtain;
(ii) **Sensor uncertainty**, as GPS jitter, gauge drift, and flow-sensor noise complicate modeling; and
(iii) **Distribution shift**, because driving conditions, routes, and loading patterns vary widely between training and deployment.
To address these challenges, simulation-based approaches offer a principled path: by generating realistic, parameterized tanker trips under controlled scenarios, one can systematically evaluate detection models under class imbalance and domain shift without requiring costly labeled datasets.

In this work, we present a simulation-driven framework for tanker fuel-theft detection. Our contributions are threefold:

- We develop a configurable trip simulator that generates realistic multi-sensor streams (geofence, speed, tank level, and flow), including authorized delivery and unauthorized theft regimes with sensor noise and operational variability.
- We design two complementary modeling strategies: (a) an end-to-end sequence model with attention (Seq-Attn), which integrates LSTM-based temporal encoding with anomaly-sensitive auxiliary features, and (b) a hybrid approach (LSTM+RF) that leverages LSTM embeddings as inputs to a Random Forest classifier for robust tri-class decisions.
- We evaluate these approaches across three scenarios—a balanced baseline, a class-imbalance regime, and a distribution-shift regime—using precision, recall, F1, PR-AUC, ROC-AUC, and confusion matrices, with emphasis on performance for the *theft* class.

Our results show that both methods achieve near-perfect discrimination of theft events in the baseline case. Under more challenging conditions, Seq-Attn consistently preserves higher recall while maintaining perfect or near-perfect precision, whereas LSTM+RF remains competitive but exhibits slightly more missed thefts and occasional false positives.

These findings suggest prioritizing Seq-Attn as the primary detector to minimize missed thefts, while leveraging LSTM+RF for complementary ensemble-based verification.

2. RELATED WORK

Prior research on detecting anomalous fuel usage in vehicular and energy systems has increasingly leveraged machine learning techniques. For instance, Barbado and Corcho [5] applied unsupervised anomaly detection combined with interpretable machine learning to flag and explain unusual fuel consumption events in fleet vehicles.
In the domain of fuel tank and generator monitoring, **Mulongo et al. [6]** applied supervised machine learning techniques to model fuel consumption patterns, employing four classifiers: Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Logistic Regression (LR), and Multi-Layer Perceptron

(MLP) for anomaly detection. The results showed that MLP achieved the best performance, with an accuracy of 96% and an AUC score of 98%.

Building on this, **Atemkeng et al. [7]** proposed a framework based on a *label-assisted autoencoder* for generator fuel data, which leverages a limited number of labeled anomalies to adjust detection thresholds. Their method achieved an anomaly detection accuracy of 97.2%, outperforming the baseline LSTM-autoencoder model with 96.1% accuracy on the same dataset. Notably, the anomalies detected in this domain (such as irregular fuel drops) correspond to real-world issues like fuel theft or leakages, underscoring the importance of time-series models and artificial intelligence in detecting fuel losses.

Transformer-based architectures have shown great promise. **Shi et al. [8]** introduced a Transformer network equipped with a convolutional-attention module, which was able to extract electricity consumption features more effectively and significantly improve the accuracy of electricity theft detection.
Similarly, **Lilhore et al. [9]** proposed a hybrid model combining CNN and bidirectional LSTM layers with a two-way attention module optimized using a Multi-Objective Particle Swarm Optimization (MPSO) algorithm, which enhanced prediction and detection performance in smart grids.

Attentional sequence models have also proven effective in anomaly-related scenarios; for example, **Zhang et al. [10]** integrated an attention mechanism into an LSTM-based pipeline monitoring system to highlight sensor signals near leakage locations, thereby improving real-time leak detection with an accuracy of **AUC=0.99**.

Khattak et al. [11] proposed a hybrid model in which a Gated Recurrent Unit (GRU), for extracting temporal consumption patterns, was combined with a CNN (for extracting abstract features). This model was evaluated on electricity consumption data and, by employing ADASYN and TomekLinks to address class imbalance, achieved superior performance compared to conventional classifiers, obtaining **PR-AUC=0.985** and **ROC-AUC=0.987**.

Even more lightweight designs have maintained strong performance; for example, **Collier and Guha [12]** developed a lightweight LSTM (“watchdog”) that activates full sequence modeling only for suspicious inputs. This approach preserved high recall for energy theft events while reducing inference energy consumption by more than **64%**.

The success of these sequence-based and attention-augmented models in the energy domain suggests that they can also be applied to analogous problems such as detecting fuel theft in vehicle fleets.

Beyond purely end-to-end neural network approaches, researchers have explored hybrid strategies in which deep sequential feature extractors are combined with traditional ensemble classifiers or other hybrid learning methods. For example, in the domain of cyber intrusion detection (a task analogous to sequential anomaly detection), **Hashmi et al. [13]** proposed a BiLSTM with a feature-weighted attention layer, and the extracted temporal features were then fed into a Random Forest (RF) classifier. This model achieved over **99% accuracy** on benchmark datasets and significantly reduced false alarms.

Similarly, **Zeekflow+ [14]** integrated a deep LSTM autoencoder with an RF classifier and effectively identified anomalous network flows in both binary and multi-class settings with accuracy exceeding **99%**.

In the context of the Industrial Internet of Things (IIoT), **Takele and Villányi [15]** introduced a hybrid framework that combines federated and centralized learning, in which LSTM autoencoders were employed. The results demonstrated that this combination improved detection performance and reduced false positive rates in IIoT anomaly scenarios.

Collectively, these studies motivate our two complementary modeling strategies: (i) an end-to-end BiLSTM+Attention model (Seq-Attn) for capturing rich temporal dependencies, and (ii) a hybrid LSTM+RF approach that combines deep sequential embeddings with ensemble-based decision making. This dual design balances accuracy, recall, and interpretability in detecting subtle fuel-theft events

3. METHODOLOGY

Our framework integrates (i) a configurable trip simulator that generates multivariate sensor streams for tanker operations, (ii) a preprocessing pipeline for window segmentation and labeling, and (iii) two modeling strategies: an end-to-end sequence model with attention (Seq-Attn) and a hybrid model combining LSTM embeddings with a Random Forest (LSTM+RF). An unsupervised LSTM Autoencoder (AE) is further employed to provide reconstruction-based anomaly features.

3.1. Data Simulation

Each trip is simulated as a multivariate time series $X \in \mathbf{R}^{T\times F}$ with $\mathrm{F} = 6$ features sampled at $\Delta \mathrm{t} = 1\mathrm{s}$:

$$x_t = \{\mathrm{geofence}, \mathrm{speed_{kmh}}, \mathrm{flow_{lpm}}, \mathrm{fuel_l}, \Delta\mathrm{fuel_{lpm}}, \mathrm{stop_{sec}}\} \tag{1}$$

The trip structure consists of three geofence segments (depot–road–depot), and the dynamics alternate between three operational regimes:

- **Normal:** smooth driving on the road with random stops. In some trips a small leak of 5—8 LPM is added to the entire flow profile.
- **Delivery (authorized unloading):** occurs only in the final depot segment, under long stops. Outflow rates are modeled as Gaussian with mean 120 LPM and standard deviation 20 LPM.
- **Theft (unauthorized discharge):** occurs only in the road segment when the vehicle speed is low ($\leq$ 3 km/h). No minimum stop duration is enforced in the current configuration. Outflow rates are sampled uniformly in the range [10,90] LPM.

For both delivery and theft events, the fuel level decreases consistently with the outflow rate. Empirically, the fuel gauge drop follows the discharged volume closely.

Noise sources include:

- fuel gauge noise (about $\pm$ 6 L),
- sloshing dynamics modeled as an AR(1) process driven by acceleration.

The fuel balance is explicitly enforced as

$$\mathrm{fuel_{t+1}} = fuel_t - \mathrm{flow}_t \times \frac{\Delta\mathrm{t}}{60} + \epsilon_t, \tag{2}$$

where ϵ_t represents gauge noise and sloshing disturbance.

3.2. Windowing and Labeling

Trips are segmented into overlapping windows of length W (e.g. W=60s, stride 10s). For a window $X_i = (x_t, \dots, x_{t+W-1})$, the label $y_i \in$ {normal, delivery, theft} is assigned by a *presence-based priority rule*: if any sample in the window is labeled theft, the window is labeled theft; else if any sample is delivery, it is labeled delivery; otherwise it is normal. This yields a dataset $\{(\mathrm{X}_i, \mathrm{y_i})\}$ for model training.

3.3. Unsupervised LSTM Autoencoder

To capture nominal behavior, an LSTM Autoencoder is trained solely on normal windows. The following is the encoder, decoder, and the reconstruction loss:

- Encoder dynamics:

$$h_t = \mathrm{LSTM_{enc}}(x_t, h_{t-1}), \qquad z = h_W. \tag{3}$$

- Decoder dynamics:

$$\hat{x}_t = \mathrm{LSTM}_{\mathrm{dec}}(z, \hat{x}_{t-1}). \quad (4)$$

- Reconstruction loss:

$$\mathcal{L}_{\mathrm{AE}} = \frac{1}{WF} \sum_{t=1}^{W} \sum_{j=1}^{F} |x_{t,j} - \hat{x}_{t,j}|. \quad (5)$$

From reconstruction errors, auxiliary features $\mathrm{f}_{\mathrm{aux}}$ are computed: per-feature MAE, mean/variance, and percentiles. These augment downstream classifiers with anomaly-sensitive signals.

3.4. Seq-Attn End-to-End Model

The Seq—Attn model integrates sequential features and auxiliary AE features. First, hidden states are extracted with a bidirectional LSTM:

$$h_t = \mathrm{BiLSTM}(x_t, h_{t-1}), \qquad h_t \in \boldsymbol{R}^d. \quad (6)$$

Attention weights:

$$e_t = v^T \tanh(W h_t + b), \qquad \alpha_t = \frac{\exp(e_t)}{\sum_{\tau=1}^{W} \exp(e_\tau)} \quad (7)$$

Context vector:

$$c = \sum_{t=1}^{W} \alpha_t h_t \quad (8)$$

Final representation concatenates attention context with auxiliary features:

$$z = [c; f_{\mathrm{aux}}] \quad (9)$$

Prediction via softmax:

$$\hat{y} = \mathrm{softmax}(W_o z + b_o), \qquad \hat{y} \in \Delta^{K-1}, K = 3 \quad (10)$$

Loss function:

$$\mathcal{L}_{CE} = -\sum_{k=1}^{K} y_k \log \hat{y}_k \quad (11)$$

3.5. LSTM+RF Hybrid Model

The hybrid model decouples sequence encoding and classification.

Step 1: LSTM encoder.

$$h_t = \mathrm{LSTM}_{\mathrm{enc}}(x_t, h_{t-1}), \qquad z = h_W. \quad (12)$$

Step 2: Random Forest
Given embeddings $\{z\}$, a Random Forest with M trees $\{T_m\}$ predicts:

$$\hat{y} = \operatorname*{argmax}_k \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}\{T_m(z) = k\}, \quad (13)$$

where $\mathbf{1}\{T_m(z) = k\}$ is the indicator function, returning 1 if the $T_m(z) = k$, and 0 otherwise. This approach provides interpretability through feature importance and resilience to class imbalance.

3.6. Evaluation Protocol

We evaluate under three scenarios: (i) Baseline (balanced data, matched distributions), (ii) Scenario A (Imbalance) where theft probability is reduced, and (iii) Scenario B (Shift) with altered test speed distributions. Models are trained on trip-wise splits (no leakage). Performance is reported with precision, recall, F1 per class, confusion matrices, and PR/ROC AUC for the theft class.

4. RESULTS

In this section, we present the evaluation results of two modeling approaches—the end-to-end **Seq-Attn** model and the hybrid **LSTM+RF** method—under three experimental scenarios: the *Baseline*, *Scenario A* with class imbalance, and *Scenario B* with distribution shift. The main focus is on the performance for the *theft* class, which is the minority class and also the most critical from an application perspective. The comparison between models is conducted using standard metrics including Precision, Recall, and F1-score for the theft class, along with PR-AUC and ROC-AUC, as well as confusion matrices to provide an intuitive understanding of errors. These results allow us to assess model performance both in ideal conditions and under more challenging scenarios with class imbalance or distribution shift.

Table 1: Performance comparison on theft class across scenarios.

2*Scenario	Seq-Attn			LSTM+RF		
	Precision	Recall	F1	Precision	Recall	F1
Baseline	1.0000	1.0000	1.0000	1.0000	0.9986	0.9993
Scenario A (Imbalance)	1.0000	0.9886	0.9943	1.0000	0.9830	0.9914
Scenario B (Shift)	1.0000	0.9972	0.9986	0.9957	0.9957	0.9957

Table 1 presents the performance of the two models, **Seq–Attn** and **LSTM+RF**, on the theft class across three different scenarios. Since fuel theft detection is the primary objective of this study, the analysis focuses on **Recall** (avoiding missed thefts) and **Precision** (reducing false alarms).

Baseline (normal conditions). Both models are nearly flawless. **Seq–Attn** achieves Precision=1.0000, Recall=1.0000, and F1=1.0000. **LSTM+RF** also attains Precision=1.0000, Recall=0.9986, and F1=0.9993. Therefore, under ideal conditions, both models are highly reliable, with Seq–Attn being perfectly accurate and LSTM+RF falling short by only a negligible margin.

Scenario A (Imbalance). When theft incidents are very rare, the results favor **Seq–Attn**: it reaches Precision=1.0000, Recall=0.9886, and F1=0.9943, while **LSTM+RF** achieves Precision=1.0000, Recall=0.9830, and F1=0.9914. In imbalanced settings, Seq–Attn misses fewer thefts without introducing additional false alarms, which is operationally significant since even a single missed theft can be costly.

Scenario B (Distribution shift). Under distributional changes at test time (e.g., altered speed patterns), both models maintain high performance. **Seq–Attn** achieves Precision=1.0000, Recall=0.9972, and F1=0.9986, while **LSTM+RF** yields Precision=0.9957, Recall=0.9957, and F1=0.9957. Thus, Seq–Attn provides slightly higher Recall and better F1 and Precision in this setting.

Takeaway. By prioritizing Recall (to avoid missed thefts) and then Precision (to reduce false alarms), the results show that **Seq–Attn is the more reliable choice in Scenarios A and B**, while in the Baseline setting both models are nearly saturated in performance.

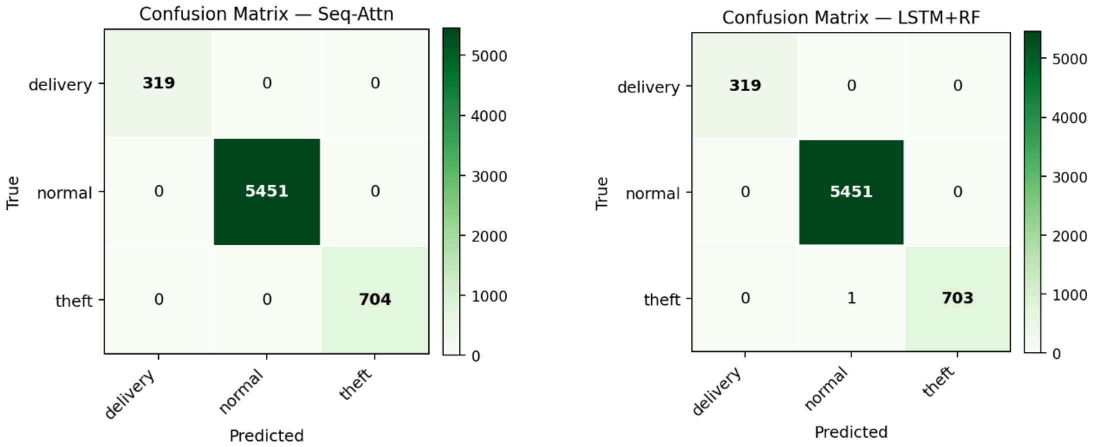

Figure 1: Confusion matrices for **Baseline;** left: Seq–Attn, right: LSTM+RF.

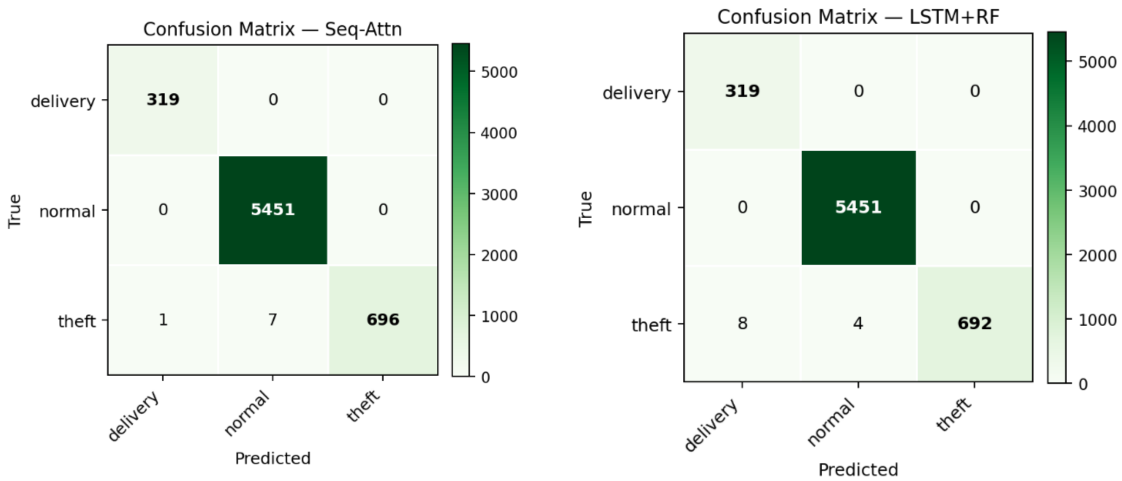

Figure 2: Confusion matrices for Scenario A (Imbalance); left: Seq–Attn, right: LSTM+RF.

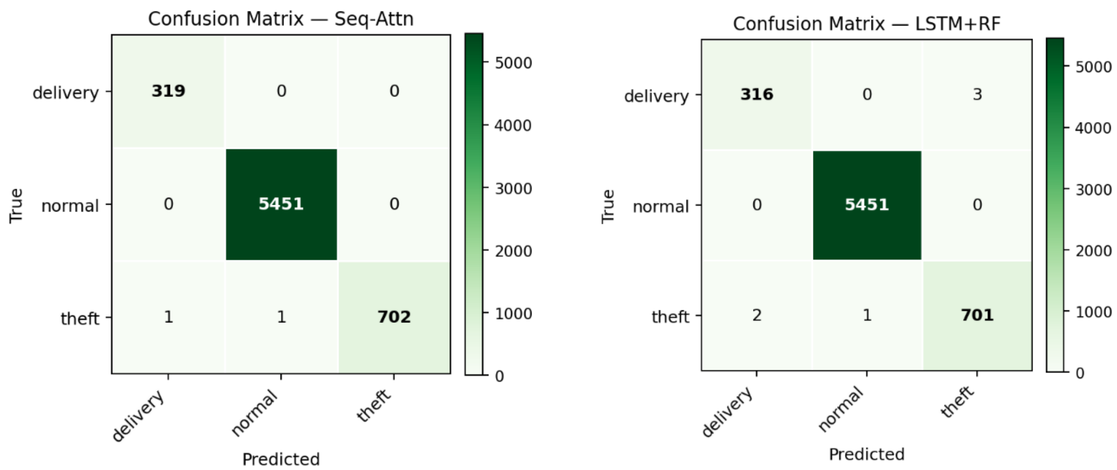

Figure 3: Confusion matrices for Scenario B (Shift); left: Seq–Attn, right: LSTM+RF.

The confusion matrices in Figures 1—3 provide a qualitative view of the models' behavior across scenarios.

Scenario A. Under severe class imbalance, Seq—Attn misses fewer theft events (higher recall) while maintaining perfect precision. LSTM+RF remains competitive but exhibits slightly more missed thefts and a few delivery false positives. Overall, Seq—Attn provides a safer operating point with fewer misses in this regime (Figure 2).

Scenario B. Under distributional shifts (e.g., different speed ranges), both models remain robust. Seq–Attn shows only one misclassification of a theft instance (to delivery), nearly maintaining perfect performance. LSTM+RF makes slightly more errors, misclassifying 2 thefts as delivery and 1 as normal, reflecting a small drop in recall under shift.

Summary. The visualizations confirm the updated metrics: Seq–Attn achieves flawless classification in the Baseline and near-flawless performance in the Shift scenario, while LSTM+RF performs better in highly imbalanced settings but is more prone to occasional false positives and theft misses.

Table 2: AUC metrics on theft class (PR-AUC and ROC-AUC).

2*Scenario	Seq-Attn		LSTM+RF	
	PR-AUC	ROC-AUC	PR-AUC	ROC-AUC
Baseline	1.0000	1.0000	1.0000	1.0000
Scenario A (Imbalance)	0.9999	0.9999	0.9992	0.9992
Scenario B (Shift)	1.0000	1.0000	0.9999	0.9999

Analysis of AUC metrics Table 2 presents the performance of the models Seq–Attn and LSTM+RF in terms of AUC metrics (including PR-AUC and ROC-AUC) for the theft class across three scenarios.

Baseline (normal conditions). Both models are saturated, achieving PR-AUC=1.0000 and ROC-AUC=1.0000. These results indicate that under normal conditions, the separation between theft and non-theft samples is perfectly accurate.

Scenario A (Imbalance). Under imbalanced conditions where theft events are very rare, Seq–Attn nearly preserves perfect separability (PR-AUC=0.9999, ROC-AUC=0.9999). In contrast, LSTM+RF shows a slightly larger drop (PR-AUC=0.9992, ROC-AUC=0.9992), consistent with its lower recall in this scenario.

Scenario B (Distribution shift). Under distributional changes in the test data (e.g., altered speed patterns), both models remain close to the ceiling. Seq–Attn achieves PR-AUC=1.0000 and ROC-AUC=1.0000, while LSTM+RF performs slightly lower at PR-AUC=0.9999 and ROC-AUC=0.9999.

Takeaway. The AUC values are nearly saturated in all scenarios, showing that distinguishing theft from non-theft samples is generally very easy. The key differences between the models emerge when focusing on the thresholded operating point: in this case, Seq–Attn provides higher recall (fewer missed thefts) without sacrificing precision.

5. CONCLUSION AND FUTURE WORK

This paper presented a reproducible, end-to-end framework for detecting fuel theft in road-tanker fleets, combining a geofence-aware trip simulator with two modeling approaches: an end-to-end sequence model with attention (Seq-Attn) and a hybrid LSTM embedding with Random Forest (LSTM+RF).

Across three evaluation scenarios (baseline, class imbalance, and distribution shift), both methods achieved near-perfect discrimination for the *theft* class. In line with the operational priority of not missing thefts, the updated results show that Seq-Attn consistently attains higher recall in challenging conditions (Scenario A: imbalance; Scenario B: shift), while also preserving perfect or near-perfect precision. By contrast, LSTM+RF remains competitive but exhibits slightly more missed thefts and occasional false positives in imbalanced or shifted regimes. Under the ideal baseline, both models saturate, though Seq-Attn is strictly perfect.

Considering the PR-AUC and ROC-AUC results, which are effectively at ceiling for both approaches, the decisive differences emerge at the thresholded operating point. Here, Seq-Attn provides a safer balance between recall and precision in non-ideal conditions.

Practical Implications. A viable deployment strategy is to use Seq—Attn as the primary detector, ensuring fewer missed thefts even under imbalance or distributional shift, and to employ LSTM+RF as a supplementary model for diversity or ensemble-based verification.

Future Work. Potential extensions include: (i) extending the simulator with richer operational patterns (multi-stop deliveries, varied theft tactics), (ii) driver training and real-world data collection, and (iii) exploring on-device lightweight models for real-time deployment in fleet monitoring systems.

REFERENCES

[1] Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.

[2] Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, *27*(2), 83-85.

[3] Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, No. 1, pp. 9—11). Cambridge: MIT press.

[4] Mohammadpour, N., Fozi, M., Ebadzadeh, M. M., Azimi, A., & Kamali, A. (2024). Proximal Policy Optimization with Adaptive Generalized Advantage Estimate. In *Proceedings of the First International Conference on Machine Learning and Knowledge Discovery (MLKD 2024)*. (pp. 453—458).

[5] Barbado, A., & Corcho, Ó. (2022). Interpretable machine learning models for predicting and explaining vehicle fuel consumption anomalies. *Engineering Applications of Artificial Intelligence*, *115*, 105222.

[6] Mulongo, J., Atemkeng, M., Ansah-Narh, T., Rockefeller, R., Nguegnang, G. M., & Garuti, M. A. (2020). Anomaly detection in power generation plants using machine learning and neural networks. *Applied Artificial Intelligence*, *34*(1), 64-79.

[7] Atemkeng, M., Osanyindoro, V., Rockefeller, R., Hamlomo, S., Mulongo, J., Ansah-Narh, T., ... & Fadja, A. N. (2023). Label assisted autoencoder for anomaly detection in power generation plants. *arXiv preprint arXiv:2302.02896*.

[8] Shi, J., Gao, Y., Gu, D., Li, Y., & Chen, K. (2023). A novel approach to detect electricity theft based on conv-attentional Transformer Neural Network. *International Journal of Electrical Power & Energy Systems*, *145*, 108642.

[9] Lilhore, U. K., Dalal, S., Radulescu, M., & Barbulescu, M. (2025). Smart grid stability prediction model using two-way attention based hybrid deep learning and MPSO. *Energy Exploration & Exploitation*, *43*(1), 142-168.

[10] Zhang, X., Shi, J., Yang, M., Huang, X., Usmani, A. S., Chen, G., ... & Li, J. (2023). Real-time pipeline leak detection and localization using an attention-based LSTM approach. *Process Safety and Environmental Protection*, *174*, 460-472.

[11] Khattak, A., Bukhsh, R., Aslam, S., Yafoz, A., Alghushairy, O., & Alsini, R. (2022). A hybrid deep learning-based model for detection of electricity losses using big data in power systems. *Sustainability*, *14*(20), 13627.

[12] Collier, C., & Guha, K. (2025). Lightweight LSTM Model for Energy Theft Detection via Input Data Reduction. *arXiv preprint arXiv:2507.02872*.

[13] Hashmi, A., Barukab, O. M., & Hamza Osman, A. (2024). A hybrid feature weighted attention based deep learning approach for an intrusion detection system using the random forest algorithm. *Plos one*, *19*(5), e0302294.

[14] Arapidis, E., Temenos, N., Giagkos, D., Rallis, I., Kalogeras, D., Papadakis, N., ... & C. Messinis, S. (2024, June). Zeekflow+: A Deep LSTM Autoencoder with Integrated Random Forest Classifier for Binary and Multi-class Classification in Network Traffic Data. In *Proceedings of the 17th International Conference on PErvasive Technologies Related to*

Assistive Environments (pp. 613-618).

[15] Takele, A. K., & Villányi, B. (2022, May). Anomaly detection using hybrid learning for industrial IoT. In *2022 IEEE 2nd Conference on Information Technology and Data Science (CITDS)* (pp. 262-266).