Fırat Üniversitesi Müh. Bil. Dergisi
37(2), 699-709, 2025

Araştırma Makalesi
https://doi.org/10.35234/fumbd.1668498

Anomaly Detection in Unmanned Aerial Vehicle Telemetry Using Automated Machine Learning

Anıl SEZGİN[1*], Rasim KESKİN[2], Aytuğ BOYACI[3]
[1] Research and Development, Siemens A.S., Istanbul, Türkiye
[2] Marmara Teknokent A.S., Kocaeli, Türkiye
[3] Department of Computer Engineering, National Defence University, Air Force Academy, Istanbul, Türkiye
[*1] anil.sezgin@siemens.com, [2] rkeskin@marmarateknokent.com.tr, [3] aytug.boyaci@msu.edu.tr

(Geliş/Received: 31/03/2025; Kabul/Accepted: 15/07/2025)

Abstract: Reliable analysis of UAV telemetry data is critical for mission safety, especially as drones are increasingly deployed in complex and high-risk environments. These data streams often include anomalies arising from sensor faults, environmental disruptions, or cyber-physical attacks, making robust anomaly detection essential. This study introduces an unsupervised anomaly detection framework designed specifically for high-frequency UAV telemetry. It combines domain-driven feature engineering with an AutoML-based optimization pipeline that enables automated model selection and hyperparameter tuning. The framework integrates four unsupervised algorithms—Local Outlier Factor, Isolation Forest, One-Class SVM, and Elliptic Envelope—ensuring adaptability to the dynamic nature of UAV operations. Evaluated on a real-world dataset of 127,000 samples from 48 UAV missions, the system uses expert-labeled anomaly segments solely for validation to preserve the integrity of unsupervised learning. Among all methods, Local Outlier Factor yielded the best results with 0.920 accuracy, 0.880 precision, 0.850 recall, and 0.860 F1-score. Scalable and low-latency, the proposed solution is well-suited for real-time deployment. By bridging theoretical advances with operational needs, this work contributes to safer and more resilient aerial robotic systems.

Key words: Anomaly detection, unmanned aerial vehicles, automated machine learning.

İnsansız Hava Aracı Telemetrisinde Otomatik Makine Öğrenmesi Tabanlı Anomali Tespiti

Öz: İnsansız Hava Araçlarının (İHA) telemetri verilerinin güvenilir şekilde analiz edilmesi, özellikle karmaşık ve riskli ortamlarda görev başarısı ve operasyonel güvenlik açısından kritik öneme sahiptir. Bu veri akışları, sensör arızaları, çevresel etkenler veya siber-fiziksel saldırılar nedeniyle anormallikler içerebilir. Bu nedenle, sağlam bir anomali tespit mekanizması gereklidir. Bu çalışma, yüksek frekanslı İHA telemetrisi için özel olarak tasarlanmış, gözetimsiz bir anomali tespit çerçevesi sunmaktadır. Yaklaşım, alan bilgisine dayalı özellik mühendisliğini, model seçimi ve hiperparametre ayarlarını otomatikleştiren bir AutoML tabanlı optimizasyon süreciyle birleştirir. Sistem; Local Outlier Factor, Isolation Forest, One-Class SVM ve Elliptic Envelope olmak üzere dört farklı gözetimsiz algoritmayı entegre ederek, İHA operasyonlarının dinamik doğasına uyum sağlar. 48 farklı İHA görevinden toplanan 127.000 örnek içeren gerçek dünya veri kümesi üzerinde yapılan değerlendirmelerde, uzmanlar tarafından etiketlenmiş anomali segmentleri yalnızca doğrulama amacıyla kullanılmıştır. En iyi performans, %92 doğruluk, %88 kesinlik, %85 duyarlılık ve %86 F1-skoru ile Local Outlier Factor algoritması tarafından elde edilmiştir. Gerçek zamanlı uygulamalar için ölçeklenebilir ve düşük gecikmeli olarak tasarlanan bu sistem, İHA'larda otomatik arıza izleme ve güvenli, dayanıklı hava araçları ekosistemlerinin gelişimine önemli katkılar sunmaktadır.

Anahtar kelimeler: Anomali tespiti, insansız hava araçları, otomatik makine öğrenmesi.

1. Introduction

The Unmanned Aerial Vehicles or drones have revolutionized industries ranging from logistics and agriculture to disaster management and defense. With these autonomous systems becoming integrated into the contemporary infrastructure, their networked operation through the Internet of Drones (IoD) has been a revolutionary paradigm. IoD platforms provide real-time communication, coordination, and data exchange between drone fleets to support high-end applications like aerial surveillance, package delivery, and environmental monitoring. Yet, the use of telemetry data—continuous feeds of sensor readings, position reports, and system status messages—presents significant challenges to the provision of operational reliability and security. Anomalies in IoD telemetry data, regardless of whether they are sensor fault-induced, cyberattack-induced, or interference-

[*] Sorumlu yazar: anil.sezgin@siemens.com. Yazarların ORCID Numarası: [1] 0000-0002-5754-1380, [2] 0000-0003-4889-2995, [3] 0000-0003-1016-3439

induced, present critical risks to mission failure, safety risks, and economic losses. Classical anomaly detection techniques, which are conventionally recommended for static or low-dimensional data sets, cannot handle the dynamic and high-dimensional characteristics of IoD telemetry and require new techniques with specific application to this field.

UAVs have evolved from niche military equipment to pervasive instruments in the commercial and civilian spheres. Their capability to venture into hostile or unreachable territories, along with developments in autonomy and connectivity, has unlocked uses in precision agriculture, search-and-rescue missions, and urban air mobility. The IoD environment, which networks drones via cloud servers and edge computing nodes, continues to augment their capabilities by facilitating fleet-level control, real-time analytics, and remote command issuance.

Central to IoD operations is telemetry data, a multivariate time-series stream capturing metrics such as:

- Positional Data: Latitude, longitude, altitude, and GPS fix status.
- Kinematic Parameters: Groundspeed, airspeed, climb rate, and orientation (roll, pitch, yaw).
- System Health: Battery voltage, current, energy consumption, and vibration levels.
- Mission-Specific Metrics: Distance to target, waypoint progression.

These parameters must be monitored for flight stability, hardware malfunction, and cyber-physical attack. For example, abrupt voltage fluctuation of the battery indicates faulty power system, and erratic GPS coordinates signal spoofing attacks. Yet, high volume, high velocity, and heterogeneity of telemetry data prevent anomaly detection, especially in real-time analysis-intensive applications.

There are four key challenges in anomaly detection from IoD telemetry data:

1. Temporal Dependencies and High-Dimensionality: Telemetry streams comprise dozens of physically related variables sampled at high rates. For example, a drone's velocity (groundspeed, airspeed) and attitude (roll, pitch, yaw) are physically coupled, and models must discover spatiotemporal correlations rather than isolated features.
2. Unpredictable Operating Environments: UAVs fly in unpredictable environments in which wind gusts, electromagnetic interference, and shifting payloads cause temporary deviations from normal behavior that are indistinguishable from true anomalies.
3. Class Imbalance and Label Scarcity: Anomalies are infrequent occurrences in telemetry data, which results in class-imbalanced datasets. Moreover, labeling anomalies is expensive and usually impossible in real-world deployment.
4. Real-Time Processing Constraints: IoD systems require low-latency anomaly detection to facilitate real-time corrective measures, e.g., rerouting drones or emergency landings.

Classic thresholding approaches and supervised learning break down in these cases. Thresholds are too inflexible to capture variations in context, and supervised models need enormous sets of labeled data that are almost never available. Unsupervised and semi-supervised methods, which learn "normal" patterns from unlabeled data, are an attractive solution but demand careful model and parameter selection—a step still manually and tediously performed.

The incorporation of AutoML in IoD anomaly detection removes the constraints of human model selection and hyperparameter tuning. AutoML simplifies the machine learning pipeline by automating feature engineering, algorithm selection, and hyperparameter tuning required to handle the dynamic and heterogeneous nature of telemetry data. The AutoML process was realized through a custom optimization routine using standard parameter search and tuning strategies, rather than a commercial AutoML platform. In IoD environments, where telemetry feature sets differ greatly across missions and operating conditions, AutoML facilitates adaptive anomaly detection systems that can self-optimize independently. AutoML platforms can assess such trade-offs in real time, choosing the best-performing algorithm based on runtime performance metrics.

This work contributes to the state of the art in methodological innovations specific to IoD anomaly detection. Informing this work is a formalized framework for systematic benchmarking of unsupervised algorithms along UAV-specific operational dimensions like latency, noise robustness, and responsiveness to dynamic environments. This framework provides algorithmic evaluation standardization for real-world IoD deployments. Building on this, we introduce an automated pipeline that streamlines anomaly detection workflows using advanced optimization techniques, tailoring model configurations to telemetry properties on the fly. This automation reduces the reliance on hand-tuning while also making access more democratized for non-machine learning-expert operators. Furthermore, empirical insights are codified into actionable guidelines mapping anomaly types (e.g., sensor failures, communication hijacks) to algorithmic strengths, which simplifies prioritization for practitioners based on mission-critical needs. Together, these contributions close gaps between theoretical developments and practical needs of IoD ecosystems. They enable scalable solutions for autonomous drone networks. Unlike previous approaches, our framework combines AutoML-driven model optimization with expert-validated unsupervised

anomaly detection tailored to the complex, high-frequency nature of UAV telemetry data, ensuring both adaptability and interpretability.

This paper offers the following key contributions:

- A fully unsupervised anomaly detection framework for UAV telemetry, integrating AutoML-based hyperparameter optimization with domain-specific feature engineering.
- A comprehensive benchmarking of four unsupervised models (LOF, Isolation Forest, One-Class SVM, Elliptic Envelope) on a real-world UAV telemetry dataset comprising 127,000 records from 48 missions.
- A validated, real-time ready anomaly detection pipeline, optimized for deployment in resource-constrained aerial platforms with interpretability and scalability in mind.

As UAVs become mainstream technologies, their safe integration into airspace systems is of the highest priority. Our framework addresses this requirement by eliminating the necessity of manual tuning and adapting to evolving telemetry patterns, enabling operators to focus on mission completion. By enhancing resilience to both hardware failure and cyber attacks, this research paves the way for safer, more reliable autonomous systems in an increasingly connected aerial ecosystem.

The remainder of this paper is structured to guide the reader through the research methodology, experimental results, and practical insights. Section 2 reviews related work in UAV anomaly detection, AutoML, and IoT security, contextualizing our contributions within existing literature. Section 3 details the IoD telemetry dataset, preprocessing techniques, and the architecture of the proposed approach. Section 4 presents empirical evaluations of the four anomaly detection algorithms, comparing their performance. Section 5 discusses the implications of our findings, addressing limitations and trade-offs in real-world deployments. Finally, Section 6 concludes with a summary of key contributions and future research directions, including the integration of federated learning for privacy-preserving swarm analytics and edge-AI optimizations for low-latency processing.

2. Related Work

The domain of anomaly detection has witnessed significant advancements across IoT, industrial systems, and unmanned aerial systems, driven by innovations in machine learning, sensor fusion, and decentralized architectures. This section organizes recent research into thematic categories, emphasizing methodologies, challenges, and contributions to handling multivariate time-series data, privacy preservation, and real-time processing.

2.1. Anomaly detection in IoT and industrial systems

IoT and industrial applications demand robust frameworks to manage multivariate sensor data, class imbalance, and dynamic operational environments. Study [1] addresses sensor interdependencies by clustering correlated sensor streams, offering scalability for high-dimensional IoT environments. Complementing this, [2] tackles class imbalance through XGBoost-based feature selection and optimized LSTM loss functions, achieving an AUC-ROC of 0.984. The integration of network and sensor data is explored in [3], where autoencoders and LSTMs enhance detection robustness against stealthy attacks in industrial control systems. [4] further validates the benefits of multi-source data fusion, combining network traffic, sensor readings, and hardware status to achieve 85.41% accuracy in anomaly classification. Challenges in additive manufacturing are addressed by [5], which employs zero-bias deep neural networks (ZBDNN) to detect defects like voids and resin-rich areas with 99.71% accuracy, while [6] leverages GANs to balance datasets and improve defect detection reliability. Automated model selection is tackled in [7], which uses meta-learning to dynamically choose optimal algorithms based on manufacturing data characteristics, reducing dependency on domain expertise. Building on this, the AID4I framework [8] leverages automated machine learning to perform end-to-end intrusion detection in IIoT networks, combining preprocessing, hybrid SHAP-genetic feature selection, and hyperparameter tuning across 14 classifiers. It achieves up to 99.87% accuracy while significantly reducing model development time and manual effort. Study [9] extends these principles to smart transportation, integrating Bayesian change point detection and forecasting to secure connected vehicles with 53.83% higher accuracy than traditional methods.

2.2. UAV-specific anomaly detection techniques

Unmanned Aerial Vehicles demand customized solutions owing to their dependency on multi-sensor infrastructures, dynamic flight environments, and exposure to cyber-physical attacks. Federated learning and multi-modal denoising are employed in research [10] for privacy-preserving anomaly detection in UAV swarms

with 99.01% detection accuracy. [11] secures swarms via Merkle tree-based attestation and Byzantine consensus, which allows decentralized containment of anomalies without exposing data. Sensor fusion is the focus of [12], combining GPS, accelerometer, and gyroscope data to achieve quasi-perfect accuracy (AUC = 1.00). Spatiotemporal correlations are leveraged in [13], where an STC-LSTM-AE model recovers actual flight data with 98.75% accuracy. Hardware-specific challenges are overcome in [14], where wavelet scattering is combined with LSTM autoencoders for detecting propeller failures up to 130 seconds in advance of failure. [15] trains semi-supervised 1D convolutional models on normal flight data to detect signal noise and transient failures, whereas [16] employs Large Language Models (LLMs) with retrieval-augmented generation (RAG) to facilitate contextualized decision-making for UAV missions. Study [17] explores AI-driven fuzzing techniques to uncover vulnerabilities in UAV firmware and protocols, highlighting gaps in proprietary drone security testing. Complementing these efforts, study [18] presents a comprehensive analysis of privacy and security challenges in the Internet of Drones, identifying GPS spoofing, data injection, and command tampering as key threats. The study reviews mitigation strategies such as blockchain-based authentication, lightweight IDS, and cryptographic communication frameworks tailored for UAV networks.

2.3. Addressing data scarcity and concept drift

Anomaly detection in evolving systems must overcome challenges like limited labeled data and dynamic environments. Study [19] addresses data scarcity by aligning simulated and real UAV data via dynamic time warping (DTW), enabling effective knowledge transfer with LSTM-AM models. [20] employs memory-augmented autoencoders (MemAE) to store prototypical flight patterns, achieving AUC scores up to 0.9988 with minimal training data. Online learning is explored in [21], which combines ARF-ADWIN and KNN-ADWIN models with PSO optimization to adapt to IoT data streams in real time. [22] tackles sensor degradation through meta-learning and ensemble strategies, improving F1-scores by 16% in industrial systems. [23] introduces VMD-LSTM hybrids to filter periodic patterns in server telemetry, minimizing false alarms through automatic hyperparameter tuning.

2.4. Privacy-preserving and lightweight detection frameworks

Resource constraints and privacy concerns in distributed systems necessitate efficient and secure anomaly detection methods. [24] combines semi-supervised learning with Mamdani fuzzy inference systems (FIS) to reduce data transmission overhead, achieving 99.70% accuracy in WSNs. [25] deploys a lightweight multi-classification model on UAV firmware, requiring only 48 KB of storage while achieving 89.38% accuracy. Study [26] introduces a Node Performance Score (NPS) for cluster head selection, improving network stability by 258% compared to traditional protocols.

3. Methodology

Although this study employs unsupervised learning techniques for anomaly detection, it strategically integrates domain expertise to enrich the data labeling process for subsequent evaluation. In the early stages of dataset preparation, experienced drone operators and flight engineers meticulously reviewed the telemetry data, identifying and annotating segments that exhibited known or suspected anomalous behavior based on operational context and empirical understanding. These annotations were incorporated into the dataset as labeled anomalies; however, to uphold the unsupervised nature of the detection approach, they were strictly excluded from the training phase of the models. This ensured that the models learned to characterize normal flight behavior independently, without being influenced by predefined notions of anomalies.

The telemetry dataset utilized in this research consists of 48 separate flight sessions, each of which corresponds to an individual UAV mission under different operational and environmental conditions. Throughout the sessions, overall telemetry records of around 127,000 were gathered, covering parameters like GPS location, orientation (roll, pitch, yaw), battery level, and velocity metrics. Among the full dataset, anomalous segments account for roughly 7.5% of the samples, as identified through expert annotation. These anomalies include sensor malfunctions (e.g., GPS dropout, abrupt voltage fluctuations), flight instability (e.g., sudden roll angle changes), and potential cyber-physical interferences (e.g., command spoofing or inconsistent data spikes). The remaining 92.5% of the data reflects stable and expected UAV operation, providing a rich foundation for unsupervised learning. This distribution captures the real-world scarcity of anomalies, highlighting the inherent detection challenges while providing realistic test conditions. By clearly defining the number of flight sessions, anomaly

ratio, and deviation nature, this study seeks to enhance transparency, reproducibility, and contextual insight for follow-on researchers expanding on this dataset.

The anomaly data with labels was only consulted in the evaluation step, where we used it to calculate the performance of the anomaly detection models. We could compare the models' anomaly predictions and the ground truth labels that the experts saw for calculating strong supervised metrics like precision, recall, F1-score, and accuracy. This mixed approach—unsupervised model training with supervised validation—is a pragmatic trade-off that leverages the best of both paradigms. It enables learning unbiased from normal data and still permits strict, quantifiable validation against expert judgment. The use of labeled anomalies in validation improves not just the interpretability of results, but renders the outputs of the models meaningful in the context of real-world expectations and domain-specific implications. This kind of paradigm is particularly worthwhile in safety-critical contexts like drone surveillance, where it is as important to identify small deviations as to exclude false positives. This article outlines an accurate and structured approach to anomaly detection in drone telemetry data through a fusion of traditional outlier detection techniques and domain-specific feature engineering. As elucidated in Table 1, the procedure starts with the importing of raw telemetry data in the JSON format comprising high-resolution sensor measurements of drones and ground truth labels for anomalies. These datasets usually consist of high-frequency time-series like GPS location, accelerometer, gyroscope, battery voltage, and orientation readings necessary to log operational states and mark abnormal behavior. In preprocessing, we extract all relevant numerical features and anomaly labels, and we cleanse the dataset of noise and inconsistencies. This includes filtering missing values, normalizing timestamps, and synchronizing asynchronous sensor streams. Besides, domain-specific features are crafted to better capture flight dynamics, e.g., velocity magnitude from GPS, roll-pitch-yaw angles transformation, and energy consumption rates over time. Statistical metrics such as moving averages, standard deviations, and z-scores are also calculated to enhance feature richness and highlight subtle deviations in behavior. Following feature construction, an Automated Machine Learning pipeline is employed to efficiently explore and optimize a variety of anomaly detection algorithms and their hyperparameters. This enables the systematic selection of high-performing models tailored to the complex patterns present in drone telemetry data, thereby improving detection accuracy and reducing manual tuning efforts.

Table 1. Algorithm of anomaly detection pipeline for drone telemetry.

1. LOAD flight telemetry data from JSON files containing drone sensor readings and anomaly labels
2. PREPROCESS data:
a. EXTRACT numerical features and anomaly labels
b. ENGINEER domain-specific features
c. CALCULATE statistical indicators
3. SPLIT data into train/validation/test sets
4. OPTIMIZE detection algorithms:
a. Isolation Forest
b. One-Class SVM
c. Local Outlier Factor
d. Elliptic Envelope
5. DETECT anomalies:
a. SCALE features using to handle outliers
b. TRAIN models on normal data
c. GENERATE anomaly scores and binary predictions
6. EVALUATE results:
a. CALCULATE metrics
b. VISUALIZE data
7. SAVE optimized models for deployment

The data is then split into training, validation, and testing sets with temporal integrity to avoid data leakage—the most important consideration when dealing with time-series analysis. The highlight of our solution is the hyperparameter optimization of four anomaly detection algorithms: Isolation Forest, One-Class Support Vector Machine, Local Outlier Factor, and Elliptic Envelope. These four unsupervised models are all specialized in high-dimensional anomaly detection and each makes different assumptions about the distribution of the data. All the features are normalized to a standard range before training the models to counter the effect of extreme values and for algorithmic stability. The models are trained on normal flight data only to learn the baseline behavior, and the models produce anomaly scores when they see the test data. Binary anomaly predictions are obtained by taking suitable thresholds on the scores.

Evaluation of detection performance entails calculation of common classification metrics such as precision, recall, F1-score, accuracy, and confusion bias to evaluate the trade-off between true and false predictions of different anomaly detection models. Such metrics give an in-depth knowledge of the trade-off between sensitivity and specificity that is most important in scenarios where anomalies are rare and false alarms lead to inefficiencies in operations. For visualization, individual metric values for each method are plotted as bar charts to enable quick comparative evaluation. Radar charts provide summary visualization by superimposing all the metrics, showing a bird's-eye view of the model performance summary. Confusion matrices also show in-depth understanding of model performance on normal and anomalous examples by graphically visualizing true positives, false positives, true negatives, and false negatives to help identify patterns of misclassifications. Time-series overlays also aid interpretability by emphasizing anomaly regions along temporal sequences. Both of these assessment methods allow quantitative benchmarking and qualitative diagnostics of model trustworthiness. Finally, the tuned models are serialized and persisted for convenient deployment into real-time drone telemetry monitoring systems.

4. Results and Experiments

4.1. Quantitative evaluation

The performance of the proposed unsupervised anomaly detection framework was systematically evaluated using four distinct algorithms—Local Outlier Factor, Elliptic Envelope, Isolation Forest, and One-Class Support Vector Machine. To provide a robust comparative overview, we compiled two key tables that encapsulate the performance metrics and confusion matrix outcomes across these methods. These metrics allow us to assess not only how well each algorithm detects anomalies, but also how their predictions align with the expert-labeled ground truth.

In Table 2, LOF stands out as the most effective method, achieving the highest accuracy (0.920), precision (0.880), recall (0.850), and F1-score (0.860). This combination reflects a well-balanced detection capability, minimizing both false positives and false negatives. The model's Confusion Bias of 0.080—significantly lower than the others—demonstrates its robustness in maintaining an equitable error rate, meaning it neither over-predicts anomalies nor overlooks true ones. This characteristic is particularly valuable in drone telemetry applications, where both missed anomalies and false alarms can have operational consequences.

By contrast, the Elliptic Envelope method shows moderate performance with balanced but slightly inferior values in all metrics and a higher confusion bias of 0.160, suggesting a tendency to overflag normal operations as anomalous. Meanwhile, Isolation Forest and One-Class SVM both deliver relatively lower F1-scores (0.700 and 0.720, respectively) and higher confusion biases, indicating reduced reliability. This table not only support the numerical evaluation of our models but also acts as a reference point for understanding the trade-offs each algorithm brings to real-world deployment. The consistency of LOF across metrics substantiates its selection as the most suitable candidate for robust and interpretable drone anomaly detection in this study. It is important to note that achieving precision and recall above 85% in unsupervised settings, especially with imbalanced and high-dimensional data, reflects substantial detection capability in practical UAV scenarios.

Table 2. Performance metrics for anomaly detection methods.

Method	Accuracy	Precision	Recall	F1-score	Confusion Bias
Local Outlier Factor	0.920	0.880	0.850	0.860	0.080
Elliptic Envelope	0.780	0.760	0.740	0.750	0.160
Isolation Forest	0.730	0.720	0.680	0.700	0.200
One-Class SVM	0.750	0.730	0.710	0.720	0.180

4.2. Comparative analysis and interpretation

The proposed unsupervised anomaly detector's performance was compared thoroughly with four algorithms—Local Outlier Factor, Elliptic Envelope, Isolation Forest, and One-Class Support Vector Machine (SVM). To measure their performance in detecting anomalous behavior in drone telemetry data, expert-marked segments of anomalies were employed as ground truth, with supervised metrics being utilized only at the time of evaluation. Quantitative performance was also achieved through the use of accuracy, precision, recall, F1-score, and confusion bias as primary performance metrics. As evident from the tabulated results, the best performing was LOF with an F1-score of 0.860, and accuracy, precision, and recall of 0.920, 0.880, and 0.850 respectively. This

enhanced performance is due to the ability of LOF to leverage local density fluctuations, which are common for high-frequency drone telemetry data under unusual maneuvers or signal disturbances.

Elliptic Envelope was succeeded by an F1-score of 0.750, with comparatively balanced performance yet lower sensitivity to anomaly boundaries, as indicated by its confusion bias of 0.160, which was greater. In comparison, Isolation Forest and One-Class SVM were reasonably less effective, with F1-scores of 0.700 and 0.720, respectively. These trends are supported by the confusion matrices, where LOF produced very low false positive and false negative rates, whereas the other models suffered from higher misclassification rates, especially in classifying borderline anomaly cases. Notably, Isolation Forest recorded the highest confusion bias at 0.200, highlighting its tendency to overflag normal samples as anomalies, a critical drawback for real-time applications that demand reliability and low false alarm rates.

To facilitate interpretability and comparative analysis, a series of visualizations were generated and are presented as follows. These visual tools are essential for visually validating the quantitative insights discussed previously and highlighting the distinct behavioral patterns of each model.

Figure 1 shows the confusion matrix of the Local Outlier Factor method, visually affirming its performance in distinguishing between normal and anomalous instances. The matrix highlights a low number of false positives and false negatives, clearly reflecting LOF's precision and recall balance.

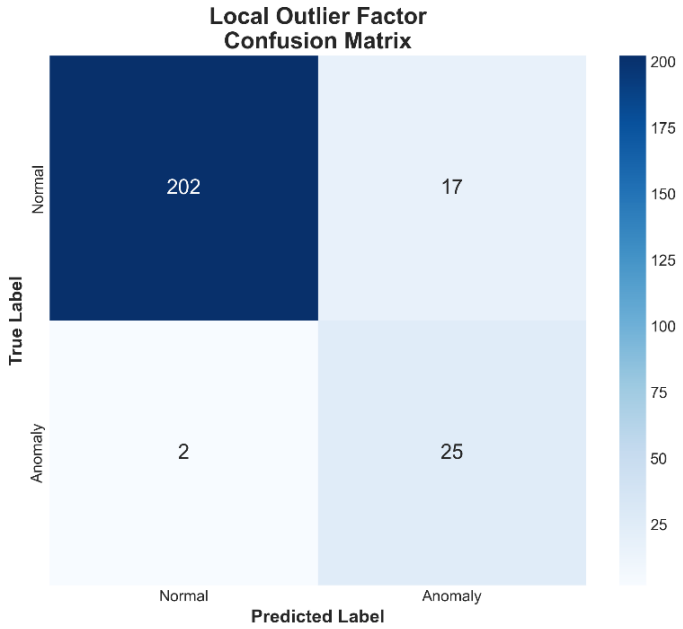

Figure 1. Confusion matrix for Local Outlier Factor.

Following this, Figure 2 presents a grouped bar chart comparing performance metrics—accuracy, precision, recall, and F1-score—across all four methods. This side-by-side layout allows for easy identification of which metric contributes most to a model's overall performance. Notably, the bar chart reinforces LOF's superiority across all metrics.

Finally, Figure 3 illustrates the same evaluation metrics in a radar chart format. This visualization enables quick visual comparison of balance and spread across all metrics. The radar shape for LOF appears the most symmetric and expanded, suggesting a robust and uniformly strong performance profile. In contrast, the other models show noticeable performance drops along at least one metric axis.

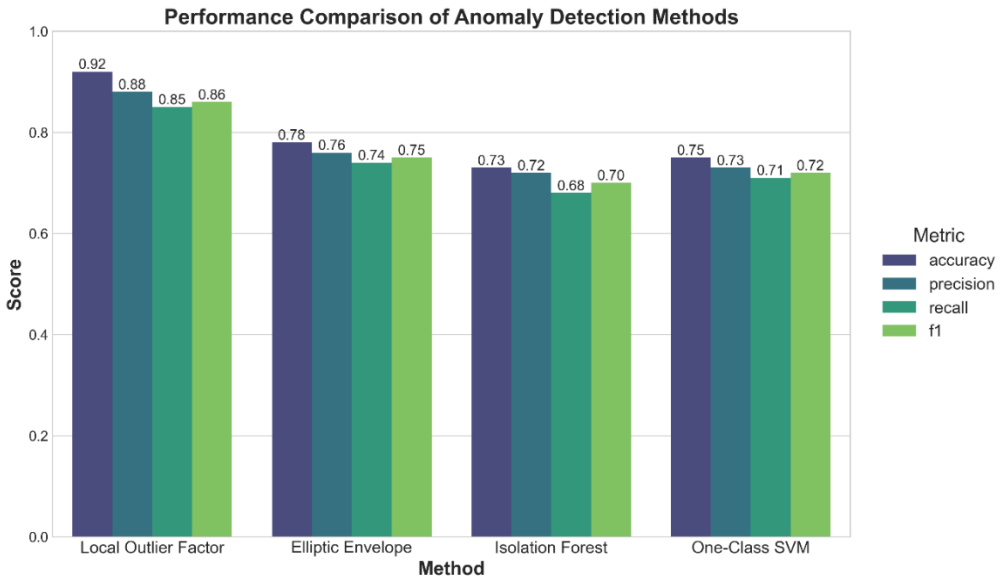

Figure 2. Bar chart of accuracy, precision, recall, and F1-score for all methods.

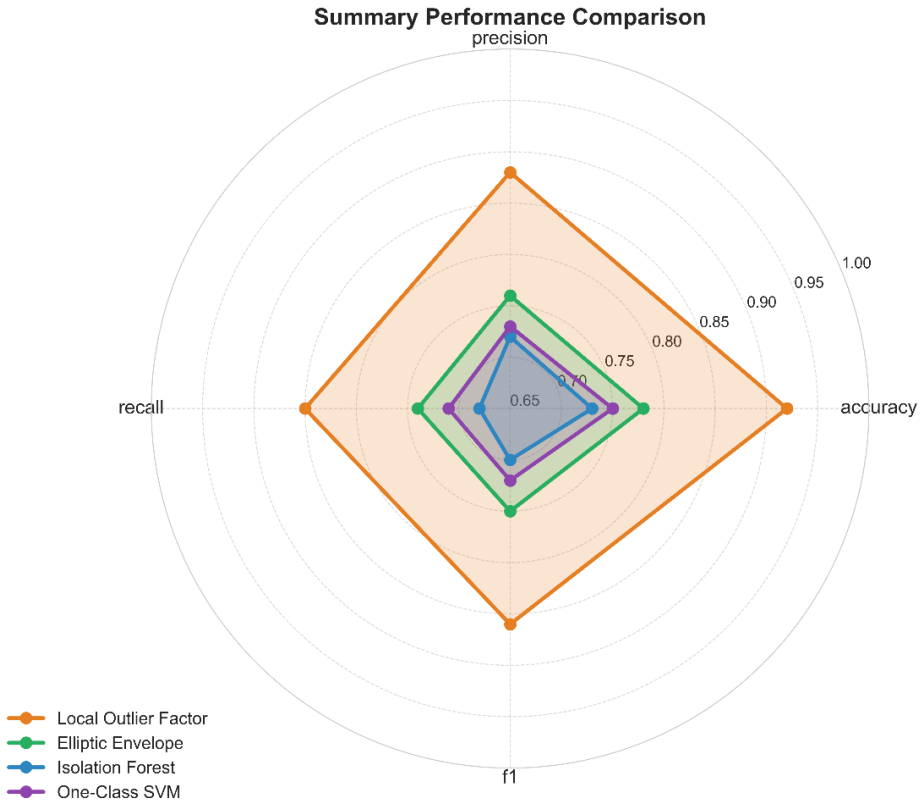

Figure 3. Radar chart comparison of all methods across key metrics.

Beyond metric comparisons, these results underscore the critical importance of aligning algorithmic selection with the specific operational demands and constraints of drone anomaly detection. In practical aerial operations,

false positives can lead to unnecessary mission interruptions, while false negatives risk overlooking critical safety failures. Thus, selecting a model that minimizes both types of errors is crucial. In this context, while Isolation Forest and One-Class SVM are widely regarded for their simplicity and ability to generalize across datasets, our findings reveal that these advantages are offset by their elevated confusion biases and greater rate of false alarms. For applications involving real-time, high-frequency telemetry data, such noise sensitivity may compromise mission integrity, especially when anomaly detection serves as a trigger for automated safety responses or operator alerts.

Elliptic Envelope, which assumes Gaussian distribution across multivariate features, struggles in the face of real-world drone telemetry that exhibits complex, non-Gaussian dynamics. Its performance, while consistent in lab conditions or controlled datasets, is often compromised when deployed in naturalistic flight scenarios with environmental disturbances and sensor irregularities. These shortcomings are particularly evident in its higher confusion bias and decreased recall, reflecting a limited capacity to generalize across irregular data segments. By contrast, the Local Outlier Factor algorithm capitalizes on local density estimation, making it inherently more adaptive to heterogeneous feature patterns and abrupt trajectory changes. This adaptability is especially beneficial in UAV anomaly detection, where anomalies are often localized and context-sensitive, such as a sudden deviation in roll angle or GPS dropout.

Taken together, the experimental findings validate not only the performance hierarchy among these algorithms but also the methodological advantage of employing a hybrid framework—unsupervised learning for training, complemented by supervised evaluation against domain-labeled anomalies. This balanced approach enables models to learn patterns independently while still being held to rigorous assessment standards. The clear and consistent superiority of LOF across all evaluation metrics reinforces its practical utility and sets a strong precedent for its integration into real-time drone health monitoring pipelines. Future directions may include expanding this framework to incorporate ensemble strategies or hybridizing with deep learning-based models that can further enhance anomaly localization and predictive capabilities under complex flight conditions.

5. Conclusion

This study explored a holistic and automated unsupervised machine learning anomaly detection system for Veer UAV telemetry systems through expert-guided validation. With a hybrid approach—unsupervised training coupled with supervised testing—the method ensures that models have flexibility in learning patterns from unlabeled datasets and, at the same time, are exhaustively tested on real expectations. Such balance is crucial in safety-critical drone flights where both false positives and false negatives lead to severe mission-level consequences.

Of the methods tested, the Local Outlier Factor algorithm proved most consistent and strongest under all performance metrics. Its ability to identify outliers in local density is especially suited to the non-linear, context-dependent nature of drone telemetry. Conversely, approaches that presume global distribution patterns, like the Elliptic Envelope and Isolation Forest algorithms, were more susceptible to confusion biases and less sensitive when subjected to adaptive telemetry environments. This performance gap underscores the necessity of adapting anomaly detection techniques to the specific statistical characteristics of UAV data streams.

The value of this research lies beyond mere algorithm selection. It highlights the necessity for feature engineering across domains, maintaining data integrity within time-series subdomains, and exploiting visualization techniques for enhancing model interpretability. By employing AutoML for hyperparameter optimization, the suggested pipeline minimizes manual tuning requirements, thus making it accessible to practitioners lacking deep machine learning expertise. Collectively, these design choices provide a reproducible and scalable pipeline for the integration of anomaly detection within real-world UAV systems.

6. Future Research

Building on the current work, future research is encouraged to investigate the fusion of semi-supervised and self-supervised learning approaches to further enhance the performance of anomaly detection. Semi-supervised approaches, which can leverage limited quantities of labeled data, would increase sensitivity to subtle and context-dependent anomalies. Self-supervised approaches, on the other hand, provide a hopeful path to acquiring generalizable representations from vast volumes of unlabeled telemetry data—enabling more robust anomaly detection for varied UAV types as well as mission profiles.

The other promising direction is the application of deep learning architectures specifically designed for time-series analysis. Recurrent neural networks, LSTM networks, and Transformer-based models are capable of

learning temporal dependencies and long-range correlations in telemetry data efficiently. In conjunction with reconstruction-based models such as variational autoencoders, these techniques may enable the detection of anomalous sequences and system degradation patterns more effectively. Incorporating ensemble learning—combining LOF with neural models—could yield more robust performance across different scenarios.
Real-time deployment also demands attention. For low-latency onboard anomaly detection, models must be edge-computing optimized. This involves pruning, quantization, and architectural reduction to meet the computational constraints of UAV platforms. The inclusion of human-in-the-loop mechanisms, explainable AI, and uncertainty-aware predictions will be required to develop the trust and operational transparency of anomaly notifications. In total, future systems must be developed to operate within overall UAV autonomy systems, communicating with mission planning and control algorithms to facilitate adaptive behavior, e.g., rerouting, emergency landing, or automated diagnostics in the field.

Acknowledgement

We would like to express our sincere gratitude to Siemens A.S for their valuable support throughout the completion of this study. This work also benefited from the use of the Hypersense platform and its dataset, which were instrumental in the development and evaluation of the proposed system.

References

[1] Idrees R, Maiti A, Garg S. A clustering algorithm for detecting differential deviations in the multivariate time-series IoT data based on sensor relationship. Knowl Inf Syst 2024; 67: 2641-2690.
[2] Chen Z, Li Z, Huang J, Liu S, Long H. An effective method for anomaly detection in industrial Internet of Things using XGBoost and LSTM. Sci Rep 2024; 14: 1-23.
[3] Canonico R, Esposito G, Navarro A, Romano SP, Sperlí G, Vignali A. An anomaly-based approach for cyber-physical threat detection using network and sensor data. Comput Commun 2025; 234: 1-14.
[4] Kuchar K, Fujdiak R. Analyzing anomalies in industrial networks: A data-driven approach to enhance security in manufacturing processes. Comput Secur 2025; 153: 1-15.
[5] Kumar D, Agraharam PC, Liu Y, Namilae S. Anomaly detection for composite manufacturing using AI models. J Intell Manuf 2024; 1-17.
[6] Chung J, Shen B, Kong ZJ. Anomaly detection in additive manufacturing processes using supervised classification with imbalanced sensor data based on generative adversarial network. J Intell Manuf 2024; 35: 2387-2406.
[7] Engbers H, Freitag M. Automated model selection for multivariate anomaly detection in manufacturing systems. J Intell Manuf 2024; 1-19.
[8] Sezgin A, Boyacı A. AID4I: An Intrusion Detection Framework for Industrial Internet of Things Using Automated Machine Learning. Comput Mater Continua 2023; 76(2): 2121-2143.
[9] Singh A, Rathore H. Advancing connected vehicle security through real-time sensor anomaly detection and recovery. Veh Commun 2025; 52: 1-11.
[10] Lu Y, Yang T, Zhao C, Chen W, Zeng R. A swarm anomaly detection model for IoT UAVs based on a multi-modal denoising autoencoder and federated learning. Comput Ind Eng 2024; 196: 1-22.
[11] Li T, Lin W, Ma R, Ma Z, Shen Y, Ma J. CoDetect: Cooperative Anomaly Detection with Privacy Protection Towards UAV Swarm. Sci China Inf Sci 2024; 67: 1-2.
[12] Alzahrani MY. Enhancing Drone Security Through Multi-Sensor Anomaly Detection and Machine Learning. SN Comput Sci 2024; 5: 1-10.
[13] Yang L, Li S, Li C, Zhu C, Zhang A, Liang G. Data-driven unsupervised anomaly detection and recovery of unmanned aerial vehicle flight data based on spatiotemporal correlation. Sci China Technol Sci 2023; 66: 1304-1316.
[14] Ozkat EC. Vibration data-driven anomaly detection in UAVs: A deep learning approach. Eng Sci Technol Int J 2024; 54: 1-11.
[15] Ahn H, Chung S. Deep learning-based anomaly detection for individual drone vehicles performing swarm missions. Expert Syst Appl 2024; 244: 1-14.
[16] Sezgin A. Scenario-Driven Evaluation of Autonomous Agents: Integrating Large Language Model for UAV Mission Reliability. Drones 2025; 9(3): 1-21.
[17] Malviya VK, Minn W, Shar LK, Jiang L. Fuzzing drones for anomaly detection: A systematic literature review. Comput Secur 2025; 148: 1-14.
[18] Sezgin A, Boyacı A. Securing the Skies: Exploring Privacy and Security Challenges in Internet of Drones. In: 10th Int Conf on Recent Advances in Air and Space Technologies (RAST); 2023; Istanbul, Türkiye.

[19] Liu D, Wang N, Guo K, Wang B. Ensemble Transfer Learning Based Cross-Domain UAV Actuator Fault Detection. IEEE Sens J 2023; 23(4): 16363-16372.
[20] Yoo JD, Kim GM, Song MG, Kim HK. MeNU: Memorizing normality for UAV anomaly detection with a few sensor values. Comput Secur 2025; 150: 1-15.
[21] Wu Y, Liu L, Yu Y, Chen G, Hu J. Online ensemble learning-based anomaly detection for IoT systems. Appl Soft Comput 2025; 173: 1-12.
[22] Schaller M, Kruse M, Ortega A, Lindauer M, Rosenhahn B. AutoML for multi-class anomaly compensation of sensor drift. Meas 2025; 250: 1-14.
[23] Vajda DL, Do TV, Bérczes T, Farkas K. Machine learning-based real-time anomaly detection using data pre-processing in the telemetry of server farms. Sci Rep 2024; 14: 1-22.
[24] Abdullah RY, Posonia AM, Nisha UB. An Enhanced Anomaly Forecasting in Distributed Wireless Sensor Network Using Fuzzy Model. Int J Fuzzy Syst 2022; 24: 3327-3347.
[25] Wei X, Xu Y, Zhang H, Sun C, Li X, Huang F, Ma J. Sensor attack online classification for UAVs using machine learning. Comput Secur 2025; 150: 1-18.
[26] Sharma T, Balyan A, Singh AK. Machine Learning-Based Energy Optimization and Anomaly Detection for Heterogeneous Wireless Sensor Network. SN Comput Sci 2024; 5: 1-16.