Anomaly and Bias Detection Techniques: Comprehensive Review and Taxonomy

Pal Amutha K[1], Shanmuganeethi V [2]
[1]Centre for Development of Advanced Computing, Chennai, India
[2]National Institute of Technical Teachers Training and Research, Chennai, India

Abstract—**At present, anomaly and bias detection have emerged as two of the most crucial problems in state-of-the-art Artificial Intelligence (AI) /Machine Learning (ML) systems. This has become more important today because there are already many new AI-based systems being used in certain fields of life, such as: health care, finance, cyber security, autonomous vehicles and smart infrastructure. Both anomaly detection and bias detection constitute areas of growing importance in AI systems deployed to high-stakes domains, where reliability, safety, and fairness are crucial. Both anomalies and biases can signal deeper problems, a data quality issue, change in distribution, or inequitable model behavior but despite such commonalities their methods of origin have similar statistical background, methods for representation learning and even procedures of model distribution. In this paper we present a com-prehensive overview of anomaly detection and bias detection methodologies and techniques employed in AI systems. As this survey takes deep dives into both anomaly detection and bias detection, it gives an extensive view on these closely related fields in parallel. A combination of conventional statistical approaches with modern Deep Learning (DL) insights, along with hybrid models that collect aspects of physics, and different framework for designing, analyzing and operating AI systems the paper presents a unified perspective that clarifies the relationships, methodologies, and challenges across these domains. It also highlights the gaps of past research, identifies new avenues for innovation, and sets out best practices to make sure AI is reliable, fair, and transparent. In doing so this survey will be useful to policy-makers, researchers or people who are working in practical positions and want the transparency of AI's near-term future.**

***Keywords*— Anomaly Detection, Bias Detection, Framework, Artificial Intelligence, Machine Learning, Internet of Things (IoT)**

I. INTRODUCTION

AI and ML are the foundational keystones for the next generation where digital ecosystems have advanced to be completely self-driving, interconnected environments. These systems shape their environment, choose actions for effects and drive the optimization of a wide variety of tasks in many fields. But the expansion in the range of tasks now transferred to these systems has made us question their trustworthiness, moral underpinnings and accountability. This kind of exceptional pattern in the data is considered an important clue to critical times for something to happen. For example, it may be fraud in financial systems, equipment failure in industrial applications, deterioration of health conditions for patients in a medical setting, security incidents such as breaches on enterprise networks or the new trend on consumer plat-forms of patterns change in behavior. Precisely detecting anomalies is crucial for both a system's continued reliability and its possibility to provide early warnings which may prevent disasters from occurring. Now-a-days, the developments of anomalous detection can range from statistical, ML models to advanced DL architectures that are able to discern tiny and/or multi-dimensional anomalies in vast complex data. Anomalies and biases are a major challenge for contemporary AI/ML. Their impact can be felt in critical characteristics: fairness, transparency and ethical acceptability. There are several causes of bias. They range from unrepresentative training data to poor decisions on model design or feedback loops that form in production infrastructures. If left unchallenged, these biases will lead to unfair advantage for groups or, at worst, result in tremendous inequalities that will come to plague society; importantly they affect people's desire for AI-based decision making because people feel increasingly suspicious of the fairness they might receive. To address these ethical issues, we must use bias detection and reduction methods for example data preprocessing to rebalance or increase volume, algorithms that incorporate principles of fairness in ML and causal modelling such as counterfactual reasoning techniques. Also, methods for examining how AI systems operate after the fact are now important to ensure that they are doing so in a fair and accountable manner.

The study is conducted to enhance the anomaly detection in video data along with studying the impact of feature reduction techniques on model's performance (accuracy and efficiency) using UFC crime dataset. It is an attempt to tackle the problem by tailoring datasets down a corridor between data dimensionality reduction and detection accuracy, overcoming issues of growing complexity and volume in the data, as well as suggesting potential improvements for the anomaly detection literature for applications across domains [1]. The present-day anomaly and bias problems are affected by many different aspects. Varying modes of operation, characteristics of data itself as well as the capabilities of systems and how users behave com-bine together to produce systems that are fragile. For instance, in finance, anomalies (or "outliers") might signal fraud and biases could lead to credit ratings which inadvertently tilt the playing field for loan applications; in healthcare normal variations from typical parameters may indicate patient deterioration but likewise abnormal changes could denote faults in devices. Indeed, it even prejudices diagnosis across an age range. Anomaly detection is also a key tool for transaction monitoring, while bias detection in retail and e-commerce settings is increasingly important. The goal of the study is to present Parameter-Sharing Graph Deviation Network with Meta-Learning (PS-GDNML), a meta-learning-based graph deviation network designed for addressing the problem of limited labeled data on multivariate time series with enhanced anomaly detection. Specifically, we emphasize efficient adaptation across a variety of detection tasks, optimal meta-learning algorithms, the construction of related graphs and the scoring of anomalies, as well as conduction of sufficient experiments on public datasets to verify our performance gains for different domains

[2]. In addition, urban environments, IoT devices and transportation networks all pose addition-al challenges through scarce sensor data feedback or biased resource allocation Net-work Allocation Vector (NAV) models. Different domains such as media, scientific research, manufacturing, telecommunications and environment monitoring, each reflect unique patterns of anomalies or bias as determined by their own data modes and operational dictates. The brisk and complex nature of these challenges means that simply understanding techniques, technologies and tools available to identify, and especially to mitigate, anomaly and other types of bias is increasingly urgent. We present the following in this survey paper: 1) Analyzing how anomalies and biases spread through fields specific to cross-domain; Evaluating their impacts and the circumstances that led to their emergence. 2) Features new models, including statistics-based model, classical ML algorithms and DL methods with its explainable deep models, fairness-aware training steps and hybrid/domain applications. 3) Evaluates the strengths and weaknesses of each approach, as well as their practical applications in real-world scenarios. 4) Focuses on promising new research directions in terms of how AI systems can be made more accurate, fairer and more transparent, and equitably managed.

II. Impacts of Anomalies and Biases

Anomaly detection and bias detection are important for ensuring fairness, interpretability, and robustness in AI systems, particularly in specialized domains such as Natural Language Processing (NLP) and Large Language Models (LLM). These are designed to expose systematic biases in model predictions or internal states, which would presumably directly reflect underlying biases in the training data. These sorts of biases demographic, language-bias, ideological bias or context-bias are often an artifact of undesired correlation or bias present in the dataset from which they were learned. Recent work in this area also seeks to understand why such biases exist, how they get transmitted up layers or networks of models, and the likelihood of being detected and mitigated effectively. In practical domains, such as healthcare or system monitoring (for e.g., cyber and finance), the detection and characterization of anomalies or bias has great influence. In the medical field such methods help to early discover anomalies in clinical data, protect patient outcome by detecting ab-normal behavior in a health monitoring system. In public health cybersecurity utilities (e.g., wi-fi systems) the main problem is finding and alerting of abnormal behavior in public IoT augmented infrastructure for public health, triggered by cyberattacks, disinformation or system sabotage. But even when domain-specific data and counterfactuals change, the overall goals remain largely the same: pinpointing irregular trends; minimizing meaningless alarms; and fighting algorithmic prejudices. These is work moving forward to ensure that we can rely upon (and trust) decisions made by smart systems in the context of being based on something other than biased data or dubious assumptions.

For example, in the IoT-interconnected systems such as power grid, vehicular net-works and smart city systems, the anomaly detection functions are not only crucial to the security of system but also, we need it to support the resilience of the system and operational efficiency. Smart water and electricity distribution systems use ab-normality detection to identify out of ordinary consumption behavior and nonstandard reading from meters and deflection elimination to prevent being highly dependent on one senor's data or a particular location. Fog-enabled Internet of Vehicles (IoV) networks focus on identifying malicious activities such as anomalous packets while minimizing feature-based bias so that the accuracy of threat detection is not compromised Smart water infrastructure should be able to manage noisy, incomplete or unstructured data so bias mitigation is essential to guarantee a proper environ-mental condition inference. Similarly, wireless IoT networks require effective mechanisms for “go/no-go” determination of whether transmitters are spoofed, and bias mitigation is needed to achieve well-balanced compromises between fairness and reliable detection. In road monitoring systems the detection of anomalies and bias is as important as it provides a more correct estimation for cracks, potholes, wear out of surface etc. Environmental changes, nonuniform sensor coverage or imbalanced data representation in different locations may carry with them these biases. Over-coming those bias, better infrastructure equity is promoted and road maintenance decisions are justified based on well considered unbiased presence observations. Finally, these are also critical in guaranteeing a more secure transport network and the sustainability of smart mobility projects in the long run. Overall, adding cutting-edge anomaly detection and bias mitigation to so many domains’ results in systems that are more trustful or at least safe while they take decisions fairly. These are crucial steps in helping create ethical, transparent and resilient societies underpinned by AI systems that respond to the fast-evolving demands of the digital age. An AI-supported framework for CBT registration with anomaly detection and trusted data prioritization. With the introduction of facial recognition, real-time ID verification and hybrid biometric fusion the solution brings increased levels of accuracy, security and scalability combined with far less human error. It also utilizes blockchain and life-long learning methodology to learn and reduce bias fraud activities, providing secure transparent registration procedures [3].

III. TECHNIQUES AND FRAMEWORKS FOR ANOMALY AND BIAS DETECTION

This section, an extensive summary was made of all methods used to detect anomalies and biases within widely different application areas. Anomaly detection methods for detecting deviations from typical patterns in the data, which generally signal errors, frauds or system failures techniques that are commonly employed in this area include models based on DL methods, statistical approaches, graph-based methods and so on. Each cater toward differing types of irregularities found within both structured and unstructured data. Bias detection, on the other hand, attempts to identify systematic unfairness latent in datasets or model predictions with respect to a sensitive attribute like gender, language or ethnicity. Methods to detect bias generally include the examination of data distributions, the assessment of model behaviors, and the analysis of discrepancies in outcomes for identifying what causes unequal patterns to occur. Such tools include fairness metrics, interpretability/explanability tools, causal inference techniques as well as end-to-end model auditing frameworks.

AI is increasingly deployed in mission-critical domains and as a result, there is an increasing interest in methods that can provide sound, principled and scalable detection. New technologies like federated learning and distributed analytics have great potential in their ability to help facilitate secure real-time detection of abnormal events and biases, supporting responsible AI deployments in sectors ranging from Health (healthcare), Wealth (finance), Secure ties (cybersecurity) and Power (smart infrastructure). A taxonomy and platform for anomaly detection: Selection and tuning of multiple methods for various applications. This platform is designed for the acquisition, processing and visualization of multi-channels data not only allows to test and compare different detection methods, but also offers a number of datasets in different fields on which models can be trained and evaluated [4].

A. Finance

Several methodologies have been developed and used to detect anomalies and biases in financial data. These approaches focus on flagging anomalies that could be indicative of fraudulent transactions, erroneous behavior or organizational bias in the decision-making process. By combining statistical analysis, ML models and domain-specific heuristics, researchers and practitioners are aiming to improve the accuracy and fairness of financial data analysis which in turn would lead to a better risk management compliance and operation cost reduction [5]. Classical approaches Classical methods both supervised and unsupervised learning (e.g., Autoencoder, Graph Neural Network GNNs) are also responsible for fraud detection, cyber security threat discovery. Group anomaly detection focuses on detecting abnormal groups of data points, and network-based approaches investigate the relationships with graph representation [6]. Text-based anomaly and bias detection utilizes DL methods like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) in combination with heavy pre-processing, to analyze structureless data more efficiently [7]. The abnormal data trends can be identified by the measures including Z-Score and Coefficient of Variation used inside the statistical models. Ensemble based techniques like Isolation Forest and One-Class Support Vector Machine (SVM) proves to be frequently used to rank as well as detect anomalies from payment systems [8]. Furthermore, new methods (federated learning, privacy-preserving models and real-time monitoring) are emerging to improve the accuracy and fairness of detection mechanism for fraud issues, opening a window for ethical and optimal financial supervision [9].

B. NLP

NLP and LLMs use a range of approaches and methodologies to diagnose and quantify systemic biases in model behavior. Predictive bias frameworks use tunable metrics such as outcome disparity, error disparity to characterize the gap between model predictions for different groups [10]. Quantitative statistical methods such as regression analysis and effect size estimation can identify the impact of biased-relevant variables [11]. Bias in Large Language Models, including political bias prediction and text continuation tasks. It studies the origin of such biases and assesses their existence, considering mitigation strategies, including prompt engineering and fine-tuning to shed light on their effectiveness. The results contribute to a better understanding of bias propagation in LLMs, and emphasize the need for building fairer and robust AI models. [12]. Furthermore, prompt-based and context-sensitive methodologies such as Bias-Aware Natural Language Classifier, Argumentation-Based Text Analysis zero-shot of ideological, informational biases in few- shot learning where also Bias Tendency Indexes are used to measure the political position in text [13]. When examined across topics and layers, we find that these biases present themselves differently in context, such as models having distinct ideological leanings relative to the structure or domain of prompts [14]. These technologies also have practical applications, revealing biases in part of speech tagging, sentiment analysis, image captioning and summarization. Taken together, these various approaches make up a holistic pipeline for us to detect and interpret the bias in all stages of NLP modeling [15], thereby contributing to fairer and more transparent language systems.

C. Healthcare

In health care imaging, unsupervised models such as Variational Autoencoders (VAEs) are extended to health care use cases where chest X-rays are analyzed by learning normal patterns and detecting anomalies such as pneumonia and nodules without labeled data [16]. In the domain of healthcare fraud detection, supervised approaches such as logistic regression, decision trees and SVM have been used to categorize insurance claims with adaptive learning and real-time analytics application that improves effectiveness by increasing accuracy and reducing human effort [17]. In manufacturing and surveillance applications, methods including predictive analytics, Model Order Reduction combined with Artificial Neural Networks and Swarm Learning can be utilized to detect anomalies in sophisticated plant data for online fault diagnosis under ethical and privacy constraints as [18]. In the field of public health IoT cyber security, a hybrid convolutional Neural Network, Bidirectional Long Short-Term Memory (CNN-BiLSTM) model were employed to analyze network traffic to take out spatiotemporal features and improved real time intrusion detection capability [19]. Collectively these various methods leverage the hierarchical learning capabilities of DL to learn complex and evolving patterns of anomalies, frauds and cyber threats across systems in multiple markets with flexible accuracy [20].

D. Industrial IoT

In the area of Industrial Internet of Things (IIoT) and Industry 4.0, various complementary AI-based solutions have been proposed to improve anomaly detection with fairness considerations, energy efficiency and scalability [21]. They are using DL models, including CNNs, LSTMs, Generative Adversarial Networks (GAN) and SVMs to enhance real-time detection precision in uses as diverse as visual quality control and sensor data analysis. Such technologies solve the problem of imbalanced and biased datasets using bias-aware training techniques as well as data augmentation [22]. Moreover, energy-efficient architectures are built to improve resource utilization in industrial data pipelines [23]. While the approaches from vision-based defect detection systems to distributed bias correction in IoT networks differ, their shared ambition is to construct trusted, balanced and effective anomaly-detection systems adapted for modern industrial settings [24].

E. Road Surface

The road surface anomaly detection spans a large number of state-of-the-art technologies, all designed to enhance accuracy, efficiency and scalability. DL network models such as Local Region CNN, YOLOv8 are based on vision-based methods and they successfully identify the surface damages like cracks, potholes etc., with high accuracy and efficient speed by employing local region segmentation, sliding window detection & dynamic cropping [25]. Sensor oriented solutions Sensor-based methods make use of crowdsourcing smartphone data and other smart city devices: accelerometers, gyroscopes, GPS, and real time connectivity to keep track on the condition of roads all the time [26]. ML classifiers like XGBoost, Random Forest (RF), SVM and CatBoost analyze these sensor signals to identify anomalies with high accuracy [27]. In addition, low-cost embedded systems, such as the SVRUM platform, which utilize a variety of sensors in combination with Arduino - based hardware at low cost can offer added benefit and provide accurate measurement to vulnerable road user groups such as cyclist/ e-bikers [28]. Altogether, such spectrum of technological advancements (ranging from vision-related DL approaches to sensor fusion and target hardware) provides a broad suite of tools for efficient real-time recognition of road anomalies in various environments [29].

F. Smart city

The techniques used to detect such anomalies in a variety of IoT settings range from conventional statistical approaches to novel DL tools developed for application to intelligent system problems. In multi-source water–electricity systems, statistical methods, such as the Distance Correlation Coefficient with DBSCAN clustering and t-Copula distributions can be used for anomaly detection on joint patterns of consumption [30]. Fog based IoV networks consist of deep leaning techniques such as convolutional autoencoders to recognize unknown attacks by learning squeezed feature set to find out the reconstruction's errors [31]. It is worth noting that the Smart water systems rely on information theoretic clustering combined with modified Hampel loss functions to reduce False Alarms and enhancement detection robustness, especially in noisy environments [32]. On the other hand, zero-bias IoT device localization systems deploy bias-mitigated dense neural layers with cosine similarity as feature functions to classify physical-layer signals and early detect spoofed/unseen devices [33]. In aggregate, these methods combine the concepts of statistic correlation with robust clustering, deep representation learning and bias-resistant neural network models that addresses the call for more interpretable while accurate and fair anomaly detection mechanisms in the context of IoT ecosystems.

IV. DETECTION METHODS: SECTORAL ANALYSIS

Hybrid ML and DL models to increase the accuracy of fraud detection and de-crease false positives. For example, it is discovered that graph-based model is particularly suitable for complex tax evasion and trade fraud or illegal cryptocurrency related activities. Regular ML techniques like SVM and Random Forest contribute to credit card fraud detection, anti-money laundering and insider trading prediction. In healthcare, unlabeled models like VAEs are used for anomaly detection in the case of little labelled images. Hybrid CNN–BiLSTM architectures can accurately predict threats in IoT systems, and AI-enabled fraud detection models yield better performance than their rule-based counterparts.

AI also enhanced environmental monitoring with the real-time and scalable anomaly detection. Industrial IoT uses CNN–LSTM hybrids for defects detection from spatiotemporal images while GANs-based methods are applied to deal with imbalanced or biased datasets. A two-layer privacy concerned trust evaluation scheme for VANETs with enhanced vehicle anomaly detection and notification. It incorporates game-theoretic incentive mechanisms to resolve privacy leakage, employs a Barycentric Lagrange interpolation-based algorithm for Nash equilibrium approximation and designs decision making as an MDP via the reward-shaping D3QN [34]. AI platforms provide network traffic, system failures and cyber threats monitoring in real time, increasing overall trust ability in the IoT-based industries. Smart cities are transitioning from rule-based systems to adaptive, accountable DL for utilities, transportation and security. GNNs can be used for modeling the relation-ships between IoT devices to perform predictive maintenance as well as clustering methods (DBSCAN, Mutual Information) are applied to analyze behavior in critical infrastructures. Vision based, CNNs, YOLOv8, dynamic cropping and smartphone-based ML classifiers can lead to fast defect detection, crowdsourced monitoring and as a result in cost-effective road and traffic management. NLP methods mitigate demographic, media and political bias in text data and language models. Over the last couple of years several studies have pointed out both imbalances in available datasets like those used for sentiment analysis and text classifications tasks and in word embedding models such as Word2Vecand GloVe, which reflect regional biases involving race/gender. More broadly, NLP techniques are being applied to study ideo-logical and emotional bias in large language models like beyond GPT-4. Table 1 provides sector-wise anomaly and bias detection methods along with key observations.

TABLE I
SECTOR-WISE ANOMALY & BIAS DETECTION METHODS & OBSERVATIONS

Sector	Core Methodology	Observations
Healthcare		
Medical Imaging (X-ray Anomalies)	Variational Autoencoder (VAE)	High accuracy in detecting structured anomalies; unsupervised model for dealing with small number of labeled instances effectively.
Healthcare Fraud Detection	Regression, Decision Trees, Ensemble Learning	Performs better than traditional rule-based and manual systems; catches evolving fraud patterns with an adaptive model.
Industrial & Environmental Monitoring	AI-enhanced MOR and Predictive Analytics	It leads to an improvement of computational cost, real-time performance and scalable anomaly detection.
IoT-based Cybersecurity	CNN–BiLSTM Hybrid Model	Obtains 98–99% detection accuracy; Models space-time correlations and reduces false positives.
Finance		
Finance & Banking	AD, GAD, SVM, Random Forest, Logistic Regression, DBN , One-Class SVM	Fraud detection, insider trading, money laundering, AML. Used in credit card, loan approval and AML systems.
Tax Evasion	Graph Theory , ML (T-EGAT, H-GCN)	Finds high-risk taxpayers and trade fraud. Specializes in relational fraud through transaction networks.
Customs Fraud	AI-driven Inspection Systems (DATE, for instance Autoencoders)	Applied in security check, customs and border inspections.
Cryptocurrency / Blockchain	Graph Convolutional Networks (GCNs)	Discovery of money laundering, deanonymization, tracing dark wallet.
NLP		

Text Classification (POS / Sentiment)	Selection & Label Bias Detection, Sampling diagnostics, outcome disparity metrics	Under representativeness causes demographic bias in Twitter and Newswire.
Word Embedding Models	Semantic Bias Detection, WEAT (Word Embedding Association Test), embedding correlation analysis	Persistent gender and racial stereotypes in GloVe, word2vec, BERT.
Media Bias Detection (LLMs)	Lexical & Informational Bias Detection, Prompt-based contextual detection (BANC, ABTA, EBTA)	GPT-4 stronger on informational bias; weaker on lexical bias.
Political Ideology Classification	Ideological Bias Measurement, Bias Tendency Index (BTI), embedding similarity analysis	LLMs continue to display the left/right skew across prompt wording.
Emotion & Coreference Resolution	Model Bias Detection, Regression models with effect size analysis	There is still bias after re-sampling datasets such as EEC and GAP.
Mental Health NLP Applications	Selection Bias in Classification, Cross-demographic error analysis	Models overfit to age/gender-specific features in PTSD and depression datasets.
Cross-Language NLP	Detection of Research Design Bias, Evaluation of Multilingual Fairness	English-centric training reduces inclusivity on XNLI, TyDiQA.
IIoT		
Visual Quality Control (Manufacturing)	CNN and LSTM Hybrid	Learns spatial-temporal characteristics from images data to recognizing product defects.
IIoT anomaly detection	Bias-Aware Collaborative GAN (C-GAN)	Collaborative adversarial learning for addressing imbalanced datasets and bias.
Imbalanced DL (Industrial IoT)	DB-CGAN (Distribution Bias-Aware Collaborative GAN)	Corrects the bias of data distribution through collaborative generator–classifier–discriminator training.
Tactile data classification	Self-Organizing Maps (SOM)	Classifies the IIoT traffic into tactile and nontactile flows according to QoS demands.
Real-time traffic anomaly detection	Support Vector Machine (SVM)	Identifies the presence or patterns of intrusive network activity.
Network control and latency optimization	Software-Defined Networking (SDN)	Dynamically reconfiguring flow tables to minimize the delay and transmit energy.
Fault detection & cybersecurity	GAN-based Data Augmentation	Produces artificial samples for the minority classes to balance training sets.
Smart City		
IoT device relationship modeling	Graph Neural Networks (GNN)	Cross correlation between devices in predictive maintenance and anomaly prediction models of the models.
Multi-Source Water–Electricity	Distance Correlation Coefficient and DBSCAN	Assessment of water-electricity compound behavior by t-Copula distribution.
Fog-Assisted IoVs (CAaDet)	Convolutional Autoencoder (CNN and AE Hybrid)	Learns compact features; anomaly = reconstruction error > the limit.
Smart Water (Noise-Resilient)	Mutual Information Clustering and Modified Hampel Loss	Unsupervised clustering plus robust thresholding using invariants and MAD.
Zero-Bias IoT Device ID	Zero-Bias Dense Layer and Cosine Similarity Metric	No-bias angular similarity-based physical layer signal classification.
Road Surface		
Transportation / Maintenance	Hand-held camera CNN, Local region analysis plus grid sampling	Cost-effective, dataset efficient.
Urban Infrastructure / Smart Cities	YOLOv8, DCS, Dynamic cropping and 2-tier detection	Real-time, lightweight.
Crowdsourced IoT / Community Sensing	Smartphone accelerometer, fuzzy logic, MAD, ML classifiers (RF, XGBoost)	Scalable, community-based
Industrial IoT / Embedded Systems	Raspberry Pi SVM (ADS), FFT, RMS, SVM (RBF kernel)	Real-time cloud IoT integration
Micro Mobility / VRU Safety	Arduino, accelerometer, sonar, Threshold, ML (RF, SVM, K-Means)	Low-cost, lightweight, replicable

V. DOMAIN-SPECIFIC CHALLENGES & SOLUTIONS

In the financial domain, detecting anomaly/bias becomes challenging because of imbalanced nature and heterogeneity in data, new patterns of fraud evolving continuously over time, subtle biases in strategies adopted and privacy issues that need to be adhered to. Solutions are advanced preprocessing, hybrid approaches of statis-tical and DL approaches, ensemble methods on robustness, federated learning on privacy preserving adaptation and explainability tools regarding fairness and regulations. Major obstacles facing implementation of blockchain-enabled supply chain finance risk management such as scalability, compliance with regulation and the importance to integrate smart contracts and secure data sharing. Ongoing assessment and thoughtful strategy development are necessary to overcome these barriers in using blockchain for better transparency and risk management [35]. NLP challenges include labeling and selection bias, exaggeration of subtle biases in training, lexical vs. information-based limits on bias, and the high sensitivity of LLMs to prompts. These concerns are addressed via stratified sampling, Bayesian annotation, fairness-aware/debiasing at the embedding level, ensemble debiasing and MDL-based probing as well as in prompt engineering (e.g., BANC and ABTA) and ethical evaluator framework. RLVAL: A high precision and recall model with reduced false positives and true anomaly resolution. It shows competitive performance on Yahoo and KPI with different proportion of active queries when compared against state-of-the-art methods. Moreover, RLVAL has potential for benchmark dataset creation when there few numbers of labels to train on, comparison with SPOT also suggest that difference in calibration and data splitting strategy [36].

AI is challenged by the scarcity of labeled data, imbalanced class distributions, changing patterns over time, lack of interpretability, high rate of false positives and privacy. The solutions are based on unsupervised or self-supervised learning (e.g., VAEs), SMOTE and bias audits, hybrid CNN–BiLSTM models, Explainable AI (XAI) integration, federated learning, model-order reduction, fusion methods, and lightweight edge AI for realistic deployment. IIoT is challenged by data im-balance, latency, energy inefficiency and manual inspection issues. Thus, serving the former, synthetic data generation (C-GAN, DGAN), SDN-based low-latency architectures, CNN–LSTM inspection models, GAN-based augmentation and light-weight feature extraction techniques are employed to build scalable industrially efficient and reliable AI systems. In a generalized setting of road anomaly detection, we have to deal with constraints such as label scarcity, environmental variations at large, sensor noise as well hardware limitations and generalization. Real time system for integrating objects' detection based on YOLO Advanced model and demonstrates that better hardware leads to the significant decrease in object detection. The study also emphasizes the importance of different road-condition datasets to improve model performance and proposes an autonomous real-time detection system for field deployment [37]. Critical solutions are synthetic and region-based augmentation (YOLOv8, DCS), signal filtering (FFT, Butterworth), fuzzy logics for speed variation, learning edge lightweight models (YOLOv8-S, embedded SVMs), dynamic thresholds and more complex datasets with various vehicles and road conditions. Challenges of the vision-based pothole detection, such as the trade-off between processing speed and on-device computing limitations in edge environment, requirement of power-efficiency and portability solutions for real-time practical application and hardware limitation that affect model performance especially when it is applied to resource-constrained devices adopted in autonomous driving and road status monitoring [38]. The smart city systems come encounter the data bias, node latency, noisy data, class imbalance as well as real-time scalability constrains. These are resolved by multi-source fusion, CAaDet-based distributed detection, robust loss merging/learning, zero-bias layers, MI clustering and lightweight CNNs for truthful, scalable and fair urban monitoring. Table 2 presents domain-specific challenges and solutions.

TABLE II
DOMAIN-SPECIFIC CHALLENGES AND SOLUTIONS

Domain	Challenges	Solutions
Finance	Data imbalance. Heterogeneous data types. Evolving fraud patterns. Subtle biases- Privacy concerns	Advanced preprocessing. Hybrid models (statistical & DL). Ensemble learning for robustness- Adaptive algorithms (e.g., federated learning). Explainability tools to ensure fairness and compliance
NLP	Labeling bias & selection bias. Amplification of subtle bias. Prompt sensitivity in LLMs. Lexical vs. informational bias. Ethical gaps	Stratified sampling, bias documentation. Bayesian annotation models. Fairness-aware training, ensemble debiasing. Embedding-level debiasing (hard/soft). Probing, MDL analysis. Prompt engineering (BANC, ABTA, EBTA)- Qualitative audits and fairness frameworks
Healthcare	Limited labeled data. Data imbalance & bias. Complex evolving patterns. Overfitting & lack of interpretability- Privacy/security. High false positive. Computational complexity	Unsupervised / self-supervised learning (e.g., VAEs). SMOTE, bias audits, fairness-aware learning. CNN–BiLSTM, Model Order Reduction. KL divergence, dropout, XAI integration. Swarm & federated learning. Threshold optimization, fusion models. Lightweight models, pruning, edge AI
Industrial IoT	Data imbalance & bias. Latency, energy inefficiency. Manual inspection limitations. High complexity in modeling. Inefficient oversampling methods	C-GAN, DGAN, DB-CGAN for synthetic data. Tactile-aware SDN, SOMs, binary tree flow mapping. CNN–LSTM for visual inspection. Lightweight feature extraction. GAN-based augmentation for better training data quality
Road Surface	Limited labeled data. Lighting & camera variability. Sensor noise & signal drift. Speed-related signal distortion. Hardware constraints- Transmission errors. Sensor bias. Environmental interference. Dataset generalization issues.	Synthetic data, local-region augmentation (YOLOv8, DCS, etc.). Segmentation, brightness augmentation. Signal filtering (Butterworth, FFT), sliding window std- MAD, fuzzy logic for speed variation. YOLOv8-S, embedded SVMs (Raspberry Pi/Arduino). 4G IoT modules for data continuity. Calibration protocols for uniform sensor data. Dynamic thresholds, feature tuning. Broader datasets with diverse vehicle/terrain types
Smart City / IoT	Single-source data bias. Fog node latency/vulnerability. Noisy/unsafe training data. Class imbalance in device identification. False alarms & missed detections. Data heterogeneity & missing values. Scalability & real-time limits	Multi-source integration (Copula, distance correlation). CAaDet model for distributed detection. Hampel loss, robust thresholding. Zero-bias dense layers. Weighted error minimization, adaptive thresholds- Mutual information clustering , time-invariant features. Lightweight CNNs, compact fingerprint representations

VI. METHODS AND MODELS ACROSS DOMAINS

Advancements in generative models for Time Series Anomaly Detection (TSAD), progress, from classical to deep generative models. It juxtaposes mainstream TSAD approaches, discusses representative works, and emphasize the role of causal inference in making them interpretable. The work focuses on high quality data preprocessing, and advises for domain-specific application of generative models [39]. Ap-plication of ML approach for detection and prediction of anomalies in satellite telemetry. It presents computational methods to detect anomalies in the telemetry data and proposes a first application kernel which carries out detection and prediction tasks. Sophisticated ML techniques are used to improve reliability and operational effectiveness in maintaining the health of critical satellite subsystems [40]. With financial domain, technique differences fall into two main categories, Statistical outlier detection, for example z-scores and Mahalanobis distances DL models like LSTMs and GNNs to recognize patterns examine relationship-based inspection. For example, hybrid and ensemble models are developed, such as: Feedforward Neural Network (FFNN)s and SVMs to improve real-time accuracy. Frameworks that are both integrated and aware of the layout integrate Optical Character Recognition (OCR) free pipelines with transformer model. Secure federated learning and XAI can confront transparent anomaly detection. NLP employs such methods to expose stereotypes in the models' representations as the Predictive Bias Framework, regression and effect sizes from the US Bureau of Statistics analysis and embedding probing. As for zero- or low-shot learning settings, methodologies are introduced that can be used such as prompt-based classification, semantic context augmentation (BANC, ABTA, EBTA). METAGAD, A meta-learning approach of few-shot graph anomaly detection effectively adapts self-supervised node representation to supervised learning mitigating overfitting and enhancing generalization. Experiments on real-world datasets demonstrate that METAGAD well capture anomalies, which further confirms the practical effect of METAGAD and demonstrates the active of complementing the limited labeled anomalies with plenty of unlabeled data [41]. Model-based solutions exist for debiasing of weak learners, Bayesian weighting annotations and debiasing strategies applied to embeddings. Through documenting and giving ethical oversight (and thus performing fairness auditing), there is potential on all counts of the NLP pipeline to argue fairness. To achieve these ends, health systems use VAEs for anomaly detection based upon medical imaging; logistic regression towards fraud and diagnosis with structured data are also followed by logistics Hybrid deep architectures like CNN – BiLSTM bring together spatial and temporal aspects where adaptive learning (e.g., online retraining) ensures that models remain as effective over time. Model order reduction (MOR) reduces orders of magnitude, fair AI frameworks provide transparency and fairness guarantees. The anomaly detection method utilizes the structure information by graph basis. Without the joint learning strategy in GANs, a one-layer network for both domain distinction and anomaly classification are presented with better detection performance on all data sets using an appropriate parameterization and learning rates [42]. IIoT applies CNN-LSTM hybrids and GAN-based synthetic data generation (C-GAN, DGAN, DBCGAN) to identify manufacturing anomalies as well as address data imbalance. Self-Organizing Maps (SOM) and SVMs are also used for real-time traffic monitoring in energy-saving SDN-networks. An interpretable rule-mining approach to real-time ECG anomaly detection with biased-trained ANN. Through a rapid tree-based search for derivable rules and training on the MIT-BIH Arrhythmia Database, the method is optimized for high-accuracy detection of abnormal beats with high sensitivity [43]. By employing computationally efficient training strategies that include adversarial training, feature weight sharing and batch normalization, the model can be made reliable. Some of the visual tools (for example UMAP) to determine the reduction in bias and measure precision, F-measure and false alarm rate. SeqVAE-CNN. Unsupervised model that combines VAEs and CNNs for learning temporal and spatiotemporal features in multivariate time-series. It deals with problems such as imbalanced and complexity data well, and shows the superior anomaly detection performances with the highest AUROC score and F1 score in most datasets. (Deep) One-Class Anomaly Detection: complements the DL based methods for better anomaly detection which are applicable in a wide variety of domains [44].

A generic ML based pipeline for anomaly detection in IoT networks that consists of data preprocessing, feature engineering, model selection and training and deployment using both labeled and unlabeled data with alert mechanism for abnormal behavior prediction [45]. Anomaly detection key approaches are presented in Table 3.

TABLE III.

ANOMALY DETECTION: KEY APPROACH

Key Approaches	Issues/Challenges Addressed	Use Case
Statistical models (Z-score, Mahalanobis distance)	Basic anomaly detection, outlier identification	Fraud detection in banking transactions; sensor outlier detection
DL (LSTM, GNN, CNN–BiLSTM Hybrids)	Capturing temporal, spatial, and relational patterns; subtle anomalies	Time-series fraud detection; network intrusion detection; video surveillance; IoT anomaly detection
Federated Learning	Privacy preservation, secure collaborative learning	Cross-bank fraud detection; multi-hospital medical anomaly detection
Hybrid Models	Combining strengths of different algorithms for accuracy	Combining CNNs and rule-based methods; hybrid visual and sensor data anomaly detection
OCR-free pipelines	Efficient document processing without OCR errors	Automated invoice processing; financial statement analysis
AutoML	Automating model selection and hyperparameter tuning	Rapid prototyping of anomaly detection models
Quantum Computing	Accelerated computation and optimization	Emerging research in fraud detection algorithms
Variational Autoencoders (VAEs)	Unsupervised anomaly detection; limited labeled data	Medical image anomaly detection; industrial sensor fault detection
Model Order Reduction (MOR)	Reducing computational complexity	Large-scale sensor dataset analysis in healthcare or energy grids
Adaptive Learning & Online Retraining	Handling evolving data patterns and concept drift	Fraud detection adapting to new attacks; real-time sensor anomaly monitoring
Logistic Regression & Decision Trees	Transparent classification and anomaly detection	Insurance claim fraud detection; credit risk assessment
Generative Models (C-GAN, DGAN, DB-CGAN)	Addressing data imbalance via synthetic data	Minority class oversampling in fraud detection; sensor data augmentation
Self-Organizing Maps (SOM), SVM	Unsupervised clustering, supervised classification	Industrial IoT traffic anomaly detection; cybersecurity
Energy-aware Software Defined Networking (SDN)	Reducing latency and energy consumption	Real-time industrial network management; IoT communication optimization
Batch Normalization & Feature Weight Sharing	Improving training stability, generalization	DL for financial fraud or healthcare anomaly detection
Low Latency Architectures	Real-time detection in resource-constrained environments	Embedded anomaly detection on edge devices; autonomous vehicle sensor analysis
Local Region CNN & YOLOv8 with Dynamic Cropping	Enhancing localization and sensitivity	Urban pothole detection; manufacturing defect localization
Smartphone Sensor Analytics (MAD, FFT)	Handling noise and signal variability	Crowdsourced road anomaly detection; wearable health monitoring
Embedded ML (Arduino, Raspberry Pi)	Lightweight anomaly detection on edge devices	Low-cost embedded sensor anomaly detection; micro-mobility safety applications

Cloud-hosted SVMs	Scalable centralized anomaly detection	Centralized IoT anomaly detection; cloud-based network security
Real-time GPS Mapping	Spatial visualization and localization	Smart city infrastructure monitoring; fleet vehicle condition tracking
Fuzzy Logic	Handling uncertainty and imprecision	Adaptive sensor fusion; anomaly detection with noisy inputs
CNN-Autoencoder (AE) Hybrids (CAaDet)	Combining reconstruction error and classification	Smart water system anomaly detection; industrial quality inspection
Modified Hampel Loss	Reducing sensitivity to noisy data	Resilient anomaly detection in contaminated sensor data
DBSCAN Clustering	Detecting clusters of anomalies and noise	Group anomaly detection in network logs; unsupervised sensor fault detection
Copula & Distance Correlation	Modeling complex dependencies	Multi-source anomaly detection in smart grids and industrial systems
Unified Multi-Source Architecture	Integrating heterogeneous data streams	Smart city anomaly detection; industrial IoT monitoring
Signal Demodulation	Extracting meaningful features from raw signals	Communication signal anomaly detection; sensor waveform analysis
Mutual Information Clustering	Identifying relevant feature groups; reducing dimensionality	Feature selection in multi-sensor data; network traffic anomaly grouping

Road surface monitoring, on the other hand, employs vision-based solutions such as Local Region CNN and YOLOv8-DCS to precisely detect potholes and cracks economically. Wearable and sensor-based techniques classify smartphone vibrations, accelerometer/gyroscope data in real-time using FFT with the help of a cloud based SVM. Affordable embedded systems such as SVRUM, with Arduino sensors and ML for micro-mobility users). Shared analytical features (RMS, MAD, FFT) across systems allow for cross- system validation and hyper-parameter tuning to ensure accuracy and generalization. A transformer-based unsupervised anomaly detection model for V2X communication, and TV2XFormer its transfer learning-boost version that improves the adaptability. Evaluate both models over the VDoS-LRS V2X data set in terms of precision, recall and F1 score and compare with five state-of-the-art unsupervised algorithms, where ours is achieving the best F1 scores showing good robustness in dynamic VE [46].

A unified workflow is adopted to achieve multi-source data acquisition, preprocessing and correlation modeling for heterogeneity in smart city infrastructure. These models are CNN-AE hybrids, zero-bias DNNs, and noise-robust loss functions (MHLP). Methods, such as DBSCAN, AE reconstruction error and similarity response are efficient for anomaly detection. Results on datasets such as SGCC, NSL-KDD, and ADS-B demonstrate that the model achieves high accuracy and reduces a large number of false positives. This trend of a unified architecture, of combining signal-level modeling, statistical inference and bias free DL, promises the high scalability and flexibility to meet the requirements of any IoT-enabled ecosystem that can lead to secure smart environments. Some of the key approaches for bias detection are shown in Table 4.

TABLE IV.

BIAS DETECTION: KEY APPROACH

Key Approaches	Issues/Challenges Addressed	Use Case
Explainable AI (XAI)	Model transparency, interpretability, fairness, trust building	Regulatory compliance in finance; clinical decision support
Embedding Probing & Prompt-based Bias Detection	Identifying and quantifying encoded biases	Bias detection in NLP models; social bias mitigation in chatbots
Bayesian Annotation Weighting	Handling unreliable or biased annotations	Improving labeling quality in medical imaging datasets
Context Augmentation (BANC, ABTA, EBTA)	Expanding semantic context to improve bias detection	Fairness enhancement in sentiment analysis or news classification
Embedding Debiasing (Hard & Soft)	Mitigating biases in feature representations	Reducing gender/racial bias in word embeddings
Fairness Audits	Systematic fairness and ethical compliance evaluation	AI-driven loan approval auditing
Fair/Ethical AI Frameworks	Ensuring ethical standards, fairness, and transparency	Governance frameworks in finance and healthcare AI
Zero-Bias Deep Neural Networks (DNNs)	Eliminating classifier bias	Fair device identification in IoT; unbiased intrusion detection
Zero-Bias Dense Layer , Cosine Similarity Metric	Removing class bias in deep models	Fair device ID classification in IoT

VII. TECHNICAL TOOLS AND FRAMEWORKS ACROSS DOMAINS

A full set of tools and frameworks are used across heterogeneous areas to solve anomaly detection problems, bias mitigation, etc. In ML/DL frameworks such as TensorFlow, Keras, PyTorch, Scikit-Learn and XGBoost make programming languages very easy to work with for model building, training and hyper-parameter optimization for both supervised and unsupervised tasks. An anomaly-detection predictive analytics methodology for imbalanced datasets comprising, fused with ensembles of data resampling techniques, feature engineering and ML algorithms. It supports three sampling strategies and six different feature selection algorithms to improve predictive capabilities; this turns the model into a flexible architecture that can be easily customized, such as in other evolving patterns, for example online fraud detection [47]. Graph-based and network analysis Graph-based approaches leverage techniques such as Markov Chains, Social Network Analysis and GCNs to identify sophisticated fraud patterns and relational abnormal behavior in areas such as finance systems or decentralized systems. TDSRL, a time series anomaly detection model which uses dual tasks in self-supervised manner learning both time and frequency domain information. The methodology simulates the anomalies using data degradation methods and evaluates it across multimeric tools, including FA-R, R-AUC-ROC, VUS-ROC, classical metrics such as MAE, MSE and MAPE. [48]. An-other class of privacy-preserving platforms such as TensorFlow Federated and PySyft facilitate private federated learning between institutions without requiring them to share confidential data.

In NLP, cutting-edge bias detection and mitigation are implemented with Bayesian annotation models, Inter-Annotator Agreement measures, Word Embedding Association Tests (WEAT), hard and soft debiasing algorithms, context augmentations and fairness loss regularization as well as cross-lingual fairness frameworks for better model interpretability and openness. Interpretability frameworks and attention visualization such as Explainable AI (XAI) tools offer important explanations to the model's predictions, enabling trust and meeting regulatory requirements.

W-HGAD, a heterogeneous graph neural network that learns distributional node embeddings in Wasserstein space. By modeling nodes through Gaussians and by the use of Wasserstein distance, it models uncertainty and maintain graph transitivity at the same time. With the help of structural, attribute, and type-level reconstruction losses, the framework demonstrates superior performance with a great margin compared to state-of-the-arts methods [49]. There are also purpose-built tools, such as Fairlearn, AIF360 that enable easy deployment of these techniques to healthcare systems, alongside big data processing frameworks, which can support high-accuracy anomaly detection in medical images showing cancer or fraud detection in claims processing. IIoT and manufacturing sectors also incorporate TensorFlow, Keras, SOM, SVM and GANs for defect detection, fault diagnosis and data augmentation to resolve class imbalance. In this context, both SDN and binary tree mapping is used in order to cut down the latency and improve network flow in industrial networks. A meta-learning approach to connect dataset properties with algorithm performance. It includes data collection and preprocessing, uses Pearson, Spearman, and Kendall correlation functions for the selection of meta-features and lays down a model search space on top of ASAD framework. The meta-learner is trained to predict the optimal ML model based on dataset characteristics, thereby ensuring systematic and automated ML model selection [50].

For sensor noise reduction and feature extraction, we can utilize signal processing techniques such as (but not limited to) Butterworth filters, FFT, RMS, Cosine Similarity Response while real-time local data processing is achieved by edge computing platforms like Raspberry Pi/Arduino. Geospatial visualizations such as Google Maps and GIS dashboards enable spatial anomaly detection and city-scale infrastructure monitoring. Algorithms such as DBSCAN, KDE, and UMAP, enable unsupervised anomaly detection and distributions analysis. Road anomaly detection method based on vehicle using Haar-like feature classifiers for vehicle detection and HMMs for trajectory modeling. 2D spatiotemporal movement of vehicles are analyzed and compared with the smooth road trajectories to find out that exists rough surface like potholes, speed breakers. The reliability and versatility of this method are verified compared with other visual- and nonvisual–based methods [51].

In the realm of smart city and IoT, copula functions and distance correlation coefficients as well as metrics for mutual information capture non-linear associations be-tween heterogeneous data streams whereas hybrid models like CAaDet leverage a combination of convolutional and autoencoder structures to enable autonomous edge detection. A second type of use case is the one on cybersecurity, where using benchmark datasets like NSL-KDD, UNSW-NB15 and model types such as ensemble learning and SEADer are employed to identify network intrusions and social engineering attacks. Finally, common metrics such as NIST's metric allow robust evaluation of false alarm rates and can deliver confidence in the performance of anomaly detection systems for critical infrastructure. Table 5 presents End-to-End Workflow for Anomaly and Bias Detection.

TABLE IV

END-TO-END WORKFLOW FOR ANOMALY AND BIAS DETECTION

Step	Tools / Frameworks	Purpose / Application
1. Data Acquisition & Preprocessing	TensorFlow, PyTorch, Scikit-Learn, Raspberry Pi, Arduino, Hadoop, Spark	Data collection, cleaning, preprocessing, noise reduction
2. Feature Extraction & Signal Processing	FFT, RMS, MAD, Butterworth Filters, Cosine Similarity	Extract meaningful features from raw sensor/signals
3. Exploratory Data Analysis & Clustering	DBSCAN, Kernel Density Estimation (KDE), UMAP	Identify clusters, understand data distribution, detect anomalies
4. Model Development & Training	TensorFlow, Keras, PyTorch, Scikit-Learn, XGBoost, Markov Chains, GCNs, T-EGAT, H-GCN	Develop models for anomaly detection, fraud detection, bias mitigation
5. Privacy-Preserving Collaborative Learning	TensorFlow Federated, PySyft	Enable training across institutions without sharing raw data
6. Data Imbalance & Augmentation	DB-CGAN, C-GAN, DGAN	Generate synthetic data to balance classes
7. Bias Detection & Mitigation	Bayesian Annotation Models, WEAT, Hard/Soft Debiasing Algorithms, Context Augmentation (BANC, ABTA, EBTA), Fairlearn, AIF360	Identify, quantify, and mitigate biases in training data and model predictions
8. Explainability & Transparency	SHAP, LIME, Custom Attention Heatmaps	Provide model interpretability and transparency
9. Real-time & Edge Deployment	Raspberry Pi, Arduino, SDN Controllers	Deploy models for real-time anomaly detection in resource-constrained environments
10. Visualization & Reporting	Google Maps, GIS Dashboards	Visualize spatial anomalies, report findings
11. Evaluation & Metrics	E(Tfa) Metric (NIST), Custom False Alarm Rate Metrics	Quantify model performance, false positive/negative rates

VIII. LIMITATION / AREAS FOR IMPROVEMENT

Main limitations related to existing approaches for anomaly and bias detection need to be overcome for better applicability in different domains. Interpretability and transparency still are crucial tasks and require the use of XAI technologies, especially when dealing with sensitive domains such as health care, finance, or NLP to explain model decisions, attribute bias and ensure regulatory compliance. Privacy-and security-driven considerations/frustrations may spur the adoption of federated and swarm learning setting where multiple parties can train a model collaboratively without sharing raw data, especially in banking, health care mining. GANs have gained much attention as an effective solution for data imbalance and minority class representation problem. They are widely used to generate synthetic dataset in industrial IoT and fraud detection. Real-time processing requirements in edge systems including embedded devices and smart infrastructure, require low latency architecture and energy-aware networking for the balancing of efficiency against detection accuracy. And when processing multimodal or mixed stream data from smart cities and industrial IOT move demand unified architectures advanced warehousing for relevant combination of different streams/heterogeneous data.

Moreover, for complex models, such as DL and LLMs, the process of detecting and mitigating bias is even more challenging. Tools such as embedding probing and primer-based methods are required. Such systems must also be robust and extensible, as well as interoperable with different networks both on premise and in the cloud. Quality of measurement problems, like obscured and contaminated sensor data, are tack-led using fuzzy logic control or through more robust loss functions such as Modified Hampel Loss Function. Instead of creating extensive regulatory barriers for any new venture, the government should be responsive to the society's changing needs by becoming an enabling and flexible supervisor that provides proper guidance. In this way, it can also effectively help direct investment to where these funds are most needed in order to maintain or increase overall economic growth. This kind of situation calls for adaptive learning and online retraining that can avoid corruption loss as models evolve in dynamic environments such as fraud monitoring, IoT management etc. In anomaly detection, the lack of labeled data forces the use of unsupervised learning, including such approaches as variational autoencoders and clustering based methods. Environmental variations and contextual incompleteness, that are natural in road surface and smart city monitoring are handled by context augmentation and hybrid sensor-vision models. Finally, compliance with regulatory and ethics governance frameworks maintains AI for finance, healthcare and smart cities within legal and societal patterns; complex relational anomaly detection in networks also benefits from GNN being used to perform social network analysis in identifying nonobvious group behaviors. Detractors to hybrid AD are putative: the lack of transparency and interpretability that is crucial for user trust, need for informed decisions. Balancing detection effectiveness and explanatory power across cybersecurity, finance, and industrial applications is a large main-standing challenge towards practical deployment [52].

IX. CONCLUSION AND FUTURE WORKS

The experiment shows that hybrid models and state-of-the-art deep neural network model (LSTM, GNN) can be well applied to handling the complexity of anomaly detection tasks. The introduction of XAI methods can reduce opacity in models and help make models more accessible to regulatory authorities. Federated learning systems are important ways to maintain data privacy and allow secure collaboration among organizations. Also, dedicated equipment for a particular field provides increased range ability and effectiveness compared with general-purpose equipment. In the same spirit, embedding mechanisms for bias detection into NLP or computer vision pipelines advances the causes of fairness and ethical AI. Furthermore, both cloud edges and embedded systems can complement each other to achieve real-time anomaly detection under conditions of limited resources.

In future, adaptive learning techniques which can learn changing anomaly pattern and shift in concepts. Improved interpretability of complex models, such as GNN and G AN, is still a critical area open for investigation. Need to build fair interventions into standardized debiasing procedures and make the choice of domain-agnostic methods. Along these lines we must not be limited to looking inward at our biases, but also be expansive across languages and cultures. Further, the development and encouragement of federated or decentralized learning mode can strengthen both privacy preservation and scaling for distributed environments. Exploration of new technologies, including quantum computing as well as future edge hardware are good opportunities to enhance detection capabilities. Finally, false positives are erased and over-all robustness of detection is improved by combining multi-modal data fusion with context augmentation methods.

REFERENCES

[1] V. Gandhi, S. Shah, M. Shah, A. Mehta and K. Srivastava, "Optimisation of Anomaly Detection in Video Processing Using Efficient Feature Engineering," 2024 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2024, pp. 1-6, doi: 10.1109/ESCI59607.2024.10497234.

[2] S. Zhu, Y. Li, K. Xu and J. Xu, "Partially-Supervised Graph Derivation Network With Meta-Learning for Time-Series Anomaly Detection," in IEEE Internet of Things Journal, vol. 12, no. 13, pp. 25472-25486, 1 July1, 2025, doi: 10.1109/JIOT.2025.3558273.

[3] Pal Amutha K, Sethukkarasi C, Mehanathen N, Ethirajan D, "A Holistic AI Approach for Secure and Scalable Exam Registration", International Conference on Image Processing and Artificial Intelligence, IPAI, Paris, France 2025

[4] Sebestyen, A. Hangan, Z. Czako and G. Kovacs, "A taxonomy and platform for anomaly detection," 2018 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), Cluj-Napoca, Romania, 2018, pp. 1-6, doi: 10.1109/AQTR.2018.8402710.

[5] J. G. Thomas, S. P. Mudur and N. Shiri, "Detecting Anomalous Behaviour from Textual Content in Financial Records," 2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI), Thessaloniki, Greece, 2019, pp. 373-377.

[6] J. Nicholls, A. Kuppa and N. -A. Le-Khac, "Financial Cybercrime: A Comprehensive Survey of Deep Learning Approaches to Tackle the Evolving Financial Crime Landscape," in IEEE Access, vol. 9, pp. 163965-163986, 2021, doi: 10.1109/ACCESS.2021.3134076.

[7] M. M. Fuad et al., "Financial Voucher Analysis with LVMs and Financial LLMs," 2025 International Conference on Computing Technologies (ICOCT), Bengaluru, India, 2025, pp. 1-6, doi: 10.1109/ICOCT64433.2025.11118347.

[8] R. D. Camino, R. State, L. Montero and P. Valtchev, "Finding Suspicious Activities in Financial Transactions and Distributed Ledgers," 2017 IEEE International Conference on Data Mining Workshops (ICDMW), New Orleans, LA, USA, 2017, pp. 787-796, doi: 10.1109/ICDMW.2017.109.

[9] Z. Umamah, J. Triloka, S. Y. Irianto and R. A. Aziz, "Systematic Review of Artificial Intelligence Techniques Datasets and Weaknesses in Finance Cybersecurity," 2025 4th International Conference on Creative Communication and Innovative Technology (ICCIT), Kota Cirebon, Indonesia, 2025, pp. 1-6, doi: 10.1109/ICCIT65724.2025.11166739.

[10] M. Mendelson and Y. Belinkov, "Debiasing Methods in Natural Language Understanding Make Bias More Accessible," Online and Punta Cana, Dominican Republic, pp. 545–1557, Nov. 2021. [Online]. Available: https://aclanthology.org/2021.emnlp-main.116/

[11] I. Maab, E. Marrese-Taylor, S. Pad´o, and Y. Matsuo, "Media Bias Detection Across Families of Language Models," Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pp. 4083–4098, 2024, conference Name: Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers) Place: Mexico City, Mexico Publisher: Association for Computational Linguistics. [Online]. Available: https://aclanthology.org/2024.naacl-long.227

[12] L. Lin, L. Wang, J. Guo, and K.-F. Wong, "Investigating Bias in LLM-Based Bias Detection: Disparities between LLMs and Human Perception," Mar. 2024. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2403.14896

[13] D. Hovy and S. Prabhumoye, "Five sources of bias in natural language processing," Language and Linguistics Compass, vol. 15, no. 8, p. e12432, 2021, eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/lnc3.12432. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/lnc3.12432

[14] E. Dayanik, N. T. Vu, and S. Pad´o, "Bias Identification and Attribution in NLP Models With Regression and Effect Sizes," in Northern European Journal of Language Technology, Volume 8, L. Derczynski, Ed. Copenhagen, Denmark: Northern European Association of Language Technology, 2022. [Online]. Available: https://aclanthology.org/2022.nejlt-1.4

[15] D. S. Shah, H. A. Schwartz, and D. Hovy, "Predictive Biases in Natural Language Processing Models: A Conceptual Framework and Overview," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault, Eds. Online: Association for Computational Linguistics, Jul. 2020, pp. 5248–5264. [Online]. Available: https://aclanthology.org/2020.acl-main.468/

[16] D. Kumar, C. Verma, Z. Illes, A. Mittal, B. Bakariya and S. B. Goyal, "Anomaly Detection in Chest X-Ray Images using Variational Autoencoder," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 216-221, doi: 10.1109/IC3I59117.2023.10397595.

[17] M. K. Mishra, L. Mohan Trivedi, M. Farooq, D. Nimma, M. B. Alazzam and B. Kiran Bala, "Cybersecurity Enhancement in IoT-Enabled Public Health Information Systems Using Deep Learning Techniques," 2025 AI-Driven Smart Healthcare for Society 5.0, Kolkata, India, 2025, pp. 25-30, doi: 10.1109/IEEECONF64992.2025.10963029.

[18] P. Ashok and A. S. Durge, "Fraud Detection and Prevention in Healthcare Insurance Claims Using Machine Learning Regression Models," 2025 International Conference on Data Science and Business Systems (ICDSBS), Chennai, India, 2025, pp. 1-7, doi: 10.1109/ICDSBS63635.2025.11031751.

[19] F. Arena et al., "Leveraging Data Science and Artificial Intelligence for Enhanced Monitoring Systems," 2025 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv), Venezia, Italy, 2025, pp. 408-413, doi: 10.1109/MetroLivEnv64961.2025.11107156.

[20] S. Kundu, D. Mishra, N. Tarasia, R. K. Lenka and R. K. Barik, "QEA-FAD: Quantum Enhanced Autoencoders for Anomaly Detection in Finance and Healthcare," 2025 International Conference on Networks and Cryptology (NETCRYPT), New Delhi, India, 2025, pp. 1280-1285, doi: 10.1109/NETCRYPT65877.2025.11102801.

[21] A. J. Mary, V. Umesh, P. N. Khairnar, R. Sahana, S. SN and S. H. J, "Anomaly Detection in Industrial Quality Control with Computer Vision and Deep Learning," 2024 International Conference on Distributed Systems, Computer Networks and Cybersecurity (ICDSCNC), Bengaluru, India, 2024, pp. 1-6, doi: 10.1109/ICDSCNC62492.2024.10939633.

[22] R. K. Gupta, N. N, B. S. Rao, R. RS, S. Sarkar and K. R. Kumar, "Deep Learning for Uneven Data in Industrial IoT Using a Distributed Bias-Aware Adversarial Network," 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2023, pp. 1435-1440, doi: 10.1109/ICIRCA57980.2023.10220695.

[23] X. Zhou, Y. Hu, J. Wu, W. Liang, J. Ma and Q. Jin, "Distribution Bias Aware Collaborative Generative Adversarial Network for Imbalanced Deep Learning in Industrial IoT," in IEEE Transactions on Industrial Informatics, vol. 19, no. 1, pp. 570-580, Jan. 2023, doi: 10.1109/TII.2022.3170149.

[24] L. Tan et al., "Energy-Efficient Tactile-Driven Rule Configuration and Anomaly Detection in Industrial IoT Systems," in IEEE Internet of Things Journal, vol. 12, no. 17, pp. 34679-34686, 1 Sept.1, 2025, doi: 10.1109/JIOT.2025.3541641.

[25] T. Kamitani, S. Fujimoto, H. Yoshimura, M. Nishiyama and Y. Iwai, "Anomaly detection using local regions in road images acquired from a hand-held camera," 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE), Nara, Japan, 2018, pp. 375-378, doi: 10.1109/GCCE.2018.8574660.

[26] T. Y. Er and S. Selçuk, "Enhancing Road Anomaly Detection with Dynamic Cropping System: A YOLOv8 Integrated Approach," 2024 International Conference on Smart Systems and Technologies (SST), Osijek, Croatia, 2024, pp. 43-47, doi: 10.1109/SST61991.2024.10755318.

[27] Y. T. Gamage, T. A. I. Thotawaththa and A. Wijayasiri, "Identification of Road Surface Anomalies Using Crowdsourced Smartphone Sensor Data," 2022 22nd International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, Sri Lanka, 2022, pp. 118-123, doi: 10.1109/ICTer58063.2022.10024097.

[28] J. -C. Wang et al., "Research on Monitoring Road Surface Anomalies Using an IoT-Based Automatic Detection System: Case Study in Taiwan," in IEEE Transactions on Industrial Informatics, vol. 20, no. 9, pp. 11404-11417, Sept. 2024, doi: 10.1109/TII.2024.3404052.

[29] M. Pasti, E. Ridolfo and A. Zanella, "Road Anomalies Detection Using Low-Cost Sensors and Machine Learning," 2024 IEEE 35th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Valencia, Spain, 2024, pp. 1-6, doi: 10.1109/PIMRC59610.2024.10817220.

[30] W. Zhou et al., "Anomaly Usage Behavior Detection Based on Multi-Source Water and Electricity Consumption Information," in IEEE Access, vol. 13, pp. 12215-12224, 2025, doi: 10.1109/ACCESS.2025.3525726.

[31] S. Yaqoob, A. Hussain, F. Subhan, G. Pappalardo and M. Awais, "Deep Learning Based Anomaly Detection for Fog-Assisted IoVs Network," in IEEE Access, vol. 11, pp. 19024-19038, 2023, doi: 10.1109/ACCESS.2023.3246660.

[32] A. Oluyomi, S. Abedzadeh, S. Bhattacharjee and S. K. Das, "Unsafe Events Detection in Smart Water Meter Infrastructure via Noise-Resilient Learning," 2024 ACM/IEEE 15th International Conference on Cyber-Physical Systems (ICCPS), Hong Kong, Hong Kong, 2024, pp. 259-270, doi: 10.1109/ICCPS61052.2024.00030.

[33] Y. Liu et al., "Zero-Bias Deep Learning for Accurate Identification of Internet-of-Things (IoT) Devices," in IEEE Internet of Things Journal, vol. 8, no. 4, pp. 2627-2634, 15 Feb.15, 2021, doi: 10.1109/JIOT.2020.3018677.

[34] G. Zheng, Q. Ni and Y. Lu, "Privacy-Aware Anomaly Detection and Notification Enhancement for VANET Based on Collaborative Intrusion Detection System," in IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 12, pp. 21172-21182, Dec. 2024, doi: 10.1109/TITS.2024.3479426.

[35] J. Qin, "Research on Supply Chain Finance Risk Management Based on Blockchain Technology," 2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS), Kalaburagi, India, 2023, pp. 1-6, doi: 10.1109/ICIICS59993.2023.10421010.

[36] B. Golchin and B. Rekabdar, "Anomaly Detection In Time Series Data Using Reinforcement Learning, Variational Autoencoder, and Active Learning," 2024 Conference on AI, Science, Engineering, and Technology (AIxSET), Laguna Hills, CA, USA, 2024, pp. 1-8, doi: 10.1109/AIxSET62544.2024.00007.

[37] S. Kulkarni, N. Mittal, R. R. Gupta and P. R N, "Investigation of YOLO models in the detection and classification of multiple negative road anomalies," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-7, doi: 10.1109/ICCCNT56998.2023.10306347.

[38] A. Ruggeri, A. Ficara, G. Sollazzo, G. Bosurgi and M. Fazio, "A Comparative Analysis of Deep Learning Approaches for Road Anomaly Detection," 2024 IEEE Symposium on Computers and Communications (ISCC), Paris, France, 2024, pp. 1-6, doi: 10.1109/ISCC61673.2024.10733559.

[39] J. Cao et al., "Generative Models for Time Series Anomaly Detection: A Survey," in IEEE Transactions on Artificial Intelligence, doi: 10.1109/TAI.2025.3614213.

[40] B. El Habib and B. Nasri, "Detection and Prediction of Satellite Telemetry Anomalies," 2024 4th International Conference on Embedded & Distributed Systems (EDiS), BECHAR, Algeria, 2024, pp. 310-313, doi: 10.1109/EDiS63605.2024.10783288.

[41] X. Xu, K. Ding, C. Chen and K. Shu, "MetaGAD: Meta Representation Adaptation for Few-Shot Graph Anomaly Detection," 2024 IEEE 11th International Conference on Data Science and Advanced Analytics (DSAA), San Diego, CA, USA, 2024, pp. 1-10, doi: 10.1109/DSAA61799.2024.10722838.

[42] K. Ding, K. Shu, X. Shan, J. Li and H. Liu, "Cross-Domain Graph Anomaly Detection," in IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 6, pp. 2406-2415, June 2022, doi: 10.1109/TNNLS.2021.3110982.

[43] G. Sivapalan, K. K. Nundy, A. James, B. Cardiff and D. John, "Interpretable Rule Mining for Real-Time ECG Anomaly Detection in IoT Edge Sensors,"

in IEEE Internet of Things Journal, vol. 10, no. 15, pp. 13095-13108, 1 Aug.1, 2023, doi: 10.1109/JIOT.2023.3260722.

[44] T. Choi, D. Lee, Y. Jung and H. -J. Choi, "Multivariate Time-series Anomaly Detection using SeqVAE-CNN Hybrid Model," 2022 International Conference on Information Networking (ICOIN), Jeju-si, Korea, Republic of, 2022, pp. 250-253, doi: 10.1109/ICOIN53446.2022.9687205.

[45] G. R. Kumar, A. D. Kulkarni, B. S. Kumar, N. Singh, V. Revathi and T. C. A. Kumar, "Machine Learning Approaches for Anomaly Detection in IoT Networks," 2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2024, pp. 1-5, doi: 10.1109/ACCAI61061.2024.10601954.

[46] H. S. Mavikumbure, V. Cobilean, C. S. Wickramasinghe, D. Drake and M. Manic, "V2XFormer: Transformer-Based Anomaly Detection for Vehicle-to-Everything Communication," 2025 IEEE 101st Vehicular Technology Conference (VTC2025-Spring), Oslo, Norway, 2025, pp. 1-7, doi: 10.1109/VTC2025-Spring65109.2025.11174553.

[47] J. Wang, R. Martins de Moraes and A. Bari, "A Predictive Analytics Framework to Anomaly Detection," 2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService), Oxford, UK, 2020, pp. 104-108, doi: 10.1109/BigDataService49289.2020.00023.

[48] Y. Dai, I. Spence, K. Rafferty, B. Quinn, J. Huang and H. Wang, "TDSRL: Time Series Dual Self-Supervised Representation Learning for Anomaly Detection From Different Perspectives," in IEEE Internet of Things Journal, vol. 12, no. 17, pp. 35078-35096, 1 Sept.1, 2025, doi: 10.1109/JIOT.2025.3577931.

[49] C. Chen, Y. Li, B. Jiao, G. Zhao and W. Li, "Wasserstein Heterogeneous Graph Neural Networks for Uncertainty-Aware Anomaly Detection," ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hyderabad, India, 2025, pp. 1-5, doi: 10.1109/ICASSP49660.2025.10890733.

[50] N. Rashid, R. Mehmood, F. Alqurashi, S. Alqahtany and J. M. Corchado, "ASAD: A Meta Learning-Based Auto-Selective Approach and Tool for Anomaly Detection," in IEEE Access, vol. 13, pp. 4341-4367, 2025, doi: 10.1109/ACCESS.2024.3524908.

[51] S. Barnwal, "Vehicle Behavior Analysis for Uneven Road Surface Detection," 2015 IEEE 18th International Conference on Intelligent Transportation Systems, Gran Canaria, Spain, 2015, pp. 1719-1722, doi: 10.1109/ITSC.2015.279.

[52] R. Mohite and L. Ouarbya, "Interpretable Anomaly Detection: A Hybrid Approach Using Rule-Based and Machine Learning Techniques," 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India, 2024, pp. 1-10, doi: 10.1109/I2CT61223.2024.10543396.