A Systematic Review of Anomaly Detection Using Machine and Deep Learning Techniques

Sarfaraz Natha[1,*], Mehwish Leghari[2], Muhammad Awais Rajput[2], Syed Saood Zia[1], Jawaid Shabir[3],

[1]Department of Software Engineering, SSUET, Karachi, Pakistan.
[2]Department of Artificial Intelligence, QUEST, Nawabshah, Pakistan.
[3]Department of Computer Engineering, SSUET, Karachi, Pakistan.
[*]Corresponding author: sasattar@ssuet.edu.pk

Abstract

Anomaly detection identifies objects or events that do not behave as expected or correlate with other data points. Anomaly detection has been used to identify and investigate abnormal data components. Detecting anomalous activities is challenging due to insufficient data size of anomalous reality, ground training data, factors related to differences in environmental conditions, working position of capturing cameras, and illumination situations. Anomaly detection has enormous applications that include (but not limited to) industrial damage prevention, sensor network, health-care services, traffic surveillance, and violence prediction. Machine learning techniques, particularly deep learning has enabled tremendous advancements in the area of anomaly detection. In this paper, we sort out an all-inclusive review of the up-to-date research on anomaly detection techniques. We seek to serve as an extensive and comprehensive review of machine and deep learning anomaly detection techniques throughout the foregoing three years 2019-2021. Particularly, we discuss both machine learning and deep learning anomaly detection applications, performance measurements, and anomaly detection classification. We also point out various datasets that have been applied in anomaly detection along with some fairly new real-world datasets. Finally, we investigate current challenges and future research prospects in this area.

Keywords—— Machine Learning, Anomaly detection, Anomaly localization, Deep Learning, Convolutional Neural Network

1 Introduction

ANOMALLY detection has been known as the process of detecting abnormalities and outliers in data. Essentially anomaly detection aims to recognize data instances that differ considerably from the bulk of data instances, hence the name anomaly detection or sometimes novelty discovery. For decades, anomaly detection has been a popular topic in research [1]. Historically, it remained an important process in a much broader range of domains e.g., artificial intelligence, computer vision, and statistics, just to name a few. The necessity of detecting anomalies in a wide range of application sectors stems from the possibility that unprotected data might include important, relevant, and critical information. For example, detecting irregularities in credit card transaction data may suggest theft [2]. Another use case could be to identify and differentiate unusual behavior in a network of computers to find a pattern that might reveal a compromised computer's assault [3]. Broadly speaking, an anomaly is described as a pattern that deviates from expected behavior. The term anomaly itself covers a wider spectrum of irregularities and accordingly, anomaly detection refers to the set of techniques to detect these irregularities. Various specific forms of anomalies such as those arising from the triggering of a Trojan (purposely inserted malicious code/logic to perform illegitimate actions) can be considered under the umbrella of malware and anomaly detection techniques can be applied to detect such triggers using state-of-the-art methods. Anomalies are broadly divided into three categories [4,5,9]:

1) **Point anomalies:** In point anomalies, a single data instance can be identified as unusual from the rest of the data and is the simplest form of

ISSN: 2523-0379 (Online), ISSN: 1605-8607 (Print)
DOI: https://doi.org/10.52584/QRJ.2001.11

anomaly.

2) **Contextual anomalies:** This type of anomalous behavior occurs when an observation could be taken as anomalous in one context but not in another. Contextual anomalies have two different types of attributes: contextual and behavioral. There are different contextual properties found in longitude, spatial, and latitude datasets of a location. Furthermore, in time-series data, time is a contextual property that indicates an instance's location in the arrangement. The second feature is regarded as a behavioral characteristic.
3) **Collective anomalies:** A group of related instances of data that can emerge as anomalous when seen together with respect to the whole dataset is termed collective anomalies [6].

To detect anomalies, statistical algorithms were among the earliest methods. To determine whether an instance corresponds to this model, a statistical inference test might be used. Statistical anomaly detection is accomplished using a variety of approaches. This comprises approaches that are based on proximity, as well as parametric, non-parametric, and semi-parametric methods [45].

Recently, an increasing amount of interest has risen to use machine learning methods as a way of detecting abnormalities [9]. A few popular machine learning-based techniques such as clustering-based, distance-based, statistical, and classification-based detection have been used widely by various researchers. In cluster-based anomaly detection, normal observations are assumed to be part of the same cluster in cluster-based detection (s). If a new remark is further away from the cluster centroid(s), it will be considered an anomaly. In a distance-based approach, an object placed remote from the neighboring data which do not have enough points will be considered an outlier. Other methods including one-class support vector machine (1-SVM), neural network, and k-NNs have also been employed extensively by researchers. Some of the most popular methods for anomaly detection are briefly described in the following.

1) **k-nearest neighbors (k-NNs)**
The k-NNs algorithm provides a non-parametric method for either grouping or regression. In an anomaly detection scenario, it is suited since it is a simple yet powerful classification method to identify and group abnormalities.
2) **Support vector machines (SVMs)**
Another popular technique for anomaly detection is via SVMs. Several applications including spacecraft, aviation, and electrical systems use SVMs (particularly one-class SVM and least squares SVM) [17,18]. However, since SVM is a supervised learning model for classification and regression issues, it is better suited as a classifier for classification-based anomaly detection.
3) **Neural networks (NNs)**
Anomaly detection systems have used neural networks (NNs) in multi-class or one-class scenarios. Basic anomaly detection using NNs is essentially carried out as a two-step process: First, a neural network is trained on multi-class normal examples; in the second step, the trained neural network finds anomalies by accepting (i.e., regular) or rejecting (i.e., irregular) test instances [15]. Neural networks have also been employed to detect one-class anomalies [16, 14].
4) **Decision tree**
A decision tree is a predictive model in machine learning in which each internal node represents a predictive variable (feature), a child node represents a variable's possible value range, an external node (leaf) rep-resents the target variable's predicted value, and a classification or judgment node reflects the target variable's predicted value [17].
5) **Deep learning-based approaches**
Deep learning has evolved in recent years up to an extent that it possesses enhanced capability in learning expressive demonstrations of difficult data such as a high-dimensional, temporal, graph, and spatial data. Deep learning is used to extract low-dimensional feature representations from high-dimensional or/and non-linearly diverse instances, which may then be used to identify irregularities [21]. Typically, deep learning for anomaly detection, or deep abnormality detection, is employed via neural networks that are first trained to learn feature representations or anomaly scores and later invoked to detect anomalies.
6) **Generative Adversarial Networks**
The purpose of this method is to create a latent feature space for a generative network G that accurately captures the data's normalcy [26]. The purpose of anomaly measure-dependent feature learning is to learn feature representations that are specifically customized for a single anomaly measure [42]. Deep distance-based anomaly detection tries to learn feature representations that are tailored to a particular type of distance-based anomaly measure. Many effective distance-based anomaly measures have been proposed, including k-closest neighbor distance [27] and random near-

est neighbor distance [28], and relative distance [29, 30].

The main aim of the review presented in this paper is to systematically evaluate machine learning and deep learning algorithms for anomaly detection and their applications. To give a comprehensive literary analysis of the research evolution, we examine a huge number of relevant works published in prestigious conferences and journals in numerous important areas of applications. To this end, this paper first offers the procedure that we adopted to gather and organize the relevant research. We then discuss the key assumptions, objective functions, important intuitions, and capabilities of various techniques in handling mainstream challenges. We also discuss several scenarios where the mentioned techniques would embrace a challenging setup. We also gather a thorough collection of accuracy metrics of the relevant works.

This survey paper is organized as a two-stage systematic review of relevant literature. In stage one, we establish a strategy for obtaining research articles linked to the issue that answers the research question, and the second stage is to find answers to the study questions based on the review's goal. In particular, the following objectives are to be considered:

- **Objective 1:** To present a detailed review of various machine learning and deep learning models that are used in anomaly detection.
- **Our method:** We explain in detail machine and deep learning models, along with their weaknesses and strengths.
- **Objective 2:** To evaluate state-of-the-art works in the domain in terms of accuracy and performance metrics.
- **Our method:** We devote an extensive section with details on the estimation accuracy of the machine and deep learning models.

To gather the relevant literature for the survey, we have used multiple online literature search tools and digital libraries e.g., Google Scholar, Elsevier, Springer, Digital Library of the ACM, and IEEE Xplorer with the search term "Anomaly detection". The list of related papers was then further narrowed down using Boolean operators (ANDs and ORs).

The rest of the paper is organized as follows: Section 2 compares some of the other recent survey papers and out-lines the key differences by highlighting the contributions of our paper. In section 3, we present an overview of various techniques proposed for anomaly detection in related literature along with the strengths and weaknesses of each work with a category-wise description. Section 4 provides categorical details on the accuracy and estimation metrics of the recent works with their datasets in the domain of machine learning-based anomaly detection. Cur-rent challenges and research directions are explained in section 5 and finally, in section 6, the paper is concluded.

2 Contribution of This Paper

With an increasing number of works dealing with anomaly detection dictates that anomaly detection is becoming of interest to researchers. In addition, there has been an increasing number of research that provides a review of the latest techniques in this domain. This section serves as a brief literature overview and a comparison of various other survey papers and outlines a clear distinction between this paper with other recent review papers. In the following, we describe some of the most related survey papers on anomaly detection and highlight key differences and issues that this paper addresses as compared to the other related surveys.

Habeeb et al. [6] provide a detailed survey covering anomaly detection in real-time big data processing application domains. Our investigation, on the other hand, is broader in scope and includes an assessment of accuracy for both model types.

In another recent survey by Kwon et al. [40], the authors provide an overview of anomaly detection and deep learning algorithms and attempt to figure out if deep learning can be used to detect network anomalies. In comparison, our paper also takes into account deep learning techniques for identifying abnormalities in network intrusion systems along with anomaly detection applications, performance metrics, and anomaly detection classification using machine learning and deep learning.

Fernandes et al. [41] in their detailed investigation, focus on various aspects of anomaly detection including intrusion detection, network data, and aberrant network traffic. On the other hand, our investigation is broader in scope, and it includes a thorough review of the estimation and accuracy of ML and DL models in addition to the detection of network anomalies in particular.

In a more recent survey by Guansong et al. [42], some of the unique issues and unresolved problems that deep learning for anomaly detection presents, are discussed, and related literature that attempts to resolve the issues is referred to. Nevertheless, the formulation of the study, which was based on three principled frameworks, our study incorporates a larger number of machine learning and deep learning applications to determine their strengths and weaknesses.

Another recent survey work from Rashmiranjan et

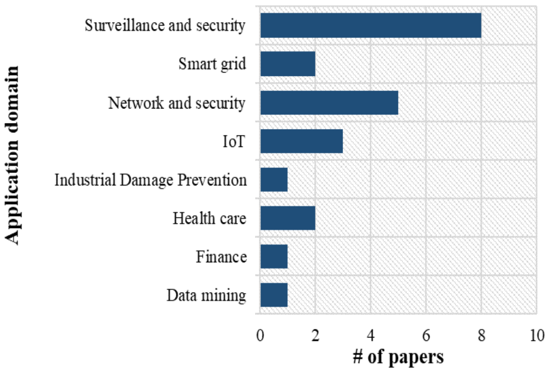

Fig. 1: Share of various application domains of papers included in this survey

al. [43], introduces accuracy and real-time processing algorithms for video anomaly identification based on deep-learning methods. Despite the work being up-to-date and addressing the latest works in the domain, the scope of the survey is limited.

Al-amri et al. [44] provide a survey of several state-of-the-art solutions for addressing the primary issues and challenges in anomaly detection in IoT data. In comparison, our paper also considers anomaly detection methods that are not covered by their paper e.g., statistical techniques.

Another recent review from Nassif et al. [45] investigates works that describe ML models to detect anomalies in multiple applications. Our survey, in addition to discussing detection techniques, further provides details on other aspects such as applications, performance evaluations, and anomaly detection classification.

3 Machine and Deep Learning Methods

This section deals with the first objective of this paper. We particularly discuss the strengths and weaknesses of machine learning and deep learning methods utilized to detect abnormalities. The provided list (see Table 1) is categorized based on various ML and DL methods employed to carry out the research. An overview of various application domains of the works presented in Table 1 is depicted in Figure 1. A large share of the papers that we present in this survey deal with anomaly detection in the surveillance domain.

4 Accuracy and Estimation of ML and DL Models

This section looks at the second objective of this paper, which is concerned with the accuracy of ML and DL mod-els. Estimation and accuracy are the most important parameter for machine and deep learning models. Table 2 provides a detailed categorical description of the related works and the estimation metrics along with reported ac-curacy results.

5 Challenges and Future Directions

The survey of the state-of-the-art research presented in this paper demonstrates an enormous growth of interest in the field of machine learning-based anomaly detection approaches. However, despite significant advancements in the estimation models, the addition of new datasets, and novel frameworks, still, a few avenues possess room for improvement. In the following, we list some of the open challenges and future directions of the research in the do-main of machine learning-based anomaly detection.

5.1 Challenges

- **Nature of input dataset:** Many experiment setups have been dealing with a system's normal functioning. The most advanced approaches are based on typical behavior training, and anything that deviates from the regular labeled data is deemed abnormal. To cope with complex datasets of real-world circumstances, more precise and trustworthy procedures are necessary [24]. Furthermore, the availability of an appropriate dataset for public anomaly detection is a major concern for training and verifying real-time anomaly detection systems.
- **Data streams pose external detection:** External detection challenges arise from data streams, such as detecting in limited memory and time, updating data once it enters, and dynamically managing data to capture fundamental changes while detecting them [93]. Novel algorithms that can adjust their setup and settings over time when new knowledge emerges can greatly improve accuracy. Detection algorithms, unlike static data, fail to adapt to complicated settings, such as the ever-changing IoT domain.
- **Noise and data complexity:** Noise and data complexity are one of the key obstacles in developing a model for anomaly detection. Data complexity includes unbalanced datasets, unexpected sounds, and redundancy within the data [94]. To acquire meaningful information and knowledge, well-developed methodologies for curating datasets are essential.
- **Unknown nature of anomaly:** The unknown nature of the anomaly is the fundamental problem

TABLE 1: TABLE 1. MACHINE AND DEEP LEARNING MODEL'S STRENGTHS AND WEAKNESSES

ML / DL Methods	**Author, Year and Reference**	**Strength**	**Strength**
Data clustering module	Chen et al., 2019 [67]	Experimental data from two usage scenarios show an anomaly detection accuracy of up to 99\% with a false positive rate of less than 1\%.	If end-to-end detection is only adapted to distinguish a few conspicuous abnormalities , it might not be capable to generalize to anomalies that are not visible but have different aberrant properties than the ones that have been detected.
Decision Trees	Zhang et al., 2019 [70]	The work achieved high accuracy by manipulating Adaboost and basic feedforward neural networks as two methods for discovering network performance anomalies based on split-sample categorization.	Adaboost is highly vulnerable to uniform noise, in case classifiers are too weak, resulting in overall poor margins and overfitting. Also, ANN requires more computing resources to prepare datasets.
	Quatrini et al., 2020 [47]	A decision forest and decision jungle-based approach is used, which is a well-known anomaly detector using industrial data.	Large execution times are required to train the models.
Random Forests	Alrashdi et al., 2019 [74]	The work proposes an Anomaly Detection-IoT (AD-IoT) system with a highly accurate model i.e., 99.34\% classification accuracy.	By combining a lot of decision trees to identify the class, the random forest learning algorithm is computationally expensive.
Swarm Intelligence	Qasim et al., 2019 [83]	The work presents a unique swarm advection approach to calculate the histogram of swarms (HOS) descriptor for abnormal event identification. The results show that the proposed approach outperforms current methods for identifying anomalous occurrences in congested situations.	Although the model supports frame-level images, it is not ideal for pixel-level image series per frame.
	Selvaraj et al., 2019 [54]	The Swarm Intelligence approach for the proposed problem achieved a high accuracy of 95.70\%.	Data preparation necessitates more computing power. Moreover, for extracting low-dimensional features, the performance of swarm intelligence is suboptimal.
K- Nearest Neighbors	Punmiya, et al., 2019 [76]	Using a Cluster-based local outlier detection using the k-nearest neighbors' histogram -based outlier technique, the work provides an automatic anomaly identification framework.	Since a large number of decision trees to identify the class are to be integrated, an enormous amount of training time is required.

Long Short-term memory	Liu et al., 2020 [77]	The LSTM and auto-encoder (AE) based models outperform state-of-the-art models.	In an auto-encoder-based approach, can uncommon regularities and the existence of outliers or anomalies in the training data can skew the learned feature representations.
	Khokhlov et al., 2019 [75]	The work effectively shows how an anomaly-based detection of innovative and complex threats can be implemented effectively and efficiently in monitored data streams via LSTM to capture the most essential Android system parameters on mobile devices.	LSTM necessitates a significant amount of time and bandwidth resources.
Convolutional Neural Networks	Ganokratanaa et al., 2020 [82]	By offering Edge Wrapping, the suggested approach improves the efficiency of anomaly localization at pixel-level assessment and automatically learns the normal samples without modifying any settings.	More computation power is required for the proposed since it uses Encode and Decodes Recurrent Neural Network as architecture.
	Mehta et al., 2020 [87]	The suggested model may be utilized commercially on any GPU-based system to reliably identify fire and pistols in regions monitored by cameras with a rapid detection rate.	The proposed anomaly detection system is applicable only for fire detection and can be improved and diversified.
	Ilyas et al., 2021 [81]	The work suggests a handmade feature to encode a high-level change at the frame level and an ML + DL model further provides a better outcome.	The proposed model is not suitable for pixel-level feature extraction.
Support Vector Machine	Liu et al., 2020 [79]	The one-class support vector machine (1-SVM) approach is used to train the model, and the model and algorithms are evaluated using real data from Colorado Water Watch.	With a noisier dataset, target classes overlap and a one-class SVM does not perform well. When the number of features per data point is greater than the number of training data samples, one-class SVM underfits the data.
	Aziz et al., 2021 [80]	Because of camera jitter and object motions in improbable motion zones, the suggested technique can decrease false motion anomaly detection and localization alarms.	The proposed model is not suitable for spatial and temporal scenarios. One class SVM model is not suitable for a dataset where complex distribution within the normal class is present.
	Shriram et al., 2019 [68]	Anomaly detection is performed well and is quite effective.	The model necessitates a vast amount of data.
	Hasan et al., 2019 [69]	Using decision trees and ANNs, the proposed strategy achieved a test accuracy of 99.40\%. The work is capable of detecting attacks and anomalies in a virtual environment.	Real-time data large datasets are not supported.
Generative adversarial network	Alfie et al., 2021 [84]	The proposed model outperforms previous techniques in small-scale crowd videos using benchmark datasets and has an acceptable accuracy rate in detecting abnormal behavior in the HAJJ widespread crowd dataset.	The accuracy still needs to be improved for large-scale crowd datasets.

Hierarchical Temporal Memory	Barua et al., 2020 [78]	The work provides a real-time anomaly detection approach using hierarchical temporal memory (HTM) with continual unsupervised learning without human intervention.	Unsupervised learning produces less accurate outputs because the input data is unknown and not classified in advance.
DADT-PW model	Pustokhina et al., 2021 [85]	The suggested technique successfully identifies and classifies the anomalies that occur in the frame based on their superior characteristics.	Two-stage detection model requires high computation power for feature extraction.
Attention networks	Zhao et el., 2020 [105]	The work is based on parallel graph attention layers to tackle different time series dynamically.	One major limitation of the work is the inability to learn the topological the structure among sensors, leaving large relational data in a highly populated and connected sensor environment.
	Koizumi et al., 2020 [106]	The work proposes an attention process that deals with time-frequency stretching. The results demonstrate the superiority of the approach with significantly better performance as compared to conventional methods.	Although a promising approach, the proposed framework SPIDERNET is not flexible enough to deal with dynamic domain shift.
Multiple models	Aboah et al., 2021. [86]	The work combines different techniques such as video sorting and anomaly candidate filtering to enhance the model's ability to detect abnormalities across most videos.	The effect of a small change in the data could potentially generate a considerable change in the decision tree.
	Bhatia et al., 2019 [71]	The work proposes DDoS attacks prevention system using inputs reconstruction that nearly mimics typical network traffic efficiently.	The quality of input reconstructions for assault inputs is suboptimal.
	Wang et al., 2019 [72]	The work provides a financial IT solution for increasing productivity and is highly precise in predicting KPI features system failure.	The work, however, lacks in predicting KPI data on fine-grained KPI time series.
	Yihunie et al., 2019 [73]	The work attempts to find a better-suited predictor among five models for detecting anomalous traffic from the NSL-KDD dataset with high efficiency and a low error rate.	The proposed models miss an important task of classifying distinct class kinds or attacking strategies.

with auto anomaly detection [95] which has a variety of applications including fraud detection, defect identification, and event detection systems in sensor networks, among others. As a result, because there are no labels for time series containing anomalies, typical machine learning algorithms cannot be utilized to train the model.

- **Time complexity:** A key feature of a data stream is the huge volume of data that arrives in real-time, requiring the algorithm to perform in real-time. However, due to a reciprocal relationship between temporal complexity and accuracy, discovering the anomalies in that setting would be a huge challenge [96].
- **Complicated environment:** Indoor climatic complexity differs from one building to the next and from one character to the next [97]. Environmental anomalies pose a different kind of challenge due to various factors. For instance, capturing the readings via a univariate sensor, and the use of a single machine learning model in a complex environment has an impact on the accuracy of anomaly detection due to a clear variation pattern for some indoor climate parameters that cannot be easily observed.
- **Lack of clarity on the subject:** A clear char-

TABLE 2: COMMONLY USED QUALITY METRICS AMONG SELECTED PAPERS

ML / DL Model	Dataset	Author, Year and Reference	Performance Metrics	Resulting Value
Neural network	Distributed Smart Space Orchestration System (DS2OS)	Hasan et al., 2019 [69]	Accuracy	0.990
	Real World Dataset	Zhang et al., 2019 [70]	Accuracy	0.920
		Selvaraj et al., 2019 [54]	Precision Recall	0.957 N/A
Decision Tree, Random Forest	NSL-KDD dataset	Yihunie, et al., 2019 [73]	Precision Recall	0.999 0.999
	UNSW-NB15	Alrashdi et al., 2019 [74]	Precision Recall	0.990 0.980
	Real-World Dataset	Quatrini et al., 2019 [47]	Precision Recall	0.996 0.997
		Wang et al., 2019 [72]	Precision Recall	0.888 0.700
Generative Adversarial Network	CUHK Avenue	Ganokratanaa et al., 2020 [82]	Accuracy	0.980
	Real-world data (Hajj)	Alafif et l., 2021 [84]	Accuracy	0.796
Support Vector Machine	Benign IoT traffic	Bhatia et al., 2019 [71]	Accuracy	0.999
	UMN datasets	Aziz et al., 2021 [80]	Accuracy	0.970
	PETS 2009	Ilyas et al., 2021 [81]	Accuracy	0.990
	NSL-KDD dataset	Pu et al., 2021 [88]	Accuracy	0.890
	Synthetic data and public domain data	Liu et al., 2020 [79]	Accuracy	0.890
Convolution Neural Networks	Real-world dataset (live CCTV)	Aboah et al., 2021 [86]	Accuracy	0.850
kNN, histogram-based outlier detection	Real-World Dataset	Punmiya et al., 2019 [76]	Accuracy	0.910
Long short-term memory (LSTM) encoder decoder	Real-World Dataset	Liu et al., 2020 [77]	Accuracy	0.978

acterization of each area of the anomaly detection problem is lacking in the arrangement [98]. This is partly owing to the ambiguous nature of the terminology involved. This ambiguity makes it simpler to group different sub-areas within these fields of study. Despite their proximity, these distinct sub-tasks are not similar, and hence the approaches to deal with them might not be the same.

- **Dataset complications:** Images form an essential part of the datasets and to train a model to recognize anomalous behavior, a proper description and inclusion of finer details are required. Images with unnecessarily low-intensity regions, poor viewing angles, skewed and rotated, and lower resolutions are all potential complications in images that have a significant impact on the accuracy of the anomaly detection model [99].

5.2 Future Directions

- **Improving model performance:** A potentially effective improvement strategy in anomaly detection models can be augmented via the accumulation of more detailed datasets that considers more diverse failure scenarios. Furthermore, using more effective machine learning architectures and training techniques, the improvement might be multi-fold.
- **Multiple validations:** Another exciting avenue in the future direction is to focus on cross-validation to yield higher performance metrics. Consider for example the application areas of AD-IoT where enhanced detection rates with a reduced false-positive rate could greatly improve the system accuracy.
- **Improvement in limited resources:** The success of machine learning models is strongly reliant on training data, which is impacted by several parameters including sensor quality, sampling rate, and dataset size. As a result, it is worth considering how to improve anomaly detection performance by utilizing low-cost sensors and a small dataset on less often sampled data.
- **Proposed quality datasets:** It would be incredibly valuable to have global benchmark datasets for every field of anomaly detection dedicated to comparing all of the approaches provided for visual tracking or abnormality identification. The first and most difficult step should be to offer a uniform measure of tracking quality that accounts

for most of the issues that come with visual tracking (variation in appearance, illuminations, occlusions, blur, and so on). The benchmark should then be divided into subsets of sequences, each addressing a distinct problem. It is also possible to suggest ranking the sequences' complexity.

- **Improve interdisciplinary approaches:** Close collaboration between academics and concerned government agencies (such as law enforcement agencies) could be beneficial to both. Researchers would be able to put their models to the test in real-world scenarios, while social forces would be able to experiment with substantial technological developments.

6 Conclusion

This survey paper attempts to provide a systematic review of the machine and deep learning techniques for anomaly detection. We build our systematic review around two objectives which we then attempt to achieve in the remaining sections. We include the most recent and up-to-date detailed review of the state-of-the-art studies in the years 2019 to 2021 (objective 1). Further, we devote a complete section to accuracy metrics used in the literature (objective 2). In addition, we provide a list of application areas where anomaly detection is employed. We put a special focus on the latest techniques of anomaly detection that are driven by the most advanced applications of the ma-chine and deep learning. Due to the fast-evolving nature of the field, we include only the most recent papers i.e., between 2019 and 2021.

We also mention several datasets that were utilized in experiments of relevant research publications, with a majority of the experiments using real-world datasets as training or testing datasets for their models. Our review reveals that many avenues are still in the infancy phase and require significant research. Moreover, many datasets are be-coming obsolete and are being replaced with newer and most relevant real-world datasets and hence are more valuable. We believe that this review could be a valuable starting point for researchers and the AI community to get up-to-date and relevant information on anomaly detection using machine learning techniques.

References

[1] Grubbs, Frank E. ”Procedures for detecting outlying observations in samples.” Technometrics 11, no. 1 (1969): 1-21.

[2] Agrawal, Shikha, and Jitendra Agrawal. ”Survey on anomaly detection using data mining techniques.” Procedia Computer Science 60 (2015): 708-713.

[3] Gogoi, Prasanta, Dhruba K. Bhattacharyya, Bhogeswar Borah, and Jugal K. Kalita. ”A survey of outlier detection methods in network anomaly identification.” The Computer Journal 54, no. 4 (2011): 570-588.

[4] Chandola, Varun, Arindam Banerjee, and Vipin Kumar. ”Anomaly detection: A survey.” ACM computing surveys (CSUR) 41, no. 3 (2009): 1-58.

[5] Chandola, Varun, Arindam Banerjee, and Vipin Kumar. ”Anomaly detection for discrete sequences: A survey.” IEEE transactions on knowledge and data engineering 24, no. 5 (2010): 823-839.

[6] Habeeb, Riyaz Ahamed Ariyaluran, Fariza Nasaruddin, Abdullah Gani, Ibrahim Abaker Targio Hashem, Ejaz Ahmed, and Muhammad Imran. ”Real-time big data processing for anomaly detection: A survey.” International Journal of Information Management 45 (2019): 289-307.

[7] Bhuyan, Monowar H., Dhruba Kumar Bhattacharyya, and Jugal K. Kalita. ”Network anomaly detection: methods, systems and tools.” Ieee communications surveys tutorials 16, no. 1 (2013): 303-336.

[8] Fiore, Ugo, Francesco Palmieri, Aniello Castiglione, and Alfredo De Santis. ”Network anomaly detection with the restricted Boltzmann machine.” Neurocomputing 122 (2013): 13-23.

[9] Kang, Myeongsu. ”Machine learning: Anomaly detection.” Prognostics and health management of electronics: fundamentals, machine learning, and the internet of things (2018): 131-162.

[10] Cortes, Corinna, and Vladimir Vapnik. ”Support-vector networks.” Machine learning 20, no. 3 (1995): 273-297.

[11] Sutrisno, Edwin, Qingguo Fan, Diganta Das, and Michael Pecht. ”Anomaly detection for insulated gate bipolar transistor (IGBT) under power cycling using principal component analysis and K-nearest neighbor algorithm.” Journal of the Washington Academy of Sciences (2012): 1-8.

[12] Xiong, Long, Hao-Dong Ma, Hong-Zheng Fang, Ke-Xu Zou, and Da-Wei Yi. ”Anomaly detection of spacecraft based on least squares support vector machine.” In 2011 Prognostics and System Health Managment Confernece, pp. 1-6. IEEE, 2011.

[13] Das, Santanu, Bryan L. Matthews, and Robert Lawrence. ”Fleet level anomaly detection of aviation safety data.” In 2011 IEEE Conference on Prognostics and Health Management, pp. 1-10. IEEE, 2011.

[14] Williams, Graham, Rohan Baxter, Hongxing He, Simon Hawkins, and Lifang Gu. ”A comparative study of RNN for outlier detection in data mining.” In 2002 IEEE International Conference on Data Mining, 2002. Proceedings., pp. 709-712. IEEE, 2002.

[15] De Stefano, Claudio, Carlo Sansone, and Mario Vento. ”To reject or not to reject: that is the question-an answer in case of neural classifiers.” IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 30, no. 1 (2000): 84-94.

[16] Hawkins, Simon, Hongxing He, Graham Williams, and Rohan Baxter. ”Outlier detection using replicator neural networks.” In International Conference on Data Warehousing and Knowledge Discovery, pp. 170-180. Springer, Berlin, Heidelberg, 2002.

[17] Ho, Tin Kam. ”Random decision forests.” In Proceedings of 3rd international conference on document analysis and recognition, vol. 1, pp. 278-282. IEEE, 1995.

[18] Aggarwal, Charu C. ”An introduction to outlier analysis.” In Outlier analysis, pp. 1-34. Springer, Cham, 2017.

[19] Akoglu, Leman, Hanghang Tong, and Danai Koutra. ”Graph based anomaly detection and description: a survey.”

Data mining and knowledge discovery 29, no. 3 (2015): 626-688.

[20] Boukerche, Azzedine, Lining Zheng, and Omar Alfandi. ”Outlier detection: Methods, models, and classification.” ACM Computing Surveys (CSUR) 53, no. 3 (2020): 1-37.

[21] Pang, Guansong, Chunhua Shen, Huidong Jin, and Anton van den Hengel. ”Deep weakly-supervised anomaly detection.” arXiv preprint arXiv:1910.13601 (2019).

[22] Simonyan, Karen, and Andrew Zisserman. ”Very deep convolutional networks for large-scale image recognition.” arXiv preprint arXiv:1409.1556 (2014).

[23] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. ”Imagenet classification with deep convolutional neural networks.” Communications of the ACM 60, no. 6 (2017): 84-90.

[24] Jiang, Xinwei, Junbin Gao, Xia Hong, and Zhihua Cai. ”Gaussian processes autoencoder for dimensionality reduction.” In Pacific-asia conference on knowledge discovery and data mining, pp. 62-73. Springer, Cham, 2014.

[25] Theis, Lucas, Wenzhe Shi, Andrew Cunningham, and Ferenc Huszár. ”Lossy image compression with compressive autoencoders.” arXiv preprint arXiv:1703.00395 (2017).

[26] Schlegl, Thomas, Philipp Seeböck, Sebastian M. Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. ”Unsupervised anomaly detection with generative adversarial networks to guide marker discovery.” In International conference on information processing in medical imaging, pp. 146-157. Springer, Cham, 2017.

[27] Ramaswamy, Sridhar, Rajeev Rastogi, and Kyuseok Shim. ”Efficient algorithms for mining outliers from large data sets.” In Proceedings of the 2000 ACM SIGMOD international conference on Management of data, pp. 427-438. 2000.

[28] Pang, Guansong, Kai Ming Ting, and David Albrecht. ”LeSiNN: Detecting anomalies by identifying least similar nearest neighbours.” In 2015 IEEE international conference on data mining workshop (ICDMW), pp. 623-630. IEEE, 2015.

[29] Zhang, Ke, Marcus Hutter, and Huidong Jin. ”A new local distance-based outlier detection approach for scattered real-world data.” In Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 813-822. Springer, Berlin, Heidelberg, 2009.

[30] Tax, David MJ, and Robert PW Duin. ”Support vector data description.” Machine learning 54, no. 1 (2004): 45-66.

[31] Sarker, Iqbal H. ”Machine learning: Algorithms, real-world applications and research directions.” SN Computer Science 2, no. 3 (2021): 1-21.

[32] Chalapathy, Raghavendra, and Sanjay Chawla. ”Deep learning for anomaly detection: A survey.” arXiv preprint arXiv:1901.03407 (2019).

[33] Zhang, Mingyang, Tong Li, Yue Yu, Yong Li, Pan Hui, and Yu Zheng. ”Urban Anomaly Analytics: Description, Detection, and Prediction.” IEEE Transactions on Big Data 8, no. 3 (2020): 809-826.

[34] Song, Hongchao, Zhuqing Jiang, Aidong Men, and Bo Yang. ”A hybrid semi-supervised anomaly detection model for high-dimensional data.” Computational intelligence and neuroscience 2017 (2017).

[35] Mohammadi, Mehdi, Ala Al-Fuqaha, Sameh Sorour, and Mohsen Guizani. ”Deep learning for IoT big data and streaming analytics: A survey.” IEEE Communications Surveys Tutorials 20, no. 4 (2018): 2923-2960.

[36] Kiran, B. Ravi, Dilip Mathew Thomas, and Ranjith Parakkal. ”An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos.” Journal of Imaging 4, no. 2 (2018): 36.

[37] Zheng, Yu, Licia Capra, Ouri Wolfson, and Hai Yang. ”Urban computing: concepts, methodologies, and applications.” ACM Transactions on Intelligent Systems and Technology (TIST) 5, no. 3 (2014): 1-55.

[38] Yu, Yingbing. ”A survey of anomaly intrusion detection techniques.” Journal of Computing Sciences in Colleges 28, no. 1 (2012): 9-17.

[39] Tsai, Chih-Fong, Yu-Feng Hsu, Chia-Ying Lin, and Wei-Yang Lin. ”Intrusion detection by machine learning: A review.” expert systems with applications 36, no. 10 (2009): 11994-12000.

[40] Kwon, Donghwoon, Hyunjoo Kim, Jinoh Kim, Sang C. Suh, Ikkyun Kim, and Kuinam J. Kim. ”A survey of deep learning-based network anomaly detection.” Cluster Computing 22, no. 1 (2019): 949-961.

[41] Fernandes, Gilberto, Joel JPC Rodrigues, Luiz Fernando Carvalho, Jalal F. Al-Muhtadi, and Mario Lemes Proença. ”A comprehensive survey on network anomaly detection.” Telecommunication Systems 70, no. 3 (2019): 447-489.

[42] Pang, Guansong, Chunhua Shen, Longbing Cao, and Anton Van Den Hengel. ”Deep learning for anomaly detection: A review.” ACM Computing Surveys (CSUR) 54, no. 2 (2021): 1-38.

[43] Nayak, Rashmiranjan, Umesh Chandra Pati, and Santos Kumar Das. ”A comprehensive review on deep learning-based methods for video anomaly detection.” Image and Vision Computing 106 (2021): 104078.

[44] Al-amri, Redhwan, Raja Kumar Murugesan, Mustafa Man, Alaa Fareed Abdulateef, Mohammed A. Al-Sharafi, and Ammar Ahmed Alkahtani. ”A review of machine learning and deep learning techniques for anomaly detection in IoT data.” Applied Sciences 11, no. 12 (2021): 5320.

[45] Nassif, Ali Bou, Manar Abu Talib, Qassim Nasir, and Fatima Mohamad Dakalbab. ”Machine learning for anomaly detection: A systematic review.” Ieee Access 9 (2021): 78658-78700.

[46] Liu, Jiangguo, Jianli Gu, Huishu Li, and Kenneth H. Carlson. ”Machine learning and transport simulations for groundwater anomaly detection.” Journal of Computational and Applied Mathematics 380 (2020): 112982.

[47] Quatrini, Elena, Francesco Costantino, Giulio Di Gravio, and Riccardo Patriarca. ”Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities.” Journal of Manufacturing Systems 56 (2020): 117-132.

[48] Hasan, Mahmudul, Md Milon Islam, Md Ishrak Islam Zarif, and M. M. A. Hashem. ”Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches.” Internet of Things 7 (2019): 100059.

[49] Alrashdi, Ibrahim, Ali Alqazzaz, Esam Aloufi, Raed Alharthi, Mohamed Zohdy, and Hua Ming. ”Ad-iot: Anomaly detection of iot cyberattacks in smart city using machine learning.” In 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), pp. 0305-0310. IEEE, 2019.

[50] Bhatia, Randeep, Steven Benno, Jairo Esteban, T. V. Lakshman, and John Grogan. ”Unsupervised machine learning for network-centric anomaly detection in IoT.” In Proceedings of the 3rd acm conext workshop on big data, machine learning and artificial intelligence for data communication networks, pp. 42-48. 2019.

[51] Wang, Jingwen, Jingxin Liu, Juntao Pu, Qinghong Yang, Zhongchen Miao, Jian Gao, and You Song. ”An anomaly prediction framework for financial IT systems using hybrid

machine learning methods.” Journal of Ambient Intelligence and Humanized Computing (2019): 1-10.

[52] Chen, Xiaoliang, Baojia Li, Roberto Proietti, Zuqing Zhu, and SJ Ben Yoo. ”Self-taught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks.” Journal of Lightwave Technology 37, no. 7 (2019): 1742-1749.

[53] Rejeb, Ridha, Mark S. Leeson, and Roger J. Green. ”Fault and attack management in all-optical networks.” IEEE Communications Magazine 44, no. 11 (2006): 79-86.

[54] Selvaraj, Aravinthkumar, Rizwan Patan, Amir H. Gandomi, Ganesh Gopal Deverajan, and Manjula Pushparaj. ”Optimal virtual machine selection for anomaly detection using a swarm intelligence approach.” Applied soft computing 84 (2019): 105686.

[55] Zhou, Joey Tianyi, Jiawei Du, Hongyuan Zhu, Xi Peng, Yong Liu, and Rick Siow Mong Goh. ”Anomalynet: An anomaly detection network for video surveillance.” IEEE Transactions on Information Forensics and Security 14, no. 10 (2019): 2537-2550.

[56] Lu, Cewu, Jianping Shi, and Jiaya Jia. ”Abnormal event detection at 150 fps in matlab.” In Proceedings of the IEEE international conference on computer vision, pp. 2720-2727. 2013.

[57] Mahadevan, Vijay, Weixin Li, Viral Bhalodia, and Nuno Vasconcelos. ”Anomaly detection in crowded scenes.” In 2010 IEEE computer society conference on computer vision and pattern recognition, pp. 1975-1981. IEEE, 2010.

[58] Unusual Crowd Activity Dataset of University of Minnesota, Available From. Accessed: Sep. 2016. [Online]. Available: http://mha.cs.umn.edu/ movies/crowdactivity-all.avi

[59] Pustokhina, Irina V., Denis A. Pustokhin, Thavavel Vaiyapuri, Deepak Gupta, Sachin Kumar, and K. Shankar. ”An automated deep learning based anomaly detection in pedestrian walkways for vulnerable road users safety.” Safety science 142 (2021): 105356.

[60] Murugan, B. S., Mohamed Elhoseny, K. Shankar, and J. Uthayakumar. ”Region-based scalable smart system for anomaly detection in pedestrian walkways.” Computers Electrical Engineering 75 (2019): 146-160.

[61] Ilyas, Zirgham, Zafar Aziz, Tehreem Qasim, Naeem Bhatti, and Muhammad Faisal Hayat. ”A hybrid deep network based approach for crowd anomaly detection.” Multimedia Tools and Applications 80, no. 16 (2021): 24053-24067.

[62] Mehta, Parth, Atulya Kumar, and Shivani Bhattacharjee. ”Fire and gun violence based anomaly detection system using deep neural networks.” In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 199-204. IEEE, 2020.

[63] Aboah, Armstrong. ”A vision-based system for traffic anomaly detection using deep learning and decision trees.” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4207-4212. 2021.

[64] Pramanik, Anima, Sobhan Sarkar, and J. Maiti. ”A real-time video surveillance system for traffic pre-events detection.” Accident Analysis Prevention 154 (2021): 106019.

[65] Cambridge, U., 2007. Camseq01 Dataset: Cambridge Labeled Objects in Video. University of Cambridge, UK. http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamSeq01/

[66] Islab-pvd: Illegally Parked Vehicle Dataset. http://mi.eng.cam.ac. uk/research/projects/VideoRec/-CamSeq01/

[67] Chen, Xiaoliang, Baojia Li, Roberto Proietti, Zuqing Zhu, and SJ Ben Yoo. ”Self-taught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks.” Journal of Lightwave Technology 37, no. 7 (2019): 1742-1749.

[68] Shriram, S., and E. Sivasankar. ”Anomaly detection on shuttle data using unsupervised learning techniques.” In 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), pp. 221-225. IEEE, 2019.

[69] Hasan, Mahmudul, Md Milon Islam, Md Ishrak Islam Zarif, and M. M. A. Hashem. ”Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches.” Internet of Things 7 (2019): 100059.

[70] Zhang, James, Robert Gardner, and Ilija Vukotic. ”Anomaly detection in wide area network meshes using two machine learning algorithms.” Future Generation Computer Systems 93 (2019): 418-426.

[71] Bhatia, Randeep, Steven Benno, Jairo Esteban, T. V. Lakshman, and John Grogan. ”Unsupervised machine learning for network-centric anomaly detection in IoT.” In Proceedings of the 3rd acm conext workshop on big data, machine learning and artificial intelligence for data communication networks, pp. 42-48. 2019.

[72] Wang, Jingwen, Jingxin Liu, Juntao Pu, Qinghong Yang, Zhongchen Miao, Jian Gao, and You Song. ”An anomaly prediction framework for financial IT systems using hybrid machine learning methods.” Journal of Ambient Intelligence and Humanized Computing (2019): 1-10.

[73] Yihunie, Fekadu, Eman Abdelfattah, and Amish Regmi. ”Applying machine learning to anomaly-based intrusion detection systems.” In 2019 IEEE Long Island Systems, Applications and Technology Conference (LISAT), pp. 1-5. IEEE, 2019.

[74] Alrashdi, Ibrahim, Ali Alqazzaz, Esam Aloufi, Raed Alharthi, Mohamed Zohdy, and Hua Ming. ”Ad-iot: Anomaly detection of iot cyberattacks in smart city using machine learning.” In 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), pp. 0305-0310. IEEE, 2019.

[75] Khokhlov, Igor, Michael Perez, and Leon Reznik. ”Machine learning in anomaly detection: Example of colluded applications attack in android devices.” In 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), pp. 1328-1333. IEEE, 2019.

[76] Punmiya, Rajiv, Olga Zyabkina, Sangho Choe, and Jan Meyer. ”Anomaly detection in power quality measurements using proximity-based unsupervised machine learning techniques.” In 2019 Electric Power Quality and Supply Reliability Conference (PQ) 2019 Symposium on Electrical Engineering and Mechatronics (SEEM), pp. 1-6. IEEE, 2019.

[77] Liu, Yu, Zhibo Pang, Magnus Karlsson, and Shaofang Gong. ”Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control.” Building and Environment 183 (2020): 107212.

[78] Barua, Anomadarshi, Deepan Muthirayan, Pramod P. Khargonekar, and Mohammad Abdullah Al Faruque. ”Hierarchical temporal memory based machine learning for real-time, unsupervised anomaly detection in smart grid: WiP abstract.” In 2020 ACM/IEEE 11th International Conference on Cyber-Physical Systems (ICCPS), pp. 188-189. IEEE, 2020.

[79] Liu, Jiangguo, Jianli Gu, Huishu Li, and Kenneth H. Carlson. ”Machine learning and transport simulations for groundwater anomaly detection.” Journal of Computational and Applied Mathematics 380 (2020): 112982.

[80] Aziz, Zafar, Naeem Bhatti, Hasan Mahmood, and Muhammad Zia. ”Video anomaly detection and localization based

on appearance and motion models.” Multimedia Tools and Applications 80, no. 17 (2021): 25875-25895.

[81] Ilyas, Zirgham, Zafar Aziz, Tehreem Qasim, Naeem Bhatti, and Muhammad Faisal Hayat. ”A hybrid deep network based approach for crowd anomaly detection.” Multimedia Tools and Applications 80, no. 16 (2021): 24053-24067.

[82] Ganokratanaa, Thittaporn, Supavadee Aramvith, and Nicu Sebe. ”Unsupervised anomaly detection and localization based on deep spatiotemporal translation network.” IEEE Access 8 (2020): 50312-50329.

[83] Qasim, Tehreem, and Naeem Bhatti. ”A hybrid swarm intelligence based approach for abnormal event detection in crowded environments.” Pattern Recognition Letters 128 (2019): 220-225.

[84] Alafif, Tarik, Bander Alzahrani, Yong Cao, Reem Alotaibi, Ahmed Barnawi, and Min Chen. ”Generative adversarial network based abnormal behavior detection in massive crowd videos: a hajj case study.” Journal of Ambient Intelligence and Humanized Computing 13, no. 8 (2022): 4077-4088.

[85] Pustokhina, Irina V., Denis A. Pustokhin, Thavavel Vaiyapuri, Deepak Gupta, Sachin Kumar, and K. Shankar. ”An automated deep learning based anomaly detection in pedestrian walkways for vulnerable road users safety.” Safety science 142 (2021): 105356.

[86] Aboah, Armstrong. ”A vision-based system for traffic anomaly detection using deep learning and decision trees.” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4207-4212. 2021.

[87] Mehta, Parth, Atulya Kumar, and Shivani Bhattacharjee. ”Fire and gun violence based anomaly detection system using deep neural networks.” In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 199-204. IEEE, 2020.

[88] Pang, Guansong, Cheng Yan, Chunhua Shen, Anton van den Hengel, and Xiao Bai. ”Self-trained deep ordinal regression for end-to-end video anomaly detection.” In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 12173-12182. 2020.

[89] Xu, Jingxin, Simon Denman, Clinton Fookes, and Sridha Sridharan. ”Unusual scene detection using distributed behaviour model and sparse representation.” In 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance, pp. 48-53. IEEE, 2012.

[90] Lu, Cewu, Jianping Shi, and Jiaya Jia. ”Abnormal event detection at 150 fps in matlab.” In Proceedings of the IEEE international conference on computer vision, pp. 2720-2727. 2013.

[91] Kumar, Vijay S., Tianyi Wang, Kareem S. Aggour, Pengyuan Wang, Philip J. Hart, and Weizhong Yan. ”Big data analysis of massive PMU datasets: A data platform perspective.” In 2021 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT), pp. 1-5. IEEE, 2021.

[92] [Online] https://www.kaggle.com/datasets/divyansh22/intel-berkeley-research-lab-sensor-data, accessed on 20.09.2022

[93] Peng, Yuhuai, Aiping Tan, Jingjing Wu, and Yuanguo Bi. ”Hierarchical edge computing: A novel multi-source multi-dimensional data anomaly detection scheme for industrial Internet of Things.” IEEE Access 7 (2019): 111257-111270.

[94] Qiu, Juan, Qingfeng Du, and Chongshu Qian. ”Kpi-tsad: A time-series anomaly detector for kpi monitoring in cloud applications.” Symmetry 11, no. 11 (2019): 1350.

[95] Fahim, Muhammad, and Alberto Sillitti. ”Anomaly detection, analysis and prediction techniques in iot environment: A systematic literature review.” IEEE Access 7 (2019): 81664-81681.

[96] Santos, José, Philip Leroux, Tim Wauters, Bruno Volckaert, and Filip De Turck. ”Anomaly detection for smart city applications over 5g low power wide area networks.” In NOMS 2018-2018 IEEE/IFIP Network Operations and Management Symposium, pp. 1-9. IEEE, 2018.

[97] Wang, Hongzhi, Mohamed Jaward Bah, and Mohamed Hammad. ”Progress in outlier detection techniques: A survey.” Ieee Access 7 (2019): 107964-108000.

[98] Dong, Yue, and Nathalie Japkowicz. ”Threaded ensembles of autoencoders for stream learning.” Computational Intelligence 34, no. 1 (2018): 261-281.

[99] Lee, In. ”Big data: Dimensions, evolution, impacts, and challenges.” Business horizons 60, no. 3 (2017): 293-303.

[100] Fossaceca, John M., and Stuart H. Young. ”Artificial intelligence and machine learning for future army applications.” In Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IX, vol. 10635, pp. 8-25. SPIE, 2018.

[101] Bao, Zhenjie, and Ru Xue. ”Survey on deep learning applications in digital image security.” Optical Engineering 60, no. 12 (2022): 120901.

[102] Yuan, Fuh-Gwo, Sakib Ashraf Zargar, Qiuyi Chen, and Shaohan Wang. ”Machine learning for structural health monitoring: challenges and opportunities.” Sensors and smart structures technologies for civil, mechanical, and aerospace systems 2020 11379 (2020): 1137903.

[103] Wang, Ge. ”X-ray imaging meets deep learning.” In Developments in X-Ray Tomography XIII, vol. 11840, p. 1184002. SPIE, 2021.

[104] Li, Xuefei, Hongyun Cai, Zi Huang, Yang Yang, and Xiaofang Zhou. ”Spatio-temporal event modeling and ranking.” In International Conference on Web Information Systems Engineering, pp. 361-374. Springer, Berlin, Heidelberg, 2013.

[105] Zhao, Hang, Yujing Wang, Juanyong Duan, Congrui Huang, Defu Cao, Yunhai Tong, Bixiong Xu, Jing Bai, Jie Tong, and Qi Zhang. ”Multivariate time-series anomaly detection via graph attention network.” In 2020 IEEE International Conference on Data Mining (ICDM), pp. 841-850. IEEE, 2020.

[106] Koizumi, Yuma, Masahiro Yasuda, Shin Murata, Shoichiro Saito, Hisashi Uematsu, and Noboru Harada. ”Spidernet: Attention network for one-shot anomaly detection in sounds.” In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 281-285. IEEE, 2020.