Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 05 Issue 02
2023

Research Article
https://doi.org/10.56979/502/2023

Integrating Machine Learning and Deep Learning Approaches for Efficient Malware Detection in IoT-Based Smart Cities

Shah Hussain Bangash[1*], Daud Khan[1], Atif Ishtiaq[1], Muhammad Imad[2], Mohsin Tahir[1], Waqas Ahmad[1], Ghassan Husnain[3], and Latif Jan[1]

[1]Iqra National University, Peshawar, Pakistan.
[2]Ulster University, London, United Kingdom.
[3]Cecos University, Peshawar, Pakistan.
*Corresponding Author: Shah Hussain. Email: shahhussain@inu.edu.pk

Received: July 29, 2023 **Accepted:** September 06, 2023 **Published:** September 17, 2023

Abstract: Smart cities have gained popularity because they promise to address some of the biggest challenges facing urban areas today, such as: traffic congestion, air pollution, energy consumption, waste management, and public safety. A comprehensive study is conducted to enhance the malware detection performance in smart cities by integrating machine learning and IoT-based approaches with deep learning. The study aims to address future challenges in malware detection and improve the effectiveness of strategies used in smart cities. Machine learning algorithms are applied to analyze and classify models' performance, enhancing computation time and categorial attacks. Deep learning techniques are commended to improve the accuracy and efficiency of malware detection in smart cities. The integration of IoT-based approaches and deep learning enables the detection of various types of malwares in smart cities. The study emphasizes the need for continuous research and development to enhance the performance of malware detection methods in the dynamic ecosystem of smart cities. The dataset was developed in a Unix/Linux-based virtual machine for classification purposes and is safe to use with malware software for Android devices based on the characteristics of the observations. 35 features and 100,000 observation data make up the data set. The results show promising results in terms of detecting malware in smart cities IoT devices.

Keywords: Malware Detection; IoT Malware; Support Vector Machine; K-nearest neighbor's; Decision Tree; Deep Learning CNN Model.

1. Introduction

Malware detection facing various challenges such as behavioral analysis, cloud and mobile malware, automated incident response, human machine collaboration, explainable AI, adversarial machine learning, privacy and preserving detection, multi-model detection [1-3]. Malware detection today is a big challenging paradigm of computing ability to the advancement in Information and Communications Technology (ICT) has changed the entire paradigm of computing [4-5]. Malware detection research challenges many different types of malwares such as, variability of malware, evasion techniques, false positives, complexity of modern systems, machine learning, multi-platform, Zero-day attack, advanced persistent threats (APT) [6-8]. The above mention malwares its own unique behaviors and characteristics to detect all the types of malwares based various methods [9]. Variability of malware detection is a difficult task to make single detection techniques [10]. The Evasion methods still developing new software to detect evade detection through security software [11]. These methods consistent anti-debugging, code obfuscation and packing [12]. Another useful method false positives malware leads legitimate and flagged by the software [13]. The modern system complexity is a challenging task to detect the system behavior and understand the main issues of malware [14]. The modern world researchers and developers using machine learning techniques to determine unknown malware detection [15]. The malware attackers using multiple platforms like IOS, Android, Mac, Linux and windows and so on [16]. The advanced persistent threats (APT) are a cyber-attack that is spotlight highly targeted people [17]. The malware attacks have many types but some of them highlighted and popular throughout the world such as worm, trojan, virus, ransomware, adware, rootkit, spyware, banking malware, fileless malware, crypto jacking malware [18-20]. Worm malware attacks spreads with the help of network often break the security vulnerabilities [21]. Another malware attack trojan allows the attackers to access legitimate software infected system [22]. Virus is a small program or file effected your software and data replicates the infected files is executed 23]. Ransomware is also a type of malware that encrypts the data, files and software restore access [24]. Another type of malware adware shows the unwanted advertisements on your system infected [25]. Most of the attackers used spyware attack collects the personal information of the users and infected the id or system [26]. Rootkit is also type of advanced malware attack on the kernel level to control the whole system [27]. Banking malware specially designed to break the security credentials and information of the banking infected system [28]. Fileless and crypto jacking malware attack with the help of legitimate tools and processes cryptocurrency system resources infected system [29]. Heuristic-based malware detection focuses on detecting intrusions by monitoring system activity and classifying it as normal or abnormal [30]. Machine learning algorithms, rather than patterns or signatures, are frequently used in classification [31]. One of its shortcomings is that it has a high false positive rate, causing many legitimate actions to be classified as intrusive, and that it requires useful training data, which is typically difficult to obtain in large IT environments [32]. Modern host-based malware detection products concentrate on in-memory patterns. Aside from heuristics, they employ techniques such as block-hashing, which computes hashes of portions of the suspicious file rather than the entire file, or are capable of detecting polymorphic encrypted payloads in memory [33]. These malware detection products, on the other hand, are usually designed to look for exploitation and malware behavior, such as code patterns that exploit a vulnerability in a software product. Although (this partially) mitigates the risk of automatic/no interactive malware infections, such as a drive-by download or watering hole attack (which are typically triggered by exploitation of either a zero-day or well-known vulnerability), these products are less effective against malware launched with human interaction, such as by tricking a targeted user into starting the malware code himself during a (spear)phishing attempt [34]. If the guerilla band discovers that this type of host-based malware detection is being used, it is advised that general (automated) Mohurle, S., & Patil, M. (2017). A brief study of wannacry threat: Ransomware attack 2017. International Journal of Advanced Research in Computer Science, 8(5), 1938-1940 [35]. Through the integration of IoT devices, smart cities are altering urban settings and enabling effective resource management and services. The growth of networked gadgets, however, also makes these cities vulnerable to security flaws, notably malware attacks. Malware can damage the entire ecosystem by taking advantage of flaws in IoT devices. Therefore, effective malware detection methods are essential to guaranteeing the security of smart cities [36]. An overview of the most recent malware detection methods used in smart cities may be found in this section. Malware has been detected and categorized using a variety of ML algorithms, including neural networks, decision trees, and support vector machines [37]. In order to find harmful activity within IoT networks, additional techniques such as behavioral analysis, anomaly detection, and signature-based algorithms have been investigated. Researchers have looked into a variety of performance improvement techniques to boost malware detection in smart cities. Techniques such as feature engineering, ensemble learning, and transfer learning have been used to enhance the precision and effectiveness of detection models. Additionally, utilizing block chain technology for tamper-proof data exchange and edge computing for real-time analysis have showed promising results in improving malware detection performance [38].

2. Literature Review

The approaches of Malware detection recently a challenges task to analysis the presence of problems. The technology advancement changing communications information of the entire paradigm of the computing. Muhammad Wazid, and Ashok Kumar Das, investigated certain drawbacks in security and privacy with the help of internet of medical things (IoMT) such as, password guessing, impersonation, remote hijacking and other malware attacks. In presence the cyber-attack altered not easily accessible authorized users depends on the architecture of IoT environment and security protocols [39]. Dukka KarunKumar Reddy and Himansu Sekhar Behera, proposed to secure in the future smart cities applied deep learning techniques to detect tracking of applications in internet of things (IoT). The exposure system categorizes the behavior of authorized users' activities and untruthful actions [40]. Jueun Jeon, Jong Hyuk Park, applied neural network model to detect the dynamic analysis cloud-based malware detection. The exiting models detecting accurately attacks of the malicious new variant through static analyze the IoT code determination [41]. Nataliia Neshenko and Christelle Nader, proposed the survey to spread the awareness and supporting smart cities in the context of cyber-attacks with various malicious [42]. Seungyeon Baek, Jueun Jeon, studied hybrid malware detection analysis through deep learning methods in internet of things (IoT). In the presence, Internet of Things (IoT) devices fully functional wide range of services such as smart houses, smart factories, smart transportation, smart cities received the cybersecurity threats [43]. Ms. Purnima Ahirao and K J Somaiya, analyzed the malware attacks proactive methods to protect the smart cities based on static and dynamic analysis. The researcher emphasized to used different tools and techniques to detect and integrated the cloud to analyze the malware [44].

Table 1. Literature Review on highlighting the importance of different Malware types, Attacks, Countermeasures, Limitations and Future Challenges.

Authors	**Malware Type**	**Attacks Performed**	**Countermeasure**	**Weakness / Limitation**	**Future Challenges**
Muhammad Wazid, and Ashok Kumar Das et.al [9]	Spyware, Keylogger, Trojan Horse, Virus, Worm, Rootkit	Malware attacks in IoT/IoMT environment. Malware attacks launched by Mirai, Reaper, Echobot, Emotet, Gamut and Necurs botnets are active these days.	Network-based anomaly detection (N-BaIoT) to extract behavior snapshots by using deep auto-encoders	Drawbacks several security and privacy issues, such privileged-insider, remote hijacking, denial of service (DoS) attacks, and malware attacks.	Malware detection in IoT/IoMT environment should be improvement
Dukka KarunKumar Reddy and Himansu Sekhar Behera et.al [10]	Seven categorical attacks found in the Distributed Smart Space Orchestration System traffic traces data set	DoS Attack, Data Probing, Malicious control, Malicious Operation, Scan, Spying, Wrong Setup	Simulation model report that deep neural network architecture improvement.	Due to the mounting complexity of the infrastructure framework of IoT, it is lifting undesirable weakness to their systems.	Improvement in most of the categorical attack.
Jueun Jeon[1], Jong Hyuk Park et.al [11]	Malicious code detection	Distributed denial of service (DDoS), cryptocurrency malicious mining, and botnet activities	Implementation of a model that can detect IoT malware using the hybrid analysis technique,	Limitation of hardware resources	Utilizing both static and dynamic techniques, will be conducted in the future.
Nataliia Neshenko and	Computer viruses, remote breaks, eavesdropping,	Attacker spreads malware with the intent to infect smart	A survey of methods supporting cyber	Allocate time and budget effectively	Technological architecture of smart cities as

Christelle Nader et.al [12]	software hijacking, injection of malicious	sensors, IoT devices, or data servers	situational awareness		well as their cyber security challenges, requirements, cyber threats, and respective countermeasures.
Seungyeon Baek, Jueun Jeon et.al [13]	Distributed denial of service attacks	Static Malware Detection, Dynamic Malware Detection, Hybrid Malware Detection	Artificial intelligence, Various evasion techniques, Trained EfficientNet-B3 model	Limited memory	It is difficult to protect the myriads of IoT devices from advanced cyberattacks using conventional security methods
Tanzila Saba et.al [14]	Viruses and network intrusion	Denial of Service (DoS) Attack, Sinkhole Attack, Jamming,	Ensemble-based classifiers are applied to determine intrusion attacks on the networks.	Further need investigation the optimal techniques to cyber safety and security	The future work may extend to guarantee the cyber safety and security of smart city projects and ensure that the IoT devices used in smart city projects are secure.
Fadi Al-Turjman and **Hadi** Zahmatkesh et.al [15]	Malicious software	Malicious attackers may produce false data during the manipulation of sensing data, which results in the loss of control over the highly intelligent systems.	Intrusion detection and countermeasure of virtual cloud systems-state of the art and current challenges	Internal weaknesses of the system aiming to steal, change, and ruin physical system components and related information	The security of these cameras is a challenging task as some of them lack encryption algorithms and others are vulnerable to attack by malware

Pranshu Bajpai and Richard Enbody et.al [16]	Ransomware based extortion attacks	Elucidate the new ransomware strategies that the attackers	Checked for violation of a hard constraint	Hard constraint limits imposed cannot be circumvented	Improves the feasibility of the attack
Md Mamunur Rashid, Joarder Kamruzzaman et.al [17]	Stacking ensemble model, ensemble approach	Attacker who compromises these IoT devices may obtain sensitive data such as information of credit card, stream video and similar personal information.	A denial of access or privacy intrusion within an automated system can greatly harm individual citizens and carry a substantial cost at both individual and jurisdiction levels.	Limited onboard functionality for security operations and send captured data to cloud servers for processing.	Our future work will explore deep learning techniques to further enhance IoT attack detection performance.
Sancheng Peng, Lihong Cao et.al [18]	Mobile malicious software, McAfee mobile threat	It draws attackers who have delivered a large amount of malwares to unsuspecting users, due to its open nature.	To design more and more effective mechanisms to detect smartphone malware	Not best for the big data	Future of the application of deep learning for smartphone security.
Yuhan Chai and Jing Qiu et.al [19]	LGMal framework to realize the malware detection	Malicious attacks can evade detection through adversarial learning	Improve the robustness of the model.	Insufficient sampling and observation of malicious behaviors will inevitably limit the detection ability	Improve the prediction results of malware
Hamad Naeem , Farhan Ullah et.al [20]	Hybrid Image Visualization and Deep Learning Model	Mostly use certain signatures to smell the malware attacks.	Color image visualization and deep convolution neural network	The state-of - the-art malware identification methods are not better in terms of computational complexity.	In the future, we will try to develop a combined blockchain and machine learning memory less malware

					detection model.
Nan Zhang and Yu-an Tan et.al [21]	Static analysis, Dynamic analysis, Hybrid analysis	N/A	The use of TextCNN to auto-detect feature representations (word sequences of reports).	TC-Droid, does not require hand-engineered feature selection in the domain of Android malware detection.	In the future, the approach will be extended to both dynamic and static features. It is worth mentioning that dynamic features could be extracted in real devices.
Seyed Mehdi Hazrati Fard , Hadis Karimipour et.al [22]	Viruses, worms, Trojans, backdoors, and rootkits.	The attacker can easily create many polymorphic/metamorphic variants of any given malware sample to evade signature-based defense systems.	Ensemble MVL-based scheme based on SRC that shows a robust ability to detect malware compared with existing state-of-the-art learning-based models and other baselines	The signature-based methods are only good for detecting known malware.	Another potential direction is to extend this method to the other ML applications. Various approaches in real-world settings suffer from the nature of imbalanced data and the proposed algorithm can help us to cope with them.
Muhammad Shafiq and Zhihong Tian et.al [23]	Malicious, anomaly and intrusion	Bot-IoT Attacks	Proposed a new framework model and a hybrid algorithm to solve this problem.	The study is limited to machine learning techniques for IoT security such as IoT authentication anomaly and intrusion detection	Future research work with the recommendation for IoT security based on machine learning techniques.

Naercio Magaia, Ramon Fonseca et.al [24]	Malicious softwar, injecting malicious code	Attacker usually is to steal or tamper sensors' data and denial of service (DoS)	Using Blockchain Technology	Heterogeneous environment, there are still many challenges in defining new protocols, authentication methods as well as to keep privacy awareness at the same time.	The envision that the outcome will facilitate future research efforts in spreading new methods.
Mohammad Shari Aliabadi et.al [25]	Viruses, botnets	perform attacks such as DDoS against network services. DDoS attacks make network services inaccessible to users.	Using machine learning methods is that they have the ability to learn and recognize the pattern of attacks and provide good accuracy.	Centralized approaches have limited ability to process large volumes of tra7c and are vulnerable to DDoS attacks.	Meta-heuristic methods, and using multiple deep learning classiers in majority voting to detect an intrusion is more accurate than our future works
PRIMA Bouchaib [1], BOUHORMA Mohamed et.al [26]	Malicious binary files	N/A	Compared these image-based (DL) results to a simpler convolutional neural network (CNN) approach trained from scratch.	Dataset features are limited	Improve automatic detection and classification of the malwares.

2.1 Previous Studies Challenges

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn malware and other cyber threats, it's critical to be updated on cybersecurity best practices [45].

Table 2. Description of different algorithms used Malware Attacks

Attacks & Date	**Description**
ILOVE YOU (2000)	A worm that targeted Microsoft Windows computers transmitted by email. By overwriting files and propagating itself to other users, it created enormous disturbance.
Code Red (2001)	A worm that caused website defacement and the disruption of online services by taking advantage of a flaw in Microsoft Internet Information Services (IIS) web servers.

Slammer (2003)	A quick-moving worm that took use of a flaw in Microsoft SQL Server. It affected numerous online services and caused significant internet slowdowns.
Conficker (2008)	A computer virus that propagated over Windows systems by taking advantage of security holes and poor passwords. It presented severe security risks and built up a sizable botnet.
Stuxnet (2010)	An advanced worm that is thought to be a state-sponsored cyberattack aimed at Iran's nuclear facilities. Industrial control systems (ICS) and vital infrastructure were intended to be disrupted.
WannaCry (2017)	A ransomware assault that infected systems and encrypted files by using a Windows SMB vulnerability. Numerous systems, including those used by governments and healthcare providers, were compromised.
NotPetya (2017)	A malware attack that pretended to be a ransomware strike but was ultimately damaging. It expanded globally and affected Ukraine's crucial systems.
Equifax Data Breach (2017)	Although not a typical malware assault, this incident entailed the exploitation of a flaw in the Equifax system, which exposed millions of people's sensitive personal information.
SolarWinds Supply Chain Attack (2020)	An extremely sophisticated attack that involved compromising software upgrades in the Orion platform from SolarWinds. Threat actors were able to access multiple public and private sector networks as a result.
Colonial Pipeline Ransomware Attack (2021)	The Colonial Pipeline, a significant petroleum pipeline in the United States, was the victim of a ransomware attack. It caused fuel supply difficulties along the East Coast.
Kaseya Supply Chain Attack (2021)	Numerous firms were affected by a ransomware assault that targeted managed service providers (MSPs) and exploited a flaw in Kaseya's software.
PrintNightmare (2021)	A serious bug that could allow remote code execution in the Windows Print Spooler service. It raised worries about possible large-scale attacks.

The paper follows a systematic review methodology to examine the current malware detection methods in the context of securing smart cities. The study integrates machine learning algorithms and IoT-based approaches with deep learning to enhance malware detection performance in smart cities. Machine learning algorithms are utilized to analyze and classify models' performance, improving computation time and categorical attacks. The paper emphasizes the use of deep learning techniques to improve the accuracy and efficiency of malware detection in smart cities. The study also highlights the importance of multi-model detection techniques and robust evasion methods to defend against malware attacks in smart cities. The research methodology for malware detection involves a systematic approach, including developing trials, gathering data, and analyzing outcomes.

2.2 Contributions of the Paper:

The paper presents a comprehensive study on advancing malware detection performance in smart cities through the integration of machine learning and IoT-based approaches with deep learning.

- The study identifies future research challenges and suggests ways to enhance the performance of malware detection strategies in smart cities.
- It emphasizes the use of machine learning algorithms to analyze and classify models' performance, improving computation time and categorical attacks.
- The paper highlights the importance of leveraging deep learning techniques to improve the accuracy and efficiency of malware detection in smart cities.

2.3 Organization of the paper

The paper follows a systematic review approach to examine the current malware detection methods in the context of securing smart cities. It begins with an introduction section by highlighting the importance of malware detection in smart cities and the potential risks associated with malware attacks. The next section highlights the need for enhancing the performance of these strategies and identifies future research challenges in malware detection. Followed by the result section which emphasizes the use of deep learning and machine learning techniques and its multi-model detection approaches to improve the accuracy and efficiency of malware detection in smart cities. The paper concludes by emphasizing the importance of raising awareness about malware attacks and implementing robust defense mechanisms in smart cities. Overall, the paper provides a comprehensive examination of malware detection methods, future research directions, and strategies to enhance the performance of malware detection in smart cities.

3. Research Methodology

In order to discover and prevent harmful software (malware), research methodology for malware detection requires taking a systematic approach to developing trials, gathering data, and analyzing outcomes. Real-time monitoring and user engagement are provided by the Control Center, while improved city services are provided via External Integration.

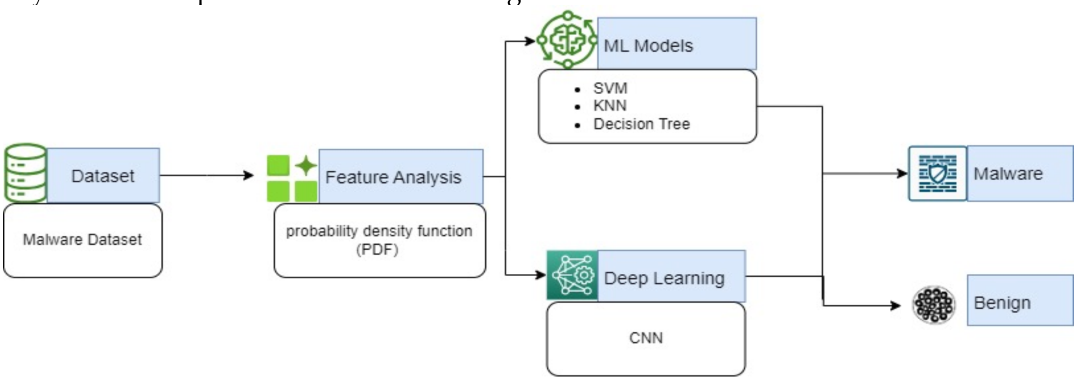

Figure 1. Process Architecture

Figure 1. represents data collection and preprocessing, where raw data is cleansed and modified, are part of the architecture of machine and deep learning processes. Model accuracy and efficiency are maintained by monitoring and feedback loops, and the process is continuously improved to increase overall performance.

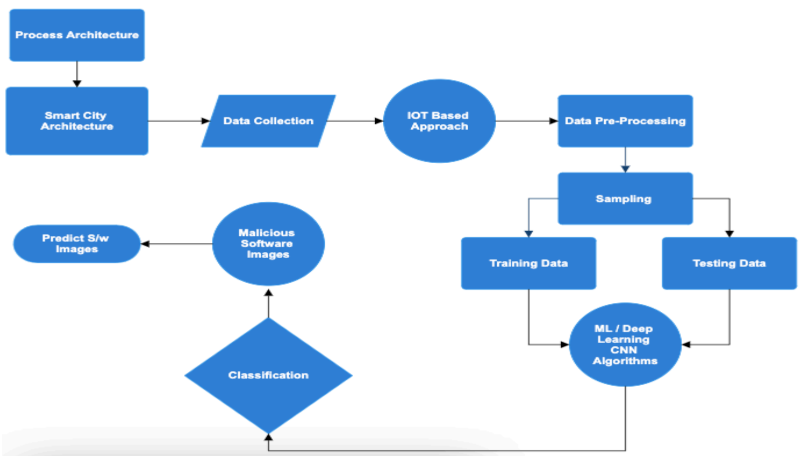

Figure 2. Block Diagram Smart City IOT Based Approach

The Figure 2. components of a Smart City IoT block diagram includes data collection, internet of things (IOT), data processing, and other phases working together. It's important phase to process and analyze the sensitive data through machine and deep learning algorithms. To assist in decision-making, central processing and analysis of smart city data is performed.

Malware Detection Dataset: Based on the characteristics of the observations, the dataset was created in a Unix / Lunix-based virtual machine for classification purposes, which are harmless with malware software for Android devices. The data set consists of 100,000 observation data and 35 features. Below is a table of specifications and descriptions.

Table 3. Features of Malware Dataset

Features	Description
Hash	APK/ SHA256 file name
Classification	malware/beign
State	flag of unrunable/runnable/stopped tasks.
usage_counter	task structure usage counter
Prio	keeps the dynamic priority of a process
static_prio	static priority of a process
normal_prio	priority without taking RT-inheritance into account
Policy	planning policy of the process
vm_pgoff	the offset of the area in the file, in pages.
vm_truncate_count	used to mark a vma as now dealt with
task_size	size of current task.
cached_hole_size	size of free address space hole.
free_area_cache	first address space hole

3.1 Research Process

The primary objective of the entire research effort was to carry out a study and thoroughly comprehend the readily available facts and data presented by earlier scholars. With the research project, we can examine the issues with malware detection by using creative qualitative research study from a variety of prior research articles on Google Scholar. The fundamental studies involved in the research process are data gathering and data analysis. However, discovering high-quality data leads to amazing research outcomes.

3.2 Data Process

The continuing study's primary goal was to investigate and deal with problems related to data processing methods. The targeted literature review set the goal of examining the field of malware detection and researching the sparse empirical diagnosis, along with qualitative research study approaches. The chosen methodology can be utilized to clarify and comprehend the cultural experiences, social experiences, and behavioral patterns of the related community. The goal of this research study is to probe extensively into the upcoming problems with plant diseases.

3.3 Data Collection

Every research methodology requires data collecting in order to thoroughly assess other research information and use the best data collection methods depending on the facts at hand. Data is gathered using a variety of methods, including surveys, telephone interviews, surveys, Wikipedia, transactional tracking, forms, social media monitoring, Kaggle, Google Scholar, and Google. On the other hand, figuring out the nature of an uncharted territory is a difficult undertaking.

3.4 Data Analysis

The procedure for analyzing data Approaching phrases or text is an essential phase in the transcribing and content analysis process. Sentences, single words, and paragraphs were all included in the qualitative content analysis, which typically used a single topic to communicate the entire document. In addition to being able to evaluate the theoretical hypotheses, the data analysis process aids in understanding and developing the procedures for data collection.

3.5 Dataset Description

Based on the characteristics of the observations, the dataset was created in a Unix / Lunix-based virtual machine for classification purposes, which are harmless with malware software for Android devices. The data set consists of 100,000 observation data and 35 features as shown in table 3.1. The malware dataset is available online via the Kaggle repository consists of malware images using for detection. In addition, 70% of the data was used for training purposes, with the remaining 30% used for testing [46-48]. We have selected malware datasets, and the dataset includes various features like shape, texture, smoothing etc.

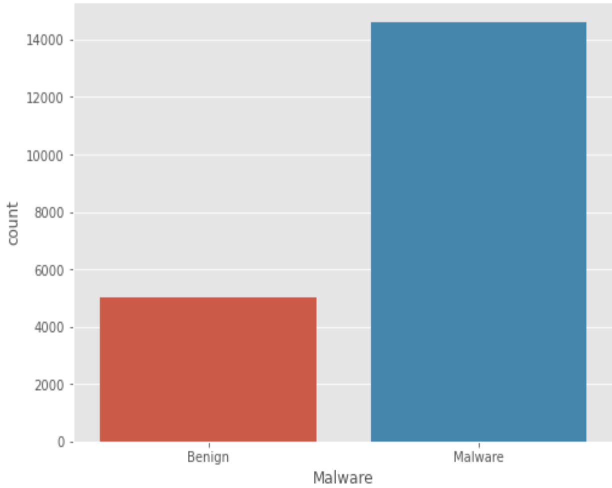

Figure 3. Comparative Benign and Malware the Performance

3.6 Feature Analysis for Malware Detection

A set of features, such as 'Major Subsystem Version' and 'Size of Code,' systematically generates probability density function (PDF) plots for each feature's distribution within the two classes [47] as shown in Figure 3. This analysis serves to visualize the differences in feature distributions between malicious ("Malware") and non-malicious ("Benign") software samples [49]. The PDF plots provide insights into potential discriminative features for malware detection, which is essential in cybersecurity research and threat analysis. Additionally, by using Seaborn's kernel density estimation with a specified bandwidth ('bw': 0.1), the code ensures that the PDF curves accurately represent the data's underlying distribution, facilitating more informed feature selection and classification model development for cybersecurity applications [50-51].

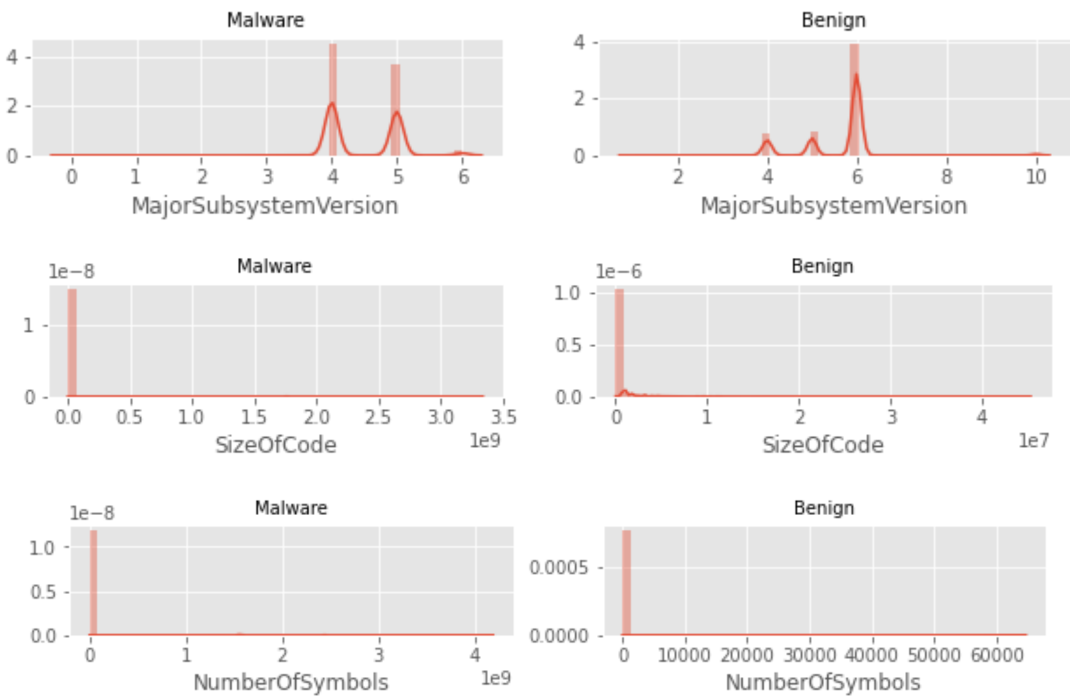

Figure 4. Feature Analysis Malware Detection

4. Results and Discussion

The paper focuses on enhancing malware detection performance and addressing future challenges in smart cities through the fusion of machine learning and IoT-based approaches, leveraging deep learning techniques. The study highlights the importance of analyzing and classifying models' performance using machine learning algorithms to improve computation time and categorical attacks. Deep learning algorithms are utilized to analyze malware images based on texture, shape, and smoothness, achieving high accuracy in classification results. The performance evaluations matrix reveals that out of 10,029 malware images, the deep learning algorithms correctly classified with high accuracy. The paper emphasizes the need for multi-model detection techniques and robust evasion methods to detect multiple types of malwares and defend against malware attacks in smart cities.

In this work, datasets of Malware to extract several features, such as shape features, texture features, smoothing features, etc. We divided the data into testing and training phases. The training data included 70% while testing selected data 30%. Deep learning has quickly become a game-changing technique with far-reaching ramifications in a variety of fields and applications. The capabilities of computers have been pushed by the outstanding findings produced by this potent branch of machine learning in a variety of fields. Here, we'll examine several noteworthy deep learning findings and their effects on various industries. The main objective of this work is to increase the accuracy while reducing computing costs, test and compare the performance of the algorithms and method further needed to investigate.

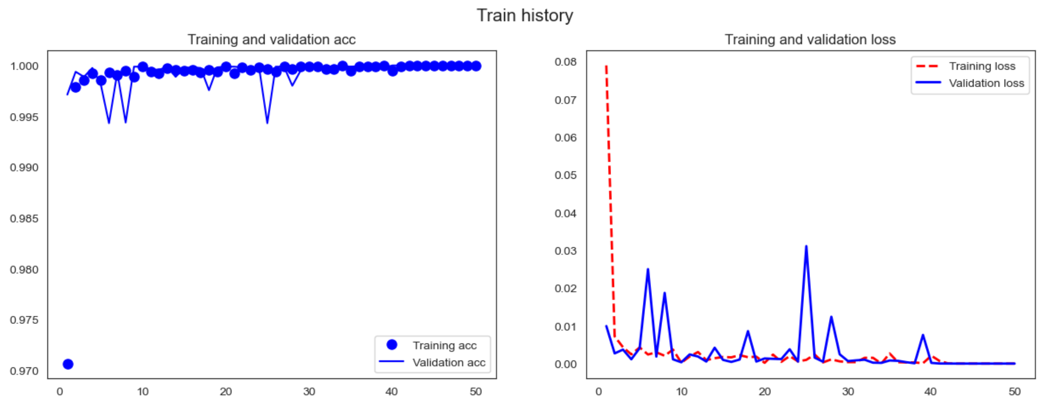

Figure 5. Training and validation accuracy of algorithms

The confusion matrix is most commonly used in machine learning to evaluate the performance of classification of the model. The confusion matrix specifies the most common matrices such is accuracy, precision, recall and F1-score.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad 1)$$

$$\text{Precision} = \frac{TP+TN}{TP+FP} \quad 2)$$

$$\text{F1-Score} = 2 * \frac{Precision*Recall}{Precision+Recall} \quad 3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad 4)$$

Table 4. Results of Deep Learning Classification Using CNN Model

Deep Learning Classification Report				
Classification	Accuracy	Precision	Recall	F1- score
Malware	1.00	1.00	1.00	1.00
Benign	1.00	1.00	1.00	1.00

Average Score	1.00	1.00	1.00

In this section, the deep learning algorithm was recycled to analyse malware images detection occupying on texture, shape, and smoothness. The table 4. depicts the confusion matrix for classification results achieved using deep learning algorithms. The performance evaluations matrix reveals that out of the 10029 malware images, the deep learning correctly classified with high accuracy.

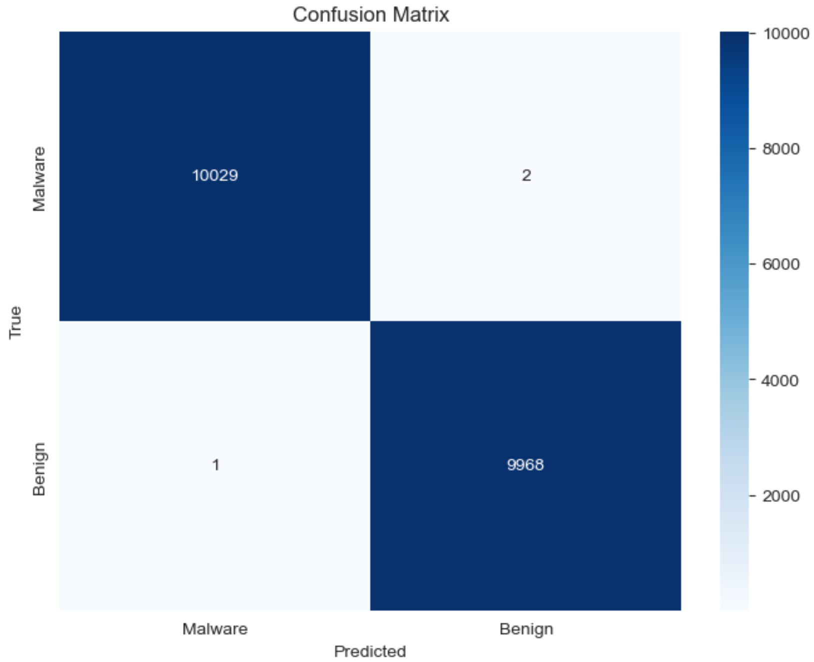

Figure 6. Performance evolution of Deep Learning using CNN

4.1 Deep learning utilizing Convolutional Neural Networks (CNN)

Figure 6 depicts the confusion matrix for categorization outcomes achieved using deep learning algorithms. The performance evaluations matrix reveals that, out of 10029 malware images, the deep learning correctly classified 10027 and incorrectly classified 2. alike, out of 1000 benign images, the deep learning successfully intimates 9968 of them.

4.2 Support Vector Machine (SVM) Performance Evolutions

This code segment encapsulates a structured workflow for building and evaluating an SVM (Support Vector Machine) classifier in binary classification tasks. Commencing with the essential step of importing libraries, it underscores the pivotal role of scikit-learn, seaborn, and matplotlib for machine learning, data visualization, and performance assessment. Subsequently, the code initiates an SVM classifier, signifying its adaptability by specifying a linear kernel for linearly separable data and setting the `C` parameter to balance margin maximization and classification error, ensuring reproducibility through the `random state` parameter.

Table 5. Results of SVM algorithm Classification

SVM Classification Report				
	Accuracy	Precision	Recall	F1-score
Malware		0.93	0.94	0.93
Benign	0.93	0.94	0.93	0.93
Average	0.93	0.93	0.93	

The values for accuracy, precision, recall, and F1-score for the SVM algorithm outcomes are shown in Table 5. In order to more accurately evaluate the effectiveness of the suggested model, we also evaluated it in a number of different ways as part of the deep learning algorithms. By combining all classifiers and giving it the name deep learning Classifier, we have added a fresh element to the deep learning model that has been proposed. We have also computed performance evaluation once again. The results of the comparison demonstrate that the deep learning classifier outperforms the other classification approaches.

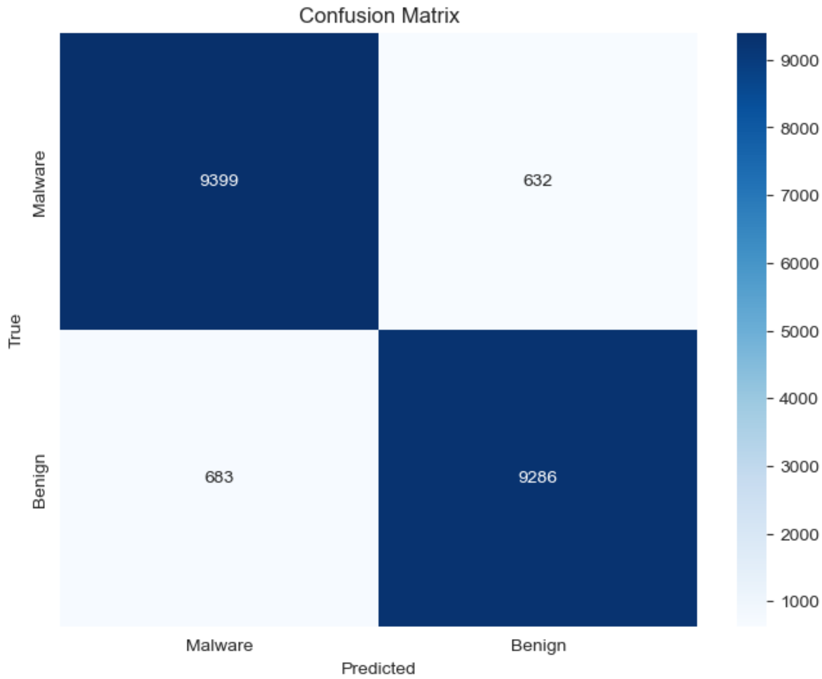

Figure 7. Performance evaluation matrix for SVM Algorithm

Figure 7 depicts the confusion matrix for classification results achieved using SVM classifiers. The performance evaluations matrix reveals that, out of 1000 malware images, the SVM correctly classified 683 and incorrectly classified 8. Similarly, out of 1000 benign images, the LDA successfully identified 9268 of them. The core training phase is emphasized, where the SVM classifier learns from labelled training data, establishing the foundation for robust predictions. Subsequent prediction generation on test data follows this training step, and the code proceeds to compute a confusion matrix. This matrix furnishes essential insights into the classifier's performance, delineating true positives, true negatives, false positives, and false negatives. Moreover, a comprehensive classification report is meticulously crafted, encompassing critical metrics like precision, recall, F1-score, and support for each class.

4.3 KNN Algorithm Performance Evolutions

The initial step involves importing crucial libraries like scikit-learn, seaborn, and matplotlib, emphasizing their significance in machine learning, data visualization, and model assessment. Subsequently, the code defines and initializes a KNN classifier, with flexibility to modify the number of neighbors for optimal performance.

Table 6. Results of KNN algorithm Classification

KNN Classification Report				
	Accuracy	**Precision**	**Recall**	**F1-score**
Malware	1.00	1.00	1.00	1.00
Benign		1.00	1.00	1.00
Average		1.00	1.00	1.00

Table 6. performance was improved in the KNN of the proposed work by combining various algorithms. As previously mentioned, the KNN classifier model compare the results with deep learning algorithm. The data above shows that we were successful in obtaining the outcomes of both features malware and benign. As a result of our success to achieve the highest accuracy of the models.

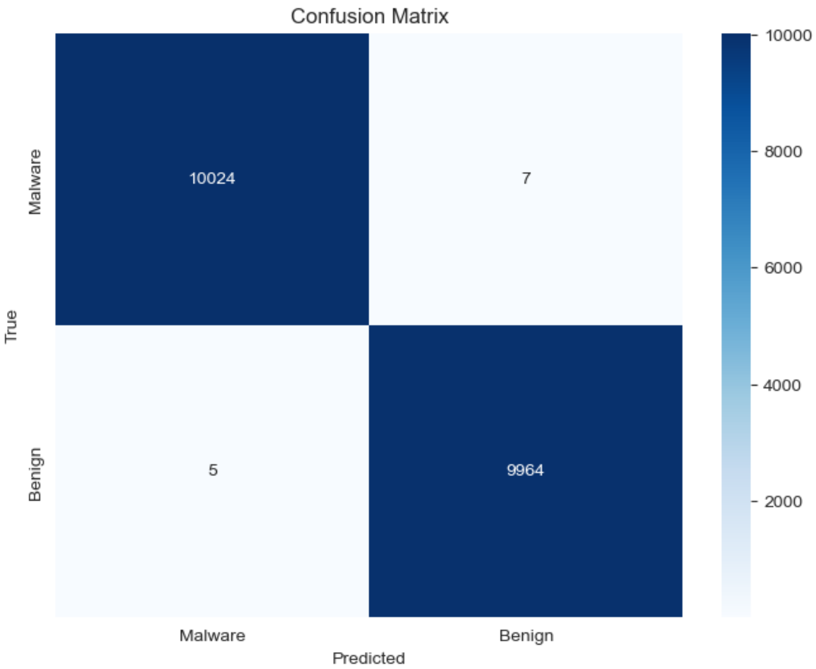

Figure 8. Results of KNN algorithm Classification

Table 7. Results of DT algorithm using Classification

DT Classification Report				
	Accuracy	**Precision**	**Recall**	**F1-score**
Malware	1.00	1.00	1.00	1.00
Benign		1.00	1.00	1.00
Average		1.00	1.00	1.00

Confusion matrix for classification outcomes obtained using KNN is shown in Figure 4.4. According to the performance assessments matrix, the KNN correctly categorized 10024 malware images while wrongly classifying 5. Similar to that, of 9964 benign photos were correctly classified by the KNN. The training phase is pivotal, where the classifier learns from labeled training data to discern patterns and relationships between features and labels, laying the foundation for accurate predictions.

4.4 Decision Tree Algorithm Performance Evolutions

Decision Tree classifier's creation and training, where it learns to make informeddecisions based on the provided training data. This step is crucial in building a predictive model that can later be used for classifying new, unseen data. Additionally, the code's capability to generate a confusion matrix and classification report is vital for assessing the model's performance. The table 7 confusion matrix offers a detailed breakdown of true positive, true negative, false positive, and false negative predictions, facilitating a thorough evaluation of the classifier's accuracy and potential for misclassification.

In Figure 9, the accuracy of the results obtained for each method, the performance evaluation of the malware Image Classification is significantly better in [Deep learning algorithms] with a higher accuracy rate of 100% precession. The confusion matrix is most commonly used in deep learning to evaluate the performance of the classification model. The confusion matrix specifies the most common matric such is accuracy, precision, recall and F1-score. The classification report provides an extensive overview of key metrics like precision, recall, F1-score, and support for each class, offering researchers valuable insights into the classifier's strengths and weaknesses.

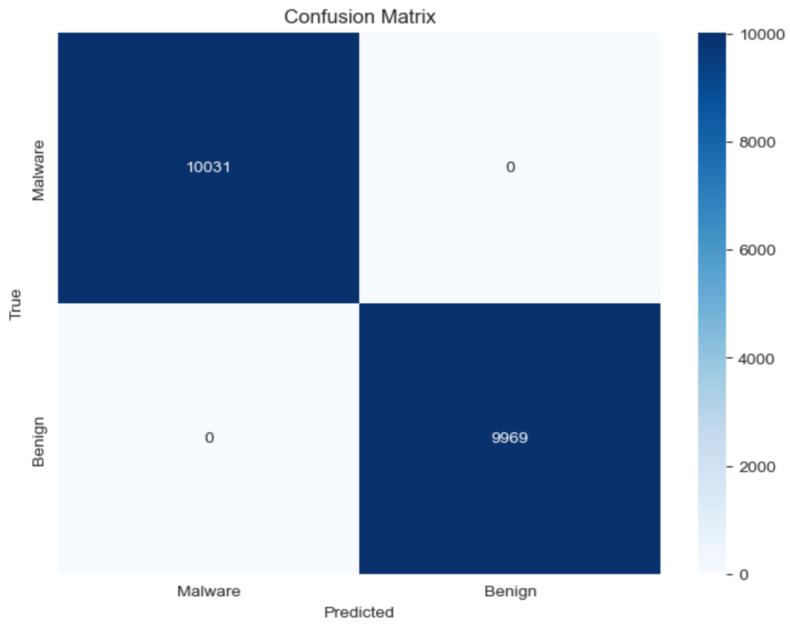

Figure 9. Results of Decision Tree Algorithm Classification

Efficiency of the method is carried out more precisely, as indicated in Table 4.5. Overall algorithms results were expanded to include the best outcome. The [Deep learning algorithm] have the other categorization methods' greatest accuracy rates as compare to other classifiers.

Table 8. Overall Algorithms Performance Evaluation Results

Techniques	Accuracy	Precision	Recall	F1-score
Deep Learning	1.00	1.00	1.00	1.00
SVM	0.93	0.93	0.93	0.93
KNN	1.00	1.00	1.00	1.00
DT	1.00	1.00	1.00	1.00

In the classification stage, we used a variety of deep learning techniques. After that, we combined them into a sequence and gave them distinctive names. The findings demonstrate that, despite taking little computational time, our suggested strategies produce effective results in terms of all performance matrices. As a result, as shown in Table 8, we evaluated the precision of our suggested models and contrasted them to a number of well used categorization techniques. During the testing and training phase machine and deep learning algorithm comparative study with each other. If you focused whole results deep learning using CNN model accuracy high as compare to machine learning algorithm. Figure 10 shows that Deep Learning model analysis performs better in terms of its accuracy, precision, recall and F-1 score using highest accuracy rate.

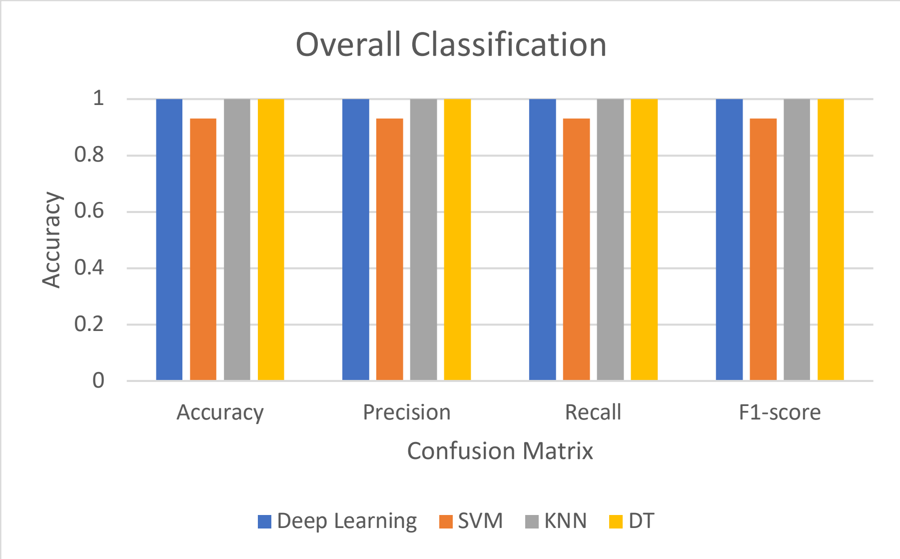

Figure 10. Overall algorithms results comparing

5. Conclusion & Future Work

The paper concludes that malware detection in smart cities can be enhanced through the fusion of machine learning and IoT-based approaches, leveraging deep learning techniques. It emphasizes the importance of analyzing and classifying models' performance using machine learning algorithms to improve computation time and categorical attacks. The study highlights the effectiveness of deep learning algorithms in analyzing malware images based on texture, shape, and smoothness, achieving high accuracy in classification results. The paper suggests the need for multi-model detection techniques and robust evasion methods to detect multiple types of malware and defend against malware attacks in smart cities. The study emphasizes the importance of raising awareness about malware attacks among administrators of smart cities and implementing measures to protect sensitive data and prevent unauthorized access. Further research is needed to enhance the computation time and categorical attacks through the analysis and classification of models' performance using machine learning algorithms .Future studies can investigate the usage of multi-model detection techniques and robust evasion methods to detect and defend against multiple types of malware in smart cities .Finally the paper suggests the need for increased usage of deep learning techniques and robust evasion methods to improve the accuracy and efficiency of malware detection algorithms.

References

1. Gaurav, Akshat, Brij B. Gupta, and Prabin Kumar Panigrahi. "A Comprehensive Survey on Machine Learning Approaches for Malware Detection in IoT-based Enterprise Information System." *Enterprise Information Systems* 17, no. 3 (2023): 2023764.
2. Rangelov, Denis, Philipp Lämmel, Lisa Brunzel, Stephan Borgert, Paul Darius, Nikolay Tcholtchev, and Michell Boerger. "Towards an Integrated Methodology and Toolchain for Machine Learning-Based Intrusion Detection in Urban IoT Networks and Platforms." *Future Internet* 15, no. 3 (2023): 98.
3. Hazman, Chaimae, Azidine Guezzaz, Said Benkirane, and Mourade Azrour. "lIDS-SIoEL: Intrusion Detection Framework for IoT-based Smart Environments Security Using Ensemble Learning." *Cluster Computing* 26, no. 6 (2023): 4069-4083.
4. Mirdula, S., and M. Roopa. "MUD Enabled Deep Learning Framework for Anomaly Detection in IoT Integrated Smart Building. e-Prime Adv." Electr. Eng. Electron. Energy 5 (2023): 100186.
5. Rashid, Md Mamunur, Joarder Kamruzzaman, Mohammad Mehedi Hassan, Tasadduq Imam, Santoso Wibowo, Steven Gordon, and Giancarlo Fortino. "Adversarial Training for Deep Learning-Based Cyberattack Detection in IoT-Based Smart City Applications." *Computers & Security* 120 (2022): 102783.
6. Kumar, Tamilarasan Ananth, Rajendrane Rajmohan, Muthu Pavithra, Sunday Adeola Ajagbe, Rania Hodhod, and Tarek Gaber. "Automatic Face Mask Detection System in Public Transportation in Smart Cities Using IoT and Deep Learning." *Electronics* 11, no. 6 (2022): 904.
7. Whaiduzzaman, Md, Alistair Barros, Moumita Chanda, Supti Barman, Tania Sultana, Md Sazzadur Rahman, Shanto Roy, and Colin Fidge. "A Review of Emerging Technologies for IoT-Based Smart Cities." *Sensors* 22, no. 23 (2022): 9271.
8. Kumar, Tamilarasan Ananth, Rajendrane Rajmohan, Muthu Pavithra, Sunday Adeola Ajagbe, Rania Hodhod, and Tarek Gaber. "Automatic Face Mask Detection System in Public Transportation in Smart Cities Using IoT and Deep Learning." *Electronics* 11, no. 6 (2022): 904.
9. Wazid, Mohammad, Ashok Kumar Das, Joel JPC Rodrigues, Sachin Shetty, and Youngho Park. "IoMT malware detection approaches: analysis and research challenges." IEEE access 7 (2019): 182459-182476.
10. Reddy, Dukka KarunKumar, Himansu Sekhar Behera, Janmenjoy Nayak, Pandi Vijayakumar, Bighnaraj Naik, and Pradeep Kumar Singh. "Deep neural network based anomaly detection in Internet of Things network traffic tracking for the applications of future smart cities." Transactions on Emerging Telecommunications Technologies 32, no. 7 (2021): e4121.
11. Jeon, Jueun, Jong Hyuk Park, and Young-Sik Jeong. "Dynamic analysis for IoT malware detection with convolution neural network model." IEEE Access 8 (2020): 96899-96911.
12. Neshenko, Nataliia, Christelle Nader, Elias Bou-Harb, and Borko Furht. "A survey of methods supporting cyber situational awareness in the context of smart cities." Journal of Big Data 7, no. 1 (2020): 1-41.
13. Baek, Seungyeon, Jueun Jeon, Byeonghui Jeong, and Young-Sik Jeong. "Two-stage Hybrid Malware Detection Using Deep Learning." *Human-centric Computing and Information Sciences* 11, no. 27 (2021): 10-22967.
14. Saba, Tanzila. "Intrusion detection in smart city hospitals using ensemble classifiers." In 2020 13th International Conference on Developments in eSystems Engineering (DeSE), pp. 418-422. IEEE, 2020.
15. Al-Turjman, Fadi, Hadi Zahmatkesh, and Ramiz Shahroze. "An overview of security and privacy in smart cities' IoT communications." Transactions on Emerging Telecommunications Technologies 33, no. 3 (2022): e3677.
16. Bajpai, Pranshu, and Richard Enbody. "Preparing smart cities for ransomware attacks." In 2020 3rd International Conference on Data Intelligence and Security (ICDIS), pp. 127-133. IEEE, 2020.
17. Rashid, Md Mamunur, Joarder Kamruzzaman, Mohammad Mehedi Hassan, Tasadduq Imam, and Steven Gordon. "Cyberattacks detection in iot-based smart city applications using machine learning techniques." International Journal of environmental research and public health 17, no. 24 (2020): 9347.
18. Peng, Sancheng, Lihong Cao, Yongmei Zhou, Jianguo Xie, Pengfei Yin, and Jianli Mo. "Challenges and trends of android malware detection in the era of deep learning." In 2020 IEEE 8th International Conference on Smart City and Informatization (iSCI), pp. 37-43. IEEE, 2020.
19. Chai, Yuhan, Jing Qiu, Shen Su, Chunsheng Zhu, Lihua Yin, and Zhihong Tian. "LGMal: A joint framework based on local and global features for malware detection." In 2020 International Wireless Communications and Mobile Computing (IWCMC), pp. 463-468. IEEE, 2020.
20. Naeem, Hamad, Farhan Ullah, Muhammad Rashid Naeem, Shehzad Khalid, Danish Vasan, Sohail Jabbar, and Saqib Saeed. "Malware detection in industrial internet of things based on hybrid image visualization and deep learning model." Ad Hoc Networks 105 (2020): 102154.
21. Zhang, Nan, Yu-an Tan, Chen Yang, and Yuanzhang Li. "Deep learning feature exploration for android malware detection." Applied Soft Computing 102 (2021): 107069.
22. Fard, Seyed Mehdi Hazrati, Hadis Karimipour, Ali Dehghantanha, Amir Namavar Jahromi, and Gautam Srivastava. "Ensemble sparse representation-based cyber threat hunting for security of smart cities." Computers & Electrical Engineering 88 (2020): 106825.
23. Shafiq, Muhammad, Zhihong Tian, Yanbin Sun, Xiaojiang Du, and Mohsen Guizani. "Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for internet of things in smart city." Future Generation Computer Systems 107 (2020): 433-442.

24. Magaia, Naercio, Ramon Fonseca, Khan Muhammad, Afonso H. Fontes N. Segundo, Aloísio Vieira Lira Neto, and Victor Hugo C. de Albuquerque. "Industrial internet-of-things security enhanced with deep learning approaches for smart cities." IEEE Internet of Things Journal 8, no. 8 (2020): 6393-6405.
25. Aliabadi, Mohammad Sharifi, and Afsaneh Jalalian. "Detection of attacks in the Internet of Things with the feature selection approach based on the whale optimization algorithm and learning by majority voting." (2023).
26. Bhodia, Niket, Pratikkumar Prajapati, Fabio Di Troia, and Mark Stamp. "Transfer learning for image-based malware classification." arXiv preprint arXiv:1903.11551 (2019).
27. Kumar, Prabhat, Randhir Kumar, Gautam Srivastava, Govind P. Gupta, Rakesh Tripathi, Thippa Reddy Gadekallu, and Neal N. Xiong. "PPSF: A Privacy-Preserving and Secure Framework Using Blockchain-Based Machine-Learning for IoT-Driven Smart Cities." *IEEE Transactions on Network Science and Engineering* 8, no. 3 (2021): 2326-2341.
28. Sarker, Iqbal H. "Smart City Data Science: Towards Data-Driven Smart Cities with Open Research Issues." *Internet of Things* 19 (2022): 100528.
29. Rashid, Md Mamunur, Joarder Kamruzzaman, Mohammad Mehedi Hassan, Tasadduq Imam, and Steven Gordon. "Cyberattacks Detection in IoT-Based Smart City Applications Using Machine Learning Techniques." *International Journal of Environmental Research and Public Health* 17, no. 24 (2020): 9347.
30. Challa, Sravani, Mohammad Wazid, Ashok Kumar Das, Neeraj Kumar, Alavalapati Goutham Reddy, Eun-Jun Yoon, and Kee-Young Yoo. "Secure Signature-Based Authenticated Key Establishment Scheme for Future IoT Applications." *IEEE Access* 5 (2017): 3028-3043.
31. Kumar, Rajesh, Xiaosong Zhang, Wenyong Wang, Riaz Ullah Khan, Jay Kumar, and Abubakar Sharif. "A Multimodal Malware Detection Technique for Android IoT Devices Using Various Features." *IEEE Access* 7 (2019): 64411-64430.
32. Wazid, Mohammad, Ashok Kumar Das, Joel JPC Rodrigues, Sachin Shetty, and Youngho Park. "IoMT Malware Detection Approaches: Analysis and Research Challenges." *IEEE Access* 7 (2019): 182459-182476.
33. Gatouillat, Arthur, Youakim Badr, Bertrand Massot, and Ervin Sejdić. "Internet of Medical Things: A Review of Recent Contributions Dealing with Cyber-Physical Systems in Medicine." *IEEE Internet of Things Journal* 5, no. 5 (2018): 3810-3822.
34. Wang, Xiaofan, Lei Wang, Yujun Li, and Keke Gai. "Privacy-Aware Efficient Fine-Grained Data Access Control in Internet of Medical Things-Based Fog Computing." *IEEE Access* 6 (2018): 47657-47665.
35. Yanambaka, Venkata P., Saraju P. Mohanty, Elias Kougianos, and Deepak Puthal. "PMsec: Physical Unclonable Function-Based Robust and Lightweight Authentication in the Internet of Medical Things." *IEEE Transactions on Consumer Electronics* 65, no. 3 (2019): 388-397.
36. Liu, Zhongjin, Le Zhang, Qiuying Ni, Juntai Chen, Ru Wang, Ye Li, and Yueying He. "An integrated architecture for IoT malware analysis and detection." In IoT as a Service: 4th EAI International Conference, IoTaaS 2018, Xi'an, China, November 17–18, 2018, Proceedings 4, pp. 127-137. Springer International Publishing, 2019.
37. Su, Jiawei, Danilo Vargas Vasconcellos, Sanjiva Prasad, Daniele Sgandurra, Yaokai Feng, and Kouichi Sakurai. "Lightweight Classification of IoT Malware Based on Image Recognition." In *2018 IEEE 42Nd Annual Computer Software and Applications Conference (COMPSAC)*, vol. 2, pp. 664-669. IEEE, 2018.
38. Islam, SM Riazul, Daehan Kwak, MD Humaun Kabir, Mahmud Hossain, and Kyung-Sup Kwak. "The Internet of Things for Health Care: A Comprehensive Survey." *IEEE Access* 3 (2015): 678-708.
39. Wazid, Mohammad, Ashok Kumar Das, Saru Kumari, Xiong Li, and Fan Wu. "Provably Secure Biometric-Based User Authentication and Key Agreement Scheme in Cloud Computing." *Security and Communication Networks* 9, no. 17 (2016): 4103-4119.
40. Wazid, Mohammad, Ashok Kumar Das, Neeraj Kumar, Mauro Conti, and Athanasios V. Vasilakos. "A Novel Authentication and Key Agreement Scheme for Implantable Medical Devices Deployment." IEEE Journal of Biomedical and Health Informatics 22, no. 4 (2017): 1299-1309.
41. Kumar, Pardeep, An Braeken, Andrei Gurtov, Jari Iinatti, and Phuong Hoai Ha. "Anonymous Secure Framework in Connected Smart Home Environments." IEEE Transactions on Information Forensics and Security 12, no. 4 (2017): 968-979.
42. Kumar, Pardeep, Andrei Gurtov, Jari Iinatti, Mika Ylianttila, and Mangal Sain. "Lightweight and Secure Session-Key Establishment Scheme in Smart Home Environments." IEEE Sensors Journal 16, no. 1 (2015): 254-264.
43. Wazid, Mohammad, Ashok Kumar Das, Neeraj Kumar, and Athanasios V. Vasilakos. "Design of Secure Key Management and User Authentication Scheme for Fog Computing Services." Future Generation Computer Systems 91 (2019): 475-492.
44. Wazid, Mohammad, Palak Bagga, Ashok Kumar Das, Sachin Shetty, Joel JPC Rodrigues, and Youngho Park. "AKM-IoV: Authenticated Key Management Protocol in Fog Computing-Based Internet of Vehicles Deployment." IEEE Internet of Things Journal 6, no. 5 (2019): 8804-8817.
45. Shen, Shigen, Longjun Huang, Haiping Zhou, Shui Yu, En Fan, and Qiying Cao. "Multistage Signaling Game-Based Optimal Detection Strategies for Suppressing Malware Diffusion in Fog-Cloud-Based IoT Networks." IEEE Internet of Things Journal 5, no. 2 (2018): 1043-1054.
46. Arias, Orlando, Jacob Wurm, Khoa Hoang, and Yier Jin. "Privacy and Security in Internet of Things and Wearable Devices." IEEE Transactions on Multi-Scale Computing Systems 1, no. 2 (2015): 99-109.
47. Yang, Geng, Mingzhe Jiang, Wei Ouyang, Guangchao Ji, Haibo Xie, Amir M. Rahmani, Pasi Liljeberg, and Hannu Tenhunen. "IoT-Based Remote Pain Monitoring System: From Device to Cloud Platform." IEEE Journal of Biomedical and Health Informatics 22, no. 6 (2017): 1711-1719.

48. Dua, Amit, Neeraj Kumar, Ashok Kumar Das, and Willy Susilo. "Secure Message Communication Protocol Among Vehicles in Smart City." IEEE Transactions on Vehicular Technology 67, no. 5 (2017): 4359-4373.
49. Alotaibi, Saud S. "Registration Center-Based User Authentication Scheme for Smart E-Governance Applications in Smart Cities." IEEE Access 7 (2018): 5819-5833.
50. Tripathy, Ajaya K., Pradyumna K. Tripathy, Niranjan K. Ray, and Saraju P. Mohanty. "iTour: The Future of Smart Tourism: An IoT Framework for the Independent Mobility of Tourists in Smart Cities." IEEE Consumer Electronics Magazine 7, no. 3 (2018): 32-37.
51. Rahman, Md Abdur, Md Mamunur Rashid, M. Shamim Hossain, Elham Hassanain, Mohammed F. Alhamid, and Mohsen Guizani. "Blockchain and IoT-Based Cognitive Edge Framework for Sharing Economy Services in a Smart City." IEEE Access 7 (2019): 18611-18621.