Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Enhancing Smart Camera Attack Classification with CGAN based Data Augmentation

Afnan Albalawi[1], Miada Almasre[1], Abeer Almakky[1], Alanoud Subahi[2], and Reemah Alhebshi[3]
1 Faculty of Computing and Information Technology, Department of Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia, aalbalawi0095@stu.kau.edu.sa ; malmasre@kau.edu.sa ; aalmak ky@kau.edu.sa
2 Faculty of Computing and Information Technology, Department of Information Technology, King Abdulaziz University, Rabigh, Saudi Arabia, asubahi@kau.edu.sa
3 Faculty of Computing and Information Technology, Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia, ralhebshi@kau.edu.sa

ARTICLE INFO

Received: 30 Dec 2024

Revised: 05 Feb 2025

Accepted: 25 Feb 2025

ABSTRACT

The Internet of Things (IoT) has transformed the way individuals interact with their environments, enabling automation and enhancing operational efficiency. However, the widespread adoption of IoT devices has raised significant concerns regarding privacy and cybersecurity risks. In this study, we create a custom smart cameras dataset and develop machine learning (ML) and deep learning (DL) models to classify and detect different attack types in smart camera network traffic. To address data imbalance, Conditional Generative Adversarial Networks (CGANs) were utilized to generate synthetic data. A comparative analysis of ML models, including Random Forest (RF), K Nearest Neighbors (KNN), and Support Vector Machine (SVM), and for DL models, we perform Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), and Attention DL models. The results show that DL models, particularly CNN, achieved exceptional classification accuracy, exceeding 94% for most attack types. Among the ML models, SVM with accuracy of 94.25% achieves higher performance than other ML models. This research contributes to enhancing smart cameras security by demonstrating the potential of conditioned synthetic dataset generation to improve the performance of advanced DL and ML techniques in identifying and mitigating cybersecurity threats, thus ensuring the protection of user data and privacy.

Keywords: Internet of Things; Camera Security; Machine Learning; Deep Learning; CGAN

INTRODUCTION

Analysts have projected that by 2025, the number of Internet of Things (IoT) de vices will reach approximately 75 million, raising concerns about their safety [1]. The IoT network comprises wireless and wired communication protocols [2], facilitating elec tronic data collection, analysis, and dissemination through sensors and drivers. The IoT has many applications and services, encompassing national infrastructure, domestic appliances, medical care, agriculture, security cameras, smart adapters, air conditioners, heat sensors, and fire alarms [3, 4]. The challenges reported in research concerning the IoT's dispersion are numerous and include proper administration of a wide range of data, implemented technologies, user accounts, connections, storage, security, and privacy [5]. Moreover, IoT relies on various technologies, such as cloud computing, fog computing, and software defined interaction, that raise the possibility of attack [6].

Amongst the widely used IoT devices are smart cameras, which are used in various settings, such as shops, hospitals, supermarkets, and universities [7, 8]. However, cameras need to be made more secure, since they are currently a prime target for at tackers. Numerous security vulnerabilities can compromise the security of cameras, such as using default login credentials, easily guessable keys, inadequate password se lection policies, and weaker encryption methods for protecting sensitive information or transmitting data without encryption [9].

According to the accumulating research, machine learning (ML) and deep learning (DL) algorithms have proven to be an efficient way to protect IoT systems [10,11] as these algorithms can classify attacks, provide suitable defensive

strategies, thus exhib iting promise in detecting and effectively managing threats [12]. Studies are using these techniques to counter the escalating threats presented by cybercrime, given the signif icant yearly increase in cyberattacks [13]. They have various applications in the domain of internet security, including the analysis of system security, which entails evaluating the vulnerabilities of networks. By examining incoming and outgoing packets, they can classify traffic in IoT devices and identify potentially malicious transactions.

However, one of the important challenges in building robust ML or DL models is the imbalanced datasets which are collected. In such contexts, model accuracy and performance are reduced due to the significant underrepresentation of many attack types [14]. Classical augmentation methods such as oversampling and Synthetic Mi nority Over sampling Technique (SMOTE), mostly generate synthetic data without considering the original data's distribution, which might lead to overfitting and less robust models [15].

More advanced methods, such as Generative Adversarial Networks (GANs), show a significant solution in generating synthetic data as they are competent in learning and replicating the complex patterns present in data [16]. Particularly, the Conditional Generative Adversarial Networks (CGANs) which are documented to improve upon traditional GANs by learning the distribution of the original data and offering greater control over outputs [17, 18]. They are reported to facilitate handling of class imbalance and enhance training stability. Dissimilar to GANs, which generate data without spe cific constraints, CGANs use conditioning variables (such as class labels) to output structured synthetic samples. This is useful in contexts where we deal with imbalanced datasets because CGANs ensure the representation of minority classes. This ultimately improves model performance and generalization. In addition, CGANs training stability is effective due to the structured supervision, which reduces issues like mode collapse and improvs sample diversity. In applications like anomaly detection in network traffic, this becomes an ideal augmentation method as it requires controlled data generation leading to enhanced detection accuracy and better overall performance in imbalanced settings [19, 20].

Thus, in this study we propose a CGAN architecture which is implemented in a supervised learning pipeline with the objective of classifying attack types based on features extracted from Smart Camera network traffic dataset. To test the effectiveness of the generated synthetic dataset, several ML and DL models were developed to detect four types of network attacks: ARP poisoning, TCP SYN flood, UDP flood, and de authentication attack, in addition to a normal state, then evaluated using perfor mance metrics like accuracy, precision, recall, and F score.

The paper presents the following contributions:

1. Improving the detection and classification of attack types in Smart Camera network traffic through synthetic data generation using CGAN.

2. Evaluating the performance of synthetic data through both ML and DL models to assess improvements in classification accuracy and overall effectiveness.

3. Addressing critical data imbalance issues within network traffic datasets, ultimately enhancing the ability of models to detect threats and classify attacks more accurately.

LITERATURE REVIEW

We will review several research on securing smart cameras. We will start by discussing smart camera security. After that, we briefly show CGANs studies that prove it effectiveness in handling data imbalance for anomaly detection. Lastly, we review ML and DL models that are used for enhancing smart camera security.

Smart Cameras Security

The security of IoT connected cameras has become an increasingly important re search area because of the rapid expansion of these devices and their sensitivity to dif ferent cyber threats. Studies have explored various aspects of camera security, including traffic classification and vulnerability assessment. An IoT camera traffic classification system achieved an accuracy of approximately 98% based on an analysis of 36 gigabytes of network data [21].

Previous research also identified IP cameras security vulnerabilities, using network sniffers tools that are employed to monitor and analyze camera traffic [22, 23]. In addi tion, NetCam IP camera security weaknesses were shown

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

through packet sniffing, proving their critical privacy concerns [24]. More studies have examined how malicious software targets these devices, posing significant risks to user privacy [25]. A study performed on the 'Smart Onvi YY HD' IP camera using Kali Linux identified multiple security threats that might be exploited by attackers [26].

These findings emphasize the need for improving security mechanisms in smart cameras to mitigate potential threats and protect user data.

CGANs for Class Balancing

Class imbalance, where certain attack types occur less frequently than others, re mains a significant challenge in most IoT datasets network traffic classification. This imbalance can affect the ability of ML or DL models to detect attacks. Over the past few years, a range of techniques have been proposed to address this issue. Among these, the use of data augmentation methods, such as generating synthetic data and fine tuning pre trained models, has additional consideration.

Recent studies have shown the effectiveness of CGANs in solving class imbalance in ML and DL applications, especially for network security tasks. CGANs have shown great results in generating synthetic data that helps balance the distribution of minority classes, which is crucial for improving model performance in imbalanced settings. For example, CGANs can generate synthetic attack traffic data to address class imbalance in intrusion detection systems. Their results showed that CGAN generated samples sig nificantly improved model training and detection accuracy for underrepresented attack types, reducing the bias typically seen in models trained on imbalanced datasets [27].

Similarly, CGANs have been employed to generate synthetic network traffic data, improving the detection of rare attack types in IoT networks. The CGAN based aug mentation method effectively mitigated the negative impact of class imbalance on model performance, leading to more balanced and accurate classifications [28]. In ad dition, CGANs have been explored for anomaly detection in network traffic, particularly focusing on their ability to generate high quality synthetic data for imbalanced attack types. The study showed that CGANs could significantly boost detection performance by increasing the number of rare attack samples, allowing models to better generalize to unseen attack patterns [29].

Other recent studies combined CGANs with other techniques, such as over sampling methods, to further improve model sensitivity to rare attack types. Their hy brid approach showed promising results, demonstrating that the combination of syn thetic data generation and traditional oversampling could improve the model's ability to detect both frequent and rare attacks without compromising performance on the ma jority class [30]. A CGAN based framework for synthetic data generation further im proved both class balance and classification performance for attacks that occur infre quently in real world datasets [31].

These studies confirm that CGANs address the challenges of class imbalance in network traffic datasets. It will not only improve classification accuracy but also enable the generation of synthetic data to train models for rare attack types.

ML model for securing IoT devices

Various ML models have been suggested for protecting devices. A research project that employed a dataset containing more than 450k attack instances and 30k benign in stances using KNN, SVM, Decision Trees (DT), RF, and Neural Networks (NN) and achieved an accuracy rate of over 99% in identifying DDoS attacks [32]. Another ap proach focuses on recognizing devices and spotting behavior using DT, RF, and DL, achieving a detection accuracy rate of 94.47 [33]. In addition, a study for researchers in [34], ML algorithms including Bayesian Generalized Linear Model (BGLM), Boosted Linear Model (Boost), and Extreme Gradient Boosting (XGBoost) were applied to combat spam with an accuracy rate of 91.9 % on the REFIT dataset. They also found that an ML approach for identifying ransomware surpassed KNN, NN SVM, and RF, achieving an accuracy rate of 89.85 % in their analysis [35]. Furthermore, other research has explored anomaly detection techniques such as SVM, One Class SVM (OC SVM), and Random Forest (RF) for class classification tasks, with high accuracy as 99% [36]. Another study [37] introduced an ML approach that developed supervised learning models like Shal low Neural Networks (SNN), Decision Trees (DT), ensemble methods like bagging trees, KNN, and SVM, with accuracy reaching 99%. RF based method tested on the dataset showed 99.95 % accuracy in identifying DoS and MIT attacks [38].

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

DL model for securing IoT devices

DL models have shown great promise in securing IoT devices, leveraging their ability to learn complex patterns in network traffic and detect sophisticated attacks. CNNs and RNNs have been applied to identify intrusion patterns, classify attacks, and prevent anomalies in IoT systems. These DL models are highly effective at processing and analyzing large volumes of data generated by IoT devices, enabling accurate detec tion of potential security threats. Additionally, transfer learning has been utilized to enhance the performance of DL models on smaller datasets, an important feature in IoT security where labeled data may be limited. For instance, CNN based models have been successfully used to identify abnormal behaviors in IoT traffic [39]. RNNs, specifically Long Short-Term Memory (LSTM) networks, are also employed for their ability to recognize temporal dependencies in time series data, making them suitable for de tecting ongoing attacks in IoT networks [40]. Moreover, some studies have highlighted the application of GANs to improve attack detection by generating synthetic attack data to train the models effectively [41]. These advancements underscore the growing po tential of DL techniques in enhancing the security of IoT environments, paving the way for more robust, automated defense mechanisms against emerging threats [42][43].

METHODS

The proposed methodology follows a structured pipeline, as shown in **Fig. 1**. First, IoT traffic data is collected from a testbed using various smart cameras and net work monitoring tools. The dataset is then preprocessed and analyzed, addressing class imbalance using a CGAN for synthetic data generation. The balanced dataset is used to develop ML and DL models for attack type classification. Finally, the models are eval uated using key performance metrics to assess their effectiveness.

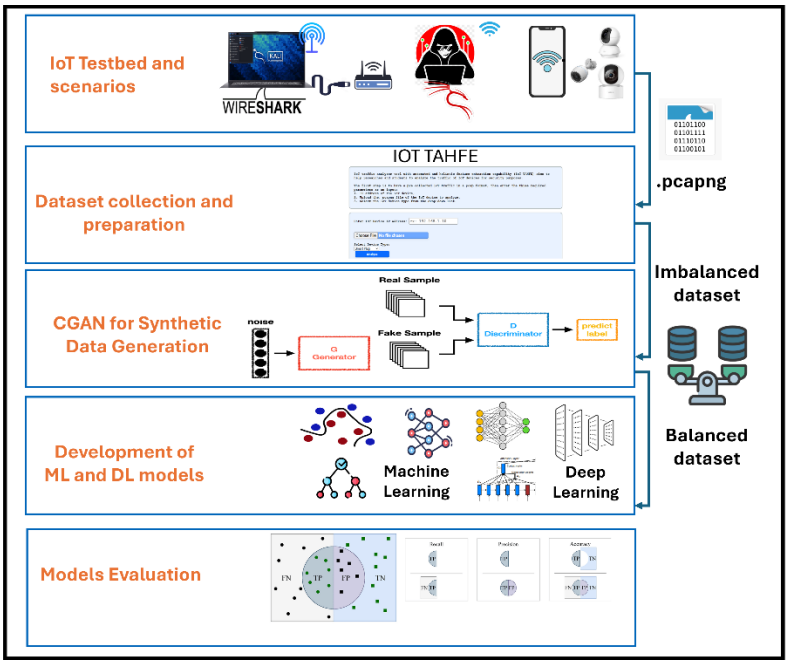

Fig. 1 Methodology Diagram.

1. The Testbed and Data Collection:

During the setup process of the testbed, a smartphone application is used to iden tify the IP addresses of three selected cameras in the network. Two laptops were used to apply different experiments. The first laptop was used to open a wireless hotspot for the cameras and to monitor network activity using Wireshark. The second laptop was used to apply various attacks targeting the cameras.

This testbed enabled the generation of realistic network traffic, including both normal and attack scenarios (The scenarios include ARP poisoning, TCP SYN flood, UDP flood, and de authentication attack). We conducted

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

experiments across 18 sce narios (6 attack scenarios × 3 cameras) and collected traffic using three meth ods—10 second intervals, 60 second intervals, and 100 packet captures—resulting in a total of 54 tested scenarios.

The dataset which was originally generated using the IoT testbed included 25 features consisting of both categorical and numerical features with 4 attack types and one normal state (5 classes). **Table 1** demonstrated the count of each class indicating an imbalance dataset.

Such an imbalance can impact classification using ML or DL models; therefore, the implementation of synthetic data generation might lead to the development of a bal anced dataset which is useful in classification tasks.

Table 1. Classes Count

Attack Type	Count
DoS TCP	58181
UDP Flood	29792
Normal	242
MITM ARP	12
De authentication	12

To look deeper into the IoT testbed dataset, we conducted a correlation analysis as demonstrated in **Fig. 2** which highlights the significant relationships which we be lieve might impact attack type detection. For example, the high correlation, between `Flow_rate` and `No_of_sent_packets_per_minutes`, introduce redundancy, which might lead to overfitting when using ML and DL models. In addition, strong correlation between `Avg_dest_SSL_payload` and `Max_dest_SSL_payload`, indicate feature de pendency, that might affect feature importance and the interpretability of the model.

These and similar correlations are to be addressed through synthetic data genera tion to enhance model accuracy in detecting various attack types.

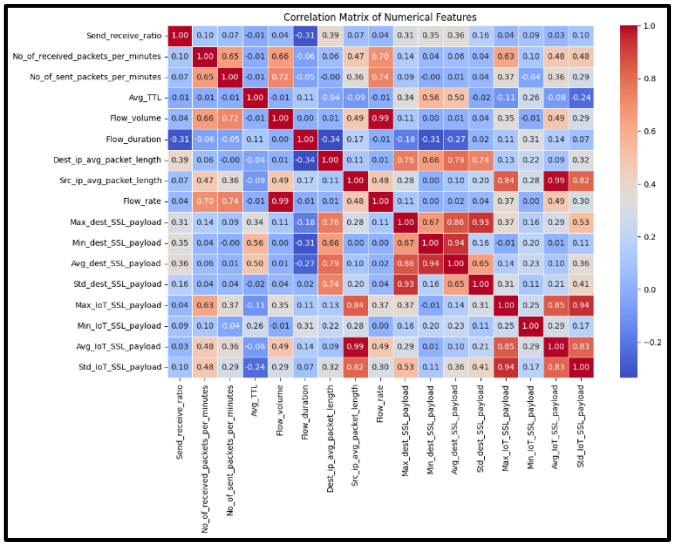

Fig. 2 Correlation Matrix.

2. CGAN for Synthetic Data Generation:

This research implements a CGAN to develop a synthetic dataset from an imbal anced smart camera network traffic dataset.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Initially, the dataset is pre-processed by encoding the categorial features. In addi tion, the numerical features are normalized to ensure a range of [0,1] for all features. The attack type labels are one hot encoded to help in the conditional learning process in CGAN.

The model includes two main components: the generator and the discriminator, which are trained competitively to generate instances based on the attack classes. The generator maps the latent space features and the class labels to the original dataset's feature space. The model's architecture is made up of an input layer that takes a latent vector concatenated with encoded class labels. This layer is followed by two hidden layers with 128 and 256 neurons. The ReLU activation function is used to introduce nonlinearity. Finally, an output layer maps the transformed features to the output di mension implementing a Tanh activation function.

The discriminator, on the other hand, differentiates between real and synthetic data instances while being conditioned on the labels of the attack types. Its input is de rived from the generator which is processed through a series of fully connected layers. In the design of the discriminator, an input layer is followed by two hidden layers with 256 and 128 neurons, each using LeakyReLU activation enforcing a negative slope of 0.2 to overcome potential vanishing gradient problems. The final layer uses a Sigmoid ac tivation function which outputs a probability score demonstrating the likelihood of the input data being real or fake.

The training of the CGAN model is carried out with specific hyperparameters in cluding setting up the number of epochs to 300 epochs with a batch size of 512. For both the generator and discriminator, the Adam optimizer is used with a learning rate of 0.0002, whereas the loss function is set to Binary Cross Entropy (BCE), which measures the model's performance. The loss during training is computed from both real and generated data features so that the discriminator can distinguish between them.

After training is completed, synthetic data is generated through sampling random latent noise vectors and concatenating them with randomly selected class labels. To ensure that the synthetic numerical values are consistent, we used np.clip() to restrict the synthetic data within the range [0,1], preventing extreme values beyond what is present in the original dataset.

3. Baseline ML Models for Classifying Attack Types:

For the ML baseline models, we have selected RF [44], SVM [45], and KNN [46] to classify network attack types considering these models' effectiveness in dealing with intricate datasets commonly found in network security scenarios.

Our implementation of the RF utilized 100 trees and a depth of 10 to balance the complexity and computational efficiency of the model. This setup reduces overfitting through using decision trees. The SVM model, on the other hand, was configured with a Radial Basis Function kernel that implemented a regularization parameter C of 0.5 to achieve a balance between model complexity and training efficiency, thus has the ability to capture non-linear relationships in high dimensional spaces which is common in network attack datasets. Similarly, KNN is set with 7 neighbors, and uses in stance-based learning. This helps in balancing the model's responsiveness to data fea tures and the possibility of overgeneralization.

4. DL Models for Classifying Attack Types:

The research utilizes three DL models CNN [47], DNN [48], and an Atten tion based model [49] to classify network attacks based on the features of a generated dataset.

Fig. 3 demonstrates CNN's model architecture. This model is used to capture spatial dependencies in network traffic datasets. The input features of the network at tack types are reshaped to be suitable for use with one dimensional convolutional lay ers, enabling the model to detect data patterns. The architecture consists of multiple convolutional layers, followed by batch normalization ones, dropout layers, and fully connected layers to support feature extraction while attempting to reduce overfitting.

Error! Reference source not found. is of the DNN model, which consists of a fully connected feedforward network. Basically, this network is designed to model the interactions of complex fea tures. It comprises multiple dense layers with a ReLU activation function. Furthermore, the network incorporates batch normalization and

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

dropout with the intention of im proving generalization. Unlike CNNs, the DNN processes input features in their orig inal vector form, allowing it to capture intricate relationships among the data attributes.

Fig. 4 depicts the Attention based Model, which is introduced to enhance the model's ability to focus on critical features. The input data is first reshaped and pro cessed through dense layers to extract relevant representations. An attention mecha nism is then applied to assign varying importance scores to different features, allowing the model to emphasize key characteristics crucial for classification.

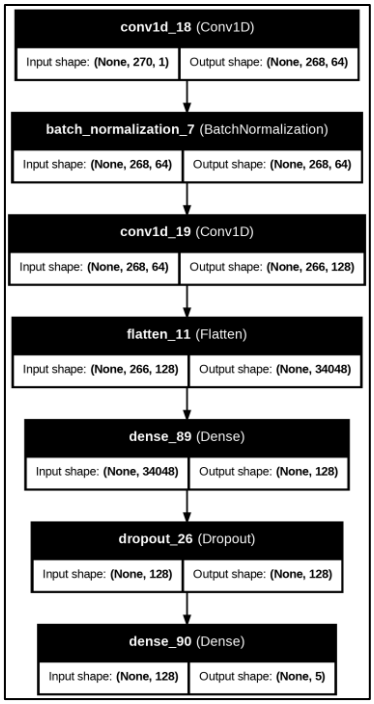

Fig. 3 CNN Model Architecture.

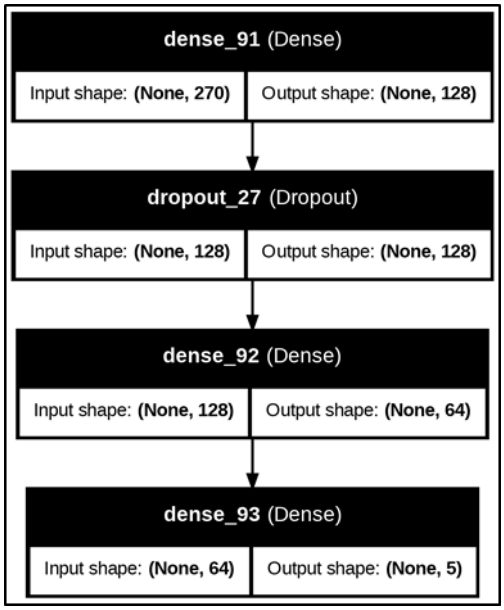

Fig. 4 DNN Model Architecture.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

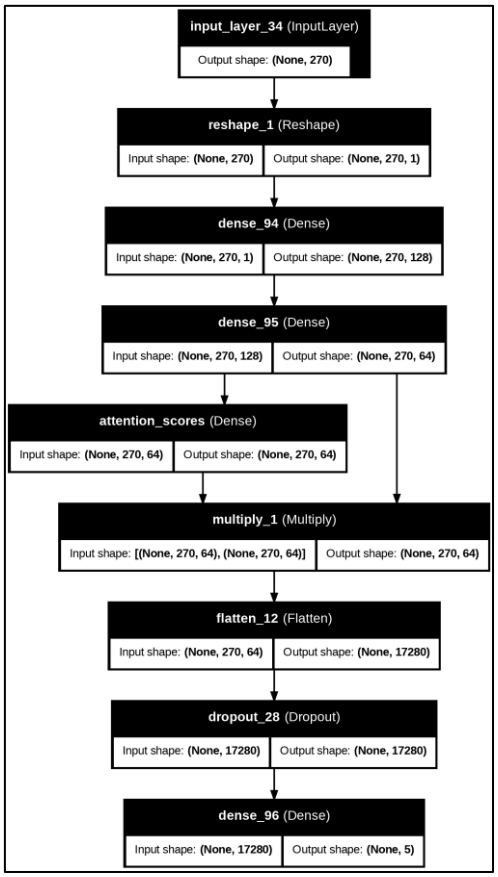

Fig. 4 Attention Model Architecture.

5. Evaluation

To evaluate the performance of the models, we utilize standard classification met rics, including accuracy, precision, recall, and the F score. Accuracy measures the ratio of correctly predicted instances to the total instances, while precision quantifies the proportion of true positive predictions among all predicted positives, and Recall, or sensitivity, assesses the proportion of actual positives correctly identified. The F score provides a balanced assessment of model performance, particularly in handling im balanced datasets.

RESULTS AND DISCUSSION

The CGAN implementation in our proposed solution yielded a dataset that con sisted of 290905 instances, with a 58181 instance per attack type, which indicates an augmented and balanced dataset. More importantly, the CGAN was conditioned to generate class instances based on the distribution of datapoints in each class.

Then, we evaluated the performance of DL and ML models for classifying smart camera network traffic into different attack types using the balanced dataset generated by the CGAN model. The results of both ML and DL models are presented in **Table 2** and **Table 3**. In addition, **Fig. 5** to **Fig. 10** illustrate the confusion matrices for the ML and DL models, providing a visual comparison of their classification performance across different attack types.

The results indicate that DL models generally outperform traditional ML ap proaches, demonstrating higher classification accuracy and better generalization across different attack types on the generated dataset.

The performance of ML models varied significantly. SVM is the best performing ML model, achieving results comparable to DL models. It has balanced precision and recall for attack classification. RF, although moderately effective, exhibited lower clas sification accuracy than SVM, indicating its limited ability to distinguish complex attack patterns. KNN demonstrated the weakest performance among all models, struggling to differentiate between various

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

attack types. This suggests that distance-based classifiers may not be well suited for high dimensional network traffic classification.

In contrast, among the DL models, CNN exhibited the highest overall performance. Its high precision and recall values indicate a strong ability to correctly identify attacks while minimizing both false positives and false negatives. The DNN achieved compa rable performance to CNN, with slightly better recall in some cases, making it particu larly effective in identifying attacks with minimal false negatives. The attention-based model, while slightly behind CNN and DNN in overall accuracy, demonstrated a well-balanced trade off between precision and recall.

When analyzing model performance for individual attack classes (by class classi fication), CNN achieved high accuracy with all attack types. It particularly detected DoS TCP and MITM ARP attacks better than the other classes. DNN performed com paratively well. Considering recall, there is an indication of its ability to capture the features distinguishing attack types. The attention-based model, on the other hand, achieved a balanced performance, with high precision value. The performance of the ML models varied as well. SVM, for example, recorded the best classification results, with high accuracy and balanced recall value with the different attack types. However, the limitation of RF and KNN were demonstrated considering the low accuracy values they achieved.

These results emphasize the suitability of DL models for Smart Camera network traffic classification, especially CNN and DNN, both outperforming other models. The precision of the attention-based model, in comparison, was higher. Nonetheless, SVM remains the best performing ML approach, as it provides competitive values with lower computational costs, whereas RF and KNN demonstrated lower results, indicating that they may not be the best choices for this classification task.

Overall, the results confirm that DL models are more effective for smart camera traffic classification, with CNN emerging as the most robust model. The study high lights the need for further research into hybrid approaches that combine the strengths of DL and traditional ML models to optimize both classification performance and com putational efficiency.

Table 2. Evaluation of ML Models.

ML / Metrics	Accuracy	Precision	Recall	F Score
RF	0.6626	0.7067	0.6626	0.6689
SVM	0.9425	0.9425	0.9425	0.9425
KNN	0.5667	0.5751	0.5667	0.5656

Table 3. Evaluation of DL Models.

Model	Class	Accuracy	Precision	Recall	F score
CNN	DoS TCP	0.9498	0.9999	0.9976	0.9987
	MITM ARP		0.9702	0.9015	0.9346
	UDP flood		0.9674	0.9190	0.9426
	De authentication		0.8945	0.9659	0.9288
	Normal		0.9252	0.9651	0.9448
DNN	DoS TCP	0.9532	0.9983	0.9999	0.9991
	MITM ARP		0.9251	0.9275	0.9313
	UDP flood		0.9616	0.9639	0.9627
	De authentication		0.9331	0.9303	0.9317
	Normal		0.9478	0.9343	0.9410
Attention	DoS TCP	0.9479	0.9992	0.9986	0.9989
	MITM ARP		0.9300	0.9176	0.9238
	UDP flood		0.9641	0.9553	0.9597
	De authentication		0.9181	0.9269	0.9225
	Normal		0.9286	0.9412	

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

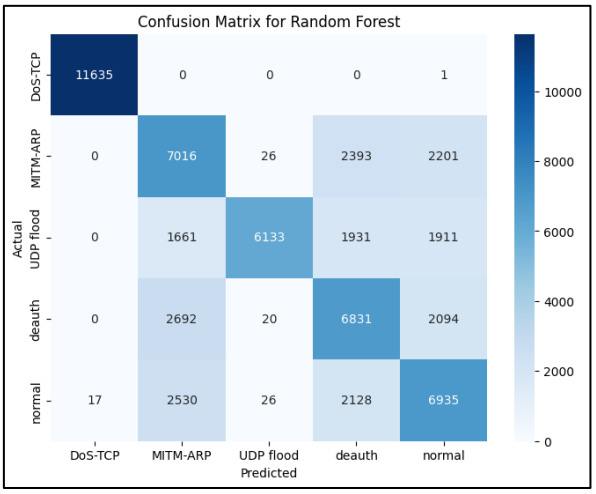

Fig. 5 RF Model Confusion Matrix.

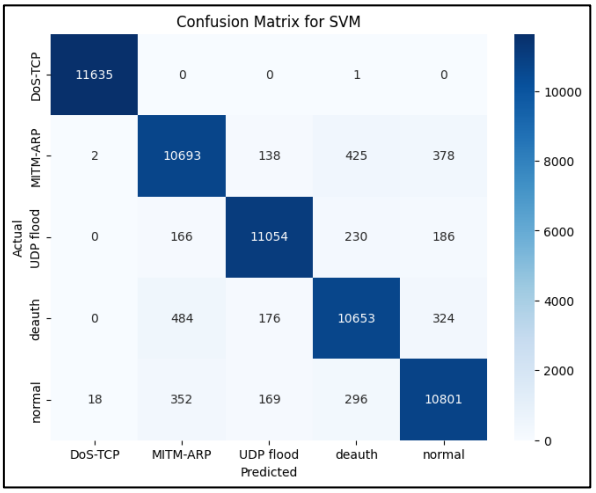

Fig. 6 SVM Model Confusion Matrix.

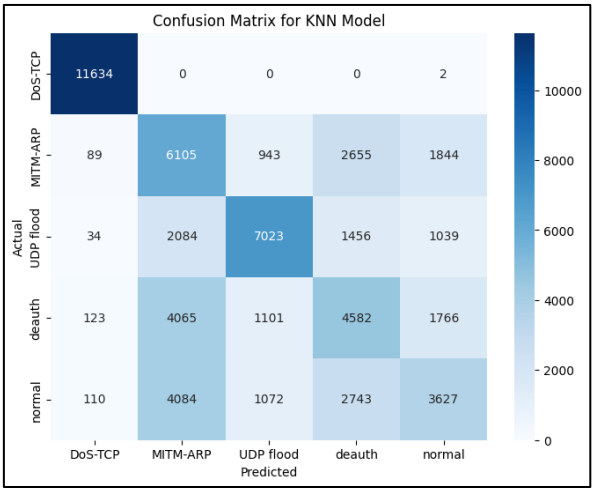

Fig. 7 KNN Model Confusion Matrix.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

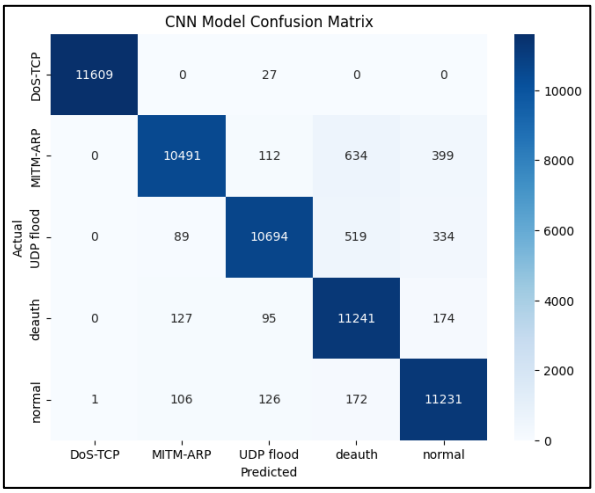

Fig. 8 CNN Model Confusion Matrix.

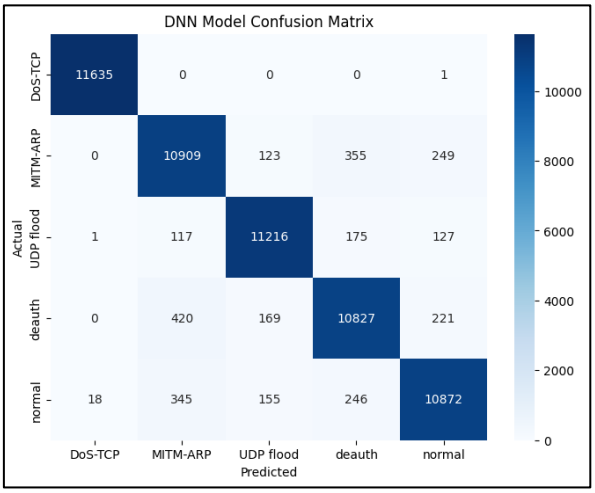

Fig. 9 DNN Model Confusion Matrix.

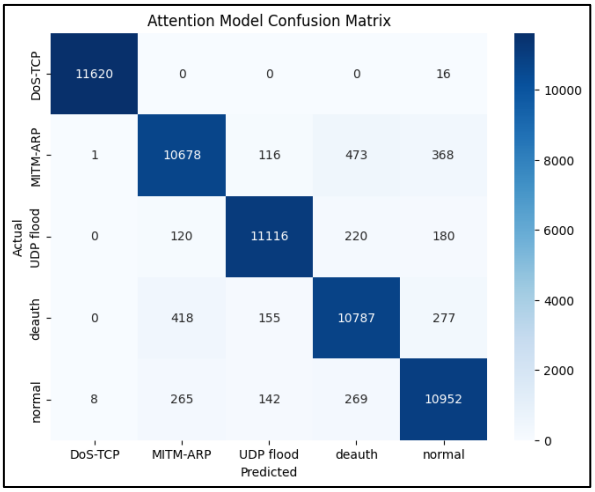

Fig. 10 Attention Model Confusion Matrix.

Journal of Information Systems Engineering and Management
2025, 10(37s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

CONCLUSIONS AND FUTURE WORKS

Numerous IoT devices lack security measures, which can make the Internet sus ceptible to cyber threats. Any secured IoT device could potentially become a gateway for attacks. ML and DL algorithms play a role in examining and enhancing the classification or grouping of network traffic. These algorithms can boost the security and authentica tion of systems by utilizing biometric based security measures and anomaly detection techniques, thereby enhancing usability and safety. However, the quality of network traffic datasets and their imbalance issues remain one of most researched topics to construct datasets which supervised learning models can classify without overfitting/underfitting. This study highlighted the role that CGAN models can have in gen erating realistic synthetic datasets that overcome the imbalance issues as well as contribute to better performance and accuracy of ML and DL in detecting network traffic anomalies. This approach improved the performance of our models by generating synthetic data for underrepresented classes, which provide a more accurate represen tation of cyber threats.

The results demonstrated that DL models, particularly CNN and DNN, achieved better classification accuracy, with CNN attaining over 94% accuracy across most attack types. The attention-based model also performed competitively, balancing precision and recall effectively. Among the ML models, SVM outperformed both RF and KNN, achieving an overall accuracy of 94.25%. These findings highlight the effectiveness of DL for network traffic classification, particularly in identifying complex attack patterns.

For future work, one can investigate how the location and brand of Smart Cameras might be a predictor of threats in combination with network traffic features. Another approach would expand on the CGAN implementation to ensure the quality of gener ated synthetic data or use Wasserstein GAN (WGAN), or even in contexts where packet data might be used implementing such models as Sequence Generative Adversarial Networks (SeqGAN).

On the level of predictive modeling on generated datasets, one might also explore the potential of hybrid approaches that integrate ML and DL models to enhance classi fication performance while optimizing computational efficiency.

REFRENCES

[1] Alam, T. A reliable communication framework and its use in Internet of Things (IoT). CSEIT1835111, 2018, pp. 450 456.
[2] Hussain, F.; Hussain, R.; Hassan, S.A.; Hossain, E. Machine learning in IoT security: Current solutions and future challenges. IEEE Communications Surveys & Tutorials 2020, 22, 1686 1721. https://doi.org/10.1109/COMST.2020.2968689.
[3] Balaji, S.; Nathani, K.; Santhakumar, R. IoT technology, applications and challenges: A contemporary survey. Wireless Personal Communications 2019, 108, 363 388. https://doi.org/10.1007/s11277 019 06965 9.
[4] Bangare, M.L.; Bangare, P.M.; Apare, R.S.; Bangare, S.L. Fog computing-based security of IoT application. Design Engineering 2021, 7, 7542 7549.
[5] Al Garadi, M.A.; Mohamed, A.; Al Ali, A.K.; et al. A survey of machine and deep learning methods for IoT security. IEEE Communications Surveys & Tutorials 2020, 22, 1646 1685. https://doi.org/10.1109/COMST.2020.2986444.
[6] Dang, L.M.; Piran, M.J.; Han, D.; et al. A survey on Internet of Things and cloud computing for healthcare. Electronics 2019, 8, 768. https://doi.org/10.3390/electronics8070768.
[7] Bugeja, J.; Jönsson, D.; Jacobsson, A. An investigation of vulnerabilities in smart connected cameras. In Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Athens, Greece, 19 23 March 2018; pp. 537 542, IEEE.
[8] Davis, B.D.; Mason, J.C.; Anwar, M. Vulnerability studies and security postures of IoT devices: A smart home case study. IEEE Internet of Things Journal 2020, 7, 10102 10110. https://doi.org/10.1109/JIOT.2020.3004222.
[9] Abdalla, P.A.; Varol, C. Testing IoT security: The case study of an IP camera. In Proceedings of the 2020 8th International Symposium on Digital Forensics and Security (ISDFS), Beirut, Lebanon, 1 2 June 2020; pp. 1 5, IEEE.
[10] Waheed, N.; He, X.; Ikram, M.; et al. Security and privacy in IoT using machine learning and blockchain: Threats and countermeasures. ACM Computing Surveys (CSUR) 2020, 53, 1 37. https://doi.org/10.1145/3400030.

[11] Awajan, A. A novel deep learning based intrusion detection system for IoT networks. Computers 2023, 12(2), 34.
[12] Tahsien, S.M.; Karimipour, H.; Spachos, P. Machine learning based solutions for security of Internet of Things (IoT): A survey. Journal of Network and Computer Applications 2020, 161, 102630.
[13] Samarakoon, S.; Siriwardhana, Y.; Porambage, P.; et al. 5G NIDD: A comprehensive network intrusion detection dataset generated over 5G wireless network. arXiv preprint arXiv:221201298, 2022.
[14] Blagus, R.; Lusa, L. SMOTE for high dimensional class imbalanced data. BMC Bioinformatics 2013, 14, 106.
[15] Shi, W.; Zhang, X. Addressing class imbalance in anomaly detection using CGANs. IET Cyber Physical Systems: Theory & Applications 2021, 6, 474 489.
[16] Iyer, O.S.; Garcia, A.M.; Rosado, P.A. Conditional Generative Adversarial Networks for Imbalanced Data Generation. IEEE Transactions on Neural Networks and Learning Systems 2020, 31(5), 1690 1703.
[17] Zhang, J.S.; Chen, Y.H.; Wu, C.X. Generative Adversarial Networks for Imbalanced Classification: A Comprehensive Survey. Pattern Recognition Letters 2020, 128, 182 195.
[18] Hong, H.S.; Wang, S.L.; Cheng, T.Z. Improved Generative Adversarial Networks for Class Imbalance. Information Sciences 2021, 542, 283 295.
[19] Liu, M.L.; Xu, Z.S.; Zhao, S.H. Addressing Class Imbalance with Conditional Generative Adversarial Networks. Computational Intelligence 2022, 37(1), 120 135.
[20] Lyra, S.; Mustafa, A.; Rixen, J.; et al. Conditional Generative Adversarial Networks for Data Augmentation of a Neonatal Image Dataset. Sensors 2023, 23(2), 999.
[21] Chaudhary, P. R., & Maiti, R. R. (2022). Detect and Classify IoT Camera Traffic. arXiv preprint arXiv:221009108.
[22] Alharbi, R., & Aspinall, D. (2018). An IoT analysis framework: An investigation of IoT smart cameras' vulnerabilities.
[23] Cowan, K. (2020). Detecting Hidden Wireless Cameras through Network Traffic Analysis. Virginia Tech.
[24] Tekeoglu, A., & Tosun, A. S. (2015). Investigating security and privacy of a cloud based wireless IP camera: NetCam. In 2015 24th International Conference on Computer Communication and Networks (ICCCN), 1 6. IEEE.
[25] Seralathan, Y., Oh, T. T., Jadhav, S., et al. (2018). IoT security vulnerability: A case study of a Web camera. In 2018 20th International Conference on Advanced Communication Technology (ICACT), 172 177. IEEE.
[26] Abdalla, P. A., & Varol, C. (2020). Testing IoT security: The case study of an IP camera. In 2020 8th International Symposium on Digital Forensics and Security (ISDFS), 1 5. IEEE.
[27] Zhang L and Li X. Using CGAN for Synthetic Data Generation to Improve Intrusion Detection in IoT Networks. Journal of Machine Learning Research, vol. 22, no. 10, pp. 2342 2356, 2021.
[28] Liu H, Wang T, and Chen Z. Anomaly Detection in Network Traffic Using Conditional Generative Adversarial Networks. Neural Networks, vol. 141, pp. 117 129, 2020.
[29] Xie F, Zhang Y, and Li W. Hybrid Oversampling and Conditional Generative Adversarial Networks for Class Imbalance in Intrusion Detection Systems. Information Sciences, vol. 542, pp. 234 245, 2021.
[30] Chen Z, Liu H, and Yang W. Improving Classification Performance with CGAN for Imbalanced Network Traffic Data. Pattern Recognition Letters, vol. 137, pp. 112 120, 2021.
[31] Iyer O S, Garcia A M, and Rosado P A. Conditional Generative Adversarial Networks for Imbalanced Data Generation. IEEE Transactions on Neural Networks and Learning Systems, vol. 31, no. 5, pp. 1690 1703, 2020.
[32] Doshi R, Apthorpe N, Feamster N. Machine learning DDoS detection for consumer internet of things devices. In: 2018 IEEE Security and Privacy Workshops (SPW) 2018, pp. 29 35. IEEE.
[33] Salman O, Elhajj IH, Chehab A, Kayssi A. A machine learning based framework for IoT device identification and abnormal traffic detection. Transactions on Emerging Telecommunications Technologies 2022; 33: e3743.
[34] Sulthana A, Mohamed IA. An Efficient Spam Detection Technique for IoT Devices using Machine Learning. 2022.
[35] Dash A, Pal S, Hegde C. Ransomware auto detection in IoT devices using machine learning. no December 2018: 0 10.
[36] Bagaa M, Taleb T, Bernabe JB, Skarmeta A. A machine learning security framework for IoT systems. IEEE Access 2020; 8: 114066 114077.

Research Article

[37] Hasan M, Islam MM, Zarif MII, Hashem M. Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet of Things 2019; 7: 100059.
[38] Alsulami AA, Abu Al Haija Q, Tayeb A, Alqahtani A. An Intrusion Detection and Classification System for IoT Traffic with Improved Data Engineering. Applied Sciences 2022; 12: 12336.
[39] A. G. Oropesa, A. Garcia Serrano, and C. L. Stiawan, "A deep learning approach to IoT security: Classification of intrusion and attack patterns," IEEE Access, vol. 9, pp. 87240 87252, 2021.
[40] P. Mohanty, S. Patra, and M. Kumar, "Intrusion detection for Internet of Things based on deep learning techniques," Journal of Information Security and Applications, vol. 58, p. 102651, 2021.
[41] Z. Liu, L. Yang, X. Zhang, et al., "IoT security: A deep learning-based approach for network intrusion detection," IEEE Internet of Things Journal, vol. 8, no. 15, pp. 12199 12208, 2021.
[42] M. A. Alqaralleh, M. M. L. Abdullah, and N. A. K. Muda, "An IoT intrusion detection system based on convolutional neural network," Computers, Materials & Continua, vol. 67, no. 2, pp. 1639 1653, 2021.
[43] S. K. P. M. R. A. P. V. D. Pradeep, "Deep learning for security in Internet of Things devices: A comprehensive review," Wireless Communications and Mobile Computing, vol. 2021, p. 2243657, 2021.
[44] Breiman L. Random forests. Machine Learning 2001; 45: 5–32.
[45] Cortes C, Vapnik V. Support vector networks. Machine Learning 1995; 20: 273–297.
[46] Cover TM, Hart PE. Nearest neighbor pattern classification. IEEE Transactions on Information Theory 1967; 13(1): 21–27.
[47] LeCun Y, Bottou L, Orr GB, et al. Gradient based learning applied to document recognition. Proceedings of the IEEE 1998; 86(11): 2278–2324.
[48] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by backpropagating errors. Nature 1986; 323: 533–536.
[49] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. Advances in Neural Information Processing Systems 2017; 30: 5998–6008.