Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Deep Learning Methodology for Detection of Multimodal Minority Class Sample Attacks

A. N. Sasikumar[1], Dr. Lilly Sheeba S[2]

[1]Research Scholar, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, Tamil Nadu, India, gprabakaransai@gmail.com
[2]Associate Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, Tamil Nadu, India, lillyshs@srmist.edu.in

Digital Attacks are expanding and to alleviate class test assaults and dangers, interruption recognition framework is presented in real time network classification. Interruption recognition frameworks are generally used to catch the going amiss examples in the organization traffic. Because of dynamic nature of changing examples of dangers and assaults, a proficient model is expected to refresh the assaults and examples present in the organization traffic information. Many AI models are sent to gain proficiency with the traffic designs yet customary models generally experience the ill effects of high traffic volume and high layered highlights. This paper proposes a profound hyper parameter assisted learning model which is strong to catch network interruptions with better ability to learn. The viability of the proposed profound learning model is exhibited utilizing CICIDS2017 dataset and the presentation of the proposed model accomplished exactness of 99.7% over other AI models.

Keywords: security attacks, multimodal, deep learning, minority class sample attacks.

1. Introduction

With the advancement of systems administration, equipment and programming abilities, class test assaults are expanding. As of now class test assaults are advancing which are convoluted and testing to distinguish. Since assailants utilize complex techniques to sidestep discovery and a clever interruption location framework is expected to forestall and catch network interruptions. Many AI models, for example, Backing Vector Machine, Irregular Woods, choice tree, Brain Organization (Zou et al., 2009) and so forth are as of now been utilized to recognize interruptions in the organization (Ahmad et al., 2019, Prachi and Sharma, 2019; Asad et al., 2020). An interruption recognition framework can be classified into have based and network based interruption identification. The host based and network based IDS are additionally assembled into abnormality based, signature based and crossover (Aydin et al.,

2009).

The huge volume of organization traffic information requires profoundly complex learning models to deal with high volume information and to deal with high layered include sets. Huge volume of information and highlights increment the computational intricacy and learning time to order traffic into ordinary and assault. Further, the presence of various kinds of assaults and interruptions in the organization traffic requests a productive model to catch subtleties of interruptions. Profound learning models are as of now utilized for its adaptability, high learning capacity, and element designing capacity in picture handling, language handling, suggestion frameworks and shortcoming location (Chhajer et al, 2022). Differentiating shallow students, profound learning models are more powerful in extricating the element data and learning of highlights. The elements of enormous information frequently influence the learning capacities prompting unfortunate identification of assaults and double-dealing from a dataset having huge number of elements. Profound learning models conquer the limits of AI models through removing highlights that map the interruption type and assaults. AI models require area specialists to cut down the element aspect intricacy through highlight choice. Profound learning models are equipped for preparing non-straight information which guarantee precise expectation of organization interruption and its sub classes through speculation of new assault variations (Mighan and Kahani, 2021).

Profound learning models have various organizations like Convolution Brain Organizations (CBO), Long Momentary Memory Organizations (LMMO), Repetitive Brain Organizations (RBO), Generative Antagonistic Organizations (GAO), Multi-facet Perceptrons (MLP), Self Getting sorted out Guides and Profound Conviction Organizations which are fit for various properties to use the elements. DL models have layers and each layer gets a contribution from its former layer. DL models process the crude elements and proliferate the significant component data to the following layers, which are at long last used to arrange or anticipate the marks. The fundamental benefit of profound learning models is the versatility, unaided self-figuring out how to further develop precision, upholds equal and dispersed preparing and programmed include age. The one hindrance is that it requires high information on network geographies and preparing boundaries.

2. Related works

Profound learning models have drawn in numerous researchers and analysts towards creating Interruption Identification Frameworks utilizing tremendous measure of information. Since customary AI models don't deal with enormous information, profound learning models are basically picked for its soundness and adaptability. (Qazi et al., 2022) proposed a one-layered CNN to order network interruptions. The proposed model has five convolution layers and five thick layers and softmax layer to foresee the probabilities of interruption types. The proposed model accomplished 98.96% precision for multiclass issue on CICIDS2017 dataset. A crossover profound learning model was proposed by (Aldallal, A. 2022) to further develop the location pace of IDS in view of GRU and LSTM. The significant highlights are chosen utilizing Pearson Relationship strategy. The proposed Cu-LSTMGRU is consolidates LSTM and GRU with a processing unit called MPDM. The data that are vital to order are passed on the following layer through the update door which data that are not generally needed are

projected out through reset entryway. The proposed model accomplished most noteworthy exactness of 97.76% on CICIDS2018 lessening the deception rate. The proficiency of IDS can be improved on the off chance that the model can deal with enormous volume of information, offer element determination and have great execution over recognition of assaults. To meet such prerequisites, (Adefemi Alimi et al., 2022) proposed a refined LSTM for refusal of-administration assault discovery. RLSTM enjoys the benefit of forestalling back proliferation mistakes and address the issue of evaporating angles. The proposed RLSTM showed precision of 99.2% for recognizing dos-assault in CICIDS2017 dataset and 98.6% in NSL-KDD dataset. (Azzaoui et al., 2022) showed the productivity of profound brain networks involving four layers engineering to identify interruptions in the organization. The model show better execution with low phony problem rate and high exactness of 99.43% on CICIDS2017 and 99.63% on NSL-KDD dataset.

(Jamil and Kim., 2021) proposed a gathering based learning and forecast of abnormality in network traffic utilizing computerized AI. The proposed model utilize Bayesian advancement of boundary tuning and Kalman channel model for expectation. The proposed strategy is contrasted with one with five distinct layers of DNN and the proposed technique show elite execution with exactness of 97.02% on CICIDS2017 and 98.8% in UNSWNB15 dataset. (Sahu et al., 2021) proposed a mixture profound learning model by joining CNN and LSTM to recognize interruption in the organization. The reason for the cross breed model is to utilize the mechanized element gaining from CNN and to characterize interruptions utilizing LSTM. The proposed model has four convolution layers with LeakyReLU as initiation capability. The chose highlights are taken care of into the LSTM layers for order of interruptions. The model accomplished a location exactness of 96% for multi assault types. Likewise a cross breed model utilizing CNN and LSTM was proposed by (Sun et al., 2020) to identify interruption in the organization. The mixture model catches the fleeting and spatial elements of the organization traffic. The organization engineering contains two convolution layers and two LSTM layers, which grouped the interruptions accurately. The model accomplished 98.67% of generally precision. The upside of utilizing LSTM layers is that settles the angle detonating and vanishing while at the same time preparing in grouping. Additionally, (Kim et al., 2020) proposed IDS in view of CNN-LSTM and the model presentation is assessed utilizing three datasets. The organization comprises of two CNN layers and three LSTM layers, the information from result of CNN layers are taken care of into LSTM. The main layer of LSTM is forward course and the second layer of LSTM is bidirectional and the last layer is DNN which amasses the forward and in reverse cells of second layer LSTM. The proposed model show low execution of 83% on CICIDS2017 dataset and 91% in CSIC-2010 dataset. The low presentation of the model is expected to overfitting as the model has 14000 teachable boundaries. (Kaur and Singh, 2020) proposed a half breed profound learning model involving RNN for network irregularity grouping. The proposed model orders network assaults as well as creates marks to induce by the IDS. Consolidating the marks and the classifier, the proposed D-sign IDS beats different models with precision of 99.1% on CICIDS2017 and 99.14% on NSL-KDD datasets. (Sethi et al., 2021) proposed a multi-specialist profound support learning based IDS for assault grouping utilizing consideration instrument. The organization of ten layers and five profound consecutive layers are contemplated and the presentation of the model on CICIDS2017 is 98.7% and 97.4% on NSL-KDD dataset.

3. Methodology

3.1 CNN Architecture

Convolution Neural Networks (CNN) fall under deep learning (DL) architecture and similar to Neural Networks with stacked layers. The term convolution involves a convolution block in the neural network. CNN represents two dimensional inputs to categorical output for classification task and to real integers for regression task. A Neural Network typically requires handcrafted features to produce accurate results while CNN uses raw data and process the input across many convolution layers. The input to CNN is a vector that represents the network traffic

X. The CNN is designed to learn a set of parameters Θ that map the input to the prediction C (attack classes) and it can be represented as,

$$C= F(X \mid \Theta) = f_h(\ldots f_2(f_1(X|\Theta 1) \mid \Theta 2)| \Theta h) \quad (1)$$

where h is the number of hidden layers and for the i^{th} layer in the convolution layer can be represented as,

$$C1=f_i(X_i|\Theta i) = A(W*X_i + b), \Theta i = [W, b] \quad (2)$$

Where * is the dot product (convolution function) with the input features, X_i is the two dimensional input matrix of N features, W is the one dimensional kernels for extracting new features from input vector and b is the bias vector and A is the activation function. Many Pooling layer is used between convolution layers that summarize the features by calculating maximum or average of the input. The output of the convolution layer is flattened and passedon to fully connected layers and it can be represented by equation 3 and for multiclass classification, softmax activation function is used on the output layer, where the each neuron points the class membership of the input samples.

$$C1=f_i(X_i|\Theta i) = A(W\ X_i + b), \Theta i = [W, b] \quad (3)$$

CNN is comprised of three fundamental layers, convolution layer, pooling layer and completely associated layer. The highlights are separated from convolution layer and pooling layer while completely associated layer relates the elements extricated to the result assault classes. Convolution alludes to a straight capability used to remove highlights. The non direct actuation capabilities utilized in the convolution layer is ReLU, sigmoid and tanh capability. The pooling layer is utilized to lessen the element aspects by collecting highlight data from various portions. Max pooling and normal pooling are the two kinds of pooling accessible where max pooling separate elements regarding portion size while disposing of others. The completely associated layer is last layer where the highlights are changed into 1D vector which is completely associated with each result class through loads. The last layer enactment capability shift as per the sort of issue i.e., parallel, multiclass or relapse. For paired arrangement and multiclass characterization sigmoid enactment and for relapse direct or personality initiation capability is utilized. The boundaries are changed as for the characterization blunder rates which are back spread while preparing through limiting the misfortune capability.

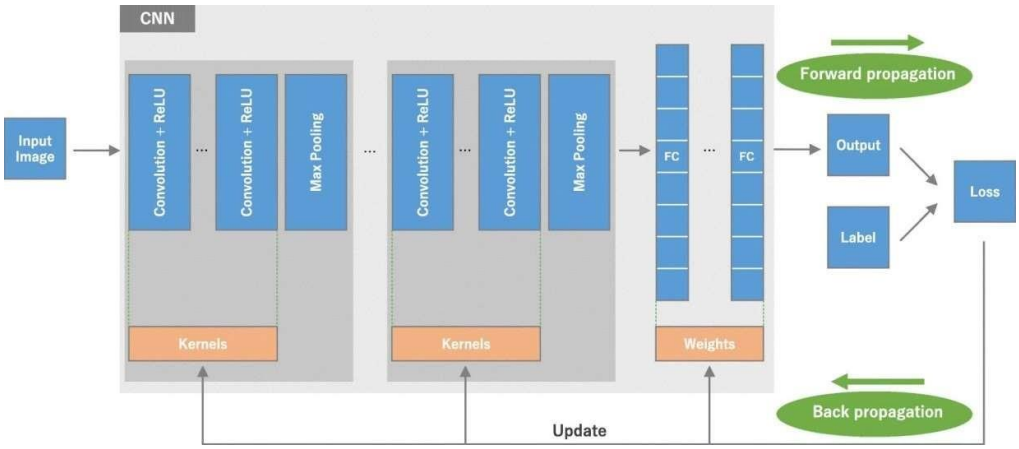

Figure 1 CNN Architecture

CNN architecture is characterized by set of hyperparameters which refers to network structure and training hyperparameters. The hyperparameters of network structure refers tolayer number, units in each layer, kernel size, strides, pooling and activation functions whilethe training hyperparameters refers to learning rate, batch size, momentum, epochs, optimizer and patience for early stopping. (Figure 1)

3.2 Proposed 1D CNN

The presentation of the CNN relies upon the engineering of the organization plan which varies as per different issue and sizes. It is hard to track down the reasonable hyperparameters for a specific issue and changing the hyperparameters esteem straightforwardly influences the model presentation. Enhancing network boundaries and finding the right mix of hyperparameters further develop the model exhibition if not the model probably won't arrange interruptions successfully. Unfortunate model execution shows that the model steadiness, handling time and it is impacted to process assets. To really order interruptions, 1D CNN is built independently for paired and multiclass issue. 1D CNN alludes to the part that slides in a single aspect. Three convolution layers are added after the information layer followed by two completely associated layers and a result layer. The primary layer has 128 neurons and second convolution layer has 64 neurons and the subsequent layer is entwined with max pooling layer with pool size of 3 and bit size of 3 x 3 with enactment capability relu. The third convolution layer has 32 neurons with piece size of 3 x 3. To keep away from overfitting a dropout layer is added to the third convolution layer to safeguard the element data. The result of the last convolution layer is straightened and given to completely associated layer. The completely associated layer has 10 neurons with softmax actuation capability predicts the organization interruption types. The misfortune capability is given in equation 4.

$$\text{Loss} = -\sum_{i=1}^{k} y_j \log (y`_i) \quad (4)$$

where k is the number of classes, y is the actual class and y` is the predicted class.

4. Experiment and analysis

4.1 Dataset

The proposed 1D convolution network execution is evaluated on CICIDS2017 dataset. The CICIDS2017 dataset contains steady association traffic data which got network traffic north of multi day stretch of time. This dataset contains latest pursues that seem to be this current reality data and moreover contain checked streams for network traffic assessment considering time stamp, source, and objective IPs, source and goal ports, shows and attack (Iman Sharafaldin et al., 2018). The normal traffic is set apart as innocuous and 15 sorts of attacks types are named. The attacks consolidates Savage Power, Heart Channel, Botnet, DoS, DDoS,Web, and Intrusion Attack. The entire dataset including multi day traffic data has an amount of 2,299,308 events which require high computational limit. To test the proposed model, Thursday early daytime working hour's dataset is utilized for the audit. The Thursday dataset has 170368 events, 78 components and 1 name segment. Around 10 components that have zero characteristics are taken out and rest of the features is associated with the audit. The last dataset set contains 68 features and 1 class name with three attack types and innocuous traffic. (Table 1)

Table 1 CICIDS2017 dataset

Day	Type	Size
Monday	Normal	11GB
Tuesday	Tuesday	11GB
Wednesday	Normal + Dos + Heartbleed Attacks	13GB
Thursday	Normal + XSS + Web Attack + Infiltration	7.8GB
Friday	Normal + Botnet + PortScan + DDoS	8.3GB

4.2 Preprocessing

The categorical features present in the CICIDS2017 dataset is converted to nominal values and the numerical features are standardized using Z-score normalization. Since each feature have different ranges which affects the training of the module. The normalization helps to keep the range between 0 and 1 which improves the training in a better way. The normalization formula is given in equation 5. Some features have infinite numbers and such features are selected and the infinite values are replaced with zeros.

4.3 Evaluation Metrics

The introduction of portrayal model is summarized and imagined using disorder structure. The chaos network for twofold request is given in Table 5. It tends to the certified characteristics and expected values. The show is unraveled using TP, TN, FP and FN where TP implies Authentic Up-sides which address the amount of positive models precisely appointed Up-sides, TN suggests Certified Negatives which address the amount of negative models requested precisely as Negatives, FP insinuates Counterfeit Up-sides which address the amount of Negative models erroneously named Positive and FN insinuates Deluding Negatives which address the amount of Positive models erroneously named Negative. Precision, Responsiveness, identity, Exactness and F-Score are the estimations that gets a handle on the portrayal model's show. The dataset is distributed into testing and planning set

in the extent 70:30 for twofold gathering and 80:20 extent for multiclass course of action.

5. Results and Discussion

5.1 Evaluation of the proposed model -Binary classification

Utilizing CICIDS2017 dataset, the exhibition of the proposed model is assessed. The measurements used to quantify the presentation are examined in segment 4.3. Fostering an interruption identification model utilizing profound learning has been the hot pattern in network security. Persuaded by the computational benefit of 1D CNN over 2D CNN and include learning capacity, this study proposed a 1D CNN for catching organization interruptions or assaults. The proposed model is approved for parallel class issue and multi-class issue. Paired order includes two classes in particular harmless and interruption while multi-class arrangement have four classes to be specific harmless, web assault Animal power, web assault XSS and web assault Sql infusion.

Table 2 Performance of the model for Binary classification

Class	Accuracy	Precision	Recall	F1-Score	FPR
0	0.997	0.999	0.997	0.998	0.018
1	0.997	0.820	0.981	0.894	0.002
Average	99.7	90.9	98.1	94.6	0.01

The disarray lattice (Figure 2) shows genuine up-sides of 33541 examples, genuine negative of 431 occasions, 94 cases of misleading positive and 8 occurrences of bogus negative. The exhibition of the model accomplished an exactness of 99.7%, accuracy of 90.9%, review of 98.1%, f1 score of 94.6% and misleading positive pace of 0.01 Table 4. Exactness of the model records for the complete examples accurately ordered, out of 34074 cases, the model accurately arranged 33972 occurrences. Accuracy alludes to the proportion of positive cases accurately delegated positive, in interruption identification harmless traffic is accurately named harmless to around 90.9% while review alludes to proportion of the positive examples accurately named positive to the absolute number of positive cases, the 98.1% of review shows that the proposed model accurately arranged 98.1% of harmless cases (33541) to the complete number of harmless cases (33635). F1 score alludes to the symphonious mean of accuracy and review, the F1 score of 99.8% alludes to the great harmony among accuracy and review. The model accomplished FPR of 0.01 for parallel order which alludes to that the assault classes are accurately grouped into negative class which brings down the erroneous characterization of assault class as harmless. Figure 3 addresses the approval misfortune and Figure 4 addresses the exactness of the proposed model for parallel arrangement.

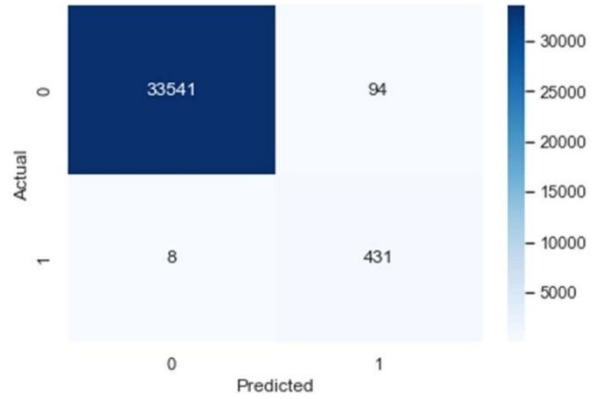

Figure 2 confusion matrix for binary classification

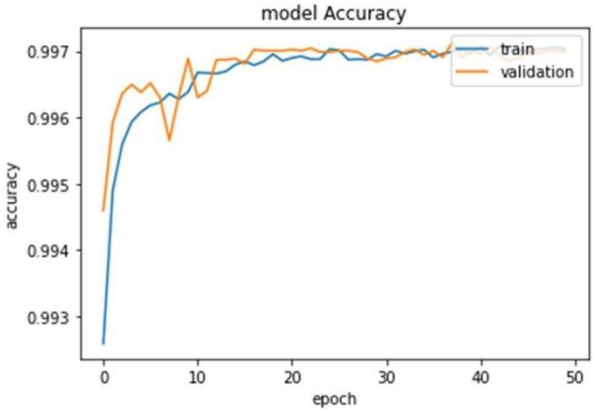

Figure 3 Accuracy of the proposed model for binary classification

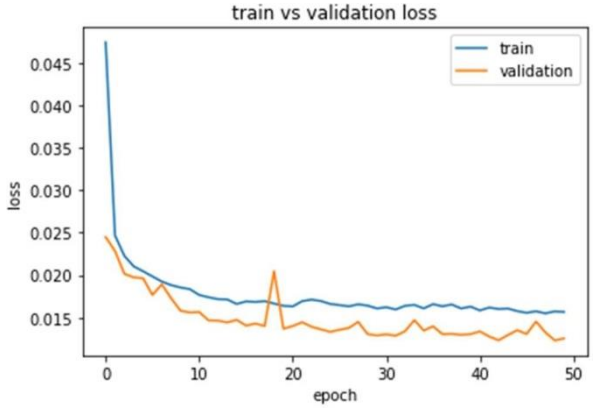

Figure 4 Validation loss of the proposed model for binary classification

5.2 Evaluation of the proposed model - Multiclass classification

The CICIDS2017 dataset has innocuous, web attack Savage power, web attack XSS and web attack Sql imbuement classes. The display of the proposed model is surveyed on different

attack classes. The introduction of the model achieved an accuracy of 99.6% for monster force, 99.9% for XSS, 99.9% for Sql mixture and for innocuous 99.6%. The normal model show for all of the classes show precision of 99.7%, exactness of 93%, audit of 86.7%, f1 score of 87.0% and deceiving positive speed of 0.001 (Table 3). Precision of the model records for the outright events precisely requested, out of 51070 models, the model precisely gathered 50909 cases and around 161 events are incorrectly portrayed. Exactness suggests the extent of positive cases precisely appointed positive, in interference ID innocuous traffic is precisely named innocuous to around 99.9%, for monster force, XSS and sql imbuement the exactness is 72.1%, 100% and 100%. Survey insinuates extent of the positive models precisely designated positive to the full scale number of positive cases, the audit for monster force, XSS and sql mixture is 99.5%, half and 97.8%. The lower survey of half for XSS is normal lower number of models. F1 score insinuates the consonant mean of precision and survey, the F1 score of 99.8% and 98.3% for innocuous and sql mixture suggests the extraordinary amicability among exactness and audit while attack class creature power have f1 score of 83.6% and XSS class have 66.6% of f1 score. The model achieved FPR of 0.003 for innocuous class and creature power which suggests that the common and attack class savage power are precisely portrayed into positive and negative class. The invalid worth of FPR for XSS and sql imbuement is attributed to the lower number of assumption tests.

Table 3 Performance of the model for Multiclass classification

Class	Accuracy	Precision	Recall	F1-Score	FPR
Benign	0.996	0.999	0.996	0.998	0.003
Brute force	0.996	0.721	0.995	0.836	0.003
XSS	0.999	1	0.5	0.666	0
Sql injection	0.999	1	0.978	0.983	0
Average	99.7	93	86.7	87.0	0.001

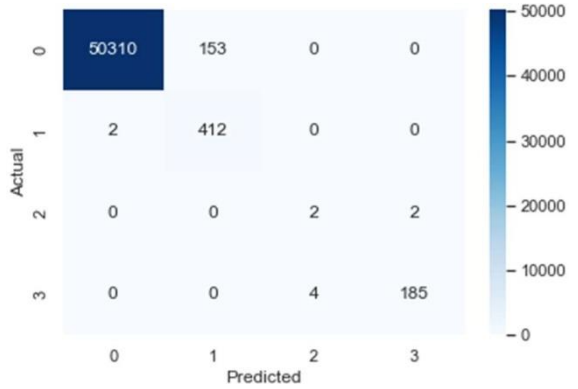

Figure 5 confusion matrix for multiclass classification

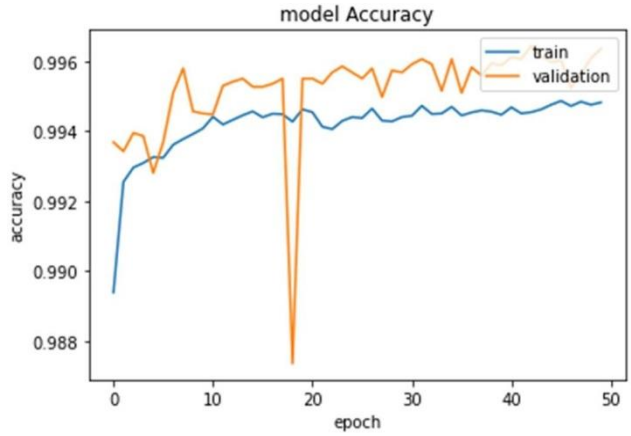

Figure 6 Accuracy of the proposed model for binary classification

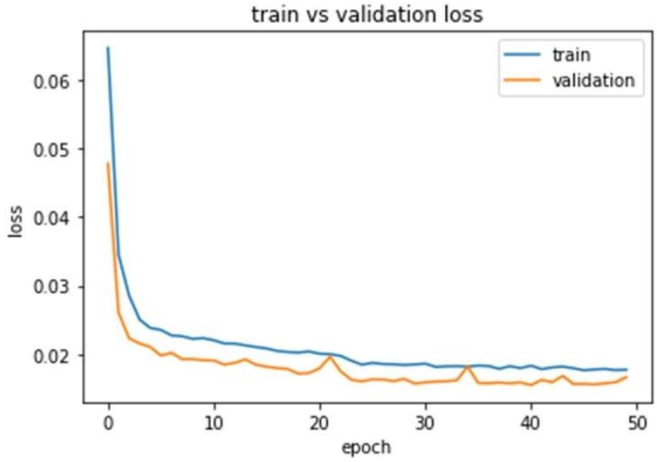

Figure 7 Validation loss of the proposed model for multiclass classification

To demonstrate the efficiency of the proposed model, the performance of model is compared with state-of-art deep learning methods. Many deep learning models are currently studied to classify network intrusions and models are trained for both binary and multiclass problems. The proposed model consists of three convolution layers and two dense layers with max pooling and single dropout (0.3) in the third convolution layer. In order to preserve the feature information the dropout layer is added to the third layer. The model achieved highest accuracy of 99.7% over 1DCNN proposed by (Qazi et al., 2022). The proposed model equally performed well with an accuracy of 99.7% with (Aldallal, A., 2022) who proposed Cu-LSTMGRU, a gated recurrent network which require high computing power. Compared to CNN, CNN-LSTM models have more parameters to handle and are sensitive to initial weights which could affect the training on spatial features. The proposed model performed good learning on spatial features as the number of parameters is less compared to CNN-LSTM, which facilitate more learning on feature maps which outperformed (Sahu et al., 2021; Sun., et al., 2020; Kim et al., 2020). RLSTM model which falls under RNN, requires more training data and due to its nature of back propagation, this model might suffer from vanishing or

exploding problem and also suffers on training long sequences. The proposed model produced higher accuracy of 99.7% than RLSTM (99.22%) and RNN network (99.1%). The proposed CNN model outperforms RNN models of (Adefemi Alimi et al., 2022) and (Kaur & Singh, 2020). The validation loss and accuracy of the model is given in Figure 6 and Figure 7 which demonstrate that 1DCNN is better in terms of accuracy, recall, precision and f1 score than other models discussed in the literature.

Table 4 comparison of proposed model with deep learning methods

Author	Year	Method	Accuracy%
Qazi et al.,	2022	1DCNN	98.96
Aldallal, A.	2022	Cu-LSTMGRU	99.70
Adefemi Alimi et al.,	2022	RLSTM	99.22
Azzaoui et al.,	2021	DNN	99.43
Sethi et al.,	2021	MARL	98.70
Jamil & Kim.,	2021	Ensemble Model	97.02
Sahu et al.,	2021	CNN-LSTM	96.0
Sun et al.,	2020	CNN-LSTM Hybrid	98.67
Kim et al.,	2020	CNN-LSTM	93.0
Kaur & Singh.,	2020	RNN	99.1
		1DCNN-binary	99.7
Proposed	2022	1DCNN-Multiclass	99.7

Table 5 Confusion Matrix Predicted

	Positive	Negative
Positive	True Positive TP	False Positive FP
Negative	False Negative FN	True Negative TN

6. Conclusion

This paper proposed a CNN model with one aspect for network interruption identification. The proposed model is assessed utilizing CICIDS2017 dataset. The pre-handled informational collection was prepared on 1DCNN model built with three convolution layers and two thick layers, max pooling and completely associated layer. The presentation of the CNN model is deciphered with execution measurements like exactness, accuracy, review and f1 score. The 1DCNN model accomplished 97% of exactness, 90.9% of accuracy, 98.1% of review, 94.6 of f1 score and 0.01 of FPR for twofold. In multiclass arrangement, the model achieved typical exactness of 99.7%, 93% of accuracy, 86.7% of review and 87.0% of f1 score with 0.001 of FPR. In view of the outcomes, the proposed 1D CNN model is proficient for network interruption recognition and subsequently it very well may be utilized for network assault identification. As a future work, the impact of component choice on network assault grouping

will be investigated with other province of CNN by tweaking the organization boundaries to accomplish ideal execution as far as huge organization traffic volume and enormous number of organization traffic highlights. Additionally, the assault types are developing and it is important to incorporate more up to date dangers and assault types to recognize interruptions really.

Conflict of Interest

The authors declare no conflicts of interest, financial or otherwise.

Authors' contributions

Sasikumar and Lilly Sheeba S, contributed on Conceptualization of introduction and existing approaches , results and discussions and overall supervision.

References

1. Adefemi Alimi, K. O., Ouahada, K., Abu-Mahfouz, A. M., Rimer, S., & Alimi, O. A. (2022). Refined LSTM Based Intrusion Detection for Denial-of-Service Attack in Internet of Things. Journal of Sensor and Actuator Networks, 11(3), 32.
2. Ahmad, I., Basheri, M., Iqbal, M. J., & Rahim, A. (2018). Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection. IEEE access, 6, 33789-33795.
3. Aldallal, A. (2022). Toward Efficient Intrusion Detection System Using Hybrid Deep Learning Approach. Symmetry, 14(9), 1916.
4. Asad, M., Asim, M., Javed, T., Beg, M. O., Mujtaba, H., & Abbas, S. (2020). Deepdetect: detection of distributed denial of service attacks using deep learning. The Computer Journal, 63(7), 983-994.
5. Atefinia, R., & Ahmadi, M. (2021). Network intrusion detection using multi-architectural modular deep neural network. The Journal of Supercomputing, 77, 3571-3593.
6. Aydin, M. A., Zaim, A. H., & Ceylan, K. G. (2009). A hybrid intrusion detection system design for computer network security. Computers & Electrical Engineering, 35(3), 517- 526.
7. Azzaoui, H., Boukhamla, A. Z. E., Arroyo, D., & Bensayah, A. (2022). Developing new deep-learning model to enhance network intrusion classification. Evolving Systems, 13(1), 17-25.
8. Chhajer, P., Shah, M., & Kshirsagar, A. (2022). The applications of artificial neural networks, support vector machines, and long–short term memory for stock market prediction. Decision Analytics Journal, 2, 100015.
9. Fernandez, G.C.; Xu, S. A Case Study on using Deep Learning for Network Intrusion Detection. In Proceedings of the MILCOM 2019—2019 IEEE Military Communications Conference (MILCOM), Norfolk, VA, USA, 12–14 November 2019; pp. 1–6.
10. Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization”, 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, January 2018
11. Jamil, F., & Kim, D. (2021). An ensemble of prediction and learning mechanism for improving accuracy of anomaly detection in network intrusion environments. Sustainability, 13(18), 10057.
12. Kaja, Nevrus, Adnan Shaout, and Di Ma. "An intelligent intrusion detection system." Applied Intelligence 49 (2019): 3235-3247.

13. Kaur, S., & Singh, M. (2020). Hybrid intrusion detection and signature generation using deep recurrent neural networks. Neural Computing and Applications, 32, 7859-7877.
14. Kim, A.; Park, M.; Lee, D.H. AI-IDS: Application of Deep Learning to Real-TimeWeb Intrusion Detection. IEEE Access 2020, 8,70245–70261.
15. Louati, F., & Ktata, F. B. (2020). A deep learning-based multi-agent system for intrusion detection. SN Applied Sciences, 2(4), 675.
16. Mhawi, D. N., Aldallal, A., & Hassan, S. (2022). Advanced feature-selection-based hybrid ensemble learning algorithms for network intrusion detection systems. Symmetry, 14(7), 1461.
17. Mighan, S. N., & Kahani, M. (2021). A novel scalable intrusion detection system based on deep learning. International Journal of Information Security, 20, 387-403.
18. Prachi, H. M., & Sharma, P. (2019). Intrusion detection using machine learning and feature selection. International Journal of Computer Network and Information security, 11(4), 43-52.
19. Qazi, E. U. H., Almorjan, A., & Zia, T. (2022). A One-Dimensional Convolutional Neural Network (1D-CNN) Based Deep Learning System for Network Intrusion Detection. Applied Sciences, 12(16), 7986.
20. Sahu, A. K., Sharma, S., Tanveer, M., & Raja, R. (2021). Internet of Things attack detection using hybrid Deep Learning Model. Computer Communications, 176, 146-154.
21. Sethi, K., Madhav, Y. V., Kumar, R., & Bera, P. (2021). Attention based multi-agent intrusion detection systems using reinforcement learning. Journal of Information Security and Applications, 61, 102923.
22. Sun, P., Liu, P., Li, Q., Liu, C., Lu, X., Hao, R., & Chen, J. (2020). DL-IDS: Extracting features using CNN-LSTM hybrid network for intrusion detection system. Security and communication networks, 2020, 1-11.
23. Zou, J., Han, Y., & So, S. S. (2009). Overview of artificial neural networks. Artificial neural networks: methods and applications, 14-22.