International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 07 | July - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

A SURVEY ON DEEP LEARNING TECHNIQUES FOR DARKNET TRAFFIC MALWARE DETECTION

Cynthia J
Research Scholar
Department of Computer Science and Engineering
Government College of Technology, Coimbatore
cynthia.phd@gct.ac.in

Dr.S.Rathi
Professor
Department of Computer Science and Engineering
Government College of Technology, Coimbatore
rathi@gct.ac.in

ABSTRACT

Encrypted data that is sent over Tor or a VPN is referred to as "darknet traffic." In order to identify network activity brought on by a cyberattack, it is crucial to be able to recognise, find, and explain darknet activities. Cyberattacks that pose a serious danger to network security and management can be effectively observed through darknet monitoring and classification. Despite the fact that many surveys were conducted on network traffic classification using machine learning techniques, only a small number of researchers have the means to review their work on classifying network traffic using deep learning techniques. In this article, we introduce a novel idea for research on the classification of malware used in darknet traffic that uses Deep Learning techniques, which will aid researchers in improving their surveys. First, based on the survey, we chose a few techniques, including Deep Neural Network (DNN), Convolution Neural Network (CNN), Recurrent Neural Network (RNN), and AutoEncoder (AE) that have proven to be more effective in recent study. Second, the dataset CIC-DarkNet-2020, which contains a variety of darknet activities including VPN and TOR traffic, is used to develop the models. Finally, after analysing the information, we discovered the best Deep Learning Model for classifying Darknet traffic, which has the potential to enhance the efficiency of malware variant detection using constrained system and network resources.

Index Terms – Darknet, DNN, CNN, RNN, AE, Deep Learning, Classifying Darknet.

1. INTRODUCTION

Darknet is the portion of the internet's address space that isn't utilised and isn't thought to interact with other computers. The term ‘Dark’ is given to it due to its cryptocurrency, online marketplace, and anonymity. Due to its passive listening nature, which accepts incoming packets but does not support outgoing packets, any communication from the dark space is regarded with scepticism. Any traffic is considered to be unsought and is typically viewed as probe, backscatter, or misconfiguration due to the absence of valid hosts in the darknet. Network telescopes, sinkholes, and blackholes are further names for darknet [1].

Stuxnet, a sophisticated malware programme discovered in 2010, was used for the first time to attack an Iranian nuclear power station. A highly developed piece of malware called Flame with significant spying capability was found in 2012. Additionally, a rising trend in DDoS attacks, particularly DRDoS (Distributed Reflection Denial of Service attack), has been observed in recent years. These attacks are designed to overwhelm and disable the services of major businesses by flooding the targeted victim with amplified network traffic [2]. For instance, the greatest DRDoS attack on the Internet ever experienced in 2014 peaked at an uncontrolled pace of 400 Gigabits per second.

YEAR	HISTORY OF DARKNET
1960	FORMATION OF ARPANET
1970	ILLEGAL ONLINE TRANSACTIONS
1980	DATA HAVENS
1990	INTERNET ACCESIBILITY AND THE RISE OF ILLEGAL MUSIC STREAMING
1999	CREATION OF NAPSTER
2000	THE LAUNCH OF FREENET
2002	THE LAUNCH OF TOR
2009	BITCOIN
2010	ARAB SPRING
2012	MARCO POLO
2013	SILKROAD SHUTDOWN
2013	SNOWDEN WHISTLEBLOWING
2015	PLAYPEN SHUTDOWN

2017	ALPHABAY
2020	DARKWEB MARKET

Table 1: History of Darknet

Network traffic monitoring and analysis is difficult because they aid in boosting network performance, reducing your attack surface, enhancing security, and more effectively managing resources. Tor is a virtual computer network that allows users to access concealed Darknet content. Darknet traffic is often seen as misconfiguration because of the host's illegal actions. A detailed examination of darknet traffic is obviously required to monitor real-time apps. [3].

IDSs now in use have demonstrated ineffectiveness in identifying various attacks, including zero-day attacks, and lowering the false alarm rates (FAR). This eventually leads to a need for a reliable, accurate, and affordable NIDS to give the network good security. The researchers have looked into the use of artificial intelligence (AI) approaches to satisfy the need of a successful IDS. [4]

Through the use of encryption methods and peer-to-peer connection networks, Darknet provides anonymous services to individual users and may successfully fend off route eavesdropping and other traffic analysis techniques. In the first quarter of 2020, around two million people connected directly to Tor services, whereas just 50,000 users did so using bridges [5].

The complete analysis of DL approaches and strategies pertinent to security in this paper fills the gap between DL and security.

In conclusion, this survey intends to:

- Identify the cybersecurity applications that utilise DL methods.
- Identify the difficulties in successfully implementing DL in cybersecurity.
- Give a thorough analysis of research that uses DL approaches for cybersecurity.
- Determine the most significant and effective research areas.

To the best of our knowledge, the research papers and books that are examined in Section II of the literature only partially address the objectives of this work. Following is the remainder of the article. Section III examines key DL methods used in Darknet or potentially applicable there. Section IV examines the metrics and datasets utilised in the evaluation of Darknet applications. Section V examines current DL research publications. Section VI lists the key takeaways, present problems, and unexplored future directions. Section VII brings the study to a culmination.

2. METHODOLOGY

This work analyses the published journal articles between 2017 and the first quarter of 2022 and conducts a thorough literature assessment of various DL-based NIDS. A systematic literature review is a strategy used to locate, analyse, and extract relevant information from the existing literature pertaining to certain study topics. This systematic review was created in two stages. To get a preliminary list of articles, Phase 1 determines the information resource (search engine) and keywords to use in a query. Phase 2 applies specific criteria to the initial list to choose the core and most pertinent items, which are then stored in the final list and discussed in this article. Figure 1 shows the entire research process employed by this paper. This review article's main goal is to provide answers to the following queries:

- What are the most recent developments in DL-based NIDS design?
- What recent DL approaches have been used for designing the traffic flow of the darknet?
- What are the advantages and disadvantages of each method that has been adopted?
- Which datasets were most recently utilized for testing?
- Which evaluation metrics are most frequently used to demonstrate performance?
- What will the research's future directions be?

2.1 SYSTEM FOR DETECTING DARKNET TRAFFIC: CONCEPT AND CLASSIFICATION

2.1.1 CONCEPT

The darknet is an encrypted area of the internet that uses non-standard communication protocols to be purposefully inaccessible on the internet and is not indexed by search engines. It also requires special configuration or authorization to access.

The majority of darknet traffic, such as that resulting from malware transmission, Internet scanning, or backscatter DDoS events falls under the second category. For a variety of applications such as IDS, packet sniffer detection tools, DDoS detection or backscatter traffic detection, traffic that enters a darknet can be exploited to uncover patterns.

The surface web, a section of the internet where webpages are easily seen using conventional web browsers and are easily indexed by prominent search engines, is the area of the internet where the majority of internet users browse content online. The majority of what the typical user see online may be on the surface, but the dark web has many more levels of concealed information. Users can access these secret websites and engage in both legal and illegal activity while hiding their IP address by using a special browser called The Onion Router (Tor).

Freenet, a University of Edinburgh student's thesis project that aimed to develop a "Distributed Decentralized Information Storage and Retrieval System," is thought to have marked the start of the dark web in 2000. Clarke set out to develop a brand-new platform for file sharing and anonymous online communication. The Tor Project, which debuted in 2002 and introduced a browser in 2008, was built on this foundation. Since the development of Tor, users have been able to access the internet anonymously and explore what is known as the "dark net."

The complete idea of the darknet is depicted in Figure 1. The surface web, deep web, and darknet are the three sections of the network. The diagram below shows how most illicit and covert services are provided on the dark web.

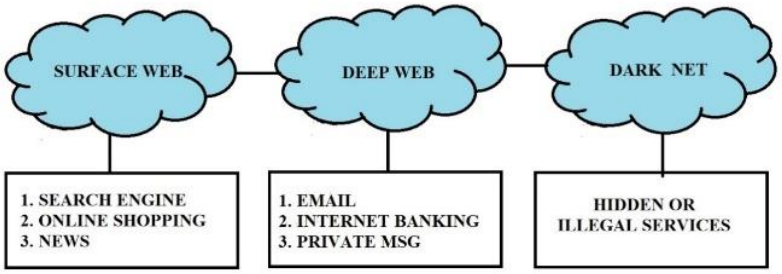

Figure 1: Concept of Darknet Traffic

2.1.2 DARKNET TRAFFIC INTRUSION DETECTION METHOD CLASSIFICATION

A system called Network intrusion detection system (NIDS) monitors network traffic for suspicious activity and sends out alerts when it is found. IDS come in a wide variety, ranging from antivirus software to hierarchical systems that keep an eye on all of the traffic in a backbone network. [7] Location-based and detection-based classifications are the most prevalent types.

Depending on where intrusion detection is used, there are two distinct types of IDSs. (1) Host-based IDSs keep an eye on every host, and if they see any suspicious activity, including changing system files or configuration settings, they notify the user. (2) Network-based IDSs investigate unusual patterns in network traffic. They are installed in a router or gateway, which is a network node. IDSs can also be categorised based on how the intrusion was detected. (1) Malicious network activity patterns are used by signature-based detection systems, also known as misuse-based systems, to identify known attacks. They offer great detection accuracy but are unable to pick up new (zero-day) attacks. (2) Contrarily, anomaly-based detection systems look for undiscovered attacks. The definition of typical and abnormal behaviour patterns serves as the foundation for the detection. The various approaches implemented in the darknet IDS are the main topic of this article.

2.1.3 RELATED WORK

A unique method to detect malicious network traffic is put forth by Le et al. [10] and is based on graph theory ideas including degree distribution, maximum degree, and distance metrics. The authors use the traffic dispersion graphs (TDG) technique to model the network traffic. As a result, they develop a method to recognise attack patterns and evaluate the variations of TDG graphs in time series to find harmful activity. The method was demonstrated with the use of actual network traces.

To monitor and validate the effectiveness of these models for IDS, Vinayakumar et al.[14] proposed a variety of supervised learning algorithms, such as multi-layer perceptrons (MLP) and hybrid CNN models like CNN-long short-term memory (CNN-LSTM), CNN-gated recurrent unit (CNN-GRU), and CNN-recurrent neural network (CNN-RNN).

Antonio Montieriet al. [8] categorise and evaluate starting with the Tranalyzer2 tool, the traffic is based on the features from the Anon17 dataset. Only a portion of these features are available in the dataset because some of them (such the ICMP and VLAN features) were eliminated as they were not relevant for fingerprinting. Five supervised classification algorithms—(i) Naive Bayes, (ii) Multinomial Nave Bayes, (iii) Bayesian Networks, (iv) C4.5, and (v) Random Forest - that have been effectively applied to the case under investigation in this study. Results show that the classifiers taken into consideration perform incredibly well (at least 97% accuracy when the amount of features is appropriately chosen) in differentiating the anonymity networks in both flow-based and early classifier.

Identifying traffic categories is a fundamental and important skill for network research, according to Cong Dong et al. [9]. Many programmes rely on the type of traffic to offer more sophisticated services. Because of their strong and abstract modelling capabilities, machine learning and deep learning are mostly used in current studies on the classification of encrypted communications to address this difficulty. The strategy put forward in this paper combines Payload Based and Flow Based two-dimensional analysis. One fully connected layer and two 1-D CNN layers are used effectively in Payload Based EE-CNN. Others include the stacked autoencoder (DP-SAE) model and the CNN model (DP-CNN), both of which exhibit great performance in their work. In the flow-based baseline, 23 features are chosen to describe the behaviour of the traffic, such as statistics of the inter-arrival time, flow duration, and others. Two models, C4.5 (C-C4.5) and KNN (C-KNN), are adopted, with the Gradient Boosting Classifier (FM-GBDT) and Random Forest (FM-RF) also being used as comparisons.

Muhammad Bilal Sarwar et al, [3] proposed Deep Learning technique has been used to provide a broad approach for classifying and detecting darknet traffic that provides excessive information about the darknet traffic and performs data pre-processing. Next, examine several feature selection strategies to choose the best features for categorising and detecting darknet traffic, and then apply fine-tuned machine learning algorithms, such as Decision Tree (DT), Gradient Boosting (GB), Random Forest Regressor (RFR) and Extreme Gradient Boosting (XGB) on selected features and compare the performance. Apply modified Convolution-Long Short-Term Memory (CNN-LSTM) and Convolution-Gradient Recurrent Unit (CNN-GRU) deep learning techniques to recognize the network traffic more accurately.

Han.Cet al, [11] proposed the model by combining three different machine learning techniques into a single framework called Dark-TRACER, they have proposed algorithms that automatically estimate and detect anomalous spatiotemporal patterns of darknet traffic in real time. They also conducted quantitative experiments to assess this framework's capacity to detect these malware activities.

A. H. Lashkari et al, [5] In order to detect and classify darknet traffic, this work introduces a unique approach called DeepImage that employs feature selection to choose the most vital features to build a grey image and feed it to a two-dimensional convolutional neural network. Here two encrypted traffic datasets are combined to produce a darknet dataset.

E. F. Fernandez et al. [12] discussed the challenges faced by security services in tracking criminal activity on the Darknet emphasised these challenges. It takes a lot of time to analyse all the images on the Darknet, and it's still not very effective. In order to address this problem, researchers investigated various automated image classification techniques based on Semantic Attention Keypoint Filtering (SAKF) and proposed a method that can *omit non-relevant topographies at the deep pixel level, effectively filtering out all pixels that are irrelevant to the forensic procedure. By combining the salience maps with Bag of Visual Words, this system is developed (BoVW). On a specially created dataset that includes the Tor image, the researchers tested their technique

To differentiate between darknet traffic and legitimate traffic, Singh et al [13] proposed deep transfer learning architecture, which consists of a trained model and a baseline classifier, was presented. Ten pre-trained models, including AlexNet, ResNet18, ResNet50, ResNet101, DenseNet, GoogLeNet, VGG16, VGG19, Inceptionv3, and SqueezeNet, were used along with three different baseline classifiers, including support vector machine, decision tree, and random forest, to identify the optimised pretrained network. Additionally, they claimed that using VGG19-based features and random forest, traffic data could be classified with 96% accuracy.

Singh et al. [20] experimentation with two models, a deep CNN model and ResNet-50, allowed them to detect malware using a deep convolutional-based approach. However, they first transformed the Malimg dataset's grayscale photos into a colour image before utilising it for categorization with 98.10% accuracy.

Using a combination of a recurrent neural network (RNN) and a convolutional neural network, Lopez-Martin, M. et al. [22] hypothesised that the raw network traffic is made up of the first 784 bytes of each session and the output contains 14 alternative network traffic flows (CNN). Highly uneven frequency distribution services were included in the dataset that was obtained. The UDP and TCP flow packet headers were mined for the statistical information. Table 2 provides a comprehensive analysis of the deep learning methods and performance reported in recent studies.

AUTHOR & YEAR	METHODOLOGY	FINDINGS	DATASET	ACCURACY
Zheng W et al.[17], 2020	RBRN model with GLOVE model	With a few shot meta learning, RBRN may achieve high classification performance and outperform state of art techniques for categorising encrypted traffic.	ISCX VPN-nonVPN traffic and ISCX 2012 IDS	0.9713
Lyu Q et al.[18], 2019	CNN and MLP	MLP has outperformed CNN in the classification of picture, text and video traffic, CNN is very good at the classification of audio traffic.	Self-dataset	0.96
Kim et al.[19] 2018	1D- CNN Model	This method can be used to predict traffic in real time using the high-velocity IDS or traffic classification system because there is no manual feature extraction.	UNB-CIC Tor network traffic dataset	0.993
Mallik et al.[16], 2022	Convolutional Recurrence model with stacked BiLSTM	Visual Representation of Malware, Augmenting the Samples, Feature Extraction, Feature Processing, Ensembling the features, and Classifying and Hyperparameter Tuning	real-life dataset,Malimg dataset, BIG2015 dataset	0.9956
Liu et al.[24], 2022	1D-CNN and WEKA tool	The general framework of the proposed heterogeneous network traffic identifier for network traffic identification and characterisation	ISCX VPN- nonVPN and ISCX TOR-nonTOR dataset	0.956
Jin et al.[25] , 2020	CNN and AutoEncoder	To investigate whether the autoencoder can detect malware by evaluating the error value and reconstruct malware images with minimal loss.	Andro-dumpsys study dataset	0.96
D. Gibert et al.[21], 2020	Hydra Model	To properly describe malware features, we suggest a baseline system that combines the advantages of feature engineering and deep learning Techniques.	Microsoft Malware Classification Challenge dataset	0.9975
K. Demertzis et al.[26], 2021	Weight-Agnostic Neural Network	Framework for managing Darknet traffic and automating the real-time detection of hazardous activity	CICDarknet-2020 dataset	0.9714
Sarwar et al.[3],2021	CNN-LSTM model and XGB feature selection Approach	A classification method with two stages and two labels that may distinguish between protocols and applications	CICDarknet-2020 dataset	0.95
Hardhik et al.[27],2022	Stacked Ensemble Model	To improve the overall effectiveness of darknet characterisation, combine the predictions of the three basic learners Random Forest, K-Nearest Neighbors, and Decision Tree in the most effective manner possible.	CICDarknet-2020	0.9889

Table 2: Summary of DL techniques in Darknet traffic detection

3. DEEP LEARNING TECHNIQUES FOR CLASSIFYING DARKNET TRAFFIC

This section outlines the overall process for classifying Darknet traffic using AI as well as specifics on the most popular deep learning (DL) algorithms that were utilised to create an effective model. Both machine learning (ML) and deep learning (DL) are broadly categorised as supervised and unsupervised algorithms, where the usable information is derived from the labelled data. In contrast, unsupervised algorithms employ unlabelled data to extract important features and information.

For the privacy protection and the detection of malware, supervised learning techniques are frequently utilised in the field of cyber security. Applications for Android have been proposed using semi-supervised learning. While in IoT scenarios and for network intrusion detection, unsupervised learning is desirable.

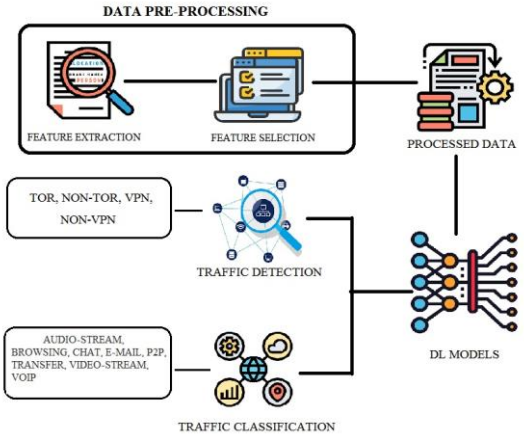

Figure 2. General Architecture to Classify Traffic in Darknet

3.1. DEEP LEARNING MODELS IN DARK NETWORK TRAFFIC CLASSIFICATION

Different modules make up deep learning (DL) techniques, which change how data is represented as it moves up the neural network hierarchy. In the DL model, the first layer gets the unprocessed input data and has many intermediate hidden layers that deal with more difficult issues before producing the output. The key benefit of using deep learning (DL) instead of classical machine learning is its higher performance on large datasets.

In order to learn data representations with various levels of abstraction, deep learning methods offer a computational architecture that integrates many processing layers. DL can be divided into three categories: reinforcement-based or hybrid learning, unsupervised learning or generalized learning, and supervised learning or discriminative learning.

3.1.1 DEEP NEURAL NETWORK (DNN):

A fundamental DL structure known as DNN enables the model to learn in layers. It is also known as Feed Forward Neural Network. It is made up of multiple hidden layers as well as an input layer and an output layer. To model intricate nonlinear functions, DNN is utilised. Since deep neural networks (DNN) have numerous hidden layers and overcome the disadvantage of shallow neural networks (SNN), which only have one hidden layer, they are better able to respond to complex datasets.

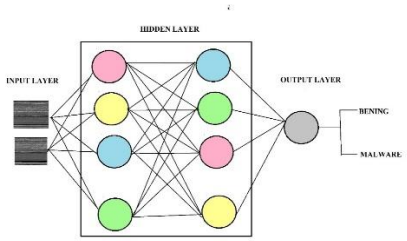

Figure 3: DNN Model

The DNN architecture was presented by M. Alimoradi et al. [14] to categorise Darknet Traffic such as TOR, Non-TOR, VPN, and Non-VPN. The raw input is taken and transferred to the hidden layer, where the rectified linear unit (ReLU) is used as an activation function, and to the output layer, where the softmax is utilised. Following is a formula for the DNN:

$$Z = \sum_{i=1}^{n} w_i \, x_i + b$$

$$\alpha = \sigma(Z)$$

$$y = softmax(\alpha)$$

Where Z is the hidden layer; the output of the hidden layer is given to ReLU activation layer (σ) max(0,x); then the output is then stored in α. The softmax layer is then offered for categorization. This model performed better than ML methods when tested on the publicly accessible dataset CICDarkNet 2020 dataset.

3.1.2 CONVOLUTIONAL NEURAL NETWORK (CNN)

CNN is a deep learning algorithm that processes images into grid-like patterns. These are made to automatically recognise and separate distinct images and to teach spatial properties at various levels of complexity. An input layer, a stack of convolutional and pooling layers used for feature extraction, and lastly a fully connected layer and a softmax classifier in the classification layer build up the structure.

Convolutional layers are used as the first layer, according to Rodriguez et al. [14], to extract various features from the input image before executing the operation between the input image and the filter.The feature map output is created by swiping the filter over the input, and it is sent into the pooling layer where the image's dimension size is decreased. The FC layer receives a flattened version of the input image from the preceding levels. The weights and biases make up the FC layer. Typically, these layers come before the output layer. When all characteristics are linked to the FC layer, the training dataset may become overfit. A dropout layer is used to get over this issue, where a few neurons are removed from the neural network. Dropout enhances performance since it makes the network less complex, preventing overfitting. The activation function is one of the most key aspects in the CNN model. It determines if the input to the work is vital for prediction or not.

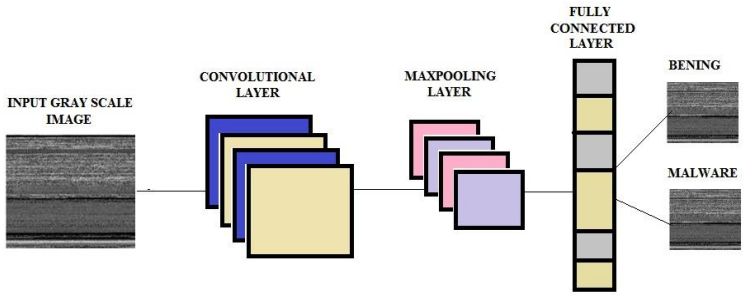

Figure 4: CNN Model

The BiLTSM was used to extract temporal features, and the CNN was utilised to extract spatial features. The CICDarkNet 2020 dataset was used for experiments. Accuracy and detection rate both perform better with the efficient implementation.

3.1.3 RECURRENT NEURAL NETWORK (RNN):

Recurrent neural networks (RNNs) are a type of neural network in which the results of one step are fed into the next step's computations. RNNs are the only type of neural network having an internal memory, making them one of the most potent and reliable ones currently being used. According to Kimmel JC et al. [15], RNN has two issues. First, RNN has a short-term memory problem, which means that lengthy inputs may make the RNN model forget past knowledge. Second, vanishing gradients can affect RNN models. The gradient value decreases at this point as the model backpropagates, which prevents the model from effectively learning. LSTM was developed to address these issues. The tanh and sigmoid (σ) activations are present in the LSTM. Inputs are compelled by sigmoid activations to fall between 0 and 1. Here, information is either kept or thrown away. The less significant a value is, the closer it is to 0. To guarantee that the output is regulated, the tanh activations maintain values between -1 and 1. Mathematically LSTM can be represented as

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

$$o_t = \sigma\left(W_o\ [h_{t-1}, x_t]\ +\ b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$

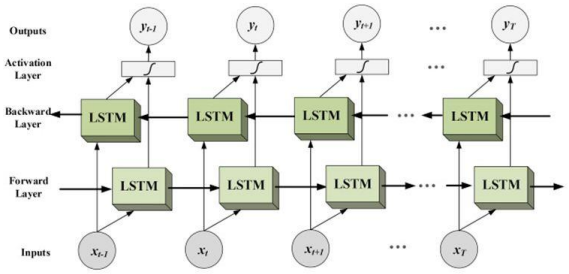

Figure 5: Bi-LSTM Model

A sequence processing model called Bi-LSTM (Bi directional LSTM) consists of two LSTMs, one of which receives input in the forward direction and the other in the backward direction. Effectively expanding the network's informational pool is Bi-LSTM technology. Mallik A et al.[16] retrieved features are processed by the BiLSTM layers, who then improve them to disclose hidden traits of the structural identities in the malware samples. Processing the derived features of an image depends on BiLSTM's ability to process the inputs consecutively. Additionally, BiLSTM aids in the better processing of the extracted features because it takes the input data from beginning to finish and end to beginning. This facilitates properly use the derived features.

3.1.4. AUTOENCODER (AE):

In order to find hidden connections between data and represent data in a more concise dimension, autoencoders are neural network-based models that are used for unsupervised learning. An autoencoder consists of an encoder that takes in the input and creates a lower dimensional encoding, a decoder that receives the encoding and reconstructs the input, and a bottleneck that is the lower dimensional hidden layer. The main applications of autoencoders are anomaly detection, noise reduction, and generative models.

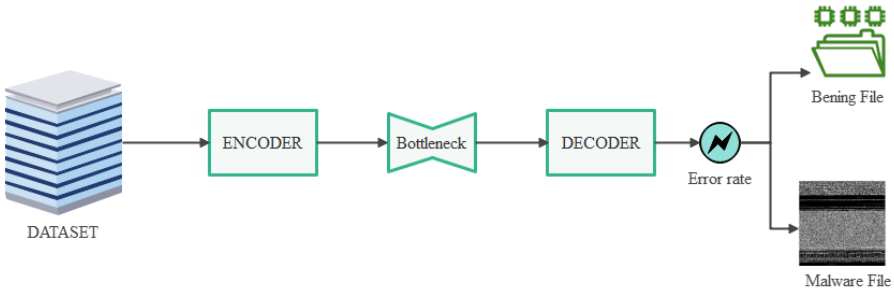

Figure 6: AutoEncoder Model

Research was conducted using the CICDarkNet 2020 dataset. Stacked AutoEncoder performs better than Deep AutoEncoder in terms of accuracy and detection rate when compared to other models because of their efficient implementation.

3.1.5. Deep Learning Model Review in Darknet traffic classification:

We examined at relevant research that investigates, analyses, and categorises darknet data and discovered a few techniques and technologies that are useful for spotting network anomalies. The comparison table is created using the darknet data shown in Figure 7., which allows for the detection and classification of anomalies.

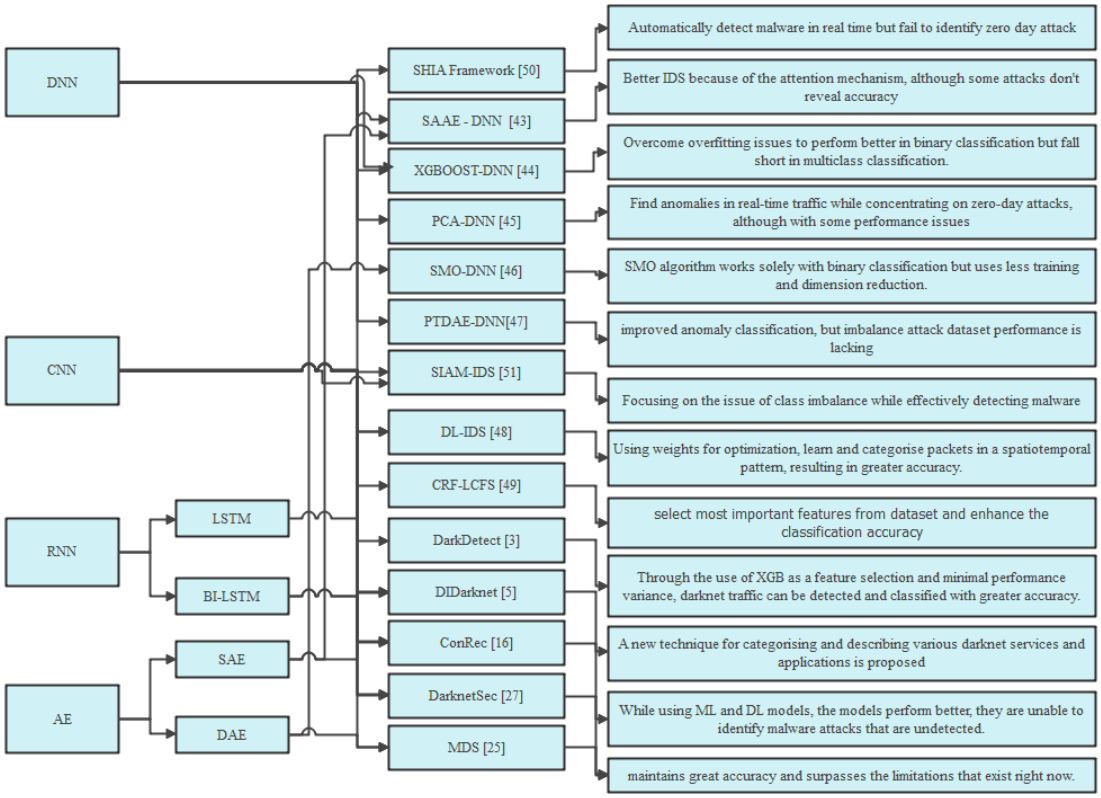

Figure 7: Pros and Cons in different Darknet framework and models

4. Evaluation Metrics and Dataset

The most important DL techniques utilised in the realm of cybersecurity were covered in the previous section. The evaluation metrics and datasets utilised in the training and testing phases are described in this section.

4.1 Evaluation Metrics

Three categories of evaluation metrics—threshold, likelihood, and ranking metrics can be made.

- When evaluated with unused data, the trained classifier's quality is measured and summarised using the evaluation metric.
- The evaluation metric task is to choose the top classifier from among many trained classifier types that concentrate on the greatest future performance (optimal model) when evaluated using illustrative data.
    - During the classification training, the evaluation metrics were used as a discriminator to identify and choose the optimal solution (best solution) out of all the generated solutions.

The accuracy (ACC), Error Rate (ERR), precision (p), false alarm rate (FAR), true positive rate (TPR), false positive rate (FPR), specificity, receiver operating characteristic (ROC) curve, area under the curve (AUC), and F1 Score metrics [4] are used to assess the DL-based cybersecurity works examined in this survey. A confusion matrix, or matrix representation of the classification results, can be used to calculate these measures.

		Actual Values: Positive (1)	Actual Values: Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Figure 8: Performance Measures

The Accuracy metric calculates the Accuracy (acc) ratio of accurate estimates over all instances considered.

$$ACC = \frac{TP+TN}{TP+TN+FP+FN}$$

The positive patterns in a positive class that are correctly predicted from the total predicted patterns are measured by precision (P).

$$P = \frac{TP}{TP+FP}$$

The number of false alerts divided by the total number of warnings or alarms in a specific study or circumstance is known as the false alarm ratio (FAR)

$$FAR = \frac{FP}{TP+FN}$$

The proportion of accurate estimates in predictions of the positive class is given by the True Positive Rate (TPR).

$$TPR = \frac{TP}{TP+FN}$$

The percentage of positive patterns that are successfully categorised is measured by recall (R).

$$R = \frac{TP}{TP+TN}$$

4.1 Datasets

This study makes use of the CIC-DarkNet-2020 dataset, which was generated by Arash et al. (2020) [1] and made available by the Canadian centre for cybersecurity (CIC). The CIC-DarkNet-2020 dataset is a unique and inclusive dataset that intelligently combines the traffic records of two publicly available datasets, (ISCXVPN2016 and ISCXTor2017), to produce a comprehensive dataset for Darknet traffic activities covering a wide range of Darknet activities, including VPN and Tor traffic.

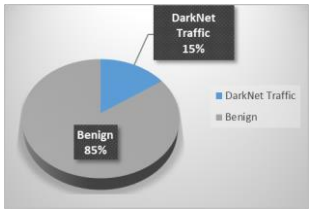

Fig. 9. No of samples of benign and Darknet Traffic

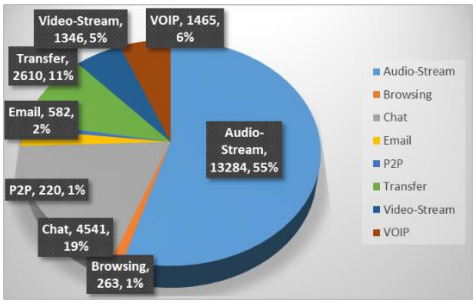

Fig.10. Number of encrypted flows in our darknet traffic

The CIC-DarkNet-2020 dataset includes chat traffic from programmes like ICQ, AIM, Skype, Facebook, and Hangouts, as well as audio- and video-stream traffic from websites like Vimeo and Youtube, browsing traffic from Firefox and Chrome, and more. Email traffic from programmes like SMTPS, POP3S, and IMAPS, as well as P2P traffic from programmes like Transmission and uTorrent (i.e., Bit Torrent), Use Filezilla to transfer traffic that is part of services like Skype, FTP over SSH (SFTP), and FTP over SSL (FTPS), as well as VOIP(**Voice over Internet Protocol**) that is part of services like Facebook, Skype, and Hangouts voice conversations. Figure 1 shows specifics regarding the number of samples of benign and darknet traffic, while Figure 2 emphasises the quantity of encrypted flows in our darknet traffic.

Using code, the dataset's details are provided.

```
[4]  subset.shape

(141530, 85)
```

```
subset.Label.value_counts()

Non-Tor    93356
NonVPN     23863
VPN        22919
Tor         1392
Name: Label, dtype: int64
```

V. PERSPECTIVES, CHALLENGES AND FUTURE TRENDS

This section first examines current patterns and perspectives in the classification of malware in encrypted network traffic using the suggested methodology, performance benchmarks, and dataset. Before describing future trends, it also highlights any potential research gaps and obstacles so that experts may create a dependable, efficient, and accurate malware detection system. Malicious hackers continue to develop new strategies to avoid detection and carry out their negative intents, despite the ongoing

advancements in countermeasure technology. This enables them to avoid detection by anti-malware engines. It is possible to use an anti-malware to understand how malware operates on a compromised node. Malware analysis and malware detection are two interrelated parts of the design of an anti-malware engine [29].

Some of the challenges faced by the researchers are: (1) To assess their findings, researchers used benchmark datasets. However, the publicly accessible datasets don't include recent network traffic's actual characteristics. Due to this, the majority of anomaly NIDSs are not suitable for use in production applications [30]. (2) Class imbalance is a problem that is highly prevalent nowadays. Despite the emphasis of many researchers, no adequate solution has been found, and as a result, it has a challenging impact on society today [31]. (3) Due to the potential harm to computer systems, malware types are the main rising risks to cybersecurity. For detecting malware variants, a variety of solutions have been put forth. However, due to the malware varieties' continual evolution, which results in concept drift, accurate identification is difficult [32]. (4) When attempting to classify malware, one of the challenges we encounter is the interpretability of models [33]. In addition, analytical methodologies for the family analysis, analysis of similarities, and analysis of variants were established. The effectiveness of the malware detection model was examined in relation to various selection and detection approaches [34].

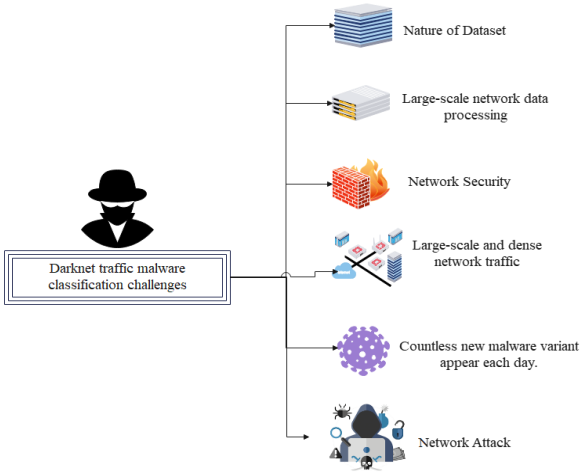

Figure 11: Challenges faced by Network Intruders

A solution has been suggested once the issue and challenges has been reviewed. Also covered in [38] was the application of CNN for the detection of hazardous code variations. The task included creating a visual, grayscale depiction of the damaging code, from which a CNN was created. However, the authors improved the neural network's efficiency by incorporating a self-attention mechanism, and as a result, their accuracy beat that of reference systems. A DRL (Deep Reinforcement Learning)-based approach for effective malware detection in a cloud setting was proposed in [39]. This technique proved successful in achieving cost-effective detection rates that were close to ideal. A DRL-based approach was also employed in the feature selection process and to learn the real-time feature distribution of the most recent malware variants [40, 41]. To increase the population of malware that may be identified and hence provide meaningful detection improvement, future research should focus on defining how to effectively combine detection techniques into hybrid solutions [37].

TECHNIQUES	PREVIOUS WORK	PROPOSED RESULT
DNN[44]	0.97	0.965
CNN+LSTM[5]	0.86	0.976
CNN+BI-LSTM[27]	0.922	0.98
AE[25]	0.93	0.934
CNN+SAE[24]	0.956	0.962
CNN+DAE[47]	0.833	0.867

Table 2: Accuracy for various methods – Existing vs Proposed

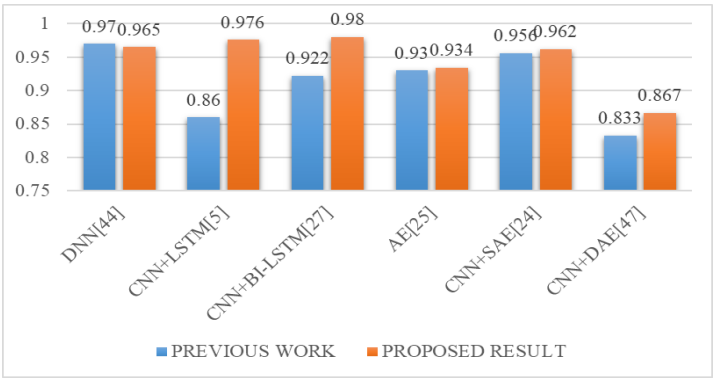

Figure 12: Graphical representation of Accuracy for various methods – Existing vs Proposed

CONCLUSION:

The classification of darknet traffic is crucial for the identification of online attacks and malicious behaviour. Users are now more at risk of their security and privacy violated due to the advancement of technology and excessive network usage. This review article offers a thorough examination of Darknet Traffic malware as well as methods for its analysis and detection. It goes into detail about the difficulties in using the tools and methods that are now available. These restrictions highlight the necessity to resolve the challenges in creating an anti-malware engine that suggests an attacker aware pro-active strategy. This study suggests a deep learning technique for classifying darknet traffic that performed better after numerous implementations. We make use of a recently released benchmark dataset, and based on various assessment criteria, Convolution Neural Network (CNN) and Bi-directional Long Short Term Memory (Bi-LSTM) outperform the related work with a significant difference of 10% in classification accuracy.

REFERENCE:

[1] Kevin Chen, Jennifer Tu, and Alex Vandiver. 2004. Analyzing Network Traffic from a Class B Darknet. MIT (2004).

[2] https://www.soscanhelp.com/blog/history-of-the-dark-web

[3] M. B. Sarwar, M. K. Hanif, R. Talib, M. Younas, and M. U. Sarwar, "DarkDetect : Darknet Traffic Detection and Categorization using Modified Convolution-Long Short-Term Memory," *IEEE Access*, vol. PP, no. Dl, p. 1, 2021, doi: 10.1109/ACCESS.2021.3105000.

[4] Hoque MS, Mukit M, Bikas M, Naser A, An implementation of intrusion detection system using genetic algorithm; 2012. arXiv preprint arXiv:1204.1336.

[5] A. H. Lashkari, "DIDarknet : A Contemporary Approach to Detect and Characterize the Darknet Traffic using Deep Image Learning," pp. 1–13.

[6] Verwoerd T, Hunt R. Intrusion detection techniques and approaches. ComputCommun. 2002;25(15):1356-1365. https://doi.org/10.1016/ S0140-3664(02)00037-3.

[7] Do Quoc Le, TaeyoelJeong, H. Eduardo Roman, and James Won-Ki Hong. Traffic dispersion graph based anomaly detection. In Proceedings of the Second Symposium on Information and Communication Technology, So ICT, pages 36–41, New York, NY, USA, 2011. ACM.

[8] A. Montieri, D. Ciuonzo, G. Aceto, and A. Pescape, "Anonymity Services Tor, I2P, JonDonym: Classifying in the Dark (Web)," *IEEE Trans. Dependable Secur. Comput.*, vol. 17, no. 3, pp. 662–675, 2020, doi: 10.1109/TDSC.2018.2804394.

[9] C. Dong, C. Zhang, Z. Lu, B. Liu, and B. Jiang, "CETAnalytics : Comprehensive effective traffic information analytics for encrypted traffic classification," *Comput. Networks*, vol. 176, no. July 2019, p. 107258, 2020, doi: 10.1016/j.comnet.2020.107258.

[10] Do Quoc Le, TaeyoelJeong, H. Eduardo Roman, and James Won-Ki Hong. Traffic dispersion graph based anomaly detection. In Proceedings of the Second Symposium on Information and Communication Technology, So ICT, pages 36–41, New York, NY, USA, 2011. ACM

[11] Han, C., Takeuchi, J., Takahashi, T., & Inoue, D. (2022). Dark-TRACER: Early Detection Framework for Malware Activity Based on Anomalous Spatiotemporal Patterns. *IEEE Access*, *10*, 13038–13058. https://doi.org/10.1109/ACCESS.2022.3145966

[12] E. F. Fernandez, R. A. V. Caro_lis, F. J. Martino, and P. B. Medina, "Classifying suspicious content in Tor Darknet," 2020, *arXiv:2005.10086*.
[13] Singh, D. , Shukla, A. , Sajwan, M. , 2021. Deep transfer learning framework for the identification of malicious activities to combat cyberattack. Future Gener. Com- put. Syst. 125, 687–697.
[13] R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Applying Convolutional Neural Network for Network Intrusion Detection," in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1222–1228, September 2017.
[14] Rodríguez E, Otero B, Gutiérrez N, Canal R. A Survey of Deep Learning Techniques for Cybersecurity in Mobile Networks. 2021;(2). doi:10.1109/COMST.2021.3086296
[15] Kimmel JC, McDole AD, Abdelsalam M, Gupta M, Sandhu R. Recurrent Neural Networks Based Online Behavioural Malware Detection Techniques for Cloud Infrastructure. *IEEE Access*. 2021;9:68066-68080. doi:10.1109/ACCESS.2021.3077498
[16] Mallik A, Khetarpal A, Kumar S. ConRec: malware classification using convolutional recurrence. *J Comput Virol Hacking Tech*. 2022;18(4):297-313. doi:10.1007/s11416-022-00416-3
[17] Zheng W, Gou C, Yan L, Mo S. Learning to Classify: A Flow-Based Relation Network for Encrypted Traffic Classification. Web Conf 2020 - Proc World Wide Web Conf WWW 2020. 2020;(Ml):13-22. doi:10.1145/3366423.3380090
[18] Lyu Q, Lu X. Effective media traffic classification using deep learning. *ACM Int Conf Proceeding Ser*. Published online 2019:139-146. doi:10.1145/3314545.3316278
[19] M. Kim and A. Anpalagan, "Tor Traffic Classification from Raw Packet Header using Convolutional Neural Network," *1st IEEE Int. Conf. Knowl. Innov. Invent. ICKII 2018*, no. Dl, pp. 187–190, 2018, doi: 10.1109/ICKII.2018.8569113.
[20] Singh, A., Handa, A., Kumar, N., Shukla, S. K.:Malware classification using image representation. In: Cyber Security Cryptography and Machine Learning, pp. 75–92 (2017)
[21]D. Gibert, C. Mateu, and J. Planes, "HYDRA: A multimodal deep learning framework for malware classification," *Comput. Secur.*, vol. 95, 2020, doi: 10.1016/j.cose.2020.101873.
[22] Lopez-Martin, M., et al.: Network traffic classifier with convolutional and recurrent neural networks for internet of Things. IEEE Access 5, 18042–18050 (2017). https://doi.org/10.1109/access.2017.2747560
[23] Yao, H., et al.: Capsule network assisted IoT traffic classification mechanism for smart cities. IEEE Internet Things J. 6(5), 7515–7525 (2019). https://doi.org/10.1109/jiot.2019.2901348
[24] F. U. Islam, G. Liu, W. Liu, and Q. M. ul Haq, "A deep learning-based framework to identify and characterise heterogeneous secure network traffic," *IET Inf. Secur.*, no. September, 2022, doi: 10.1049/ise2.12095.
[25] X. Jin, X. Xing, H. Elahi, G. Wang, and H. Jiang, "A malware detection approach using malware images and autoencoders," *Proc. - 2020 IEEE 17th Int. Conf. Mob. Ad Hoc Smart Syst. MASS 2020*, pp. 631–639, 2020, doi: 10.1109/MASS50613.2020.00009.
[26] K. Demertzis, K. Tsiknas, D. Takezis, C. Skianis, and L. Iliadis, "Darknet traffic big-data analysis and network management for real-time automating of the malicious intent detection process by a weight agnostic neural networks framework," *Electron.*, vol. 10, no. 7, 2021, doi: 10.3390/electronics10070781.
[27] J. Lan, X. Liu, B. Li, Y. Li, and T. Geng, "DarknetSec: A novel self-attentive deep learning method for darknet traffic classification and application identification," *Comput. Secur.*, vol. 116, May 2022, doi: 10.1016/j.cose.2022.102663.
[28] Hardhik Mohanty, Arousha Haghighian Roudsari, Arash Habibi Lashkari, "Robust stacking ensemble model for darknet traffic classification under adversarial settings", Computers & Security, Volume 120, 2022, 102830, ISSN 0167-4048, https://doi.org/10.1016/j.cose.2022.102830.
[29] Selvaganapathy, S. G., Sadasivam, S., & Ravi, V. (20DI ttack classification of an intrusion detection system using deep learning and hyperparameter optimization 21). A Review on Android Malware: Attacks, Countermeasures and Challenges Ahead. *Journal of Cyber Security and Mobility*, *10*(1). https://doi.org/10.13052/jcsm2245-1439.1017
[30] E. K. Viegas, A. O. Santin, and L. S. Oliveira, ''Toward a reliable anomaly-based intrusion detection in real-world environments,'' Comput. Netw., vol. 127, pp. 200–216, Nov. 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128617303225.
[31] Guan, J., Jiang, X., & Mao, B. (2021). A method for class-imbalance learning in android malware detection. *Electronics (Switzerland)*, *10*(24). https://doi.org/10.3390/electronics10243124
[32] Darem, A. A., Ghaleb, F. A., Al-Hashmi, A. A., Abawajy, J. H., Alanazi, S. M., & Al-Rezami, A. Y. (2021). An Adaptive Behavioral-Based Incremental Batch Learning Malware Variants Detection Model Using Concept Drift Detection and Sequential Deep Learning. *IEEE Access*, *9*. https://doi.org/10.1109/ACCESS.2021.3093366
[33] Lin, Y., & Chang, X. (2021). Towards Interpretable Ensemble Learning for Image-based Malware Detection. *ArXiv*, *abs/2101.0*.
[34] Li, J.; Cheng, K.; Wang, S.; Morstatter, F.; Trevino, R.P.; Tang, J.; Liu, H. Feature selection: A data perspective. ACM Comput. Surv. 2017, 50, 1–45.
[35] S. Arshad, M.A. Shah, A. Wahid, A. Mehmood, H. Song, and H. Yu, "SAMADroid: A Novel 3-Level Hybrid Malware Detection Model for Android Operating System," IEEE Access, vol.6, pp.4321–4339, 2018.

[36] A. Damodaran, F.D. Troia, C.A. Visaggio, T.H. Austin, and M. Stamp, "A comparison of static, dynamic, and hybrid analysis for malware detection," J. Computer Virology and Hacking Techniques, vol.13, no.1, pp.1–12, 2017.
[37] Caviglione, L., Choras, M., Corona, I., Janicki, A., Mazurczyk, W., Pawlicki, M., & Wasielewska, K. (2021). Tight Arms Race: Overview of Current Malware Threats and Trends in Their Detection. *IEEE Access*, . https://doi.org/10.1109/ACCESS.2020.3048319
[38] W. Li, R. Zhang, and Q. Wen, ''A malicious code variants detection method based on self-attention,'' in Proc. 6th Int. Conf. Comput. Technol. Appl., New York, NY, USA, Apr. 2020, pp. 51–56, doi: 10.1145/3397125.3397145.
[39] Y. Birman, S. Hindi, G. Katz, and A. Shabtai, ''Cost-effective malware detection as a service over serverless cloud using deep reinforcement learning,'' in Proc. 20th IEEE/ACM Int. Symp. Cluster, Cloud Internet Comput. (CCGRID), May 2020, pp. 420–429.
[40] L. Binxiang, Z. Gang, and S. Ruoying, ''A deep reinforcement learning malware detection method based on PE feature distribution,'' in Proc. 6th Int. Conf. Inf. Sci. Control Eng. (ICISCE), Dec. 2019, pp. 23–27.
[41] Z. Fang, J. Wang, J. Geng, and X. Kan, ''Feature selection for malware detection based on reinforcement learning,'' IEEE Access, vol. 7, pp. 176177–176187, 2019.
[42] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," 2018 IEEE Trans Emerg Top Comput Intell 2:41–50.
[43] C. Tang, N. Luktarhan, and Y. Zhao, "Saae-dnn: Deep learning method on intrusion detection," *Symmetry (Basel).*, vol. 12, no. 10, pp. 1–20, 2020, doi: 10.3390/sym12101695.
[44] Devan, P., Khare, N. An efficient XGBoost–DNN-based classification model for network intrusion detection system. *Neural Comput & Applic* **32**, 12499–12514 (2020). https://doi.org/10.1007/s00521-020-04708
[45] M. Al-Fawa'reh, M. Al-Fayoumi, S. Nashwan, and S. Fraihat, "Cyber threat intelligence using PCA-DNN model to detect abnormal network behavior," *Egypt. Informatics J.*, vol. 23, no. 2, pp. 173–185, 2022, doi: 10.1016/j.eij.2021.12.001.
[46] N. Khare *et al.*, "SMO-DNN: Spider monkey optimization and deep neural network hybrid classifier model for intrusion detection," *Electron.*, vol. 9, no. 4, 2020, doi: 10.3390/electronics9040692.
[47] Y. N. Kunang, S. Nurmaini, D. Stiawan, and B. Y. Suprapto, "Attack classification of an intrusion detection system using deep learning and hyperparameter optimization," *J. Inf. Secur. Appl.*, vol. 58, no. March, p. 102804, 2021, doi: 10.1016/j.jisa.2021.102804.
[48] 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*. https://doi.org/10.1155/2020/8890306
[49] B. Riyaz and S. Ganapathy, "A deep learning approach for effective intrusion detection in wireless networks using CNN," *Soft Comput.*, vol. 24, no. 22, pp. 17265–17278, 2020, doi: 10.1007/s00500-020-05017-0.
[50] R. Vinayakumar, M. Alazab, S. Member, and K. P. Soman, "Deep Learning Approach for Intelligent Intrusion Detection System," *IEEE Access*, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.
[51] P. Bedi, N. Gupta, and V. Jindal, "Siam-IDS : Handling class imbalance in Intrusion Detection Systems using Siamese Neural Network Siam-IDS : Handling class imbalance problem in Intrusion Detection Punam using *, Vinita Network Systems Siamese Neural," *Procedia Comput. Sci.*, vol. 171, no. 2019, pp. 780–789, 2020, doi: 10.1016/j.procs.2020.04.085.