# Enhancing Network Security: Pinpointing Similarities and differences of deep Learning Models for Firewall Intrusion Detection

**Ms. Asfiya Shireen Shaikh Mukhtar[1], Dr. R.N. Jugele**
[1]Research Scholar, Department of Computer Science,
Shivaji Science College, Nagpur.
[2]Professor, Department of Computer Science,
Shivaji Science College, Nagpur

**Abstract:** Firewall intrusion detection is an important part of network security, and the advent of deep learning has shown promise to improve its effectiveness. In this study, we perform an in-depth analysis to identify the similarities and differences between deep learning models used for firewall intrusion detection. Through extensive testing and evaluation, we evaluate the performance, robustness, and interpretability of popular deep learning architectures, including convolutional neural networks (CNNs), Recurrent neural networks (RNNs). Our results illuminate the strengths and weaknesses of each model and provide insight into their applicability in various intrusion detection scenarios. In addition, this research is a valuable resource for security professionals and researchers who want to use deep learning methods to strengthen network defenses.

**Keywords**: Firewall Anomaly Detection, Deep Learning Models, Cyber security Defense, Network Traffic Analysis

## 1. Introduction:

The security of systems and networks is critical in today's networked digital world. Rule-based systems and other conventional intrusion detection techniques are no longer sufficient to recognize and stop complex attacks due to the quick growth of cyber threats. Because of this, intrusion detection systems in firewalls are increasingly depending on deep learning models. A firewall monitors and regulates incoming and outgoing network traffic in accordance with pre-established security rules, serving as the first line of protection in a network. A firewall's capacity to identify anomalies and possible intrusions is improved by the real-time analysis of large volumes of network data that deep learning models provide. The present research attempts to investigate and identify commonalities and differences between several deep learning models applied to intrusion detection in firewalls. Security experts and researchers can choose and optimize DL-based solutions that best meet their security requirements by being aware of these subtleties.

## 2. Common Deep Learning Algorithm:

### 2.1 Recurrent neural networks (RNN):

RNNs are seen to be a suitable option when applying deep learning techniques to sequentially ordered data because the neural network can analyze all the network's state data. Traditionally, RNNs have a directed graph with temporal behavior formed by connections between hidden layers. When the trained data have substantial interdependencies and the output must be dependent on every previous condition, it is helpful. One of the main benefits of employing RNNs is that all the steps in an RNN have the same weights. This means that the total number of parameters the network needs to learn is reduced. Speech, text, and DNA sequences are a few real-world applications of RNNs.

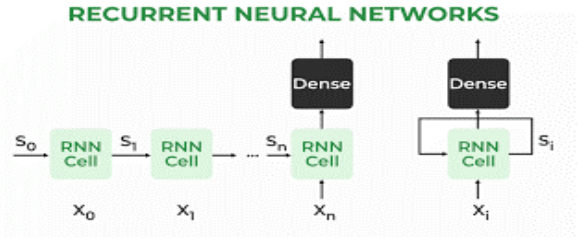


**Fig 1: Recurrent neural networks (RNN)**

### 2.2 Convolution neural networks (CNN)

Multi-dimensional connected data cannot be processed using neural arrangement models that have been studied from a distance. For example, the number of hubs and the characteristics become extremely crucial, almost unimaginable, when symbolism information is transferred as preparation information. A CNN architecture that makes use of the convolution channel veil is suggested [14]. It uses locally associated neurons and information specific components for learning instead of predetermined parts. The channel cover is linked to the complete image many times because CNN is specifically suggested for the analysis of visual information**.**

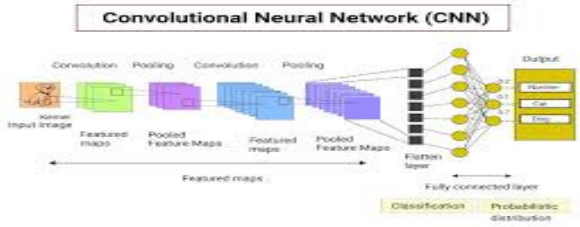


**Fig 2: Convolution neural networks (CNN)**

### 2.3 Deep Conventional Neural Network:

A normal feedforward neural network, Deep CNN uses the Back propagation (BP) algorithm to modify network parameters in order to lower the value of the cost function. (biases and weights). There are four innovative designs that set it apart from the traditional BP network significantly: the nearby responsive field, grouping, shared weight, and combination of various layers. Deep CNNs are intended to handle data using a topological network. It's widely used for object identification. In pictures, identify trends in temporal sequence data, and categorize sensor data [15]. Deep conventional neural networks are a unique kind of artificial neural network (ANN) that substitute's convolution for standard matrix multiplication in at least one of its layers.

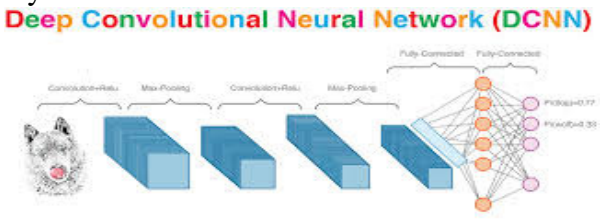


**Fig 3: Deep Conventional Neural Network**

### 2.4 Auto-encoders:

In auto encoders, unsupervised mode is used when the hidden layer reconstructs the input layer. Unlike an RNN that contains weights and offsets from an input state to a hidden state, the dimensions are specified. The dimensions of the hidden layer are smaller than those of the input layer, and a non-linear transformation function is used to calculate the stimulus of the hidden layer. At the same time as the dimensions of the hidden layer is decrease a dominant structure appears in the input. The hidden and input layer dimensions should be kept the same, the nonlinearity function should not be added to learn the detection function [17].

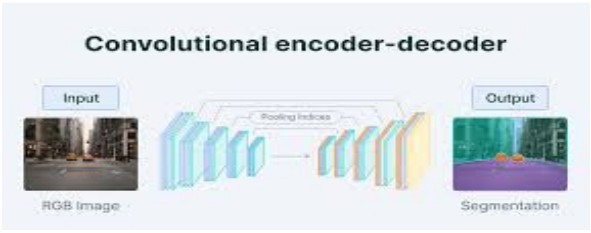


**Fig 4: Auto-encoder**

### 2.5 Deep Boltzmann Machine

Potential uses for the Deep Boltzmann Machine include improving the functionality of firewall intrusion detection systems. It is an effective tool for simulating complex distributions in data. Through the application of advanced machine learning techniques, DBMs present a promising path for enhancing network security by exploiting its capacity to build hierarchical representations and detect anomalies. Nonetheless, additional investigation and advancement are required to surmount obstacles and completely actualize their capabilities in pragmatic cyber security implementations.

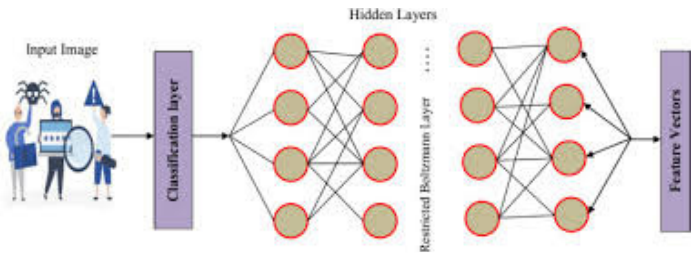


**Fig 5: Deep Boltzmann Machine**

## 3. Classification of deep learning Models:

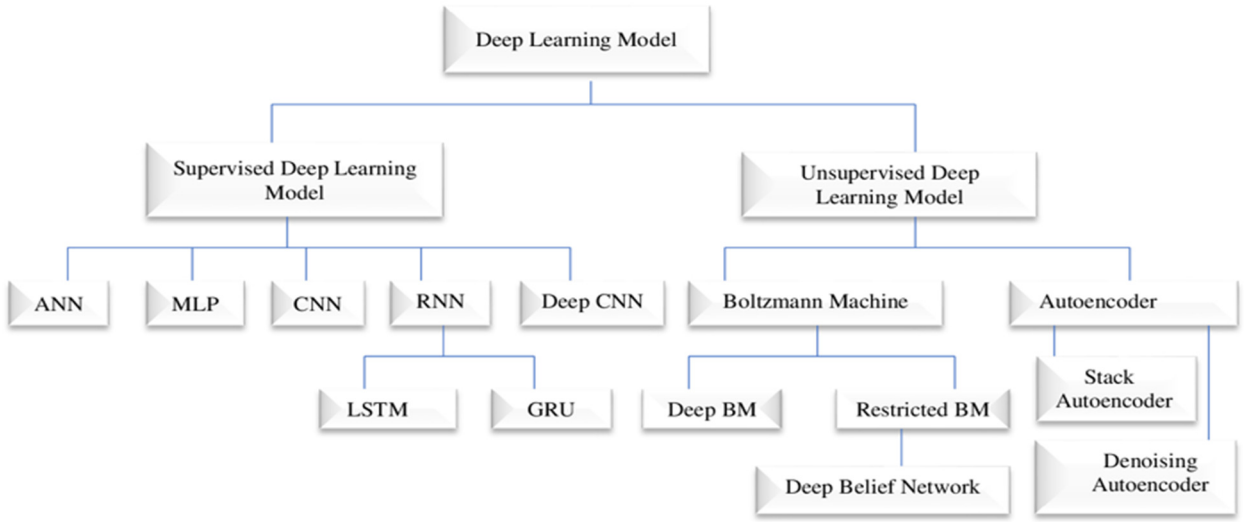


**Fig 5**: Hierarchical representation of deep learning models.

## 4. Characteristics of different Deep Learning Models

The following table shows the characteristics of different Deep Learning Models.

**Table I: Maching of different Deep Learning Algorithm**

| Characteristics | Deep Learning Algorithm | | | | | | |
|---|---|---|---|---|---|---|---|
| | CNN | RNN | Auto-encoder | GAN | LSTM | DBN | DNN |
| Supervised Learning | ✓ | ✓ | □ | □ | ✓ | □ | ✓ |
| Unsupervised Learning | □ | □ | ✓ | □ | □ | ✓ | □ |
| Classification | ✓ | ✓ | □ | □ | ✓ | □ | ✓ |
| Regression | ✓ | ✓ | ✓ | □ | ✓ | □ | ✓ |
| Scalability | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Bias-Variance Trade-off | □ | □ | □ | □ | □ | □ | □ |
| Generalization | ✓ | ✓ | ✓ | □ | □ | □ | ✓ |
| Adaptability | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Non-linearity | ✓ | ✓ | | □ | ✓ | □ | □ |
| Fault Tolerance | ✓ | ✓ | ✓ | □ | ✓ | □ | □ |
| Simple and Intuitive | ✓ | ✓ | ✓ | ✓ | □ | ✓ | ✓ |
| Non-Parametric | □ | ✓ | □ | □ | ✓ | □ | □ |
| Memory Efficient | □ | ✓ | □ | □ | ✓ | □ | □ |
| Versatile | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Robust to Overfitting | □ | □ | □ | □ | □ | □ | □ |
| Sensitive to Outliers | ✓ | ✓ | ✓ | ✓ | □ | □ | □ |
| Feature Scaling/Selection | ✓ | ✓ | □ | □ | ✓ | □ | ✓ |
| Hyper-parameter Tuning | □ | □ | □ | □ | □ | □ | □ |
| Probabilistic Model | ✓ | ✓ | □ | □ | □ | □ | □ |
| Interpretability | ✓ | ✓ | □ | ✓ | ✓ | □ | ✓ |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Robustness** | ✓ | ✓ | ✓ | ✓ | ✓ | □ | ✓ |
| **Sequential Data Processing** | ✓ | □ | □ | ✓ | □ | ✓ | ✓ |
| **Activation Functions** | ✓ | □ | ✓ | □ | □ | □ | □ |
| **Feed Forward Network** | ✓ | □ | □ | □ | □ | ✓ | ✓ |
| **Encoder-Decoder Architecture** | □ | □ | ✓ | □ | □ | □ | □ |
| **Dimensionality Reduction** | ✓ | ✓ | ✓ | □ | □ | □ | □ |
| **Regularization** | ✓ | ✓ | ✓ | □ | □ | ✓ | □ |

## 5. Pros and cons of various Deep Learning Models:

The following table summarize the pros and cons of various deep learning models.

**Table 2: Pros and cons of Deep Learning Algorithm.**

| Algorithm | Functions | Advantages | Disadvantages |
|---|---|---|---|
| **CNN** | Packet Payload Analysis<br>Malware Detection<br>Spatial and Temporal Analysis<br>Signature-based Detection | Hierarchical Feature Learning<br>Translation Invariance<br>Robustness to Local Variations<br>Feature Hierarchy Visualization | High Computational Cost<br>Large Memory -Requirements<br>Difficulty with Scale and Rotation Variance |
| **RNN** | Memory of Previous State<br>Sequential Data Modelling<br>Handling Variable-Length Sequences | Sequential Modelling<br>Temporal Dynamics<br>Backpropagation Through Time<br>Contextual Understanding | Vanishing and Exploding Gradient Problem<br>Memory Constraints<br>Difficulty with Long-Term Dependencies |
| **GAN** | Graph Structure Analysis<br>Interpretable Outputs<br>Node Embedding's<br>Transfer Learning | Generative Modelling<br>Feature Learning<br>Image Translation and Style Transfer<br>Enhanced Realism | Training Instability<br>Mode Dropping<br>Training Time and Resource Intensiveness |
| **DBN** | Generative Modelling<br>Adaptability to Unlabelled Data<br>Pre-training for Deep Architectures | Hierarchical Representation Learning<br>Feature Extraction<br>Flexible Architecture<br>Transfer Learning | Complexity and Training Time<br>High Memory Requirements<br>Hyper parameter Sensitivity<br>Data Dependency |
| **Auto-encoder** | Feature Learning<br>Noise Reduction<br>Generation of Synthetic Data<br>Visualization of Latent Space | Data Denoising<br>Interpretability<br>Robustness to Noise and Missing Data<br>Feature Learning | Difficulty with Handling Noisy Data<br>limited Generative Modelling Capacity<br>Sensitivity to Hyper parameters<br>Overfitting |
| **LSTM** | Long-Term Dependencies<br>Detection of Temporal Anomalies<br>Multivariate Time Series Analysis | Long-term Dependencies Handling<br>Memory Cell<br>Gradient Flow<br>Versatility | Overfitting<br>Hyper parameter Sensitivity<br>Training Instability<br>Computationally Intensive |

## 6. CONCLUSIONS

Our study has provided valuable insights into the similarities and differences among deep learning models for firewall intrusion detection. We have identified the strengths and weaknesses of popular deep learning models such as convolutional neural networks (CNNs), Recurrent neural networks (RNNs), Deep Conventional Neural Network, Auto-encoders and Deep Boltzmann Machine Firstly, our findings indicate that CNNs excel in capturing

spatial dependencies within network traffic data, making them well-suited for detecting patterns indicative of intrusion attempts. RNNs demonstrate proficiency in capturing temporal dependencies and sequential patterns, which are prevalent in certain types of network attacks. Additionally, we have highlighted the significance of characteristics in influencing model performance, thereby enhancing the reliability of intrusion detection systems. Overall, our study underscores the importance of selecting the most appropriate deep learning model based on the specific requirements and characteristics of the intrusion detection. By understanding the nuances of different architectures and their respective strengths, security practitioners and researchers can make informed decisions when designing and deploying firewall intrusion detection systems.

## 7. REFERENCES

**[1]** Zeyuan Fu, "Computer Network Intrusion Anomaly Detection with Recurrent Neural Network", Mobile Information Systems, vol. 2022, Article ID 6576023, 11 pages, 2022. https://doi.org/10.1155/2022/6576023

**[2]** El-Nagar, Ahmad &Zaki, Ahmad & Soliman, Fouad & el Bardini, Mohammad. (2022). Hybrid deep learning diagonal recurrent neural network controller for nonlinear systems. Neural Computing and Applications. 34. 1-20. 10.1007/s00521-022-07673-9.

**[3]** Khan, Muhammad Ashfaq. 2021. "HCRNNIDS: Hybrid Convolutional Recurrent Neural Network-Based Network Intrusion Detection System" Processes 9, no. 5: 834. https://doi.org/10.3390/pr9050834

**[4]** Ermal Elbasani, Jeong-Dong Kim, "LLAD: Life-Log Anomaly Detection Based on Recurrent Neural Network LSTM", Journal of Healthcare Engineering, vol. 2021, Article ID 8829403, 7 pages, 2021. https://doi.org/10.1155/2021/8829403

**[5]** M. Rathika, P. Sivakumar, K. Ramash Kumar, Ilhan Garip, "Cooperative Communications Based on Deep Learning Using a Recurrent Neural Network in Wireless Communication Networks", Mathematical Problems in Engineering, vol. 2022, Article ID 1864290, 12 pages, 2022. https://doi.org/10.1155/2022/1864290

**[6]** Jin Gao, Jiaquan Liu, Sihua Guo, Qi Zhang, Xinyang Wang, "A Data Mining Method Using Deep Learning for Anomaly Detection in Cloud Computing Environment", Mathematical Problems in Engineering, vol. 2020, Article ID 6343705, 11 pages, 2020. https://doi.org/10.1155/2020/6343705

**[7]** Junjie Cen, Yongbo Li, "Deep Learning-Based Anomaly Traffic Detection Method in Cloud Computing Environment", Wireless Communications and Mobile Computing, vol. 2022, Article ID 6155925, 8 pages, 2022. https://doi.org/10.1155/2022/6155925

**[8]** Hivehch, Yahya DorostkarNavaei, "Local and Deep Features Based Convolutional Neural Network Frameworks for Brain MRI Anomaly Detection", Complexity, vol. 2022, Article ID 3081748, 11 pages, 2022.

**[9]** Chao Wang, Bailing Wang, Hongri Liu, Haikuo Qu, "Anomaly Detection for Industrial Control System Based on Autoencoder Neural Network", Wireless Communications and Mobile Computing, vol. 2020, Article ID 8897926, 10 pages, 2020. https://doi.org/10.1155/2020/8897926

**[10]** Lerina Aversano, Mario Luca Bernardi, Marta Cimitile, Riccardo Pecori, Luca Veltri, "Effective Anomaly Detection Using Deep Learning in IoT Systems", Wireless Communications and Mobile Computing, vol. 2021, Article ID 9054336, 14 pages, 2021. https://doi.org/10.1155/2021/9054336

**[11]** R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," arXiv preprint arXiv:1901.03407, 2019.

**[12]** LirimAshiku, CihanDagli, Network Intrusion Detection System using Deep Learning, Procedia Computer Science, Volume 185, 2021, Pages 239-247, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2021.05.025.

**[13]** JOUR Wanjau, Stephen Wambugu, Geoffrey Oirere, Aaron 2022/06/01 16 Network Intrusion Detection Systems: A Systematic Literature Review of Hybrid Deep Learning Approaches 10 10.35940/ijese. F2530.0610722 International Journal of Emerging Science and Engineering.

**[14]** Parametrict-Distributed Stochastic Neighbor Embedding Combined with Hierarchical Neural Network for Network Intrusion Detection. Int. J. Netw. Secur., 22, 265-274.

**[15]** Elkhadir, Zyad& Khalid, Chougdali & Benattou, Mohammed. (2018). Improving Network Intrusion Detection Using Geometric Mean LDA. International Journal of Network Security. 820-826. 10.6633/IJNS.201809.20(5).02).