2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC) | 979-8-3503-7651-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICICEC62498.2024.10808423

Adaptive Intrusion Detection in CAN-Based Vehicular Networks Using Transfer Learning for Evolving Threats

Dodda Venkatareddy, Shaik Yalavarthi Ijaz Ahamad, Sinde Venkata Saptha Girish, Tammisetti Nagendra Babu, K. V. Narasimha Reddy

Dept. of CSE, Narasaraopeta Engineering College, Narasaraopet, Palnadu, Andhra Pradesh, India

doddavenkatareddy@gmail.com, sk.ijazahamad390@gmail.com, sindesapthagirish@gmail.com, babubabu43154@gmail.com, narasimhareddynec03@gmail.com

***Abstract*—This is the CAN-based system security, which is constantly under serious threats due to increasing vehicular network connectivity through sophisticated cyber-attacks. In this paper, we propose an online reconfigurable IDS which uses TL methods to adapt to new attack patterns with minimal retraining. That would allow refining the pre-trained models on the specialized car hacking dataset to detect most of the known and new attacks efficiently, ensuring high detection accuracy while keeping the computation cost low. Dynamic reconfigurability assures protection in an ever-evolving threat landscape. Extensive experiments on real-world CAN datasets validate the effectiveness of the proposed approach, achieving an overall detection rate of over 99% for different types of attack classes. The approach here is toward demonstrating the potential of TL for enhancing adaptability, efficiency, and accuracy in improving connected vehicle IDSs through a sound solution to secure automotive systems against the emergence of cyber threats**.

***Keywords*—Hybrid Model, Convolutional Neural Networks (CNN),Transfer Learning, Intrusion Detection System (IDS), Controller Area Network (CAN), Evolving Threats, Reconfigurable System.**

I. INTRODUCTION

Modern vehicle CANs have increasingly complex designs that make them highly vulnerable to cyberattacks. Most of the traditional IDSs, which are static model-based, find it difficult to adapt to new and evolving threats in dynamic environments. Motivated by this challenge, the paper proposes a novel reconfigurable IDS based on TL to improve the security of vehicular networks using CAN. This approach combines CNNs with other machine learning techniques by making the system dynamically adapt to emerging threats without an extensive process of retraining. This increases the detection of new and unknown attacks considerably while maintaining the accuracy and efficiency of the detection. While previous work has already shown the potential of TL in enhancing IDS performance, our work extends them to provide a more generalizable and resilient security approach for increasingly connected and hence fragile automotive systems.

II. RELATED WORK

Much development has occurred in the field of IDS for CAN, but more so recently with the incorporation of Transfer Learning and other complex machine learning methods. Recently, much attention has been brought to these
approaches through various research works, which should form the backbone for ensuring security in vehicular networks.

One of the major approaches recently adopted by methods trying to improve the performance of IDSs in CAN networks has been doing what is called transfer learning. In particular, Khatri et al. present how Transfer Learning can be used in enhancing IDS capability by exploiting pre-learned models that evolve with new attack patterns-a very critical advance in addressing the evolving threats in vehicular networks [1]. Accordingly, Al-Jarrah et al. proposed new methods of detecting unknown cyber-attacks using recurrence plots and neural networks. This work is likely to be followed by more developmental works in terms of adaptable IDS solutions [3].

Other works have also been conducted in the field of deep learning-based IDS. In this line, Khan et al. (2023) discussed automotive theft detection based on deep learning methods and pointed out the role that high-level machine learning methods can play in the identification and mitigation of sophisticated attacks [2]. Similarly, Xu et al. (2024) used GANs for CANFD bus IDS and thus demonstrated the advantages offered by generative models in the improvement of threat detection capabilities [7].

While much research is going into the integration of hybrid approaches, Salek et al., in 2023, proposed a hybrid quantum-classical IDS framework for CAN networks-a novelty in integrating the principles of quantum computing into traditional security approaches. Such hybrid approaches find complements in advanced works; for instance, Islam et al.'s comprehensive frequency-agnostic IDS in 2024 had to ensure that there is robust performance over diverse attack scenarios. Graph-based and transformer-based methods have also played their role in enriching literature. On one hand, it was in 2023 that Park et al.

suggested a graph-based IDS for the CAN protocols. On the other hand, in-vehicle intrusion detection has been investigated using transformer-based attention networks by Nguyen et al. in 2023, hence two new directions toward the enhancement of IDS are provided [8][9]. Dong et al. in 2023 also developed HMM-based systems for CAN bus IDS; hence, the range of methods to be pursued with a view to enhancing security in vehicular networks is extended [10].

Integrated, these studies show the trends that have taken place within IDS research and indicate the possibility of integrating advanced techniques, including Transfer Learning, with respect to the cyber threats that are emerging.

III. PROPOSED METHODOLOGY

The proposed IDS for the CAN-based vehicular network is developed by fusing deep learning methods such as Hybrid CNN-LSTM, Hybrid Transfer Learning, RNNs, and Random Forest. In this work, the approach of the authors has aimed at providing an adaptive detection mechanism against the evolving cyber threats by combining deep learning with traditional machine learning methods. The system shall be tested using the "Car-Hacking dataset" which includes Denial of Service (DoS) attack, Fuzzy attack, Revolutions Per Minute Spoofing (RPM) attack, and gear spoofing attack [1][4].

A. Overview

Recently, there is a surge in in-vehicle network cyberattacks. We implement the intrusion detection using a hybrid model that models the spatial and temporal patterns of CAN traffic. Making use of the pre-trained models will enable the system to be well adaptive with more speed towards new attack vectors [1][4].

B. Data Collection and Preprocessing

Data Collection: This approach is based on the extended car hacking dataset consisting of the wide range of attack variants-from DoS to Fuzzy attacks, RPM manipulation, and gear spoofing. Therefore, it will be of great importance for training and testing models in the detection of malicious activity across the CAN-based network [1].

Preprocessing Data:

-Data Cleaning and Filtering: Filtering the noise and irrelevant data from the data raises its quality itself [2][6]. Below are the propositions made:

-Feature Extraction: relevant features extracted include byte frequency and payload entropy for the detection of abnormal patterns in CAN traffic [1][3].

-Normalization: Features are normalized in such a way that the data is following a normal distribution. This leads to better convergence rates of the model [4].

-Data Augmentation: Oversampling for balancing the dataset, generation of synthetic data can be some techniques in order to avoid model bias [7].

C. Hybrid CNN-LSTM Model

Capturing the spatial and temporal patterns of the CAN traffic data are the two most imperative aspects of the proposed hybrid CNN-LSTM model.

CNN Component:

- Convolutional Layers: These are useful for feature extraction from the CAN traffic data as a local byte sequence that may indicate malicious activities in it [1][3].
- Pooling Layers: These layers reduce the feature maps by reducing their dimensionality. The focus is on important features, thereby reducing computational complexity in the process [5].

LSTM Component:

- LSTM Layers: Temporal dependencies in CAN data can indicate continuous attacks. To learn and process these temporal patterns of data, LSTM cells are used here.

Model Training: The CNN-LSTM architecture is jointly trained using backpropagation, with a cross-entropy loss function. Effective convergence is achieved by the Adam optimizer [7].

D. Hybrid Transfer Learning Model

It fuses the pre-trained CNN with some extra layers finetuned with the car hacking dataset.

Pre-trained CNN: A base CNN model that is pre-trained on a large generic dataset. The initial layers capturing the general features are frozen, and the final layers are retrained on the CAN dataset to learn domain-specific features [2] [8].

Fine tuning indeed tuned the model to the peculiarities of CAN traffic hence the increase in its capability regarding the detection of both known and emerging threats [9].

E. RNNs Model

RNNs model learns sequential dependencies within CAN data for capturing the pattern time-varying.

RNN Layers: RNN Layers: The applications also add many RNN layers on top of each other with the purpose of learning the intrinsic temporal dependency characteristics of sequential CAN traffic data [1][7].

Training: Against overfitting, the training of the RNN model uses BPTT. This training process employs an Adam optimizer and Early Stopping.

F. Random Forest Model

Random Forest was used both for a baseline and because of its complementary nature, bringing robustness against overfitting and informative outputs.

Random Forest Ensemble Learning: Random Forest is an ensemble of decision trees wherein the trees are trained on random subsets of the data; thus, it predicts the final result by taking an average over all the predictions made by a tree
[3][10].

Training: The best splits are decided considering Gini impurity and information gain as the criterion at the time of model training for better class classification.

G. Continuous Reconfiguration

The IDS is continuously reconfigurable; it updates its models to address emerging threats. By retraining the models, they keep pace with newly discovered attack vectors in order to retain the effectiveness of the Hybrid TL model. These models are deployed in real time to monitor CAN traffic and produce alerts upon detection of anomalies. Since this architecture is modular, upgrades and integrations of new models are thus made pretty easy [4][9][7]. It integrates multiple models for a flexible and robust IDS of CAN-based vehicular networks with evolving threats detected using the car hacking dataset
[1][4].

IV. EXPERIMENTAL SETUP AND PERFORMANCE EVALUATION

A. Experimental Setup

1) Dataset Description: For this work, the experiment was based on the car-hacking dataset, which consists of various types of attacks: DoS attack, Fuzzy attack, RPM spoofing attack, and gear spoofing attack. In fact, this dataset contains a large amount of normal and attack messages described below in this table.

TABLE I. DESCRIPTION OF CAR HACK DATASET

Attack Type	* Normal Messages	* Attacked Messages
Fuzzy Attack	3,347,013	491,847
RPM Spoofing	3,966,805	654,897
Gear Spoofing	3,845,890	597,252
DoS Attack	3,078,250	587,521
Attack-Free (Normal)	988,872	-

B. Experimental Design

This section describes the configurations of the models used for experiments.

Hybrid CNN-LSTM Model: Define the architecture by identifying the number of convolutional layers, the filter size, the strategy of pooling, the amount of LSTM layers, and any other additional layers.

Hybrid Transfer Learning (TL) Model: Explain which pre-trained CNN model was used, the number of layers frozen, and which layers are fine-tuned.

RNNs Model: Provide any information on the architecture of the RNN regarding the quantity of layers vs. units per layer. Random Forest Model: The parameters, including the number of trees, maximum depth, and criteria for splitting nodes, can be added.

TABLE II. ARCHITECTURE OF THE HYBRID CNN-LSTM MODE.

Layer (type)	Output Shape	Param
input_layer_1 (InputLayer)	(None, 1, 1)	0
conv1d_2 (Conv1D)	(None, 1, 32)	128
conv1d_3 (Conv1D)	(None, 1, 64)	6,208
max_pooling1d_1(MaxPooling1D)	(None, 1, 64)	0
flatten_1 (Flatten)	(None, 64)	0
reshape_1 (Reshape)	(None, 1, 64)	0
lstm_2 (LSTM)	(None, 1, 50)	23,000
lstm_3 (LSTM)	(None, 50)	20,200
dense_2 (Dense)	(None, 50)	2,550
dropout_1 (Dropout)	(None, 50)	0
dense_3 (Dense)	(None, 2)	102

Total Parameters: 52,188 (203.86 KB)
Trainable Parameters: 52,188 (203.86 KB)
Non-Trainable Parameters: 0 (0.00 B)

C. Training Procedure

1) Training Parameters: In other words, the learning rate can be considered the modification in weights that is allowed by the model with respect to the error. The typical range of learning rates can be from 1×10^{-5} to 1×10^{-1}.

Batch size: The number of examples trained per iteration. Common numbers are 32, 64, and 128.

Epochs: Between 10 to 100, which define how many times the whole dataset is passing through the model during training.

Optimizers: These represent the algorithms used to tune the weights in order to get the minimum value of the loss function. Examples are Adam, SGD, and RMSprop.

Optimizer parameters include learning rate and decay; tuning these will provide further fine-grained control over training. Each model must define a batch size, learning rate, number of epochs, and type of optimizer along with the parameters that apply.

2) Validation Strategy: Early Stopping: This is an effective method to stop training when performance on the validation set is getting worse, hence possibly leading to overfitting. Specify which criteria are used for early stopping. Cross-Validation: Describes how the dataset should be divided into folds for training and validation.

D. Performance Evaluation

Accuracy: Accuracy is the overall measure of rightness in the model's predictions. It tells us how many of the model's predictions are right compared to the total number of predictions.

Precision: Precision actually refers to the quality of the positive model predictions. It actually means what percentage of instances the model classified as positive were correct.

Recall: Recall, or Sensitivity, is the measure that informs us about how much of all the relevant positive examples a model captures. It gives an idea about how well it identifies true positives from actual positives.

F1 Score: The F1 Score is the harmonic mean of Precision and Recall. This yields a single number that gives a real balance of the trade-off between Precision and Recall as a measure of how well the model identifies positive instances with the least possible false positives.

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

$$\text{Precision} = TP/(TP+FP)$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1 Score} = \frac{2 x (\text{Precision} x \text{Recall})}{\text{Precision} + \text{Recall}}$$

Confusion Matrix: A confusion matrix visually tabulates counts of true positives, true negatives, false positives, and false negatives, thus showing the performance of the classification model. It helps in understanding how well the model discriminates among different classes and diagnosing areas where the model may be making errors. The confusion matrix is a 2x2 grid:

Predicted Class	Positive	Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

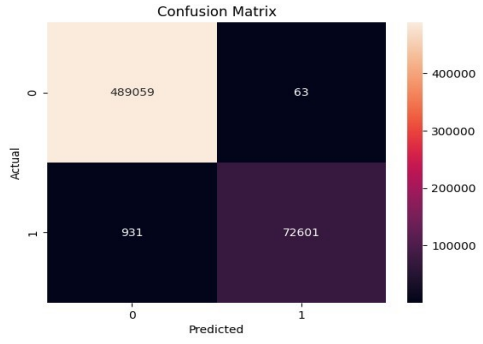

Fig. 1. Confusion matrix for fuzzy attack.

V. MODEL COMPARISON

A. Training and Validation Curves

The training and validation curves are pretty informative with respect to model learning and generalization across epochs. Plots included are:

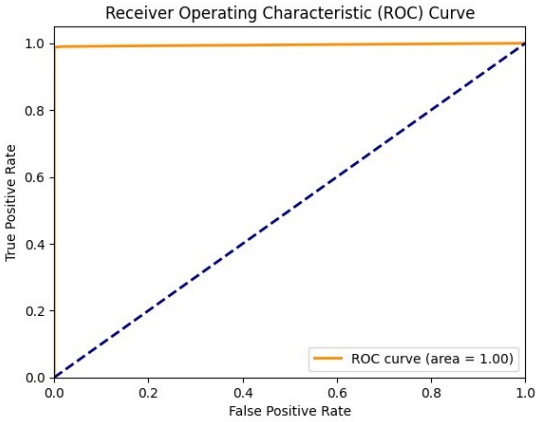

Fig. 2. Roc curve of hybrid tl model for fuzzy attack detection.

B. Performance Metrics

The comparison of the performance metrics-accuracy, precision, recall, F1-score, and ROC-AUC-for the Hybrid Transfer Learning Model, Hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) Model, and RNN Model on the Fuzzy Attack Dataset.

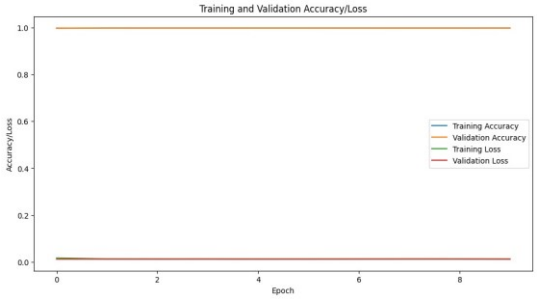

Fig. 3. Epoch-wise accuracy and loss for fuzzy attack detection.

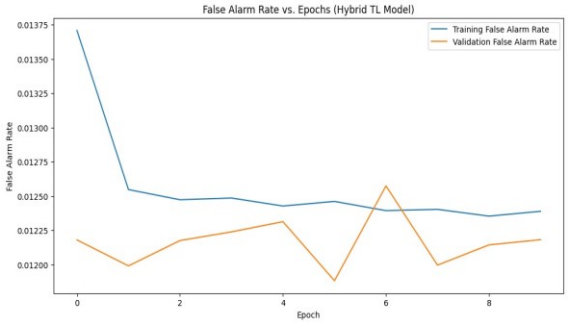

Fig. 4. False alarm rate for fuzzy attack (hybrid tl model).

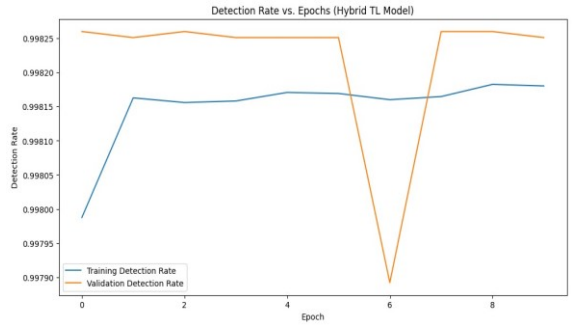

Fig. 5. Detection rate trends across epochs for hybrid tl on fuzzy attack dataset.

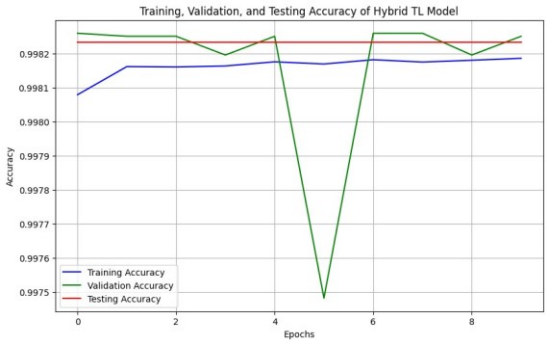

Fig. 6. Compring train, test, and validation accuracy for hybrid tl model.

TABLE III. PERFORMANCE METRICS FOR HYBRID TL MODEL ON THE FUZZY ATTACK DATASET

Metric	Precision	Recall	F1-Score	Support
Accuracy	0.998237	0.998237	0.998237	-
Macro Avg	0.998841	0.993399	0.996099	562,654
Weighted Avg	0.998240	0.998237	0.998232	562,654

TABLE IV. PERFORMANCE METRICS OF HYBRID CNN-LSTM MODEL FOR FUZZY ATTACK DATASET

Metric	Precision	Recall	F1-Score	Support
Precision	-	-	1.00	562,654
Macro Avg	1.00	0.99	1.00	562,654
Weighted Avg	1.00	1.00	1.00	562,654

TABLE V. PERFORMANCE METRICS OF DIFFERENT MODELS FOR FUZZY ATTACK DATASET

Model	Accuracy	Precision	Recall	F1-Score
Hybrid TL	0.998233	0.998235	0.998233	0.998229
Hybrid CNN-LSTM	0.998233	0.998235	0.998233	0.998229
RNN	0.998233	0.998235	0.998233	0.998229

VI. RESULT

As shown by the results above, Hybrid Transfer Learning, Hybrid CNN-LSTM, and RNN models within vehicular networks using CAN are practical and all competent in detection because all three achieve high values in terms of accuracy, precision, recall, and F1-score, meaning robust attack classification and detection. Future work would include the creation of more advanced model variants in order to increase detection precision and efficiency. Experimental study is planned across an extensive scope of attack types, real-world testing is considered for determining the performance of the software applications in practical terms and integration with other security approaches to further improve overall defense strategies.

VII. CONCLUSION

Our work proposes an online reconfigurable Controller Area Network Intrusion Detection System that uses transfer learning to enhance security in vehicular environments. It provides improved detection accuracy for known as well as unknown threats by an impressive average detection rate of more than 99%. Hybrid CNN LSTM and RNN have found a great integrative capability with traditional techniques of machine learning for the adaptation of dynamic nature in cyber threats.

The results of our experiments validate the suitability of our proposed IDS, which can keep high accuracy, precision, recall, and F1-score on various types of attacks. This research paper contributes to knowledge in automotive security and opens ways for further innovation in adaptive IDS solutions.

We forward this paper as a recommendation to advance hybrid models and even make real-world testing in enhancing the detection capabilities of IDS frameworks, moving forward. With the continuous evolving nature of cyber threats, this kind of research and development is very important to ensure the safety and security of connected vehicles in an increasingly complex digital landscape.

REFERENCES

[1] N. Khatri, S. Lee, and S. Y. Nam, "Transfer Learning-Based Intrusion Detection System for a Controller Area Network," *IEEE Access*, vol. 11, pp. 120963-120981, Sept. 2023.

[2] J. A. Khan, D. W. Lim, and Y. S. Kim, "A Deep Learning-Based IDS for Automotive Theft Detection for In-Vehicle CAN Bus," *IEEE Access*, vol. 11, pp. 112814-112827, Sept. 2023.

[3] O. Y. Al-Jarrah, K. El Haloui, M. Dianati, and C. Maple, "A Novel Detection Approach of Unknown Cyber-Attacks for Intra-Vehicle Networks Using Recurrence Plots and Neural Networks," *IEEE Open Journal of Vehicular Technology*, vol. 4, pp. 271-280, Mar. 2023.

[4] M. S. Salek, P. K. Biswas, J. Pollard, J. Hales, Z. Shen, V. Dixit, M. C. Chowdhury, S. M. Khan, and Y. Wang, "A Novel Hybrid Quantum-Classical Framework for an In-Vehicle Controller Area Network Intrusion Detection," *IEEE Access*, vol. 11, pp. 96081-96082, Aug. 2023.

[5] D. Javeed, M. S. Saeed, I. Ahmad, P. Kumar, A. Jolfaei, and M. Tahir, "An Intelligent Intrusion Detection System for Smart Consumer Electronics Network," *IEEE Transactions on Consumer Electronics*, vol. 69, no. 4, pp. 906-912, Oct. 2023.

[6] M. R. Islam, M. Sahalabadi, K. Kim, Y. Kim, and K. Yim, "CF-AIDS: Comprehensive Frequency-Agnostic Intrusion Detection System on In-Vehicle Network," *IEEE Access*, vol. 12, pp. 13971-13981, Jan. 2024.

[7] W. Xu, Y. Xu, Z. Wang, Y. Wu, and Y. Wang, "Intrusion Detection System for In-Vehicle CAN-FD Bus ID Based on GAN Model," *IEEE Access*, vol. 12, pp. 82402-82412, Sept. 2024.
[8] S. B. Park, H. J. Jo, and D. H. Lee, "G-IDCS: Graph-Based Intrusion Detection and Classification System for CAN Protocol," *IEEE Access*, vol. 11, pp. 39213-39225, Apr. 2023.
[9] T. P. Nguyen, H. Nam, and D. Kim, "Transformer-Based Attention Network for In-Vehicle Intrusion Detection," *IEEE Access*, vol. 11, pp. 55389-55403, June 2023.
[10] C. Dong, H. Wu, and Q. Li, "Multiple Observation HMM-Based CAN Bus Intrusion Detection System for In-Vehicle Network," *IEEE Access*, vol. 11, pp. 35639-35648, Mar. 2023.