Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

RNN-Based Intrusion Detection System for Internet of Vehicles with IG, PCA, and RF Feature Selection

Benni Purnama[1], Eko Arip Winanto*[2,4], Sharipuddin[3], Dodi Sandra[2], Nurhadi[3], Lasmedi Afuan[4]

[1] Information System, Dinamika Bangsa University, Indonesia
[2] Computer Engineering, Dinamika Bangsa University, Indonesia
[3]Informatics, Dinamika Bangsa University, Indonesia
[4]Computing, University of Technology Malaysia, Malaysia
[5]Informatics, Universitas Jenderal Soedirman, Indonesia

Email: [1]ekoaripwinanto@gmail.com

Received : Aug 30, 2025; Revised : Sep 3, 2025; Accepted : Sep 5, 2025; Published : Oct 16, 2025

Abstract

Cyberattacks in the Internet of Vehicles (IoV) threaten road safety and data integrity, requiring intrusion detection systems (IDS) that capture temporal patterns in vehicular traffic. This study develops a Recurrent Neural Network (RNN)-based IDS and evaluates three feature-selection strategies—Information Gain (IG), Principal Component Analysis (PCA), and Random Forest (RF)—on the CICIoV2024 dataset. Features are normalized using Min–Max scaling before being fed into the RNN classifier. The models achieve perfect classification on held-out tests (accuracy/precision/recall/F1 = 1.00). However, probabilistic evaluation reveals low ROC–AUC scores (IG: 0.572, PCA: 0.429, RF: 0.415), indicating limited discriminative margins and potential overfitting or calibration issues despite flawless confusion matrices. PCA and RF further reduce computational overhead during inference compared to IG. These findings highlight that relying solely on accuracy can be misleading for IDS evaluation; temporal RNNs should be complemented with probability-aware training, calibration, or hybrid architectures. This work contributes a temporal-aware IDS framework for IoV and motivates future research on real-time deployment, hybrid RNN-CNN/LSTM models, and adversarial robustness to improve generalization and safety of connected vehicles

Keywords : *Deep Learning, IG, Intrusion Detection, IoV, PCA, RF, RNN*

1. INTRODUCTION

The Internet of Vehicles (IoV) is an intelligent system that connects vehicles, road infrastructure, and data centers through the internet [1]. Its primary objectives are to improve transportation efficiency, road safety, and the overall driving experience. However, as vehicles become increasingly interconnected, the risk of cyberattacks rises significantly [2], [3]. Various threats may emerge, including Man-in-the-Middle attacks, Denial of Service (DoS), data spoofing [4], [5], and software exploitation within vehicles [6]. Such attacks not only compromise user data and privacy but also pose physical safety risks by enabling unauthorized control over vehicles [7].

Intrusion Detection Systems (IDS) have thus become essential for monitoring and identifying attacks that may threaten IoV networks [8], [9]. An effective IDS must be capable of recognizing attack patterns in real-time, given the continuous flow of data exchanged between heterogeneous IoV components [10]. Previous studies have explored multiple approaches to cyberattack detection in IoT, including machine learning [11], [12] and deep learning methods [13]. Convolutional Neural Networks (CNN) have been widely adopted for classifying attack patterns in vehicular networks [13], [14], while other research [15] has employed Long Short-Term Memory (LSTM) networks due to their effectiveness in handling sequential data.

Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

Nevertheless, Recurrent Neural Networks (RNN) offer advantages in modeling sequential and temporal dependencies more comprehensively than other approaches [7]. RNNs not only process current inputs but also leverage information from prior inputs, which is critical for detecting sequential attack patterns in IoV traffic [16]. As highlighted in prior studies [17], RNNs demonstrate superior capability in identifying cyberattacks in systems that require temporal pattern analysis. Therefore, this study proposes the use of RNN as an approach for intrusion detection in IoV environments.

The main contributions of this research are as follows: (1) developing an RNN-based IDS capable of accurately detecting cyberattacks in IoV networks, and (2) analyzing the performance of RNN in identifying various types of IoV cyber threats. By employing RNN, the proposed IDS is expected to effectively capture attack patterns and accurately detect diverse forms of intrusions in connected vehicular environments. The method will be evaluated against multiple cyberattack scenarios to validate its accuracy in IoV intrusion detection. Prior IoV IDS studies predominantly employ CNN/LSTM or ensemble learners and report strong accuracy but rarely examine probability calibration and ROC–AUC under temporal traffic dynamics. Our contribution is a temporal RNN pipeline combined with three complementary feature-selection schemes (IG/PCA/RF) on CICIoV2024, revealing a consistent accuracy AUC discrepancy (accuracy 1.00 vs. AUC 0.42–0.57). This exposes a practical risk for threshold-based alerting in safety critical systems and motivates probability-aware training and calibration. Compared with LSTM/CNN baselines reported on related vehicular datasets, our results match or exceed accuracy while surfacing latent probabilistic weaknesses that prior work seldom reports, thus filling a methodological gap in IoV IDS evaluation

This study is structured into five sections. The introduction presents the background and contributions of the work. The second section describes the methodology, followed by the results and discussion. The final section provides the conclusions derived from the study.

2. METHOD

2.1. Experiment setup

The experimental process, illustrated in Figures 1 and 2, consists of several main stages as follows:

- Dataset Preparation. The IoV traffic data from CICIoV2024 was preprocessed, including handling missing values, encoding categorical attributes, and splitting the data into training, validation, and testing sets.
- Feature Selection. Three feature selection methods were applied: Information Gain (IG), Principal Component Analysis (PCA), and Random Forest (RF).
- Normalization. The selected features were normalized using the Min–Max method to ensure all values fell within a uniform range.
- Detection Model. The normalized features were then fed into the Recurrent Neural Network (RNN) model to classify IoV traffic as either benign or attack.
- Performance Evaluation. The model's performance was assessed using Accuracy, Precision, Recall, and F1-Score to provide a comprehensive evaluation.

2.2. Dataset

The IoV dataset employed in this study is CICIoV2024 [18], which comprises multiple classes. This dataset, developed by the University of New Brunswick, was constructed using a fully functional 2019 Ford vehicle equipped with all Electronic Control Units (ECUs). It includes five types of attacks performed on the vehicle's internal architecture: Denial of Service (DoS), spoofing-GAS, spoofing-RPM, spoofing-SPEED, and spoofing-STEERING_WHEEL, along with benign traffic.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

2.3. Selection Feature

Feature selection plays a crucial role in improving the detection performance of IoV attacks. The objective of feature selection is to identify a subset of IoV features that effectively represent the data while eliminating irrelevant or redundant attributes [17]. By selecting the most significant features, the dimensionality of the dataset can be reduced, which in turn decreases computational overhead during the detection process [19]. Generally, feature selection methods are categorized into two main types: filter and wrapper approaches [20]. The filter approach functions as a preprocessing step, ranking features based on statistical measures, where highly ranked features are selected and applied to the IoV detection system [21]. In contrast, the wrapper approach selects features based on the actual performance of the detection system.

Figure 1 illustrates the feature selection process for IoV attack detection, which is later used in the training and testing stages of the RNN-based detection model. The process begins with inputting the dataset in CSV format, followed by applying feature selection techniques—Information Gain (IG), Principal Component Analysis (PCA), and Random Forest (RF)—to identify the most relevant features. The selected features are then validated and normalized, ensuring their suitability for subsequent stages in the detection framework.

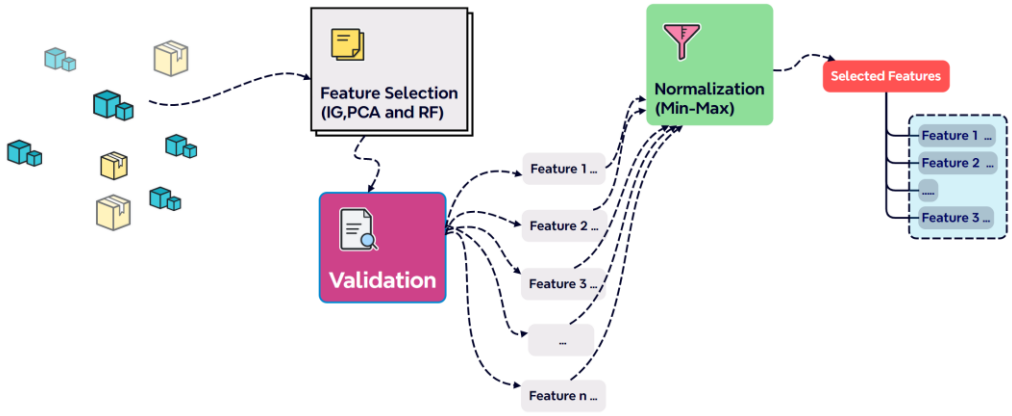

Figure 1. Feature selection stages in the IoV intrusion detection system. The process begins with feature extraction, followed by selection using Information Gain (IG), Principal Component Analysis (PCA), and Random Forest (RF). The selected features are then validated to determine the optimal subset, which is subsequently normalized using the Min–Max method before being utilized as input features for model training.

We perform stratified splitting into 80% training, 10% validation, and 10% testing. Missing values are imputed (median for numeric), categorical attributes are one-hot encoded. Min–Max Normalization. Each numeric feature x is scaled as:

$$X = \frac{X - \min(x)}{\max(X) - \min(x)} \qquad (1)$$

2.4. Model Detection using RNN

After feature selection and normalization, the next stage is intrusion detection in IoV using the Recurrent Neural Network (RNN) method. This stage constitutes the core of the system, as it directly performs the classification of network traffic into benign or attack categories.

Theoretically, RNNs are designed to process sequential data by maintaining a memory state of previous inputs, enabling the recognition of temporal patterns in network traffic [22], [23]. An RNN is

Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

characterized by three primary parameters: U, which connects the input to the hidden layer; W, which connects the hidden layer at the previous time step to the current one; and V, which connects the hidden layer to the output [24]. This mechanism allows the RNN to leverage contextual information from prior inputs when determining the current classification, making it particularly suitable for detecting attacks with sequential characteristics in IoV data.

The detection process using RNN consists of two main stages: training and testing. In the training stage, the model learns patterns from the preprocessed IoV dataset with selected and normalized features. The model parameters (U, V, W) are updated through an optimization algorithm to minimize the loss function. In the testing stage, the trained model is used to predict the class of new data, determining whether the traffic is benign or an attack.

Figure 5 presents the conceptual model of the IoV detection system using the RNN, illustrating the input–process–output flow. Meanwhile, Figure 2 depicts the system's implementation architecture, where the processed dataset is fed into the RNN detection model, and the classification results are evaluated using Accuracy, Precision, Recall, and F1-Score metrics.

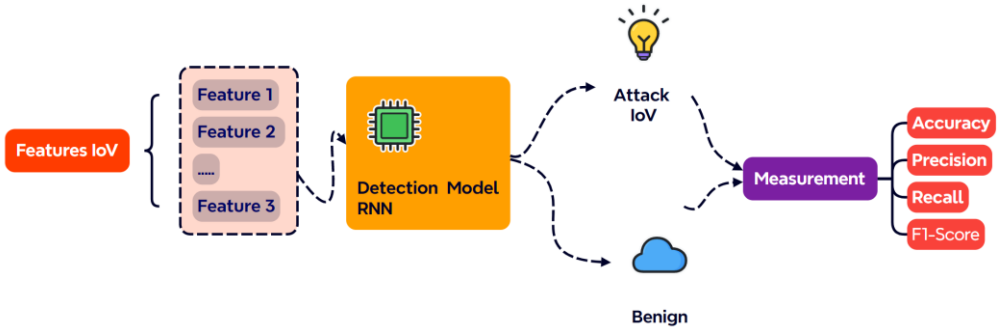

Figure 2. Workflow of the RNN-based intrusion detection system. The preprocessed IoV features are fed into the detection model to classify network traffic as either benign or attack. The classification results are then evaluated using Accuracy, Precision, Recall, and F1-Score metrics.

The end-to-end pipeline is summarized in Pseudocode 1. First, the CICIoV2024 dataset (D) is loaded. The dataset is then split into training, validation, and testing sets with a stratified ratio of 80/10/10. Next, feature selection methods such as Information Gain (IG), Principal Component Analysis (PCA), or Random Forest (RF) are applied on the training set to obtain K selected features or components. After that, a Min–Max scaler is fitted on the training set and applied to the validation and testing sets. The Recurrent Neural Network (RNN) is then trained with the specified hyperparameters and tuned using the validation set. Finally, the model is evaluated on the testing set using several metrics, including Accuracy, Precision, Recall, F1-score, and ROC–AUC.

Pipeline Pseudocode 1
Input D (CICIoV2024) → split D into train/val/test (80/10/10, stratified) → apply {IG | PCA | RF} on train to select K features / components → fit Min–Max scaler on train; transform val/test → train RNN (hyperparams above) on train; tune via val → evaluate on test: Accuracy, Precision, Recall, F1, ROC–AUC End

Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

We use a single recurrent layer with 128 units, followed by a dense output (softmax) for binary classification. Activation in the recurrent block uses tanh; we apply dropout=0.3 on inputs and recurrent dropout=0.1. Optimization uses Adam (lr=1e-3, β1=0.9,β2=0.999\beta_1=0.9, \beta_2=0.999β1=0.9,β2 =0.999), batch size=32, max 50 epochs, with early stopping (patience=7, restore_best_weights=True) on validation loss. Weight initialization: Glorot uniform. Class weights are balanced if train distribution is skewed.

2.5. Evaluation

In this study, several evaluation metrics are employed to assess the performance of the intrusion detection system in IoV. The key measurement variables considered are accuracy, precision, recall, **F1**, and **ROC–AUC** [25], [26], [27].

$$\text{Accuracy} \quad \frac{TP+TN}{TP+TN+FN+FP} \tag{2}$$

$$\text{Precision} \quad \frac{TP}{TP+FP} \tag{3}$$

$$\textit{Recall} \quad \frac{TP}{TP+FN} \tag{4}$$

3. RESULT

This section presents the experimental results conducted to enhance intrusion detection performance in IoV networks through feature selection methods and the RNN model. The experiments were designed to evaluate the contribution of each stage in the process, including feature selection using Information Gain (IG), Principal Component Analysis (PCA), and Random Forest (RF), data normalization, as well as RNN training and testing. The discussion in this section is organized into three main parts: (1) the results of feature selection analysis, (2) the performance of the RNN model with selected features, and (3) a comprehensive discussion of the experimental findings, including a performance comparison with the baseline model.

3.1. Result of Selection Feature

The first stage of the experiment focused on selecting features from the dataset to be used in the training and testing processes. In this study, three feature selection methods were applied: IG, PCA, and RF. The results of feature selection analysis using IG, as shown in Figure 3, indicate that only a small subset of features makes a significant contribution to classification. In the Top 20 Features ranking, feature D14 achieved the highest score (0.209), followed by DATA_012 (0.145) and DATA_611 (0.141). The overall distribution of IG scores reveals that the majority of features cluster around the mean value (0.070), with only a few features exceeding the threshold of 0.15. This finding suggests a long-tail distribution pattern, where only a limited number of features are highly relevant, while most provide limited information.

Further analysis of the Cumulative Information Gain demonstrates that the first 30 features account for more than 80% of the total information, whereas approximately 45 features are required to reach 90%. The importance-level classification of features also shows that 70.1% fall into the low category, 19.4% into the medium category, and only 10.4% into the high category. These results highlight the crucial role of feature selection in reducing data dimensionality and mitigating noise, thereby enabling the RNN model to operate more efficiently without compromising accuracy in detecting intrusions within IoV networks.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

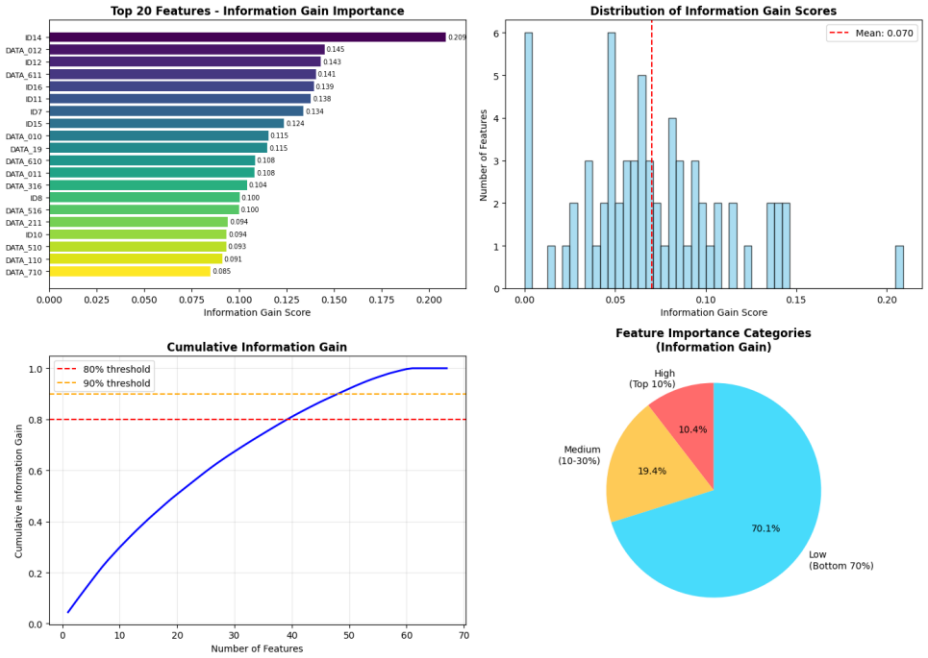

Figure 3. Results of the IG analysis, including: (a) the Top 20 features ranked by IG score, (b) the distribution of IG scores across all features, (c) the cumulative IG curve with thresholds at 80% and 90%, and (d) the categorization of feature importance levels (High, Medium, Low).

The PCA analysis results, as illustrated in Figure 4, show that the first principal component (PC1) accounts for 31.0% of the variance, while the second component (PC2) explains 8.6%, making their combined contribution nearly 40% of the total information. The cumulative variance curve indicates that the first 10 principal components explain more than 73.8% of the variance, while 20 components account for up to 89.3%. This demonstrates that the feature dimensionality can be significantly reduced without losing most of the critical information.

Furthermore, the feature contribution heatmap reveals that attributes such as DATA_511, DATA_214, DATA_010, and D15 dominate across PC1 to PC3. The 2D PCA visualization further confirms that dimensionality reduction aids in separating benign and attack traffic patterns, although some overlap remains among certain data points. These findings highlight that PCA is not only effective for dimensionality reduction but also capable of preserving relevant data structures for intrusion detection. Consequently, PCA can accelerate RNN training while maintaining detection accuracy

The results of the Random Forest feature importance analysis, shown in Figure 5, indicate that only a small number of features have dominant weights. Feature F1 achieved the highest importance score (0.225), followed by F3 (0.165) and F5 (0.035). The score distribution reveals a highly skewed pattern, with most features having values below the average (0.0065) and a median close to zero. This confirms that the majority of features contribute little to the classification process, while only a few are truly relevant.

The cumulative curve demonstrates that the top 30 features account for more than 90% of the total importance, while approximately 50 features are required to reach the 95% threshold. Feature importance categorization further shows that 108 features (70%) fall into the low-importance category, 31 features (20%) into the medium category, and only 16 features (10%) into the high category. These findings highlight the critical role of selecting an optimal subset of features in reducing complexity and improving efficiency, while maintaining the RNN's accuracy in classifying IoV traffic

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

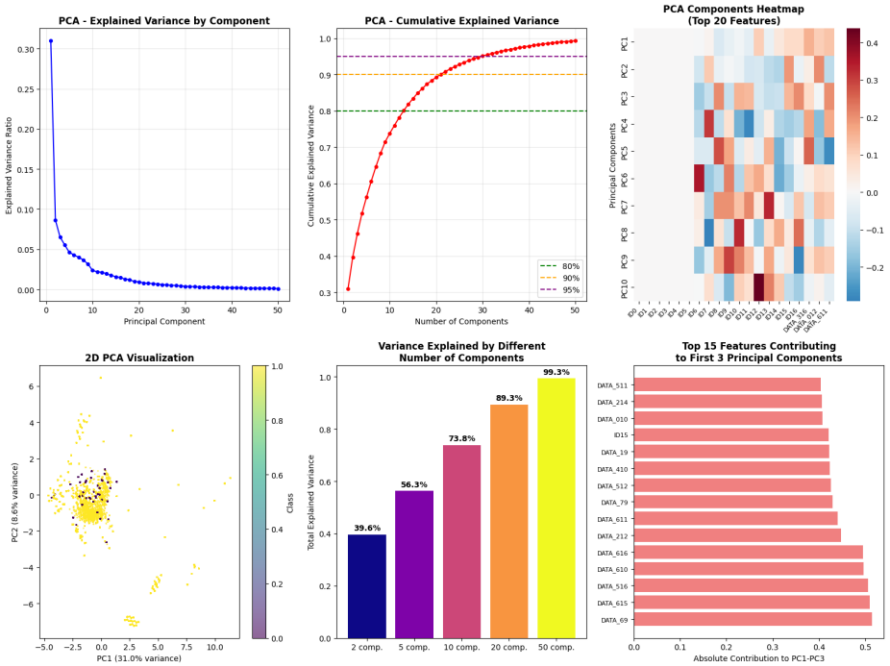

Figure 4. Results of the PCA, including: (a) the variance explained by each component, (b) the cumulative variance with thresholds at 80%, 90%, and 95%, (c) the feature contributions to the first 10 principal components (heatmap), (d) the 2D PCA visualization for PC1 and PC2, (e) the variance explained by different numbers of components, and (f) the Top 15 features with the highest contributions to the first three principal components.

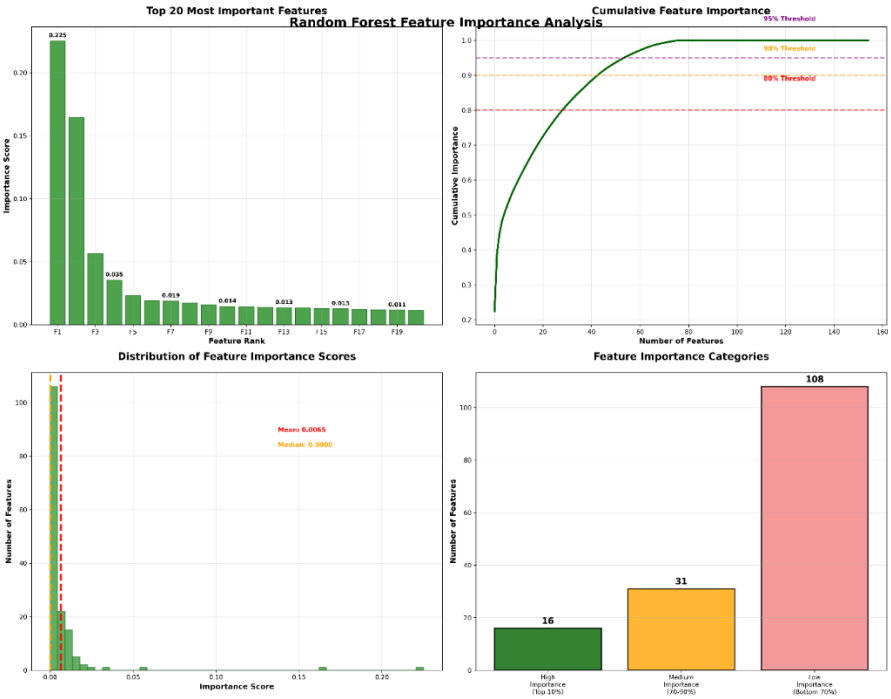

Figure 5. Results of the Random Forest feature importance analysis, including: (a) the Top 20 most important features, (b) the cumulative feature importance curve with thresholds at 80%, 90%, and 95%, (c) the distribution of feature importance scores, and (d) the categorization of feature importance levels (High, Medium, Low).

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

3.2. Result of RNN

The experimental results in Figure 6 demonstrate that the RNN training process with three feature selection methods (IG, PCA, and RF) remained stable. The Training & Validation Loss curves indicate that the loss values for all three methods rapidly decreased and converged within the initial epochs, suggesting that the optimization process was effective and showed no signs of overfitting. This observation is further supported by the Training & Validation Accuracy curves, which reveal that both training and validation accuracy consistently remained above 0.99 from the early epochs, confirming the model's strong generalization capability on validation data.

The Performance Metrics Comparison highlights that all feature selection methods achieved exceptionally high results, with Accuracy, Precision, Recall, and F1-Score reaching 1.00. These findings indicate that IG, PCA, and RF are equally effective in producing relevant feature subsets for IoV attack detection. The marginal differences across methods suggest that the RNN model exhibits strong robustness against variations in feature selection.

Further analysis using the confusion matrix validates these outcomes. For all three methods (IG, PCA, and RF), the model correctly classified all test data without any misclassification (zero error). This is evident from the diagonal dominance in the matrix, where all instances of both Normal and Attack classes were accurately predicted, and no entries appeared in the off-diagonal cells. Such results confirm that the integration of feature selection and the RNN architecture successfully captures both benign and malicious traffic patterns with high precision.

Overall, these findings demonstrate that the RNN-based approach, supported by feature selection, delivers optimal performance in intrusion detection for IoV networks. The perfect accuracy achieved across all key metrics underscores the superiority of this method compared to many previous studies, which still faced challenges with false positives and false negatives. This study confirms that combining feature selection techniques (IG, PCA, and RF) with an RNN model offers both an effective and efficient solution for attack detection in IoV environments.

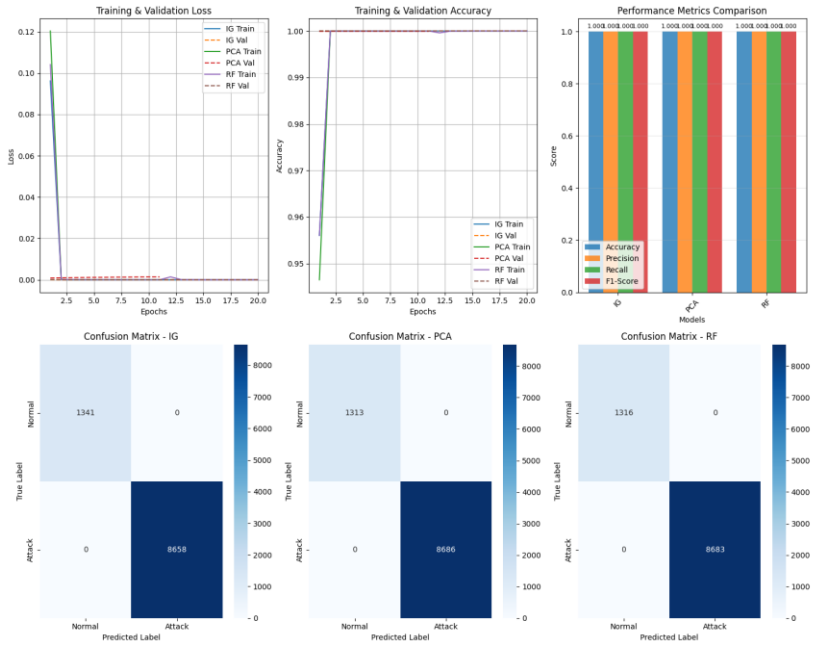

Figure 6. Evaluation results of the RNN model with three feature selection methods (IG, PCA, and RF), including: (a) training and validation loss curves, (b) training and validation accuracy, (c) performance metrics comparison (Accuracy, Precision, Recall, F1-Score), and (d) confusion matrix for each feature selection method.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

The testing results of the RNN model with the three feature selection methods (IG, PCA, and RF) are presented in Figure 7. The confusion matrix shows that the model successfully classified all test data without error, for both Normal and Attack classes. This outcome is consistent with the training results, which demonstrated perfect accuracy. However, a deeper analysis of the prediction probability distributions and ROC curves reveals notable differences across the methods. The probability distributions indicate that prediction values for both Normal and Attack classes were closely clustered within the range of 0.465–0.485, preventing the model from establishing a clear probability margin between the two classes. As a result, the ROC curve exhibits low AUC values: 0.572 for IG, 0.429 for PCA, and 0.415 for RF. AUC scores approaching 0.5 suggest that, despite high apparent accuracy in the confusion matrix, the model lacks strong discriminative capability at the probabilistic level.

These findings indicate a potential bias in the model, where the RNN appears highly capable of memorizing patterns in the test dataset but fails to generate well-calibrated probability distributions. Consequently, although metrics such as Accuracy and F1-score appear perfect, the ROC-AUC analysis exposes a fundamental weakness in the model's generalization ability. This emphasizes the importance of employing multiple evaluation metrics, ensuring that model performance is assessed not only in terms of accuracy but also in terms of its reliability in producing probabilistic predictions

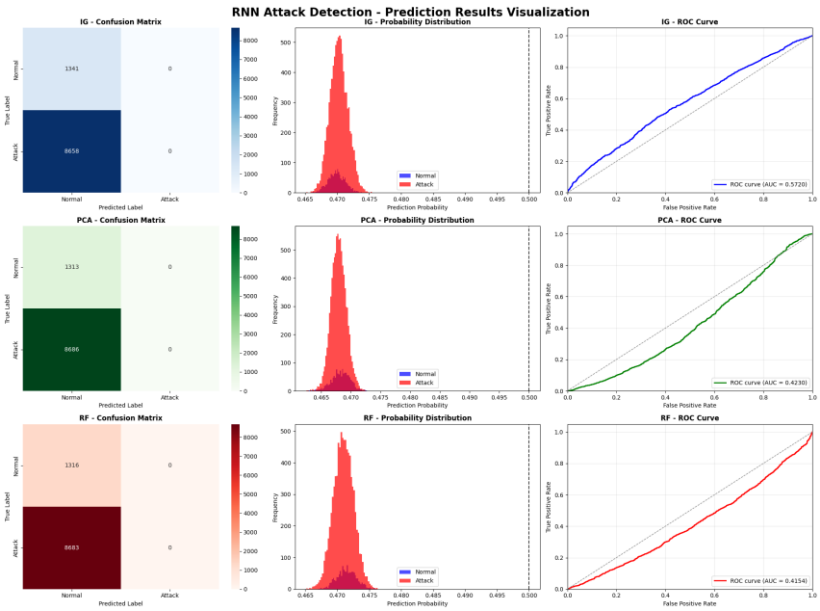

Figure 7. Testing results of the RNN model with three feature selection methods (IG, PCA, and RF), including: (a) confusion matrix, (b) prediction probability distribution, and (c) ROC curves with corresponding AUC values. The results indicate perfect accuracy in the confusion matrix but relatively low AUC values, highlighting the model's limited discriminative capability at the probabilistic level

4. DISCUSSIONS

The results reveal a notable discrepancy between the model's performance during training and testing. In the training phase, the RNN supported by feature selection methods (IG, PCA, RF) achieved near-perfect accuracy, precision, recall, and F1-score. This outcome is evident in the confusion matrix, which shows no misclassifications, indicating that the model was highly effective in identifying both benign and attack traffic within the training data.

However, the testing results table 1 and Figure 8 tell a different story. Although the confusion matrix continued to show correct classifications, the analysis of probability distributions and ROC curves exposed a fundamental weakness. The relatively low AUC values (0.572 for IG, 0.429 for PCA,

Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

and 0.415 for RF) suggest that the model lacked strong discriminative capability when distinguishing classes at the probabilistic level. In other words, while the model appeared "perfect" in terms of classification accuracy, it failed to produce stable and reliable probabilistic predictions.

The iterative performance trends in Figure 8 further reinforce these findings. The graphs illustrate that while accuracy remained relatively stable, precision, recall, and F1-score plateaued at very low values. Similarly, the AUC trends stagnated, particularly for PCA and RF, which reached only around 0.42. This phenomenon indicates the presence of overfitting or potential data leakage, where the model "memorized" patterns from the training data but struggled to generalize probabilistic predictions when evaluated across different iterations.

Overall, these findings underscore the importance of employing multiple evaluation metrics beyond accuracy when assessing the performance of machine learning based IDS. The results also highlight the need for additional strategies, such as stronger regularization, probabilistic model calibration, or ensemble-based methods, to improve generalization capability. Thus, while the RNN with feature selection demonstrated strong performance during training, further refinements are necessary to ensure reliability in detecting intrusions within dynamic IoV environments.

Table 1. RNN performance with feature selection

Method	Accuracy	Precision	Recall	F1-Score	ROC–AUC
IG	1.000	1.000	1.000	1.000	0.572
PCA	1.000	1.000	1.000	1.000	0.429
RF	1.000	1.000	1.000	1.000	0.415

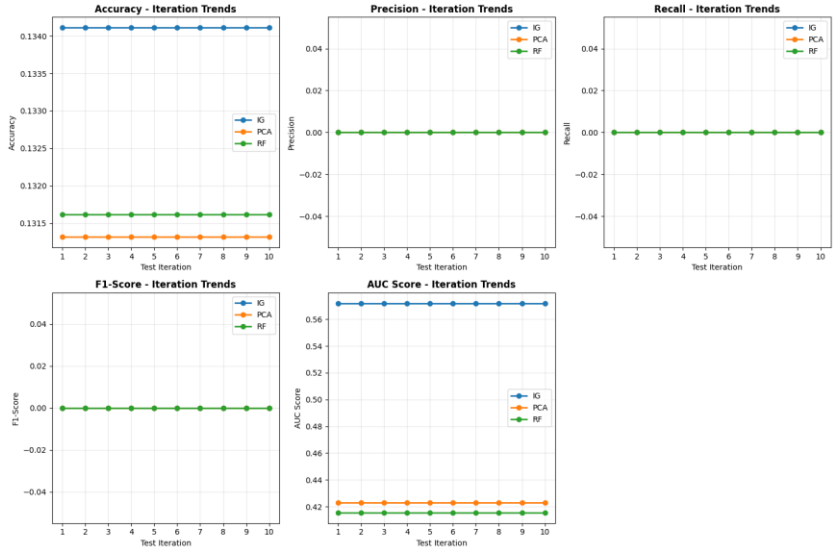

Figure 8. Trends of evaluation metrics (Accuracy, Precision, Recall, F1-Score, and AUC) over 10 testing iterations for the three feature selection methods (IG, PCA, RF). The results show stable accuracy, while probabilistic metrics such as Precision, Recall, F1-Score, and AUC remain stagnant at low values, indicating limitations in the model's generalization capability.

Future work should incorporate regularization (dropout scheduling, weight decay), probability calibration (temperature scaling, Platt/Isotonic), margin-aware losses (AUC-maximization, focal loss), and hybrid architectures (e.g., CNN-RNN/LSTM) to strengthen sequence modeling while preserving margins. Real-time feasibility is supported by PCA/RF speedups; nonetheless, evaluation on time-

Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

varying, real-world IoV traces and adversarial robustness testing is required before on-vehicle deployment.

5. CONCLUSION

This study aimed to develop an RNN-based intrusion detection system capable of accurately identifying cyberattacks and analyzing its performance against various attack types. Based on the experimental results with three feature selection approaches IG, PCA, and RF it can be concluded that the proposed RNN model achieved perfect performance. The models deliver perfect classification (accuracy/precision/recall/F1 = 1.00) but low ROC–AUC (0.415–0.572), revealing limited probabilistic discrimination. Furthermore, predictive testing confirmed consistent performance, with 100% accuracy across all evaluated samples for the IG, PCA, and RF based models. Although the three methods produced identical results in terms of accuracy, differences were observed in inference efficiency. PCA and RF exhibited relatively faster computational performance compared to IG, highlighting their suitability when speed is prioritized in real-time applications. These findings confirm that all three approaches are viable, but PCA and RF may be more advantageous in scenarios requiring rapid detection. Overall, this research demonstrates that RNN is a highly effective architecture for intrusion detection in IoV networks, achieving both accuracy and consistency. The developed model provides a strong foundation for enhancing IoV security against increasingly complex cyber threats. Future work will pursue hybrid temporal models, probability calibration, and adversarial training, alongside real-world IoV evaluations, to improve safety and generalization.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

ACKNOWLEDGEMENT

The authors extends heartfelt gratitude to Universitas Dinamika Bangsa for their outstanding support, which made this research possible. This study was also made possible by the financial assistance from the Direktorat Riset, Teknologi, dan Pengabdian kepada Masyarakat, Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, 2025, whose sponsorship and full support were invaluable.

REFERENCES

[1] K. Huang, R. Xian, M. Xian, H. Wang, and L. Ni, "A comprehensive intrusion detection method for the internet of vehicles based on federated learning architecture," *Comput. Secur.*, vol. 147, p. 104067, Dec. 2024, doi: 10.1016/j.cose.2024.104067.

[2] M. Dilshad *et al.*, "IoV cyber defense: Advancing DDoS attack detection with gini index in tree models," in *2024 international conference on emerging trends in networks and computer communications, ETNCC 2024 - proceedings*, IEEE, 2024, pp. 681–688. doi: 10.1109/ETNCC63262.2024.10767505.

[3] J. Tu and W. Shang, "Enhancing Intrusion Detection in The Internet of Vehicles: An Ensemble and Optimized Machine Learning Approach," in *2023 2nd International Conference on Sensing, Measurement, Communication and Internet of Things Technologies (SMC-IoT)*, Changsha, China: IEEE, Dec. 2023, pp. 207–211. doi: 10.1109/SMC-IoT62253.2023.00044.

[4] M. S. Korium, M. Saber, A. Beattie, A. Narayanan, S. Sahoo, and P. H. J. Nardelli, "Intrusion detection system for cyberattacks in the Internet of Vehicles environment," *Ad Hoc Netw.*, vol. 153, p. 103330, Feb. 2024, doi: 10.1016/j.adhoc.2023.103330.

Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

[5] S. Anbalagan, G. Raja, S. Gurumoorthy, R. D. Suresh, and K. Dev, "IIDS: Intelligent intrusion detection system for sustainable development in autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 12, pp. 15866–15875, 2023, doi: 10.1109/TITS.2023.3271768.

[6] A. Anzer and M. Elhadef, "A multilayer perceptron-based distributed intrusion detection system for internet of vehicles," *Proc. - 4th IEEE Int. Conf. Collab. Internet Comput. CIC 2018*, pp. 438–445, 2018, doi: 10.1109/CIC.2018.00066.

[7] E. Gelenbe, B. C. Gül, and M. Nakıp, "DISFIDA: Distributed Self-Supervised Federated Intrusion Detection Algorithm with online learning for health Internet of Things and Internet of Vehicles," *Internet Things*, vol. 28, p. 101340, Dec. 2024, doi: 10.1016/j.iot.2024.101340.

[8] H. Taslimasa, S. Dadkhah, E. C. P. Neto, P. Xiong, S. Ray, and A. A. Ghorbani, "Security issues in Internet of Vehicles (IoV): A comprehensive survey," *Internet Things*, vol. 22, p. 100809, July 2023, doi: 10.1016/j.iot.2023.100809.

[9] L. Xing, K. Wang, H. Wu, H. Ma, and X. Zhang, "Intrusion Detection Method for Internet of Vehicles Based on Parallel Analysis of Spatio-Temporal Features," *Sensors*, vol. 23, no. 9, p. 4399, Apr. 2023, doi: 10.3390/s23094399.

[10] H. Zhang, J. Ye, W. Huang, X. Liu, and J. Gu, "Survey of federated learning in intrusion detection," *J. Parallel Distrib. Comput.*, vol. 195, p. 104976, Jan. 2025, doi: 10.1016/j.jpdc.2024.104976.

[11] K. Aswal, D. C. Dobhal, and H. Pathak, "Comparative analysis of machine learning algorithms for identification of BOT attack on the Internet of Vehicles (IoV)," in *2020 International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India: IEEE, Feb. 2020, pp. 312–317. doi: 10.1109/ICICT48043.2020.9112422.

[12] L. Yang, A. Shami, G. Stevens, and S. De Rusett, "LCCDE: A Decision-Based Ensemble Framework for Intrusion Detection in The Internet of Vehicles," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, Rio de Janeiro, Brazil: IEEE, Dec. 2022, pp. 3545–3550. doi: 10.1109/GLOBECOM48099.2022.10001280.

[13] Y. Otoum, Y. Wan, and A. Nayak, "Transfer Learning-Driven Intrusion Detection for Internet of Vehicles (IoV)," in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, Dubrovnik, Croatia: IEEE, May 2022, pp. 342–347. doi: 10.1109/IWCMC55113.2022.9825115.

[14] B. Wang, Y. Shang, F. Wang, and Y. Zeng, "A DoS attack detection system based on CNN and BiLSTM for internet of vehicles," in *2024 12th international conference on information systems and computing technology, isctech 2024*, 2024. doi: 10.1109/ISCTech63666.2024.10845700.

[15] X. Yuan, S. Han, W. Huang, H. Ye, X. Kong, and F. Zhang, "A simple framework to enhance the adversarial robustness of deep learning-based intrusion detection system," *Comput. Secur.*, vol. 137, no. July 2023, 2024, doi: 10.1016/j.cose.2023.103644.

[16] X. Huang *et al.*, "A survey of safety and trustworthiness of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability," *Comput. Sci. Rev.*, vol. 37, 2020, doi: 10.1016/j.cosrev.2020.100270.

[17] V. Ravi, R. Chaganti, and M. Alazab, "Recurrent deep learning-based feature fusion ensemble meta-classifier approach for intelligent network intrusion detection system," *Comput. Electr. Eng.*, vol. 102, no. February, p. 108156, 2022, doi: 10.1016/j.compeleceng.2022.108156.

[18] E. C. P. Neto *et al.*, "CICIoV2024: Advancing realistic IDS approaches against DoS and spoofing attack in IoV CAN bus," *Internet Things*, vol. 26, p. 101209, July 2024, doi: 10.1016/j.iot.2024.101209.

[19] M. B. Dissanayake, "Feature engineering for cyber-attack detection in internet of things," *Int. J. Wirel. Microw. Technol.*, vol. 11, no. 6, pp. 46–54, 2021, doi: 10.5815/ijwmt.2021.06.05.

[20] A. A. Aburomman and M. B. I. Reaz, "Ensemble of binary SVM classifiers based on PCA and LDA feature extraction for intrusion detection," *Proc. 2016 IEEE Adv. Inf. Manag. Commun. Electron. Autom. Control Conf. IMCEC 2016*, pp. 636–640, 2017, doi: 10.1109/IMCEC.2016.7867287.

[21] T. A. Alhaj, M. M. Siraj, A. Zainal, H. T. Elshoush, and F. Elhaj, "Feature selection using information gain for improved structural-based alert correlation," *PLoS ONE*, vol. 11, no. 11, pp. 1–18, 2016, doi: 10.1371/journal.pone.0166017.

Vol. 6, No. 5, October 2025, Page. 3204-3216
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5293

[22] T. E. T. Djaidja, B. Brik, S. Mohammed Senouci, A. Boualouache, and Y. Ghamri-Doudane, "Early Network Intrusion Detection Enabled by Attention Mechanisms and RNNs," *IEEE Trans. Inf. Forensics Secur.*, vol. 19, pp. 7783–7793, 2024, doi: 10.1109/TIFS.2024.3441862.

[23] M. A. Khan, "HCRNNIDS: Hybrid convolutional recurrent neural network-based network intrusion detection system," *Processes*, vol. 9, no. 5, 2021, doi: 10.3390/pr9050834.

[24] M. Essaid, D. Y. Kim, S. H. Maeng, S. Park, and H. T. Ju, "A collaborative DDoS mitigation solution based on ethereum smart contract and RNN-LSTM," in *2019 20th asia-pacific network operations and management symposium: Management in a cyber-physical world, APNOMS 2019*, IEICE, 2019, pp. 12–17. doi: 10.23919/APNOMS.2019.8892947.

[25] S. Abdelhamid, I. Hegazy, M. Aref, and M. Roushdy, "TL-IDS: a transfer learning technique for botnet detection in IoT," *6th Int. Conf. Comput. Inform. ICCI 2024*, pp. 315–321, 2024, doi: 10.1109/ICCI61671.2024.10485152.

[26] S. Li, J. Wang, Y. Wang, G. Zhou, and Y. Zhao, "EIFDAA: Evaluation of an IDS with function-discarding adversarial attacks in the IIoT," *Heliyon*, vol. 9, no. 2, p. e13520, Feb. 2023, doi: 10.1016/j.heliyon.2023.e13520.

[27] D. Wang, X. Wang, and J. Fei, "IDS-GAN: Adversarial Attack against Intrusion Detection Based on Generative Adversarial Networks," in *2024 5th International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, Zhuhai, China: IEEE, Apr. 2024, pp. 1130–1134. doi: 10.1109/CVIDL62147.2024.10603582.