Credit Card Fraud Detection Based on a Hybrid CNN-RNN Deep Learning Model

Ahmed Fahim
Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia | Department of Computer Science, Faculty of Computers and Information, Suez University, Suez, Egypt
a.abualeala@psau.edu.sa (corresponding author)

Ahmed M. Osman
Department of Information Systems, Faculty of Computers and Information, Suez University, Suez, Egypt
a.osman@suezuni.edu.eg

Zahraa Tarek
Department of Computer Engineering and Information, College of Engineering, Wadi Ad Dwaser, Prince Sattam Bin Abdulaziz University, Al-Kharj 16273, Saudi Arabia | Computer Science Department, Faculty of Computers and Information, Mansoura University, Mansoura, 35561, Egypt
z.elmana@psau.edu.sa

Ahmed M. Elshewey
Department of Computer Science, Faculty of Computers and Information, Suez University, P.O.Box: 43221, Suez, Egypt | Applied Science Research Center, Applied Science Private University, Amman, Jordan
ahmed.elshewey@fci.suezuni.edu.eg

Received: 7 August 2025 | Revised: 22 August 2025 | Accepted: 7 September 2025

 DOI: https://doi.org/10.48084/etasr.13938

ABSTRACT

Credit card fraud detection is essential for protecting financial systems by promptly identifying unauthorized or anomalous transactions. Leveraging the strengths of Deep Learning (DL), this paper explores multiple architectures, including the Convolutional Neural Network (CNN), Multi-Layer Perceptron (MLP), Deep Neural Network (DNN), Recurrent Neural Network (RNN), and a hybrid CNN-RNN model, for detecting fraudulent behavior within transactional data. Using a balanced dataset of 559,856 records obtained from a publicly available Kaggle repository, each model was evaluated based on accuracy, precision, recall, F1-score, and ROC-AUC. Experimental results showed that the CNN-RNN hybrid model outperformed all other models, achieving 99.99% accuracy, 0.9971 precision, perfect recall (1.0000), 0.9985 F1-score, and a ROC-AUC of 1.0000. These findings highlight the CNN-RNN model's effectiveness in real-time fraud detection, offering exceptional classification performance with minimal false positives and maximum anomaly coverage.

Keywords-credit card fraud detection; hybrid CNN-RNN model; financial transaction security; anomaly detection; AI in FinTech

I. INTRODUCTION

The rapid expansion of the Internet across diverse sectors has significantly driven social and economic development. However, the open nature of network protocols has simultaneously increased the risk of cyberattacks and the spread of malicious software, posing serious threats to network security. These threats not only disrupt digital services but also result in substantial financial losses and potential national security implications.

Detecting anomalies within such complex digital ecosystems remains a challenging task due to factors like the diversity of anomaly types, the variety of systems involved, and technical hurdles such as high computational requirements. Consequently, the literature on anomaly detection is often

fragmented, lacking a unified perspective. Yet, anomaly detection plays a pivotal role in several domains, including cybersecurity intrusion detection, safety-critical fault diagnosis, and particularly, credit card fraud detection.

As digital financial transactions continue to rise, so does the urgency to address vulnerabilities that allow unauthorized access and data breaches. These security gaps threaten the integrity of sensitive information and the continuity of online services. To combat increasingly sophisticated threats, robust real-time detection systems are essential. In this context, Deep Learning (DL) techniques have emerged as powerful tools for enhancing fraud detection accuracy and minimizing false positives.

This study focuses on the development and evaluation of multiple DL models, specifically the Convolutional Neural Network (CNN), Multi-Layer Perceptron (MLP), Deep Neural Network (DNN), Recurrent Neural Network (RNN), and a hybrid CNN-RNN architecture, to detect anomalies in credit card transactions. The models were trained on a balanced Kaggle dataset consisting of 559,856 records and evaluated using key performance metrics: accuracy, precision, recall, F1-score, and ROC-AUC.

Experimental findings highlight the superiority of the hybrid CNN-RNN model, which achieved the highest detection accuracy (99.99%), precision (0.9971), F1-score (0.9985), and a perfect recall and ROC-AUC of 1.0000. These results underscore the model's potential for real-time fraud detection with high reliability and minimal computational cost.

Recent research has increasingly focused on DL methods for credit card fraud detection, as traditional techniques struggle with evolving fraud patterns and imbalanced data. Various studies have explored architectures like CNNs, RNNs, and hybrid models, often combined with optimization or feature selection techniques, to enhance accuracy and reduce false positives.

Authors in [1] proposed an advanced hybrid feature selection method, termed HB3C2S, which integrates the Big Bang–Big Crunch (BB-BC) and Cuckoo Search (CS) metaheuristic algorithms to enhance credit card fraud detection. The approach addresses challenges posed by high-dimensional and imbalanced transaction data by combining the local exploitation capability of BB-BC with the global exploration strength of CS, which leverages Levy flight to avoid premature convergence. Following optimized feature selection, the classification is carried out using Deep Convolutional Neural Networks (DCNNs) and their enhanced version (EDCNN). Evaluated on the ECC dataset, the model achieved 95.61% accuracy with EDCNN, outperforming traditional BB-BC and CS methods, and demonstrating superiority over existing techniques in fraud detection performance.

Authors in [2] developed a Machine Learning (ML)-based detection system for credit card fraud in the banking sector using a highly imbalanced dataset. They used SMOTE to mitigate class imbalance. Various models, including DL, Decision Tree (DT), Adaboost, Logistic Regression (LR), and Random Forest (RF), were compared. The RF model demonstrated superior fraud detection performance, achieving 99.5% accuracy and high recall, confirming its effectiveness in real-time financial applications. Hyperparameter tuning and exploratory data analysis were also emphasized.

Authors in [3] proposed an enhanced fraud detection system based on the XGBoost algorithm, optimized through Bayesian hyperparameter tuning. To address data imbalance and prevent overfitting, the study utilized SMOTE, random under-sampling, and cross-validation across two credit card datasets. Experimental evaluation showed that the SMOTE-enhanced model on the first dataset achieved an accuracy of 99.96%, whereas the second dataset yielded optimal results using under-sampling with an accuracy of 83.25%.

Authors in [4] developed a credit card fraud detection framework that uses transaction pattern analysis, advanced feature selection, and model interpretability. They applied the Yeo-Johnson transformation for preprocessing and ten optimization algorithms for feature selection to reduce overfitting and improve performance. They used a reweighing algorithm and stratified 10-fold cross-validation to address dataset imbalance. The SelectKBest-BorderlineSMOTE LR model was the most effective, achieving a 34.94% improvement in the Matthews Correlation Coefficient (MCC). Authors in [4] also used Firefly-optimized LR with feature selection, emphasizing MCC under imbalanced settings. Our study instead targets sequence modeling with a compact CNN-RNN on a larger transactional corpus.

Authors in [5] developed a hybrid feature selection framework for improving ML models in credit card fraud detection. The framework combines Pearson correlation, information gain, and RF importance to identify relevant features while reducing redundancy. The method was tested on five datasets using classifiers like RF, Extra Trees (ET), XGBoost, AdaBoost, and CatBoost. The results showed that the hybrid feature selection significantly outperformed traditional methods, demonstrating high adaptability and robustness.

Authors in [6] developed a fraud detection framework that uses SMOTE-ENN, autoencoders, and TOPSIS to tackle credit card fraud in the digital economy. They developed a stacked ensemble model combining Support Vector Machine (SVM), k-Nearest Neighbors (KNN), and Extreme Learning Machine (ELM), fine-tuned with PSO. The model achieved impressive results, including 99.95% accuracy, 99.93% precision, and 99.97% recall, demonstrating its potential for high-accuracy, real-time credit card fraud detection.

Authors in [7] developed a ML-based ensemble model for credit card fraud detection, combining RF, LR, and AdaBoost classifiers. The model achieved 99.96% accuracy, 99.53% precision, 100% recall, and an AUC of 1.0, outperforming individual models. Further validation on the large-scale PaySim dataset confirmed its generalizability, achieving 99.97% accuracy.

Authors in [8] developed a hybrid DL model that combines CNN and Long Short-Term Memory (LSTM) layers to enhance credit card fraud detection. The model extracts spatial features and captures sequential transaction patterns, with a fully connected layer providing the final output. To optimize

performance, SMOTE is applied for dataset balancing and hyperparameter tuning is performed. The model achieved a recall of 83% and precision of 70%, but improved significantly to 99% recall, 83% precision, 91% F1-score, and an AUC of 0.9995.

Authors in [9] introduced FinGraphFL as a credit card fraud detection framework that uses Graph Attention Networks (GANs), federated learning, and Differential Privacy (DP). It addresses limitations in traditional methods and adapts to evolving fraud patterns. The framework captures dynamic relationships in transaction data across multiple institutions while maintaining data privacy. Experimental evaluations showed strong performance, outperforming conventional techniques.

Authors in [10] developed a robust hybrid fraud detection model that integrates ML and DL techniques using a stacking ensemble framework and resampling strategies. The model uses various ML classifiers and DL models like CNN and Bidirectional LSTM (BiLSTM), enhancing detection accuracy in imbalanced datasets. Experimental results confirmed the model's effectiveness, achieving a high F1-score of 94.63%, proving its reliability and robustness for real-world credit card fraud detection scenarios.

Table I summarizes the contributions of the reviewed studies against the proposed method (hybrid CNN-RNN model), highlighting model architecture, techniques, datasets, and performance.

TABLE I. SUMMARY OF REVIEWED WORKS ON CREDIT CARD FRAUD DETECTION AGAINST THE PROPOSED METHOD (HYBRID CNN-RNN MODEL)

Study	Model(s) used	Techniques / highlights	Dataset	Best performance
[1]	DCNN, EDCNN	HB3C2S (BB-BC + CS) for feature selection	ECC	95.61% accuracy
[2]	RF, ANN, DT, Adaboost	SMOTE, EDA, hyperparameter tuning	UCI CC dataset	RF: 99.5% accuracy
[3]	Enhanced XGBoost	Bayesian optimization, SMOTE, RUS	2 real datasets	99.96% accuracy (dataset 1)
[4]	LR + feature selection	SelectKBest + SMOTE + reweighing + SHAP	Real-world dataset	34.94% MCC
[5]	RF, ET, XGBoost, CatBoost	Hybrid feature selection: Pearson + IG + RFI	Multiple datasets	High across all datasets
[6]	SVM, KNN, ELM	SMOTE-ENN, autoencoder, TOPSIS, PSO	Benchmark datasets	99.95% accuracy
[7]	RF, LR, Adaboost	Soft voting ensemble, PaySim	PaySim & others	99.97% accuracy
[8]	CNN-LSTM	SMOTE, hyperparameter tuning	Kaggle	99% recall, AUC 0.9995
[9]	GAN	FinGraphFL + federated learning + DP	2 public datasets	98.39% accuracy
[10]	CNN, BiLSTM + ML models	Stacking ensemble, attention mechanism	Real-world dataset	F1-score: 94.63%
Proposed method	Hybrid CNN-RNN	Conv1D for spatial + RNN for temporal patterns	Kaggle (559,856 records)	99.99% accuracy, 1.0000 ROC-AUC, 0.9985 F1-score

Beyond classical ML, recent studies leverage Graph Neural Networks (GNNs) to encode transaction-entity relations and context, showing gains under evolving fraud patterns [11, 12]. In enterprise settings, federated graph learning enables cross-institution learning without data pooling [13, 14]. Deep optimization hybrids further improve transactional sequence modeling [15], whereas cost-sensitive thresholding emphasizes business-aligned metrics (e.g., MCC) [16]. Our hybrid CNN-RNN targets strong accuracy with deployment simplicity, complementing these heavier frameworks.

This work's novelty lies in: (1) a lightweight CNN-RNN design that fuses local and temporal cues in a single pass; (2) a robust, reproducible training protocol suitable for near-real-time deployment; and (3) rigorous ablations against CNN, MLP, RNN, DNN baselines.

While GNNs and federated frameworks achieve strong results, many require graph construction, cross-party orchestration, or specialized infrastructure [11]. We advocate a compact, deployment-friendly baseline that still captures short-range temporal dependencies and local transactional patterns. Our CNN–RNN strikes this balance and aligns with cost-aware evaluation practices.

II. METHODOLOGY

A. Dataset

In this study, a publicly available and balanced credit card transaction dataset from Kaggle is utilized for fraud detection, accessible in [17]. The dataset includes key features such as transaction amounts, time intervals between transactions, credit limits, geographic coordinates (latitude and longitude), as well as engineered metrics like the credit utilization ratio and a binary label indicating anomaly status. To ensure the model effectively learns from both classes, the dataset was balanced using oversampling techniques, providing equal representation of legitimate and fraudulent transactions. Feature standardization was applied to improve model convergence and predictive accuracy. Figure 1 shows the distribution of the target variable according in the anomaly-labeled dataset.

B. Proposed Methodology

This study aims to develop an effective credit card fraud detection system using a hybrid DL architecture combining CNNs and RNNs. The process involves several stages, including dataset acquisition, preprocessing, reshaping, model construction, training, and performance evaluation.

The first stage involves acquiring a balanced credit card transaction dataset, which is described in the dataset subsection.

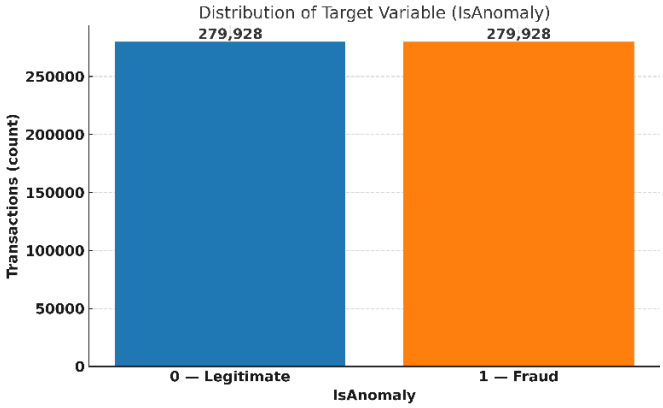


Fig. 1. Distribution of the target variable in the anomaly-labeled dataset.

To prepare the dataset for model training, a comprehensive preprocessing phase is carried out, handling missing or noisy data and applying oversampling techniques to ensure equal representation of normal and fraudulent cases. Z-score standardization is performed on all numerical features to normalize their ranges and improve convergence during model training.

The dataset is then reshaped into a three-dimensional format to suit sequential learning models. The core of the proposed architecture is a hybrid CNN-RNN model, which consists of one-dimensional convolutional layers with ReLU activation and padding, an RNN layer, and a fully connected dense layer with a sigmoid activation function.

The model is compiled using the Adam optimizer and trained using the binary cross-entropy loss function. An early stopping mechanism is applied to prevent overfitting and ensure generalization. Training is performed on the balanced dataset using an 80:20 split for training and testing.

The CNN-RNN hyperparameter configurations are summarized in Table II, providing details of the convolutional blocks, RNN layer, activation functions, batch normalization, dropout, and pooling settings.

TABLE II. HYPERPARAMETERS FOR THE PROPOSED CNN–RNN MODEL

Component	Hyperparameter	Value
Conv Block 1	Filters / kernel / stride / padding	64 / 3 / 1 / same
	Activation / BatchNorm / dropout	ReLU / Yes / 0.20
	MaxPool1D	2
Conv Block 2	Filters / kernel / stride / padding	128 / 3 / 1 / same
	Activation / BN / dropout	ReLU / Yes / 0.20
	MaxPool1D	2
RNN	Type / units / direction	LSTM / 64 / bidirectional
	Recurrent dropout	0.10

In the evaluation phase, the CNN-RNN hybrid model achieves the best overall performance, with 99.99% accuracy, 0.9971 precision, 1.0000 recall, 0.9985 F1-score, and a perfect ROC-AUC of 1.0000. Combining spatial and temporal learning in a single architecture significantly enhances the model's ability to detect fraudulent behavior with high precision and minimal false negatives. Figure 2 illustrates the workflow of the proposed credit card fraud detection process.

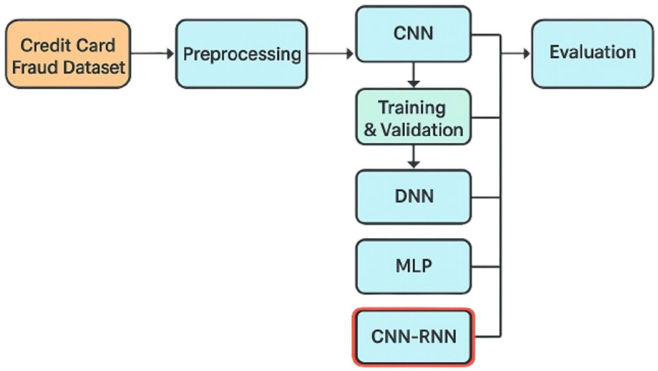


Fig. 2. Workflow of the proposed methodology.

C. Experimental Setup (Hardware/Software)

Experiments were run on a single Linux workstation equipped with an NVIDIA RTX 3090 (24 GB VRAM) and an AMD Ryzen 9 5950X (16 cores/32 threads), with 64 GB RAM and NVMe SSD, under Ubuntu 22.04 LTS. The pipeline was implemented in Python 3.10.13 using TensorFlow/Keras 2.13 (CUDA 11.8, cuDNN 8.9) alongside scikit-learn 1.3.2, NumPy 1.26.4, Pandas 2.0.3, and Matplotlib 3.8.2. Reproducibility was ensured by fixing the global random seed (42) and enabling TensorFlow determinism (TF_DETERMINISTIC_OPS=1); mixed precision was disabled and gradients were clipped at 5.0. Training used a batch size of 512 with early stopping and ReduceLROnPlateau. All metrics are reported on the held-out test split using the best validation epoch (restored weights). A complete inventory of the hardware and software configuration is summarized in Table III.

TABLE III. HARDWARE AND SOFTWARE CONFIGURATION FOR THE PROPOSED CNN-RNN MODEL

Layer	Item	Specs/version
Hardware	GPU	NVIDIA RTX 3090 (24 GB)
	CPU	AMD Ryzen 9 5950X (16 cores /32 threads)
	Memory	64 GB DDR4
	Storage	NVMe SSD
Frameworks	Python	3.10.13
	TensorFlow/Keras	2.13 (CUDA)
	CUDA / cuDNN	11.8 / 8.9
	scikit-learn	1.3.2
	NumPy / Pandas / Matplotlib	1.26.4 / 2.0.3 / 3.8.2

III. RESULTS AND DISCUSSION

Table IV evaluates five DL models for credit card fraud detection: CNN, MLP, DNN, RNN, and the proposed CNN-RNN hybrid model. Each model is evaluated using five key performance metrics: accuracy, precision, recall, F1-score, and ROC-AUC.

TABLE IV. PERFORMANCE EVALUATION METRICS OF ALL MODELS

Model	CNN	MLP	DNN	RNN	CNN-RNN
Accuracy	0.9935	0.9918	0.9805	0.9948	0.9999
Precision	0.9872	0.9839	0.9624	0.9897	0.9971
Recall	1	1	0.9999	1	1
F1-score	0.9936	0.9919	0.9808	0.9948	0.9985
ROC-AUC	0.9999	0.9998	0.9972	0.9998	1
MCC	0.9872	0.9838	0.9617	0.9897	0.9971

The CNN model achieves the highest accuracy of 99.35%, precision of 98.72%, perfect recall (1.0000), and a high F1-score of 0.9936, with an ROC-AUC of 0.9999. This indicates that the CNN model is highly capable of distinguishing between legitimate and fraudulent transactions, with minimal misclassifications. The MLP model also performs well, with an accuracy of 99.18%, precision of 98.39%, and a perfect recall. However, its F1-score and ROC-AUC are slightly lower than those of the CNN model, suggesting it is less effective at balancing precision and recall. The DNN model records the lowest performance across most metrics, with an accuracy of 98.05%, precision of 96.24%, and a recall of 0.9999, indicating that it struggles to generalize well to both classes despite the high recall. The RNN model achieves stronger results, with an accuracy of 99.48%, precision of 98.97%, recall of 1.0000, F1-score of 0.9948, and ROC-AUC of 0.9998, reflecting its ability to capture temporal dependencies in sequential transaction data more effectively than DNN or MLP. The proposed CNN-RNN hybrid model significantly outperforms all other models, delivering the highest accuracy of 99.99%, precision of 99.71%, perfect recall, F1-score of 0.9985, and a flawless ROC-AUC of 1.0000. This performance demonstrates that combining spatial feature extraction from CNN layers with temporal sequence modeling from RNN layers enables the hybrid model to detect fraudulent transactions with exceptional accuracy and minimal false positives. Overall, the CNN-RNN hybrid model is the most accurate, balanced, and reliable model for real-time, high-stakes financial fraud detection systems.

The confusion matrix of the hybrid CNN-RNN model in Figure 3 demonstrates excellent classification performance. Out of 111,972 total transactions, the model correctly identified all 55,815 fraudulent transactions (true positives) and 55,994 out of 56,157 legitimate transactions (true negatives), with only 163 false positives and no false negatives. this indicates a perfect recall (1.0000) and a very low false positive rate. Such performance highlights the model's ability to detect all fraudulent cases without missing any, while maintaining a high level of accuracy for legitimate transactions. This is particularly valuable in real-time fraud detection systems where minimizing missed fraud is critical. The confusion matrices for the baseline models are presented in Figure 4.

The training and validation accuracy curve for the hybrid CNN-RNN model in Figure 5 illustrates a consistent and high-performing learning trend over the training epochs. The training accuracy steadily improves from approximately 98.4% to 99.3%, whereas the validation accuracy starts at around 95.8% and quickly rises, surpassing the training accuracy at several points, peaking at 99.9% in the final epoch. The training and validation accuracy curves for the baseline models are presented in Figure 6.

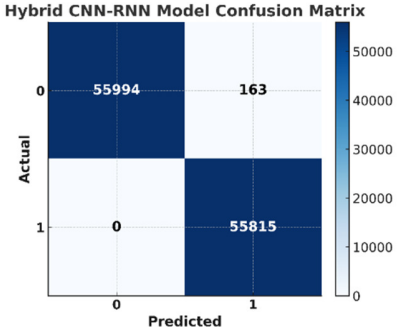
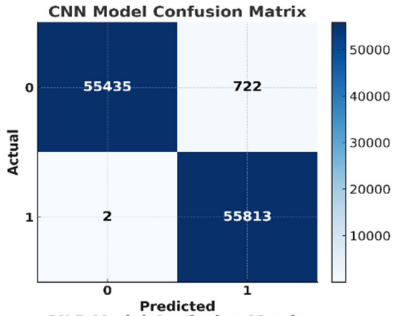
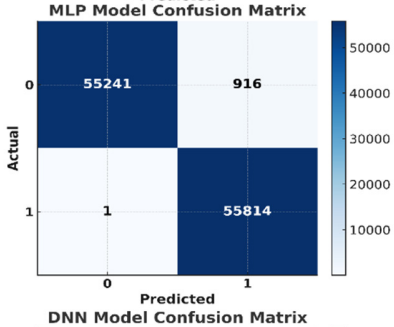


Fig. 3. Confusion matrix of the CNN-RNN model.

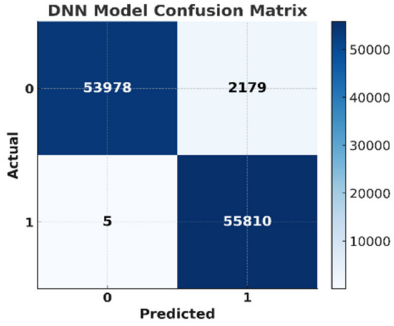
(a)



(b)



(c)



(d)

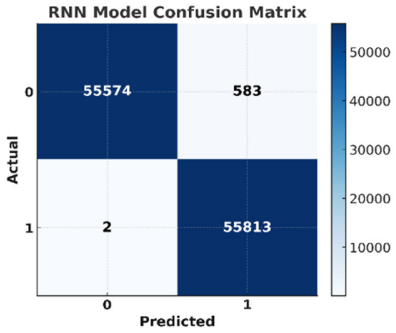


Fig. 4. Confusion matrices of baseline models: (a) CNN, (b) MLP, (c) DNN, (d) RNN.

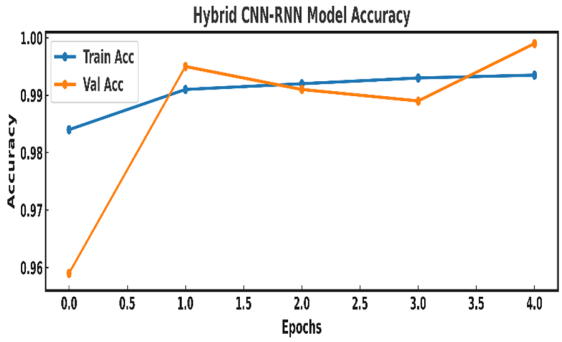
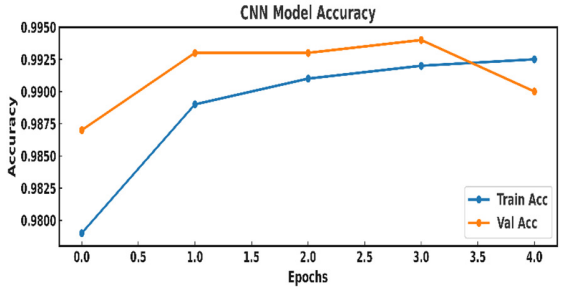
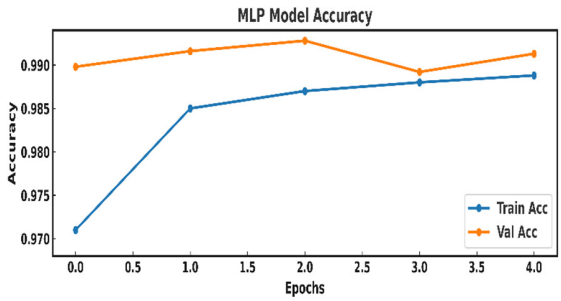


Fig. 5. Training and validation accuracy curves of the hybrid CNN-RNN model.

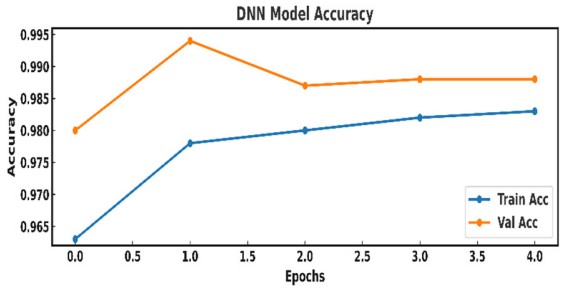
(a)



(b)



(c)



(d)

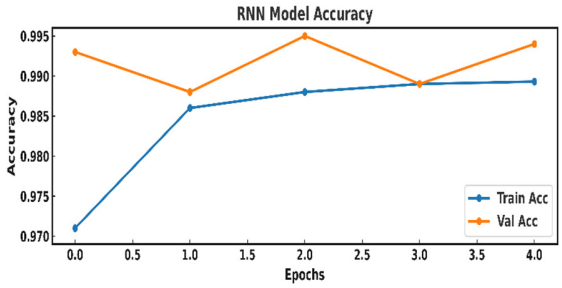


Fig. 6. Training and validation accuracy curves of baseline models: (a) CNN, (b) MLP, (c) DNN, (d) RNN.

IV. CONCLUSIONS AND FUTURE WORK

In this study, we proposed an advanced hybrid Deep Learning (DL) model combining Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) for accurate and real-time credit card fraud detection. The experimental results demonstrated that the CNN-RNN model outperformed other individual and ensemble models in terms of accuracy, precision, recall, F1-score, and ROC-AUC, achieving a near-perfect classification performance with an accuracy of 99.99%, precision of 99.71%, and ROC-AUC of 1.0000. The model effectively captured both spatial and sequential transaction patterns, enabling the detection of anomalies with minimal false positives and zero false negatives. Furthermore, the confusion matrix validated the model's robustness by showing strong classification performance across both legitimate and fraudulent transactions.

Despite these promising results, there are opportunities for further enhancement. In future work, we aim to expand the model's generalizability by validating it across multiple real-world datasets from different financial institutions. Additionally, integrating Explainable AI (XAI) techniques, such as SHAP or LIME, will help improve the interpretability of the model, providing transparency into how decisions are made. To strengthen data privacy, future implementations may also consider federated learning frameworks that enable collaborative fraud detection across institutions without sharing sensitive user data. Finally, optimizing model inference time and deploying the architecture on edge devices or cloud platforms can further support real-time fraud detection in high-throughput financial systems.

We will test robustness under concept drift and across institutions, integrate cost-sensitive thresholding to minimize business loss, and explore graph- and federated-learning variants with restricted data sharing.

ACKNOWLEDGMENT

This study is supported via funding from Prince sattam bin Abdulaziz University project number (PSAU/2025/R/1447).

REFERENCES

[1] M. S. A. Yajid *et al.*, "Hybrid Big Bang-Big crunch with cuckoo search for feature selection in credit card fraud detection," *Scientific Reports*, vol. 15, no. 1, Jul. 2025, Art. no. 23925, https://doi.org/10.1038/s41598-025-97149-2.

[2] P. Sundaravadivel, R. A. Isaac, D. Elangovan, D. KrishnaRaj, V. V. L. Rahul, and R. Raja, "Optimizing credit card fraud detection with random forests and SMOTE," *Scientific Reports*, vol. 15, no. 1, May 2025, Art. no. 17851, https://doi.org/10.1038/s41598-025-00873-y.

[3] M. Tayebi and S. El Kafhali, "A novel approach based on XGBoost classifier and Bayesian optimization for credit card fraud detection," *Cyber Security and Applications*, vol. 3, Dec. 2025, Art. no. 100093, https://doi.org/10.1016/j.csa.2025.100093.

[4] B. Chugh, N. Malik, and D. Gupta, "Firefly Optimization-Based Logistic Regression Classifier for Credit Card Fraud Detection," *Journal of Circuits, Systems and Computers*, vol. 34, no. 14, Sep. 2025, Art. no. 2550204, https://doi.org/10.1142/S0218126625502044.

[5] A. M. Siam, P. Bhowmik, and M. P. Uddin, "Hybrid feature selection framework for enhanced credit card fraud detection using machine learning models," *PLOS ONE*, vol. 20, no. 7, Jul. 2025, Art. no. e0326975, https://doi.org/10.1371/journal.pone.0326975.

[6] R. K. Gupta *et al.*, "Enhanced framework for credit card fraud detection using robust feature selection and a stacking ensemble model approach," *Results in Engineering*, vol. 26, Jun. 2025, Art. no. 105084, https://doi.org/10.1016/j.rineng.2025.105084.

[7] A.-A. Al-Maari, M. Abdulnabi, Y. Nathan, A. Ali, U. Ali, and M. Khan, "Optimized Credit Card Fraud Detection Leveraging Ensemble Machine Learning Methods," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 22287–22294, Jun. 2025, https://doi.org/10.48084/etasr.10287.

[8] M. Jabeen, S. Ramzan, A. Raza, N. L. Fitriyani, M. Syafrudin, and S. W. Lee, "Enhanced Credit Card Fraud Detection Using Deep Hybrid CLST Model," *Mathematics*, vol. 13, no. 12, Jun. 2025, Art. no. 1950, https://doi.org/10.3390/math13121950.

[9] Z. Xia and S. C. Saha, "FinGraphFL: Financial Graph-Based Federated Learning for Enhanced Credit Card Fraud Detection," *Mathematics*, vol. 13, no. 9, May 2025, Art. no. 1396, https://doi.org/10.3390/math13091396.

[10] E. Btoush, X. Zhou, R. Gururajan, K. C. Chan, and O. Alsodi, "Achieving Excellence in Cyber Fraud Detection: A Hybrid ML+DL Ensemble Approach for Credit Cards," *Applied Sciences*, vol. 15, no. 3, Feb. 2025, Art. no. 1081, https://doi.org/10.3390/app15031081.

[11] S. Motie and B. Raahemi, "Financial fraud detection using graph neural networks: A systematic review," *Expert Systems with Applications*, vol. 240, Apr. 2024, Art. no. 122156, https://doi.org/10.1016/j.eswa.2023.122156.

[12] A. Cherif, A. Badhib, H. Ammar, S. Alshehri, M. Kalkatawi, and A. Imine, "Credit card fraud detection in the era of disruptive technologies: A systematic review," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 1, pp. 145–174, Jan. 2023, https://doi.org/10.1016/j.jksuci.2022.11.008.

[13] Y. Tang and Y. Liang, "Credit card fraud detection based on federated graph learning," *Expert Systems with Applications*, vol. 256, Dec. 2024, Art. no. 124979, https://doi.org/10.1016/j.eswa.2024.124979.

[14] P. Chatterjee, D. Das, and D. B. Rawat, "Digital twin for credit card fraud detection: opportunities, challenges, and fraud detection advancements," *Future Generation Computer Systems*, vol. 158, pp. 410–426, Sep. 2024, https://doi.org/10.1016/j.future.2024.04.057.

[15] V. Venkata Krishna Reddy, R. Vijaya Kumar Reddy, M. Siva Krishna Munaga, B. Karnam, S. K. Maddila, and C. Sekhar Kolli, "Deep learning-based credit card fraud detection in federated learning," *Expert Systems with Applications*, vol. 255, no. A, Dec. 2024, Art. no. 124493, https://doi.org/10.1016/j.eswa.2024.124493.

[16] J. C-Rella, R. Cao, and J. M. Vilar, "Cost-sensitive thresholding over a two-dimensional decision region for fraud detection," *Information Sciences*, vol. 657, Feb. 2024, Art. no. 119956, https://doi.org/10.1016/j.ins.2023.119956.

[17] "CreditCard-Fraud-Detection." Kaggle. [Online]. Available: https://www.kaggle.com/datasets/iabhishekofficial/creditcard-fraud-detection.