FEDERAL UNIVERSITY OF TECHNOLOGY, AKURE
JOURNAL OF ENGINEERING AND ENGINEERING TECHNOLOGY

INVESTIGATING A TRANSPARENT AND INTERPRETABLE DEEP LEARNING MODEL FOR ENHANCED JAMMING ATTACK DETECTION IN UAV COMMUNICATION NETWORKS

Ayegun, O. M.*, Akingbade, K. F., Popoola, J. J., Ubochi, B. C.

Department of Electrical and Electronics Engineering, School of Electrical Systems Engineering, Federal University of Technology, Akure, Nigeria

**Correspondence: oluwadaremichaela@futa.edu.ng*

Ayegun, O. M., Akingbade, K. F., Popoola, J. J. and Ubochi, B. C. (2025): Investigating a Transparent and Interpretable Deep Learning Model for Enhanced Jamming Attack Detection in UAV Communication Networks. *FUTA Journal of Engineering and Engineering Technology / 19(1), 215-225*

Received Date: 10.5.25 **Accepted Date: 17.6.25**

Abstract
Unmanned Aerial Vehicles (UAVs) have become indispensable in military and civil applications, playing a crucial role in intelligence gathering for tactical military operations and logistics delivery services. The operational effectiveness of UAVs anchors on secure and reliable radio frequency communication links, which can be susceptible to Jamming Attack (JA) threats. Among the current techniques used for JA detection (JAD) in wireless communication networks, Convolution Neural Network (CNN) models have demonstrated improved detection performance across individual and multiple jammer types. However, there is insufficient information on the model's transparency and the relative contributions of the components of the CNN model in JAD. In this work, an ablation study is performed on the developed CNN model in order to examine the contributions or impact of key components on the overall detection accuracy of the model. By systematically removing each of these components and then training the modified CNN model on the same dataset under identical hyperparameter conditions with the baseline model, the JAD performance accuracy was evaluated. The results reveal that the convolutional layers have the biggest impact on the overall performance of the JAD model, degrading the detection accuracy from 97.0% to 65.3% when ablated. The study highlights the importance of key components of the CNN model, for achieving effective JAD in UAVs and similar sensitive wireless communication networks, thus ensuring secure operations.

Keywords: *Unmanned aerial vehicle, wireless communication network, radio frequency, jamming attack, jamming attack detection, convolutional neural network, ablation study*

Introduction

Unmanned Aerial Vehicles (UAVs), popularly known as drones (Laghari *et al.*, 2023), are invaluable assets for both military and civilian operations due to their flexibility, low operational costs, and wide range of capabilities (Rydberg *et al.*, 2007; Alvear *et al.*, 2017; Slimane *et al.*, 2022). In the military realm, UAVs are frequently deployed for reconnaissance, surveillance, and target acquisition missions (Abiodun, 2020). On the civilian side, UAVs play a vital role in infrastructural inspections, precision agriculture, emergence response, and logistics delivery services (Ukwueze *et al.*, 2024). The UAV market was estimated at 19.7 billion dollars in 2020 and is projected to reach about 102.4 billion dollars by 2030, representing annual compound growth of 19.6% (Hafeez *et al.*, 2023). This surge is largely driven by growing demand for automation and rapid development in supporting technologies. As the UAV technologies continues to evolve, having strong and reliable wireless communication links becomes even more crucial (Paredes *et al.*, 2016). The UAVs communicate with ground control stations or remote controllers using wireless connections over 2.4 GHz or 5.8 GHz Radio Frequency (RF) channels (Kaleem *et al.*, 2019; Yaacoub *et al.*, 2020). These links are essential for the UAV to receive navigation commands, send sensory data, and keep track of its surroundings in real time.

However, despite all the technological progress, UAV systems still face significant vulnerabilities due to their reliance on RF communication, making them prime targets for intentional signal attacks (Naqvi *et al.*, 2018; Fei *et al.*, 2019; Laghari *et al.*, 2023). One of the most concerning threats is Jamming Attacks (JAs), where RF signals are deliberately emitted to disrupt or completely block the UAV communication links, putting the mission and safety of the UAV and nearby assets art risk (Slimane *et al.*, 2022; Laghari *et al.*, 2023). Jamming threats can lead to poor performance, forced landings, or even total loss of the UAV. Hence, having effective Jamming Attack Detection (JAD) system is very important, for ensuring the safe and secure operation of the UAVs, especially in contested environments (Manesh and Kaabouch, 2019).

To address this issue, numerous studies have been published that suggest various strategies for detecting JAs in UAV wireless communication. For instance, Arthur *et al.* (2019) introduced signal irregularity-based intrusion identification strategy, specifically designed to identify JAs and GPS spoofing attack within UAV communication networks. The approach is made up of two algorithms, namely: self-taught learning algorithm and a multiclass support vector machine algorithm. The strategy achieved an impressive 90% detection accuracy. The study by Sedjelmaci *et al.* (2017), proposed a combination of intrusion detection and response schemes to protect UAV networks from jamming signals. The developed schemes operate by checking the UAV communication signals for markers of abnormal or malicious behaviour and uses this information to identify JA signals. The strategy attained 93% detection accuracy. Also, Greco *et al.* (2021) proposed multi-layer perceptron and decision trees approach for JAD in UAVs communication networks. The hybrid machine learning model strategy attained an impressive 96% JAD accuracy. Duan *et al.* (2018), suggested a joint UAV's received signal strength and packet error rate strategy for jamming signal detection. The strategy successfully localized the detected network jammers and mitigated them using a robust path planning approach.

Mowla *et al.* (2019), proposed a federated learning-based JAD in flying ad-hoc network formed by UAVs. The approach leverages on the Dempster-Shafer theory to outperform traditional JAD solutions in terms of overall detection accuracy. To address the challenge of UAV jamming electromagnetic interference in battlefield, Gao *et al.* (2020) proposed semantic analysis that is based on a fuzzy logic reasoning approach. Jamming signals are detected while monitoring the abnormal behaviour of UAVs via a tracing comparison method. Although most of these existing JAD strategies exhibited impressive detection accuracy with simple JAs, they are not efficient under smart or sophisticated JAs acting individually or together. Additionally, most of the studies focused on detection of a certain JA type and performed poorly when implemented under unfamiliar jammer types. Furthermore, most existing studies on JAD, only focused on general wireless network without considering the peculiarity of UAVs communication networks (Gecgel *et al.*, 2019; Arjoune *et al.*, 2020). Other limitations such as high detection latency, poor management of constrained resources and high computational cost made some of the existing solutions not reliable for UAV applications. The recent development in the field of Artificial Intelligence (AI) and computer vision's technology, have led to the deployment of Machine Learning (ML) algorithm for JAD in wireless communication networks (Slimane *et al.*, 2022).

Currently, the Convolutional Neural Networks (CNNs) algorithm, a Deep Learning (DL) model, has emerged as a powerful tool for RF signal intrusion detection and classification in wireless communication systems (Alrefaei, 2024). CNNs are especially advantageous in JAD because of their unique capability to automatically learn signal features representation in spectrogram datasets and other time-frequency signal features. CNN-based models have outperformed Traditional ML models that rely on human-assisted features, demonstrating high JAD accuracy across diverse wireless networks' jamming scenarios. While CNN models are known for their impressive performance in practice, they often face criticism for being "black box" models. This refers to their lack of transparency on how they make predictions and which components of their architecture are most influential on their performance (Assaduzzaman *et al.*, 2024). This opacity can be problematic in critical areas like UAV operations, where it's important to understand the decision-making process to build trust, validate results and enhance the system capability. Moreover, without a thorough evaluation of how each component of the CNN model contributes, there is a risk of overfitting, redundancy, or inferior performance stemming from poor design choices.

To tackle this challenge, this paper introduces a well-structured ablation study aimed at breaking down the performance contributions of different CNN architectural components, specifically in the realm of JAD within UAV communication networks. An ablation study is all about selectively removing or tweaking components of the CNN model to see how they individually affect overall performance of the model (Abate *et al.*, 2023; Assaduzzaman *et al.*, 2024). By carrying out this analysis in a simulated UAV communication environment created in MATLAB, this research offers empirical evidence and valuable insights into how architectural choices and hyperparameters like the number of convolutional layers, pooling strategies, activation functions, optimization algorithms impact the overall effectiveness of the CNN model. Considering that the purpose of this paper is to conduct an ablation study of the CNN structure to investigate the impacts of its key components on the overall performance of the model, the main contributions are:

a. Design and simulation of a UAV communication network under JAs conditions using a MATLAB-based environment to generate labeled spectrogram data for model training and evaluation.

b. Development of a CNN-based JAD model, followed by systematic ablation of critical components to evaluate their individual impacts on performance.

c. Comprehensive analysis of key components of the CNN-based JAD model, to identify optimal combinations for enhanced jamming signal detection accuracy.

This research work advances transparency and interpretability of the CNN models for enhanced JAD in UAV applications. By identifying the most impactful design components, it provides a foundation for deploying more effective and trustworthy DL models in real-world UAV systems operating in adversarial environments. The remaining sections of this research paper is structured as follows: section 2 focuses on related studies, while in section 3 the proposed methodology is presented offering insights into the study approach. Section 4 presents the results and discussions. Conclusions are made by summarizing the key findings and implications in section 5. Section 6 proposes future research works.

The increasing integration of UAVs into critical infrastructure and defense systems has accelerated research into safeguarding their wireless communication channels against malicious threats, particularly JA. Several studies have been conducted on different types of JAD including constant, random, and smart jammers, with smart jammers being the most sophisticated due to their ability to dynamically adapt their interference based on real-time monitoring of the UAV communication patterns (Alvear *et al.*, 2017). Traditional detection methods, such as energy detection, matched filtering, or statistical signal processing, often struggle in the presence of advanced or low-power jammers and lack robustness in dynamic environments (Rydberg *et al.*, 2007). The adoption of DL methods, particularly CNNs, has revolutionized wireless JAD by enabling automatic feature extraction from high-dimensional RF data. CNNs are especially suited for processing spectrograms visual representations of the signal's frequency content over time which provide rich spatial information that can be used to distinguish between normal and jammed signal patterns (Slimane *et al.*, 2022).

Despite the impressive JAD accuracy achieved in previous studies by CNN-based techniques, there's a clear gap when it comes to understanding how the model makes its decisions or predictions. This highlights the need for evaluations that focus more on interpretability, particularly through ablation studies. The ablation studies are commonly used in ML and DL to assess how individual components affects the overall performance of a model. By carefully removing or altering specific parts of a CNN model, like certain layers, activation functions, or regularization methods, researchers can gain valuable insights into which elements are essential and which ones might be unnecessary or even harmful to the anomaly detection process. is clearly a noticeable lack of transparency in the model's prediction or decision-making process (Abate *et al.*, 2023; Assaduzzaman *et al.*, 2024).

Ablation study has been used to investigate the critical impacts of key components of the CNN's model in MRI-based early detection and classification of Alzheimer's disease. By leveraging on different image processing techniques and systematically adjusting hyperparameters, the study achieved impressive overall accuracy performance of 99.50%, superior to existing traditional strategies, with recall, F1-score, and precision values of 99%, 99%, and 100%, respectively. The study highlighted the importance of Alzheimer's disease early detection and identification, with prospect for improving disease diagnosis's accuracy. This development has potential to positively contribute towards Alzheimer's disease patients, healthcare workers and the entire medical sector by enabling timely and accurate detection, consequently improving the lives of patients (Assaduzzaman *et al.*, 2024).

Also, in Abate *et al.* (2023), an ablation analysis was used on face-based recognition study conducted via CNN architectural model. The study investigates the part(s) of the human face that significantly impact on enhanced recognition rate. The main aim is to understand the relevance of each part of face and to know which part of face is most impactful on human identification. By using four face elements of mouth, nose, eyes, and eyebrows, predictably, the deployed CNN model attained high recognition performance accuracy of 96%. The recognition performance under three face elements of eye, mouth, and eyebrows is also satisfactory, with recognition accuracy result of 92%. As predicted, the recognition results drop to 87% and 63% accuracy under two and one face parts, respectively.

In CNN-based JAD in UAVs communication networks, however, such studies remained relatively scarce. A notable exception is the work by Wu *et al.* (2022), which conducted an ablation analysis of a specialized CNN model for cyber-attacks detection in UAVs networks. The study's results show that, the combined efforts of the model key components are responsible for the efficient detection accuracy recorded on different jamming strategies implemented. There is the need to critically investigate the impacts or contributions of individual key components of the model, to arrive at an optimized CNN architectural layout for enhanced JAD in UAV communication networks. Thus, architectural and hyperparameters ablation studies of the CNN-based JAD model have been proposed in this work. Therefore, this research aims to address the identified study gap.

Methodology

This segment details the strategy adopted for the design, implementation and evaluation of the CNN-based JAD system that works within a simulated UAV communication environment. It also explains the methodology used in the ablation study. The entire study follows five key steps: (1) simulation and data generation, (2) data preprocessing and augmentation, (3) CNN model design, (4) ablation study, and (5) performance evaluation.

UAV Communication Network Simulation and Data Generation

To accurately reflect real-world UAV communication scenarios impacted by JAs, detailed simulation environment was created using MATLAB Simulink. The setup mimics the communication link between a UAV and its remote controller under various jamming conditions. The UAV was set up to transmit data using standard wireless protocol over a 2.4 GHz frequency channel. Five jamming techniques were simulated: Barrage Jammer, which transmits a continuous interference signal across wide range of frequency band; Sweeping Jammer, which disrupts communication by continuously sweeping through a range of frequencies, emitting jamming signals as it goes; Tone Jammer, which emits a continuous jamming signal near a specific frequency to disrupt it; Broadband Jammer, creates interference over a broad spectrum, disrupting multiple communication channels; and Pulse Jammer, a smart jammer type, which emits jamming signals in short, high-intensity bursts or pulses to interrupt communication. These jammers types represent common and evolving threats in UAV missions. By tuning jammer parameters such as power, burst timing, and frequency overlap, the simulation allowed for controlled generation of diverse interference patterns.

The simulated communication signal spectrograms were captured under three UAV-to-controller communication scenarios; legitimate communication, collision or jammed communication and no signal or noise communication. The generated spectrogram datasets for the three communication conditions are effectiveness in revealing the temporal and spectral characteristics of the jamming signal interference, making it ideal for CNN-based JAD model learning. Each spectrogram displayed a 2D time-frequency representation of the signal, with jamming activity visibly distinguishable across the frequency bands. A total of 1,862 samples were generated across the three UAV's communication conditions (762 samples for legitimate communication, 650 samples for collision communication and 450 samples for no signal communication), forming the foundation of a

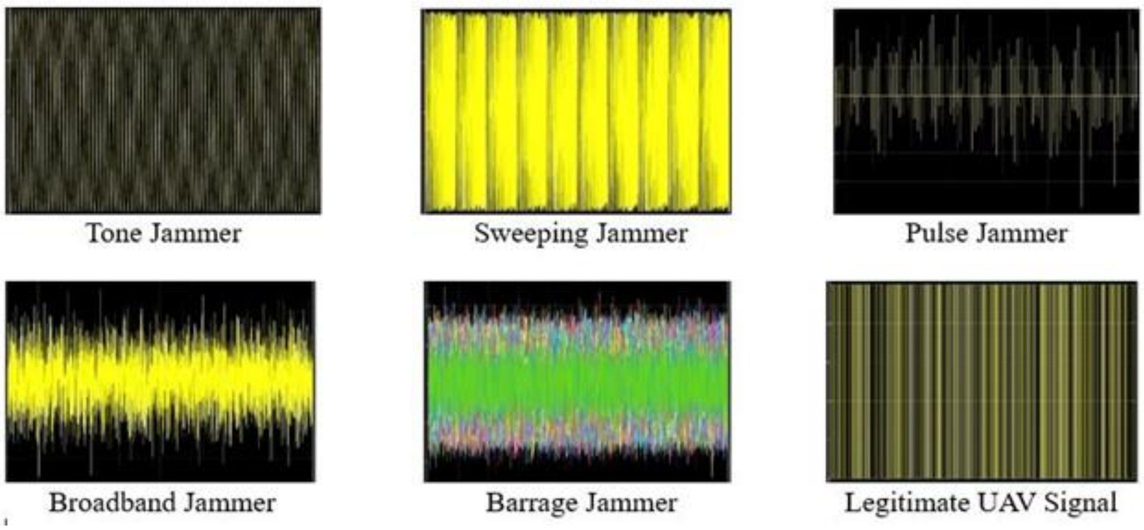

Figure 1: The Spectrograms of Selected Dataset used in CNN-based Model Development

labeled dataset suitable for supervised learning. This synthetic yet realistic dataset reflects dynamic UAV communication behaviour under adversarial conditions. Figure 1 presents the spectrograms of some of the selected dataset obtained during model simulations.

Data Preprocessing and Augmentation

The raw spectrogram images generated from the simulation were standardized to a uniform size of 145×145 pixels and normalized to ensure consistent pixel intensity distributions. This normalization step was crucial to stabilize CNN training and to reduce the influence of outlier values or signal amplitude variations. Label encoding was applied to associate each image with its respective class: 0 for legitimate communication, 1 for compromised communication, and 2 for no signal or noise signal. Before being fed into the CNN, the dataset was randomly shuffled to avoid any ordering bias and shared using ratio 70:15:15 for training, validation, and testing datasets, respectively.

To improve how well the model generalizes and to mimic the variability found in real-world signal conditions, generated data augmentation was carried out. These techniques included random rotation of up to plus or minus 10 degrees, horizontal and vertical flips, adding Gaussian noise, and random cropping. Not only did this augmentation increase the effective size of the dataset, but it also introduced slight distortions that helped the CNN model become more robust against noise and shifts in temporal signals. Consequently, the model learned to concentrate on consistent patterns that indicate JAs rather than focusing on the exact position or shapes of the signals. This step was crucial for enhancing detection performance and generalization, especially during UAV missions in challenging signal conditions.

CNN Model Architecture

The baseline CNN architecture was crafted to strike a balance between model complexity and detection accuracy. The network featured four sequential convolutional layers, each utilizing a 3 × 3 filter size, followed by Rectified Linear Unit (ReLU) activation function and 2 × 2 max pooling. This design enabled the extraction of spatial features (both high and low-levels) from the spectrogram datasets. Deeper layers captured complex frequency-time interactions induced by jamming. Following feature extraction, the feature maps were converted to one-dimensional (1D) vector by the flatten layer before passing through a Fully Connected (FC) dense layer of 64 neurons. A dropout rate of 0.5 was used to reduce overfitting. An illustration of the CN-based JAD model architecture model is as shown in Figure 2.

The signal classification layer used Softmax to output probabilities for each of the three classes classification. Adam optimizer was employed to compile the model, which adapts learning rates for each weight dynamically. The categorical cross-entropy was deployed for the loss function during model's training. The CNN model was implemented in Python version 3.9.0 using TensorFlow library and trained for 50 epochs with a batch size of 32. Early stopping and learning rate scheduling were employed to prevent overfitting and accelerate convergence. This architecture served as the foundation baseline for the ablation study, where the individual 5 key components were systematically ablated to evaluate their contributions to the overall JAD task.

Ablation Study Design

The ultimate goal of this ablation study is to investigate how the CNN model's key components (convolutional layers, pooling layers, activation functions, flatten layer and FC layer) influence its

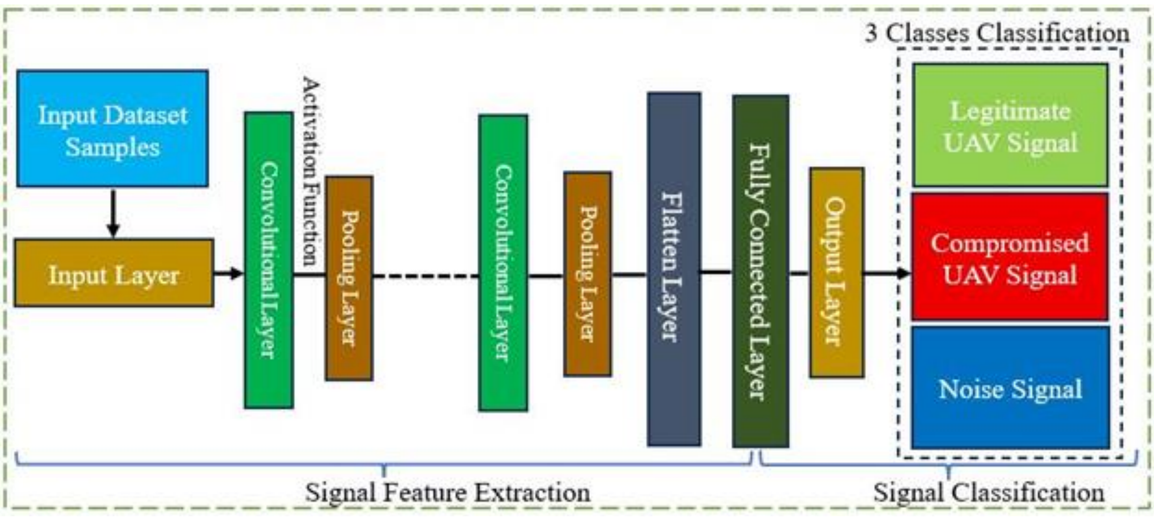

Figure 2: An Illustration of the CNN's Model Architecture for JAD Operation

overall JAD accuracy performance. By systematically removing each of these components and then training the modified model on the same dataset under identical hyperparameter conditions with the baseline model, the aspects of the CNN model that are essential for effective JAD, and which might be superfluous or detrimental can be determined. In this phase, the baseline CNN model (which includes all the 5 key components) was first trained and evaluated to obtain benchmark accuracy, precision, recall, and F1-score performance metrics. This enables the development of an optimized CNN-based JAD for securing UAVs communication under JA threats. The workflow of the ablation study design and implementation is presented in Figure 3.

The first set of experiments focused on the convolutional layers, which are the backbone of CNN model. These layers are responsible for feature extraction, capturing hierarchical patterns in the input spectrogram datasets. Here, all the CNN model's convolutional layers are ablated and the modified model is trained under identical hyperparameter conditions and the same dataset as the CNN baseline model. In the second experiment, the pooling layer was removed entirely to observe the impact of skipping the down-sampling step (the model's ability to retain key features during down-sampling) on the accuracy and computational efficiency. This experiment tested the importance of the pooling layer in the CNN-based JAD under identical hyperparameter conditions as the baseline model. The third experiment investigated the contributions of the activation functions used within the CNN layers. Initially, the model employed ReLU, a widely used activation function that helps to mitigate the vanishing gradient problem. However, to test the robustness of the architecture, ReLU was ablated to investigate the modified model's capacity to differentiate normal UAV communication from JAs signals.

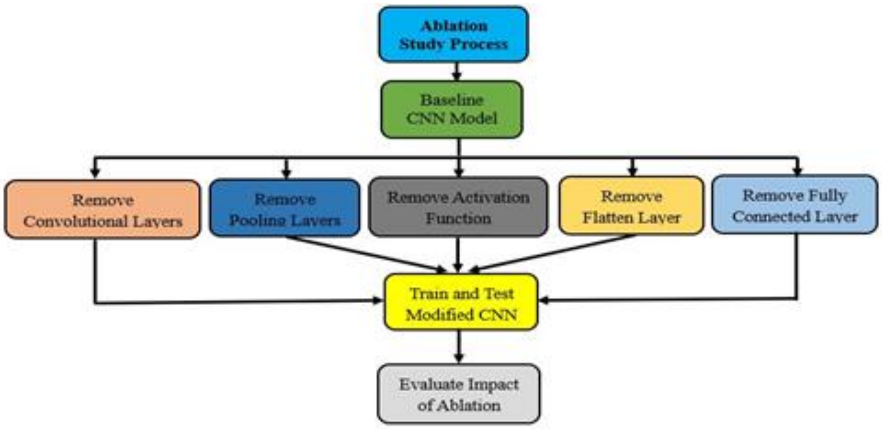

Figure 3: CNN-Based JAD Model Component Ablation Approach

The fourth experiment focused on investigating the impact of the flatten layer on the performance of the CNN model. The flatten layer in a CNN structure serves the crucial role of transitioning multidimensional feature maps into 1D vector, which is then passed into FC layers for JAD task. To assess how the flatten layer impacts CNN model for JAD, the layer was completely ablated. Instead of flattening the feature maps, a Global Average Pooling

(GAP) was used. This modification helps the model to keep the semantic information from each feature map intact without reshaping them, allowing for a more streamlined architecture. The modified model was then trained under identical conditions as the baseline model. The final phase of the ablation focused on the FC layers which are critical in CNN architectures for JAD task. All FC layers were systematically removed from the CNN model and GAP was applied after the final convolutional layer. This pooling output was then directly connected to the softmax classification layer. The rationale behind this was to test whether convolutional layers alone, coupled with global pooling, could capture sufficient spatial and semantic patterns for classification without the FC layer's contribution or involvement. The ablated model was trained and validated using the same dataset and hyperparameters.

Performance Evaluation

The confusion matrix analysis was employed to evaluate the performance of the baseline and modified CNN-based JAD models. The row and column of the confusion matrix represent the instances in actual and predicted classes, respectively. The breakdown allows a clearer view

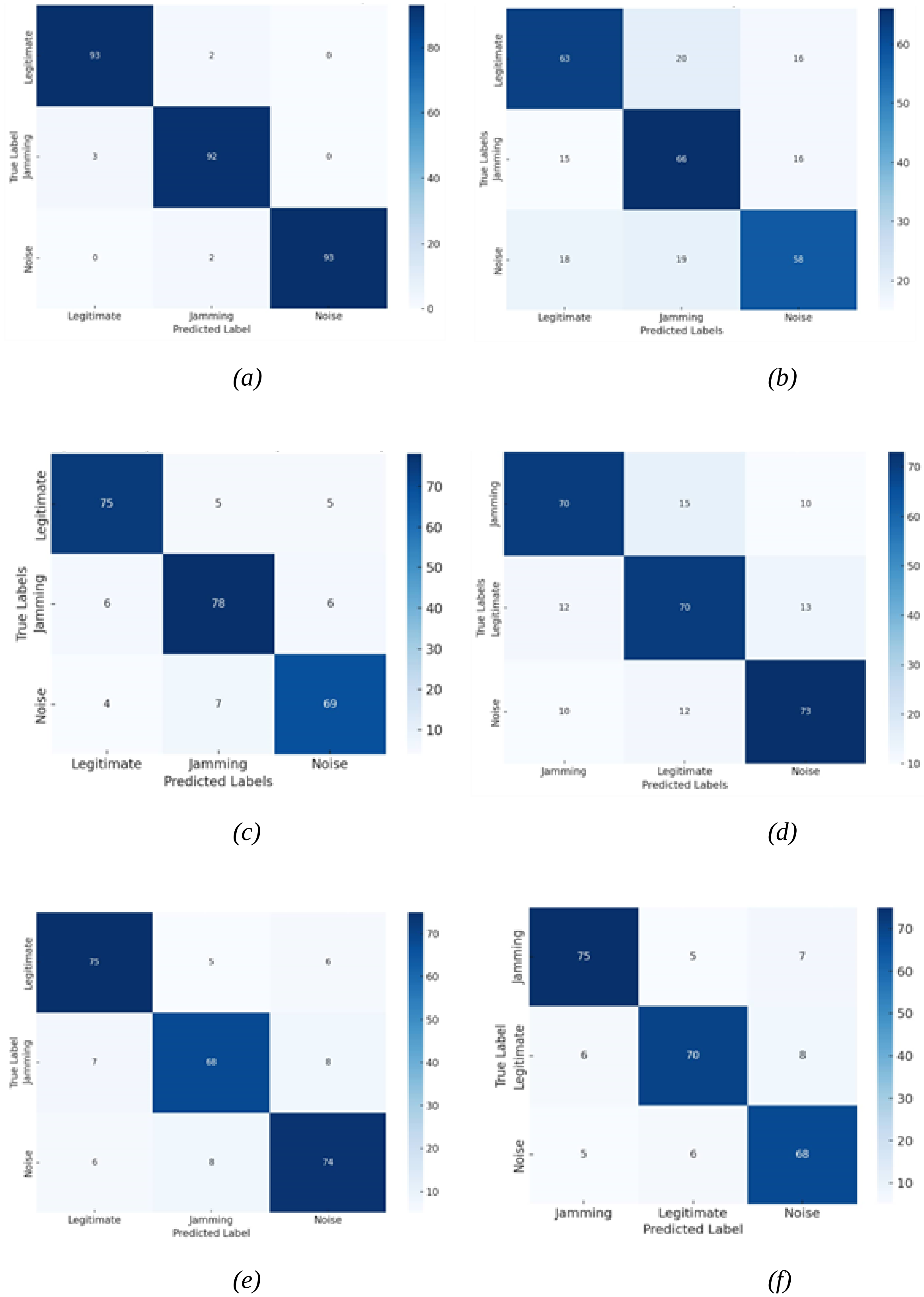

Figure 4: The Results of Confusion Matrix Analysis for the CNN-Based JAD Performance (a) Baseline Model (97%) (b) Ablated Convolutional Layers' Model (65.3%) (c) Ablated Pooling Layers' Model (85.2%) (d) Ablated Activation Function's Model (78.7%) (e) Ablated Flatten Layer's Model (82.4%) (f) Ablated FC Layers' Model (84.1%)

different UAV signals categories (legitimate UAV signal, jammed or compromised UAV signal and noise signals). The confusion matrix provides insights beyond overall accuracy, revealing where the model predictions are correct and where the predictions are incorrect. Thus, the matrix helped identify specific misclassification trends, such as confusing legitimate UAV signals with jammed signals (compromised UAV signals) under JA or noise signals. Statistical averaging over multiple training runs ensured robustness of results.

In addition to the confusion matrix evaluation, other performance metrics were deployed to further quantitatively assessed the baseline and modified CNN models. These metrics include: accuracy, precision, recall, and F1-score. Accuracy is defined as the overall percentage of correctly detected instances across all classes. Precision evaluates the proportion of predicted compromised signals that were actually JAs, while recall quantifies the model's ability to detect all true JA signals. F1-score, is the mean value of precision and recall, which provides a balanced measure that is especially useful in cases of class imbalance. The model's detection capability under different kinds of JA scenarios is comprehensively provided by this performance evaluation.

Results and Discussion

The baseline and other ablated or modified CNN-based JAD models' performances were evaluated on the dedicated 15% test dataset samples. Figure 4 presents the results of all models' confusion matrix evaluations. The results reveal that the baseline model correctly detected 92 UAV-controller signals to have been compromised by JA signals, while 3 datasets were misidentified as legitimate signals out of 95 actual compromised signals that were tested. The confusion matrix results during convolutional layers ablation in the CNN model shows that only 66 jamming signals were correctly detected, while 31 datasets were misclassified out of the 97 actual compromised UAV signals tested.

The result obtained from ablating the pooling layer shows 78 correctly identified jamming signals while 12 were erroneously classified out of the 90 actual jamming signals tested. Also, a total of 95 actual jamming signals test dataset was evaluated on the activation function ablated CNN model which returned 70 correctly identified samples and 25 samples misidentified. The confusion matrix results during flatten layer ablation show that out of 83 actual jamming samples tested, 68 samples were correctly detected while 15 samples were misclassified. Also, the results during FC layers ablation show that 75 samples were correctly detected while 15 samples were misclassified out of 87 actual jamming samples tested. The discrepancies in the models' JAD accuracy shows each ablated components' degree of impacts of on the CNN model's performance. Notably, the convolutional layer ablation shows the most significant impact, with the least number of correctly detected JA signals.

Other performance metrics were evaluated to confirm the efficacy of the CNN-based JAD baseline and the other ablated CNN models. The summary of the evaluation results is presented in Table 1. The table of results clearly shows that removing the convolutional layers significantly degrades the accuracy of the model to 65.3% when compared with the accuracy attained under the baseline or benchmark model at 97.0%. The 78.7% accuracy recorded during the activation function ablation shows drop in model performance. Also, the removal of the CNN model's pooling, flatten and FC layers resulted in 85.2%, 82.4% and 84.1% accuracies, respectively. These results indicate the detrimental effects of bypassing individual components on the overall model performance.

The performance loss recorded during each component ablation expressed in percentage is presented in Figure 5. The result shows that the ablation of the model's convolutional layers caused the most significant performance loss of 35%. This is followed by 22% loss recorded during the activation function ablation. The calculated loss when the pooling, flatten and fully connected layers are ablated is 13%, 14% and 16%, respectively. The results of this analysis underscore the impact of the convolutional layers as the most essential component of the CNN model without which the model's JAD performance is significantly degraded.

Components Contributions

From an ablation standpoint, this result confirms that convolutional layers are indispensable for capturing localized spectral variations such as sudden frequency spikes or consistent waveform patterns that characterize different UAV communication signal classes. Their removal leads the JAD model to rely solely on shallow representations, possibly from fully connected layers, which are inadequate for such nuanced discrimination. This severely hampers the ability of the model to generalize and correctly classify unfamiliar dataset with drop in detection accuracy. Therefore, this analysis supports the broader conclusion that retaining convolutional layers, particularly deeper ones, is critical for maintaining high performance in CNN-based JAD systems. Removing the pooling layers means that the model has to depend solely on convolutional and fully connected layers to compress and make sense

Table 1: CNN Architecture Component Ablation Study

Ablated Component	Accuracy (%)	Precision	Recall	F1 Score
None (Baseline)	97.0	96.0	96.5	96.0
Convolution Layers (All)	65.3	60.0	58.0	59.0
Pooling Layers	85.2	83.0	82.0	82.0
Activation Function	78.7	76.0	74.0	75.0
Flatten Layer	82.4	79.0	81.0	78.0
FC Layer	84.1	81.0	80.0	80.0

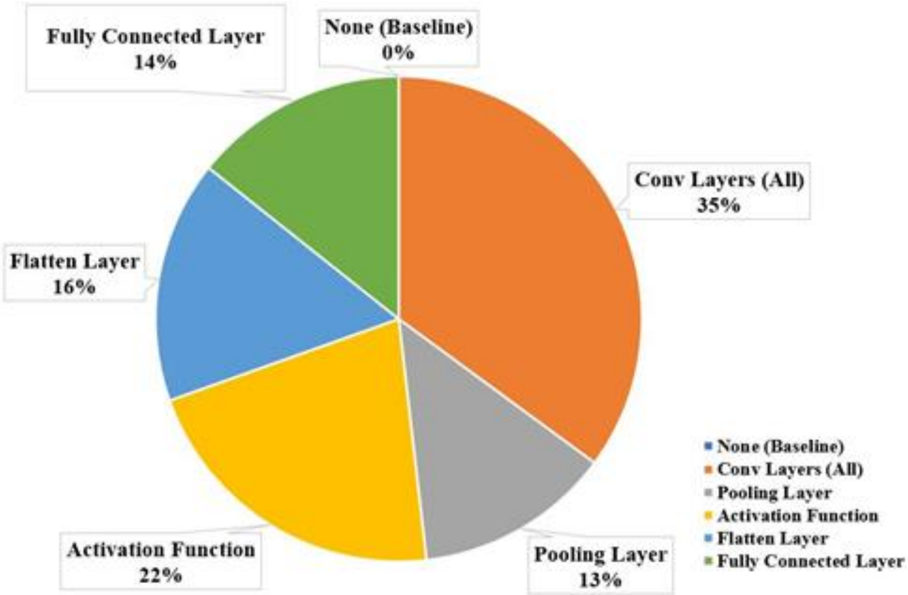

Figure 5: The Analysis of CNN-based JAD Model Performance Losses in Ablation Study

of feature hierarchies. This change leads to a confusion matrix that reveals more misclassifications, particularly in tricky situations temporal or spectral features are only slightly different. For instance, jamming signals might get misclassified as noise because the high-frequency artifacts that pooling layers would have smoothed out are still present. This reinforces the critical role of pooling layers in creating translation-invariant features and ensuring robust classification performance, especially in environments with signal interference variability like UAV communications under JA signals.

The results of ablation analysis highlights how the removal of activation functions hampers the CNN's capacity to introduce non-linearity and learn complex patterns in the spectrogram data. The model becomes more linear and less expressive, reducing its ability to distinguish overlapping or subtly different signal classes. Consequently, this results into degraded class-specific performance, evident in the decreased precision and recall values, and contributes to the overall reduced effectiveness of the JAD model. The performance degradation recorded during flatten layer's ablation underscores its importance in preserving and organizing spatial hierarchies extracted by convolutional layers. When the flattening step is skipped, the structure and intensity distributions captured in the spectrogram features become less linear and more difficult for the dense layers to interpret. As seen in the matrix, confusion among the classes, especially jamming vs noise signals, increased due to a diminished feature vector representation. The results confirm that while CNNs retain some predictive power even without flattening, the integration of this layer is critical for maximizing the fidelity of deep feature representation and ensuring high-confidence JAD tasks.

Also, by removing the FC layers and instead relying on alternatives like global pooling, the model loses some of its ability to weight and prioritize features effectively across the entire feature map. This results in a flatter decision space, leading to increased confusion between similar signal classes, as reflected by the overlap in the confusion matrix.

Although global pooling simplifies the architecture and reduces the parameter count potentially mitigating overfitting it does so at the expense of discriminative power. This outcome underscores how crucial FC layers are for achieving outstanding JAD performance, especially in sensitive systems like UAV operations.

Conclusion

This research analyzes in detail the impact of each component of a CNN-based model for JAD in UAV communication networks. Using an ablation study, contributions of the convolutional layers, pooling strategy, activation function, and flatten as well as FC layers were evaluated on the model's overall performance. The results confirmed the importance of the convolutional layers, particularly with a depth of four layers, which achieved the highest detection accuracy of 97%. Also, the study confirmed the usefulness of pooling layers, as well as activation function, flatten and FC layers to improve the efficiency of the CNN-based JAD model. These results support CNN approaches to increasing the JAs resilience of UAVs communication systems. The findings highlight that the effectiveness of CNN architecture for JAD solutions is highly dependent on the details of the designs. A structured CNN model that considers the convolutional layers, activation function, and other critical components will enhance the overall performance of the model. This study aids the efforts towards providing more robust frameworks for UAV communication networks and stress the need for JAD models that are interpretable and explainable. The integration of these techniques into real-world UAV operations could enhance operational resilience, providing more robust jamming threats detection and ensure safer and more reliable missions.

Future Research Works

The exciting findings from this study pave the way for a variety of future research opportunities in the areas of JAD in UAV communication networks. One intriguing path to explore is enhancing the CNN architecture by integrating more sophisticated techniques like Attention Mechanism or Recurrent Neural Networks (RNNs). By using Attention Mechanism, the model could identify the most pertinent aspect of the spectrogram dataset, which might boost detection accuracy even in noisy settings. Also, merging CNNs with RNNs could improve the model's capability to identify temporal dependencies in communication signals, a key factor for spotting evolving or adaptive jamming tactics over time. Another crucial area for future investigation is testing the CNN-based JAD model in real-world data often presents challenges like noise, variability, complex interference conditions that might not be fully captured in controlled simulations. Moreover, exploring transfer learning techniques to adapt models trained on one set of communication environments to another could significantly enhance model generalization, making it suitable for various UAV platforms and operational contexts. Finally, integrating CNN-based JAD models with other security measures, such as encryption or frequency hopping, could lead to even more robust defense mechanisms against JAs in UAV communication networks.

References

Abate A. F., Cimmino L. and Lorenzo-Navarro J. (2023). An ablation study on part-based face analysis using a Multi-input Convolutional Neural Network and Semantic Segmentation. Pattern Recognition Letters, 173, 45 - 49.

Abiodun T. F. (2020). Usage of Drones or Unmanned Aerial Vehicles (UAVs) For Effective Aerial Surveillance, Mapping System and Intelligence Gathering in Combating Insecurity in Nigeria. African Journal of Social Sciences and Humanities Research, 3(2), 29 - 44.

Alrefaei F. (2024). Machine Learning for Intrusion Detection into Unmanned Aerial System 6G Networks. Doctoral Thesis and Master's Dissertations.

Alvear O., Zema N. R., Natalizio E. and Calafate C. T. (2017). Using UAV-Based Systems to Monitor Air Pollution in Areas with Poor Accessibility. Journal of advanced Transportation, 1, 1 - 14.

Arjoune Y., Salahdine F., Islam M. S., Ghribi E. and Kaabouch N. (2020). A novel jamming attacks detection approach based on machine learning for wireless communication. International Conference on Information Networking, p. 459 - 464.

Arthur M. P. (2019). Detecting signal spoofing and jamming attacks in UAV networks using a lightweight IDS. 2019 International Conference on Computer, Information and Telecommunication Systems (CITS), p. 1 - 5.

Assaduzzaman M., Dutta M., Saha A. and Paul S. G. (2024). ALSA-3: Customized CNN model through ablation study for Alzheimer's disease classification. Informatics in Medicine Unlocked, 50, 1 - 17.

Duan B., Yin D., Cong Y., Zhou H., Xiang X. and Shen L. (2018). Antijamming path planning for unmanned aerial vehicles with imperfect jammer information. IEEE International Conference on Robotics and Biomimetics (ROBIO), p. 729 - 735.

Fei Q., Xuetian Z., Ge M., Michel K. and Wei L. (2019). UAV Network and IoT in the Sky for Future Smart Cities. IEEE Network, 33(2), 96 - 101.

Gao X., Jia H., Chen Z., Yuan G. and Yang S. (2020). UAV security situation awareness method based on semantic analysis. IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), p. 272 - 276.

Gecgel S., Goztepe C. and Kurt G. K. (2019). Jammer detection based on artificial neural networks: A measurement study. Proceedings of the ACM Workshop on Wireless Security and Machine Learning, p. 43 - 48.

Greco C., Pace P., Basagni S. and Fortino G. (2021). Jamming detection at the edge of drone networks using multi-layer perceptions and decision trees. Applied Soft Computing, 111, 107806.

Hafeez S., Khan A. R. and Al-Quran M. M. (2023). Blockchain-Assisted UAV Communication Systems: A Comprehensive Survey. In IEEE Open Journal of Vehicular Technology, 4, 558 - 580.

Kaleem Z., Yousaf M., Qamar A., Ahmad A., Duong T. Q., Choi W. and Jamalipour A. (2019). UAV-empowered disaster-resilient edge architecture for delay-sensitive communication. IEEE Network, 33(6), 124 - 132.

Laghari, A. A., Jumani, A. K., Laghari, R. A. and Nawaz, H. (2023). Unmanned aerial vehicles: A review. Cognitive Robotics, 3, 8 - 22.

Manesh M. R. and Kaabouch N. (2019). Cyber-attacks on unmanned aerial system networks: Detection, countermeasure, and future research directions. Computers and Security, 85, 386 - 401.

Mowla N. I., Tran N. H., Doh I. and Chae K. (2019). Federated learning-based cognitive detection of jamming attack in flying ad-hoc network. IEEE Access, 8, 4338 - 4350.

Naqvi S. A. R., Hassan S. A., Pervaiz H., and Ni Q. (2018). Drone-Aided Communication as a Key Enabler for 5G and Resilient Public Safety Networks. IEEE Communications Magazine, 56(1), 36 - 42.

Paredes J., Jacinto C., Ramírez R., Vargas I., and Trujillano L. (2016). Simplified fuzzy-pd controller for behavior mixing and improved performance in quadcopter attitude control systems. In 2016 IEEE ANDESCON, p. 1 - 4.

Rydberg A., Söderström M., Hagner O. and Börjesson T. (2007). Field specific overview of crops using UAV (Unmanned Aerial Vehicle). Precision Agriculture'07, 357-364.

Sedjelmaci H., Senouci S. M. and Ansari N. (2017). A hierarchical detection and response system to enhance security against lethal cyber-attacks in UAV networks. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 48(9), 1594 - 1606.

Slimane H. O., Benouadah S., Khoei T. T., and Kaabouch N. (2022). A Light Boosting-based ML Model for Detecting Deceptive Jamming Attacks on UAVs. 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC), p. 328 - 333.

Ukwueze, T. K., Nwafor, S. C., Ugwueze, K. O., Nnadozie, E. C., Odo, M., Ezechi, U., Chukwuma, C. K., Okafor, K. and Ani, A. O. (2024). Design and Implementation of an Unmanned Aerial Vehicle (UAV) for Image Capture in Enterprise Farming. Nigerian Journal of Technology, 43(1), 150 – 158.

Wu S., Li Y., Wang Z., Tan Z. and Pan Q. (2023). A highly interpretable framework for generic low-cost UAV attack detection. IEEE Sensors Journal, 23(7), 7288 -7300.

Yaacoub J. P., Noura H., Salman O., and Chehab A. (2020). Security analysis of drone's systems: Attacks, limitations, and recommendations. Internet of Things, 11, 100218.