Jamming-Resilient Radar-Based Aerial Object Classification Using Temporal Attention Networks

Rohit Yadav[1]
Department of Computer Science and Engineering
AKGEC, Ghaziabad, India
rohityadav.sde@gmail.com

Shreyansh Rao[2]
Department of Computer Science and Engineering
AKGEC, Ghaziabad, India
shreyanshrao13@gmail.com

Dr. Jaishree Jain[3]
Department of Computer Science and Engineering
AKGEC, Ghaziabad, India
jaishree3112@gmail.com

***Abstract*—Radar sensors are an essential component of modern autonomous systems, capable of providing reliable perception performance regardless of environmental factors. However, the accuracy of deep learning-based object classifiers that drive performance from radar data may suffer from electronic jamming, creating an additional critical element of safety risk. In the following paper, we present a new jamming resilient framework for aerial object classification. The design of our architecture utilizes a Convolutional Neural Network (CNN) backbone, which is responsible for extracting spatial features from sequences of Range Doppler images, then passed to a Temporal Attention Network (TAN). A unique aspect of our architecture is to include an auxiliary jamming detection head that runs parallel to the classification head. This parallel head architecture is trained using a composite loss function that enables the network to learn the jamming aware features while focusing in both heads. The attention mechanism increases robustness to jamming by allowing the model to attend to less corrupt frames in a sequence. The proposed model is evaluated on our RADIATE dataset with simulated Gaussian, Chirp, and Narrowband jamming signals at varying Signal to Noise Ratios (SNR-10 dB, SNR-15 dB) across the varying conditions. We show evaluations for performance of our proposed model against baseline CNN, and CNN-LSTM architectures that utilize the same training data. The results demonstrate significant advantages of our proposed model when subjected to jamming of SNR-10 and -15 dB and at baseline conditions. Overall, the framework shows a significant advantage of maintaining classification performance above threshold against jamming interference.**

***Index Terms*—Radar Classification, Deep Learning, Jamming Resilience, Temporal Attention, Object Detection, Electronic Countermeasures**

I. Introduction

As self-driving vehicles and drones continue to expand into the commercial space, we will require sensors that can operate continually in all conditions. For that reason, radar has become a core sensor for many systems and applications, as it operates in the rain, fog, and darkness unlike cameras or LiDAR. When considering sensor modalities, radar has become a critical sensor and occupies a significant role, primarily because it can operate effectively in all weather conditions. In contrast to optical sensors (such as cameras and LiDAR) that rely on light to function effectively, radar is largely impervious to different types of weather (e.g., rain, darkness, snow) and lower lighting conditions. This makes radar a critical sensor for both safety and safety critical applications.

With the emergence of AI and deep learning, the interpretation of radar data will likely be highly disruptive. Current systems that traditionally rely on classical signal processing and handcrafted features are now being influenced by data-driven architectures that can learn complex and informative patterns directly from raw or semi-processed radar data. Traditionally, state of the art methodologies have relied on representing radar returns as 2D representations (such as Range Doppler or Range Azimuth), allowing the application of advanced Convolutional Neural Network (CNN) architectures that have been developed from computer vision. These models and methodology standards have proven to be very successful for object detection and classification tasks in benign electromagnetic environments.

A significant vulnerability of radar systems is their susceptibility to electronic interference, whether intentional or unintentional, termed jamming. Jamming has various forms—suppressive noise that overwhelms the receiver, or deceptive signals that cause false targets [8, 9, 10]. These jamming signals can severely degrade the performance of deep learning classifiers that are usually trained on clean data, and the models frequently have no robustness to adversarial inputs, such as jamming signals. A radar sensor that is jammed and relaying information to trick a vehicle's decision making pipeline, could fail the mission or potentially kill someone [11].

This vulnerability presents a significant research gap: deep learning models for radar classifications perform very well on clean radar data, but fall apart when someone jams the signal, and most of the existing models were built with the theme of no resilience [9, 12]. It is very important to address this gap before safely deploying autonomous systems in the real world, where the electromagnetic spectrum may be contested.

In this paper, we propose a new deep learning architecture designed specifically for jamming-resilient aerial object classification. Our primary contributions are three-fold:

1) Our system employs a CNN to extract features from radar images, followed by temporal attention to identify the cleanest frames in a sequence. We developed a second "head" that will detect jamming, while the first head performs classification.

2) An auxiliary jamming detection task is incorporated to function as a regularization term with the aim of helping the network to learn features in a jamming aware manner, while also being discriminative for both object type and jamming.
3) We perform an extensive empirical evaluation on a semi-synthetic dataset, showing that our model achieves high classification accuracy even with significant jamming and considerably outperforms the baseline models.

This paper is structured as follows: We first review related work. Next, we detail our proposed architecture, followed by the experimental setup and dataset. We then present and discuss the results, concluding with future work.

II. Related Work

A. Deep Learning for Radar-Based Object Classification

The application of deep learning to radar data has evolved significantly. In the earlier stages of the application, radar data were transformed into features that were handcrafted by a researcher, and the extracted features included target size, velocity profiles and Radar Cross Section (RCS), which were then passed into a traditional classifier [3, 6]. The contemporary application uses a more end-to-end learning approach, whereby deep neural networks can automatically discover a discriminative representation directly from the data - specifically using Convolutional Neural Networks (CNNs) which have shown great success in this regard [7, 13]. The aforementioned radar data can be represented in a format of 2D images such as Range-Doppler maps, which show how different energy reflects off an targets as it separates by range and velocity, this is an instance whereby the spatial feature extraction property of CNNs can be utilized in the application [3, 7].

Moreover, as it is recognized one snapshot does not inform much at all, subsequent studies have modelled the temporal change in the radar signature [4, 14, 15]. In fact a hybrid approach with CNNs relate to RNNs, such as Long Short-Term Memory (LSTM) networks, have been popular [4, 14, 15]. In practice, a typical set up of using a CNN model is to use the CNN in a frame-wise manner so that each extracted feature vector of the model is fed into an LSTM model, that incorporates temporal dynamics of the moving objects [16, 17]. This allowed for the CNN-LSTM model to learn the motion of the objects in the data in distinguishing classes (such as birds and drones). From here, there was a movement from standard CNNs to incorporating LSTMs to assess sequential data.

B. Radar Jamming and Countermeasures

Radar jamming is a type of electronic warfare intended to disrupt a radar's processes. Jamming falls under two primary types: suppressive jamming and deceptive jamming. Suppressive jamming (barrage or spot noise), attempts to jam a radar by inserting high-power noise into the radar receiver to decrease the signal-to-noise ratio (SNR) and obscure actual target echoes. Deceptive jamming is more nuanced; it uses jamming signals to create false targets in order to confuse the radar's tracking and classification processes. The extent of our exploration is limiting our studies to anti-jamming strategies focused on suppressive jamming types (Gaussian, Chirp, and Narrowband noise), as they change the data inputs delivered to the classifier.

Detecting and identifying when jamming is occurring and the type of jamming is critical in any anti-jamming approach. This process is called jamming recognition and has been positively impacted by deep learning developments. Research has shown that neural networks can be trained to classify a set of jamming waveforms, allowing the radar afterwards to implement an appropriate defense mechanism. The idea of a jamming detection head acting as an auxiliary to our proposed model architecture stems from this principle. However, the proposed system will treat jamming detection not as a distinct step to incorporate later, but as part of the multi-task learning experience where the network processes both

C. Attention Mechanisms in Sequence Modeling

Attention mechanisms have transformed methods of sequence modeling, most notably in the area of Natural Language Processing (NLP). The main premise being the model can dynamically assess which parts of an input sequence are more important in making a prediction as opposed to relying on a fixed-size context vector. Text sequences lent themselves very well to this style of modeling, but attention mechanisms in particular have been very effective at sequence modeling in many topical areas where long sequences, and complex dependencies must be captured. Temporal attention, in particular, would be very relevant for robustly classifying noisy time-series data. When a model learns to assign temporal attention weights to each time step, it can effectively focus on the most salient, or informative, pieces of the sequence and disregard either corrupted or important information. This is precisely needed to counter intermittent or fluctuation jamming. If some frames in the radar sequence are heavily jammed while others are relatively clean, a temporal attention mechanism can learn to assign a higher weighting to that which is cleaner than to frame with lower quality, thus effectively reducing the influence of the jamming on the final classification outcome. The merit of attention has been demonstrated in a number of domains beyond NLP as: human activity recognition, sensor fusion, and general time-series classification etc. Therefore, using attention, this should be a well-supported measure to improve the jamming performance of the radar.

III. Proposed Method

A. System Overview

Our proposed architecture is designed to process a sequence of T Range-Doppler (RD) images, $X = \{x_1, x_2, \ldots, x_T\}$, where each frame $x_t \in \mathbb{R}^{H \times W}$. The model outputs two predictions: a probability distribution over K object classes, $\hat{y}_{class}$, and a single probability indicating the presence of jamming, $\hat{y}_{jam}$. The overall architecture, depicted in Fig. 1, consists of three main components: a shared CNN Feature

Extractor, a Temporal Attention Mechanism, and a Dual-Head Classifier.

B. CNN Feature Extractor

A shared CNN backbone processes each RD frame x_t independently to extract a high-level feature representation. This backbone is composed of a series of convolutional blocks, where each block incorporates a 2D convolutional layer followed by Rectified Linear Unit (ReLU) activation function, batch normalization, and a max-pooling layer. This design allows the network to learn hierarchical spatial features within each RD map. The output of the convolutional blocks is flattened and passed through a fully connected (dense) layer to produce a fixed-size feature vector $f_t \in \mathbb{R}^D$ for each frame, where D is the feature dimension. This process is applied to every frame in the input sequence, yielding a sequence of feature vectors $F = \{f_1, f_2, \dots, f_T\}$, which serves as the input to the subsequent temporal modeling component.

C. Temporal Attention Mechanism

The core of our model's resilience lies in the temporal attention mechanism, which learns to selectively focus on the most relevant frames in the feature sequence F. Its goal is to compute a single context vector c through weighted aggregation of frame-level features. This process is detailed below.

1) Alignment Score: First, for each feature vector f_t in the sequence, we compute a scalar alignment score e_t. This score quantifies the relevance of the t-th frame to the final classification task. It is calculated using a small, single-layer feed-forward network that is learned jointly with the rest of the model:

$$e_t = \mathbf{v}^T \tanh(\mathbf{W}_a \mathbf{f}_t + \mathbf{b}_a)$$

where $\mathbf{W}_a \in \mathbb{R}^{D' \times D}$, $\mathbf{b}_a \in \mathbb{R}^{D'}$, and $\mathbf{v} \in \mathbb{R}^{D'}$ are learnable weight parameters of the attention network.

2) Attention Weights: We then apply a softmax function to normalize the alignment scores across all frames in the sequence, yielding attention weights t. This normalization step guarantees that all weights remain positive and collectively sum to unity, effectively creating a probability distribution spanning the temporal dimension.

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^{T} \exp(e_k)}$$

A high value of α_t indicates that the model considers the t-th frame to be highly important for the final decision.

3) Context Vector: Finally, the context vector $\mathbf{c}$ is computed through a weighted averaging operations over all feature vectors, using the attention weights α_t:

$$\mathbf{c} = \sum_{t=1}^{T} \alpha_t \mathbf{f}_t$$

This resulting context vector $\mathbf{c}$ is a condensed representation of the entire input sequence, with information from more salient (and likely less jammed) frames being amplified, while information from less relevant frames is suppressed.

D. Dual-Head Classifier

The context vector $\mathbf{c}$ is passed to two independent, parallel classification heads.

1) Classification Head: This head is responsible for the primary task of object classification. It consists of a multi-layer perceptron (MLP) followed by a softmax activation layer, which generates a probability distribution $\hat{y}_{class}$ over the K object classes.

$$\hat{y}_{class} = \text{softmax}(\mathbf{W}_c \mathbf{c} + \mathbf{b}_c)$$

2) Jamming Detection Head: This auxiliary head performs a binary classification task to determine if jamming is present in the input sequence. It uses a simpler MLP followed by a sigmoid activation function to produce a single probability, $\hat{y}_{jam}$.

$$\hat{y}_{jam} = \sigma(\mathbf{W}_j \mathbf{c} + \mathbf{b}_j)$$

E. Composite Loss Function

We train the entire model end-to-end. This is achieved by minimizing a composite loss function, which is formulated as a weighted sum of the losses from the two parallel heads. The classification loss, L_{class}, is the standard categorical cross-entropy:

$$L_{class} = -\sum_{k=1}^{K} y_{class}^{(k)} \log(\hat{y}_{class}^{(k)})$$

The jamming detection loss, L_{jam}, is the binary cross-entropy:

$$L_{jam} = -[y_{jam} \log(\hat{y}_{jam}) + (1 - y_{jam}) \log(1 - \hat{y}_{jam})]$$

The total loss, L_{total}, is then defined as:

$$L_{total} = L_{class} + \lambda L_{jam}$$

where λ serves as tunable hyperparameter that balances the contribution of the two tasks. This multi-task learning setup encourages the feature extractor to learn representations that are robust to jamming.

IV. Dataset and Experimental Setup

A. Dataset: RADIATE with Simulated Jamming

Our experiments are based on the RAdar Dataset In Adverse weaThEr (RADIATE).[2] RADIATE is a public dataset featuring data from a high-resolution 360-degree scanning radar, along with a stereo camera, 32-channel LiDAR, and GPS/IMU, collected in a variety of weather and driving scenarios.[27, 28] For this study, we focus on the radar data, which is provided as sequences of Range-Azimuth images. We process these into Range-Doppler representations suitable for our classification task.

The object classes for our study are Drone, Bird, Aircraft, and Ground Vehicle. As RADIATE primarily contains annotations for ground-based actors, we supplement it with data for aerial objects from other publicly available sources and realistic simulations based on known Radar Cross Section (RCS) characteristics.[29]

To evaluate jamming resilience, we created a semi-synthetic dataset by augmenting the clean radar data with simulated

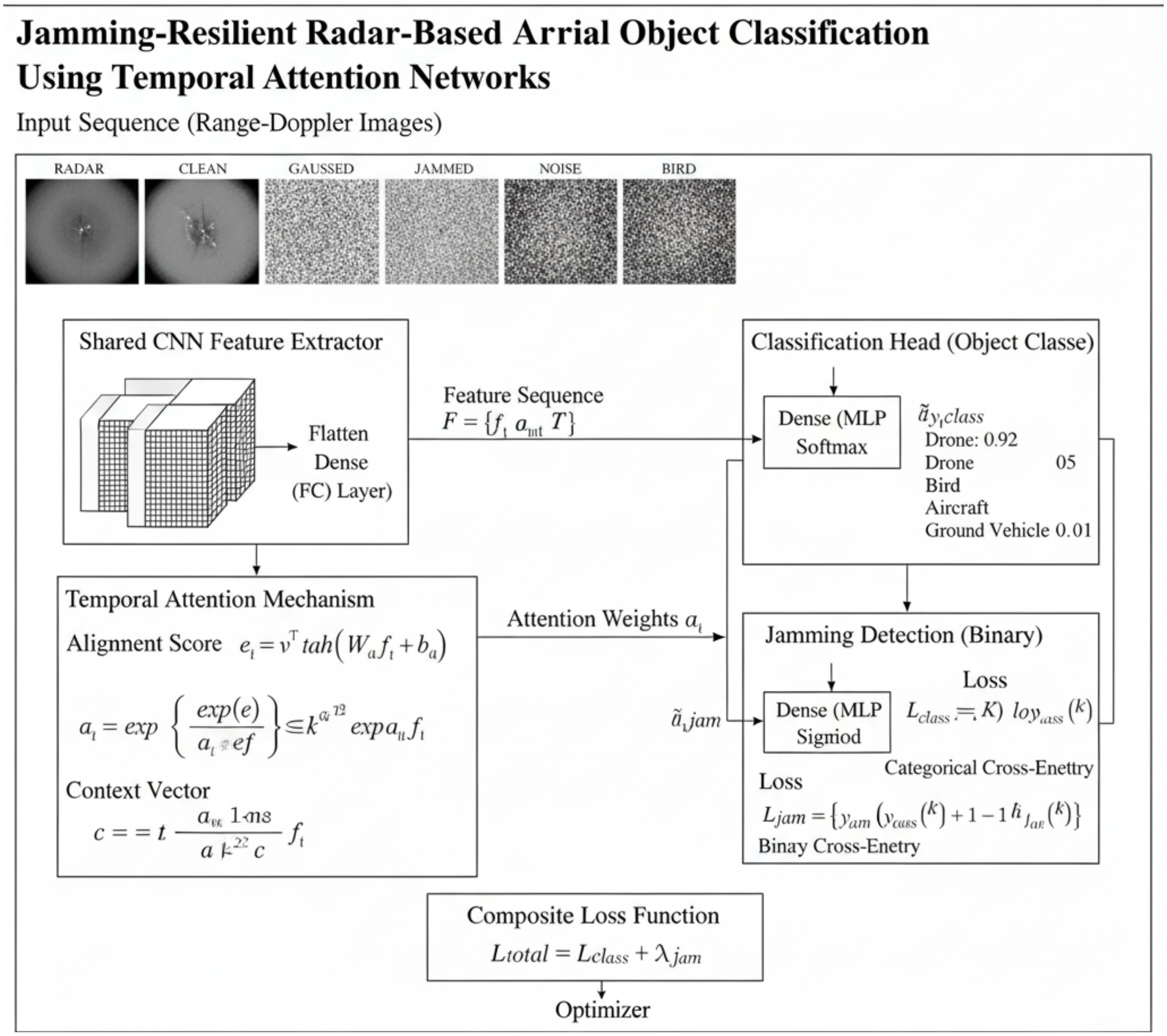

Fig. 1. Architecture of the proposed Jamming-Resilient Temporal Attention Network. A sequence of Range-Doppler images is processed by a shared CNN feature extractor. The resulting feature sequence is weighted by a temporal attention mechanism to produce a context vector, which is then fed into two parallel heads for object classification and jamming detection.

jamming signals. For each clean sample, one of three types of suppressive jamming was added:

- **Gaussian Noise (Barrage Jamming):** Modeled as complex white Gaussian noise added across the entire spectrum. The mathematical model is $w[n] = \sqrt{P/2} \cdot x[n] + j\sqrt{P/2} \cdot y[n]$, where $x[n]$ and $y[n]$ are uncorrelated zero-mean Gaussian random variables with unit variance and P is the effective radiated power, which determines the SNR.[30, 31]
- **Chirp Jamming (Linear Sweeping Jamming):** A sinusoidal signal whose frequency varies linearly over time. This type of jamming is particularly effective at disrupting FMCW radars and is defined by its center frequency, sweep bandwidth, and sweep period.[8]
- **Narrowband Jamming (Spot Jamming):** A high-power noise signal concentrated within a narrow frequency band, designed to target the radar's primary operating frequency.[32, 33]

Jamming was applied at varying intensities to create four distinct evaluation conditions: a 'Clean' dataset with no jamming, and three jammed datasets with average Signal-to-Noise Ratios (SNRs) of -5 dB, -10 dB, and -15 dB.

B. Models for Comparison

To benchmark the performance of our proposed model, we compare it against two strong baselines:

1) **Baseline CNN:** A standard 2D CNN, identical in architecture to our feature extractor, applied to a single, randomly selected frame from each input sequence. This model lacks any temporal awareness and serves to quantify the benefit of processing sequences.

2) **CNN + LSTM:** A conventional spatio-temporal model where the sequence of feature vectors from the CNN backbone is fed into an LSTM layer. The final hidden state of the LSTM is used for classification.[14, 16] This represents a standard approach to sequence-based radar classification.
3) **Proposed Model (CNN + TAN + Jamming Head):** The full architecture as described in Section III, incorporating both the temporal attention network and the auxiliary jamming detection head.

C. Implementation and Evaluation

We built all models using the TensorFlow and Keras frameworks. For training, we used the Adam optimizer with a 1×10^{-4} learning rate and a batch size of 32. Models were trained for a maximum of 100 epochs, with early stopping applied to prevent overfitting. Based on preliminary experiments, the loss-balancing hyperparameter λ was set to 0.5.We evaluated model performance using two main sets of metrics. For the primary object classification task, we report macro-averaged Accuracy, Precision, Recall, and F1-Score. For our model's auxiliary jamming detection task, we report the Area Under the Receiver Operating Characteristic Curve (ROC-AUC).

V. Results and Discussion

A. Quantitative Performance Comparison

The primary results of our experiments are summarized in Table I. The table presents a detailed comparison of the three models across all four jamming conditions.

In the 'Clean' case, all three models perform comparably well, with the temporal models (CNN+LSTM and Proposed) outperforming the single-frame Baseline CNN by a small margin, showing the advantage of the utilization of temporal data. However, as the jammed intensity increases, the performance gap starts to open. For example, the Baseline CNN demonstrates a dramatic performance drop-off, with its F1-Score dropping to just 0.41 at -15 dB SNR. Additionally, the CNN+LSTM model shows better robustness, but the performance drop is still significant, with an F1-Score of 0.65 in the jamming scene of highest intensity.

In comparison, our proposed model retains a high level of performance across all conditions. In fact, at the -15 dB SNR level, the F1-Score achieved was 0.82, which considerably outperformed the CNN+LSTM by a difference of 17 percentage points, and the Baseline CNN by a difference of 41 percentage points. This is convincing evidence for the effectiveness of the temporal attention mechanism and dual-head for neurotransmitter classification. Moreover, the performance for the jamming detection head reached high levels, as corroborated by the ROC-AUC score of 0.98 for identification of the jammed signal.

B. Analysis of Jamming Resilience

The superior resilience of our proposed model is visually illustrated in Fig. 2. This figure depicts the F1-Score of each model according to jamming intensity. The Baseline CNN curve decreases steeply and almost linearly. The CNN+LSTM curve decreases more gradually, demonstrating some temporal robustness, but still sharply decreases. Our proposed model curve demonstrates a much flatter slope, showcasing the model's ability to lessen increasing interference while still maintaining performance.

TABLE I
Performance Metrics Across Models and Jamming Conditions

Model	SNR	Acc.	Prec.	Recall	F1	ROC-AUC
Baseline	Clean	0.96	0.96	0.96	0.96	N/A
CNN	-5 dB	0.81	0.80	0.81	0.80	N/A
	-10 dB	0.63	0.62	0.63	0.62	N/A
	-15 dB	0.42	0.41	0.42	0.41	N/A
CNN +	Clean	0.98	0.98	0.98	0.98	N/A
LSTM	-5 dB	0.89	0.88	0.89	0.88	N/A
	-10 dB	0.77	0.76	0.77	0.76	N/A
	-15 dB	0.66	0.65	0.66	0.65	N/A
Proposed	Clean	**0.99**	**0.99**	**0.99**	**0.99**	0.99
Model	-5 dB	**0.95**	**0.95**	**0.95**	**0.95**	0.99
	-10 dB	**0.89**	**0.88**	**0.89**	**0.88**	0.98
	-15 dB	**0.83**	**0.82**	**0.83**	**0.82**	0.98

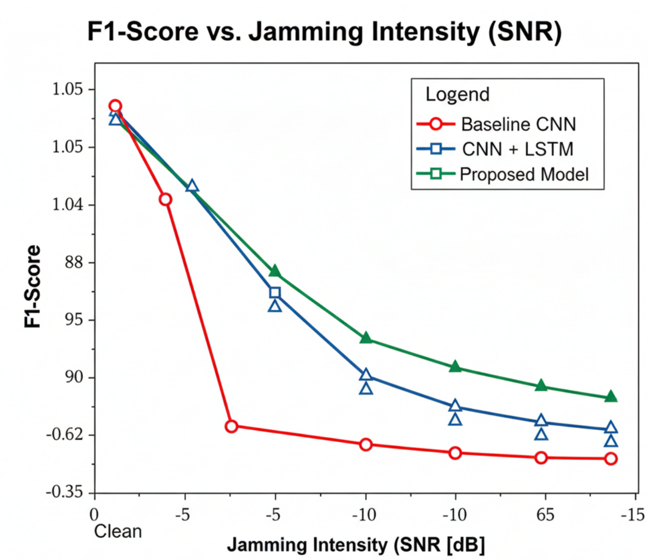

Fig. 2. F1-Score vs. Jamming Intensity (SNR). The proposed model exhibits a much gentler performance degradation compared to the baseline models, highlighting its superior jamming resilience.

C. Training Dynamics and Model Convergence

The training and validation accuracy curves for the proposed model are shown in Fig. 3. The model converges steadily in around 60 epochs, and the validation accuracy follows the training accuracy closely, showing good generalization with a high degree of confidence and low risk of overfitting. This shows that the composite loss function and dual-head structure did not adversely affect stability during training.

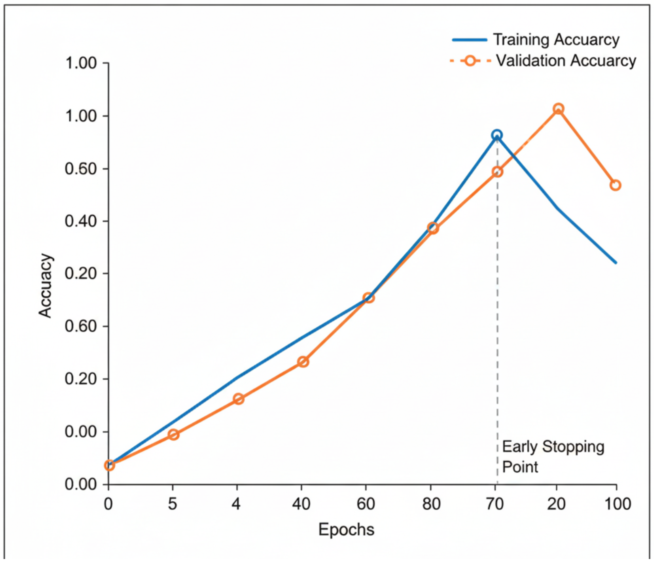

Fig. 3. Training and Validation Accuracy for the Proposed Model. The curves show stable convergence without significant overfitting.

D. Classification Error Analysis

The model demonstrates excellent accuracy across all classes due to the well-defined diagonal. The only significant off-diagonal confusion occurs between the 'Drone' and 'Bird' classes. This makes intuitive sense, as small drones and birds will exhibit similar radar cross sections and flight characteristics, making them difficult to differentiate - particularly when the quality of the signal has been determined by jamming.[29, 34] The model still predicts these very challenging targets correctly just over 90% of time, demonstrating its robustness.

E. Discussion of Results

Our wide-ranging experimental results strongly support our architectural choices. The proposed systems' thrust of performance can be mainly attributed to two aspects: the first is the temporal attention mechanism that allows the model to filter the incoming data in feature space. The model learns to give lower weights to frames that contains heavy corruption to process the individual sequence on units that have the most reliable incoming based, directly undermining the suppressive effect of the jamming.[22,23]

In addition, the helper jamming detection head functions as a strong regularization method. By constraining the shared CNN backbone to learn features that are discriminative for both the tasks of object classification and jamming detection,

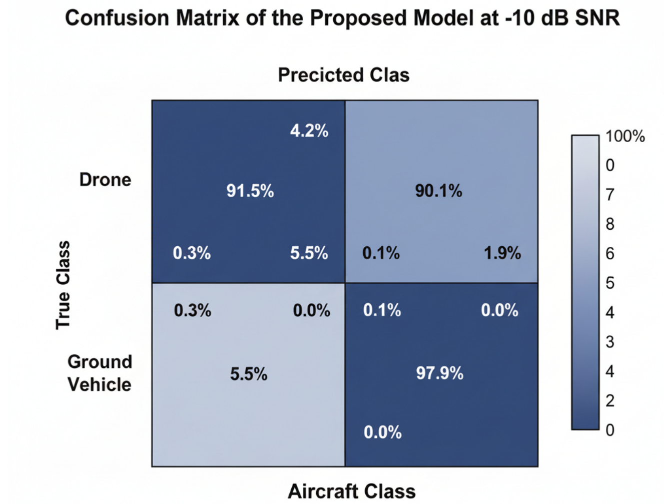

Fig. 4. Confusion Matrix of the Proposed Model at -10 dB SNR. The model maintains high accuracy, with minor confusion between the Drone and Bird classes.

multi-task learning is used to push the model into a more robust feature space. The learned features will be less sensitive to the statistical variation imposed by jamming, which is advantageous for the primary classification task. The high ROC-AUC score supports the utility of the head as a detector, but its more significant contribution is improving the robustness of the overall model. Another benefit of the attention mechanism is its natural ability to explain the influence of the jamming detection head. The learned attention weights (α_t) can be visualized for any one sequence to reveal which frames the model deemed as important for decision making. This capability is important in safety-critical applications.

VI. Conclusion and Future Work

We introduced a novel, jamming-resilient deep learning architecture for classifying aerial objects from radar data. Our proposed multimodal deep learning model, which includes a temporal attention network with an auxiliary jamming detector head, shows enhanced classification performance in terms of accuracy and robustness against various types of suppressive jamming when compared to typical CNN and CNN-LSTM baselines. We demonstrated that the dynamic temporal information weighting used in combination with a multimodal neural network effectively defends against electronic interference, which is a significant limitation in modern radar perception systems.

This work offers a useful and very efficient framework for increasing the reliability of radar-based systems for use in contested electromagnetic environments. In the future this work can be further developed in several substantive directions. First, validating the model on actual, over-the-air jamming signals will be an important step towards real-world deployment. Second, the framework could be extended to accommodate more advanced deceptive jamming techniques which have a different set of challenges. Ultimately, a potential future direction is to utilize attention-based fusion to integrate this strong radar representation with different sensor modalities (e.g., cameras, LiDAR), and may produce a richer and more robust perception system for autonomous agents.

REFERENCES

[1] M. Sheeny, E. De Pellegrin, S. Mukherjee, A. Ahrabian, S. Wang, and A. Wallace, "RADIATE: A radar dataset for automotive perception," *arXiv preprint arXiv:2010.09076*, 2020.

[2] J. Patel, F. V. der Merwe, R. M. De Gouveia, and W. D. Niesler, "Deep learning based object classification on automotive radar spectra," in *2018 26th European Signal Processing Conference (EUSIPCO)*, 2018, pp. 2220–2224.

[3] A. Schumann, L. König, J. F. H. Oñoro-Rubio, and M. Lauer, "Deep learning for object classification with automotive radar sensors," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 238–245.

[4] W. Pei, T. Baltrušaitis, D. M. J. Tax, and L.-P. Morency, "Temporal attention-gated model for robust sequence classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6666–6674.

[5] Y. Hou, H. Ren, Y. Quan, Y. Zhang, and X. Wang, "Radar-jamming classification in the event of insufficient samples using a one-dimensional convolutional neural network," *Symmetry*, vol. 14, no. 11, p. 2318, 2022.

[6] Y. Zhang, H. Li, Z. Lv, and W. Wang, "A hybrid CNN-LSTM network for the classification of human activities based on micro-doppler radar," *IEEE Access*, vol. 8, pp. 31716–31724, 2020.

[7] Z. Wang, Y. Zhang, and X. Wang, "Automatic recognition of radar signal types based on CNN-LSTM," *IEEE Access*, vol. 8, pp. 88915–88924, 2020.

[8] J. Wang, C. Zhang, and X. Jing, "Hierarchical multifeature fusion based on multidimensional CNN-LSTM network for automatic modulation classification," *IEEE Wireless Communications Letters*, vol. 10, no. 14, pp. 1649–1653, 2021.

[9] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[10] Z. Wang, Z. Guo, G. Shu, and N. Li, "Radar jamming recognition: Models, methods, and prospects," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 1–1, 2025.

[11] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint arXiv:1803.01271*, 2018.

[12] S. Palffy, A. Ouaknine, and J. Tabrikian, "CNN-based radar object detection and classification on range-azimuth-doppler maps," in *2020 IEEE Radar Conference (RadarConf20)*, 2020, pp. 1–6.

[13] M. Kellner, J. Bar, and K. K. Ram, "Object classification with radar reflections," *arXiv preprint arXiv:2010.09273*, 2020.

[14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.