Complex Deep Learning for Reliable Radio Fingerprinting in Adversarial Environments

Fahmida Afrin and Nirnimesh Ghose

School of Computing at UNL

Motivation

- Radio fingerprinting is crucial for wireless device identification in various applications.
- It works by analyzing hardware imperfections in RF signals to create unique device "fingerprints."
- Previous approaches using machine learning have shown poor performance in adversarial environments, particularly in cross-day scenarios.
- Our proposed solution is a deep-learning approach that utilizes complex-valued activation functions to capture phase information in addition to amplitude.
- We also explore different pre-processing techniques and hyperparameter tuning to improve our approach's robustness to different scenarios.

Methodology

- A complex deep-learning method for radio signal analysis was proposed in order to improve device identification accuracy in adversarial environments.
- Experiments were carried out to evaluate deep neural network models with various complex activation functions, such as modReLU, CReLU, and ZReLU.
- To enhance robustness, we investigated various pre-processing techniques and hyperparameter tuning on various parameters such as stride (s), trace length (L), and window size (w). The classification model's performance was evaluated using the accuracy rate as the performance measure. The number of layers in the neural network model was also changed to improve model performance.
- The performance metrics were used to compare the efficacy of deep neural network models with various complex activation functions and pre-processing parameters. The outcomes were examined in order to identify the best model for device identification in adversarial environments.

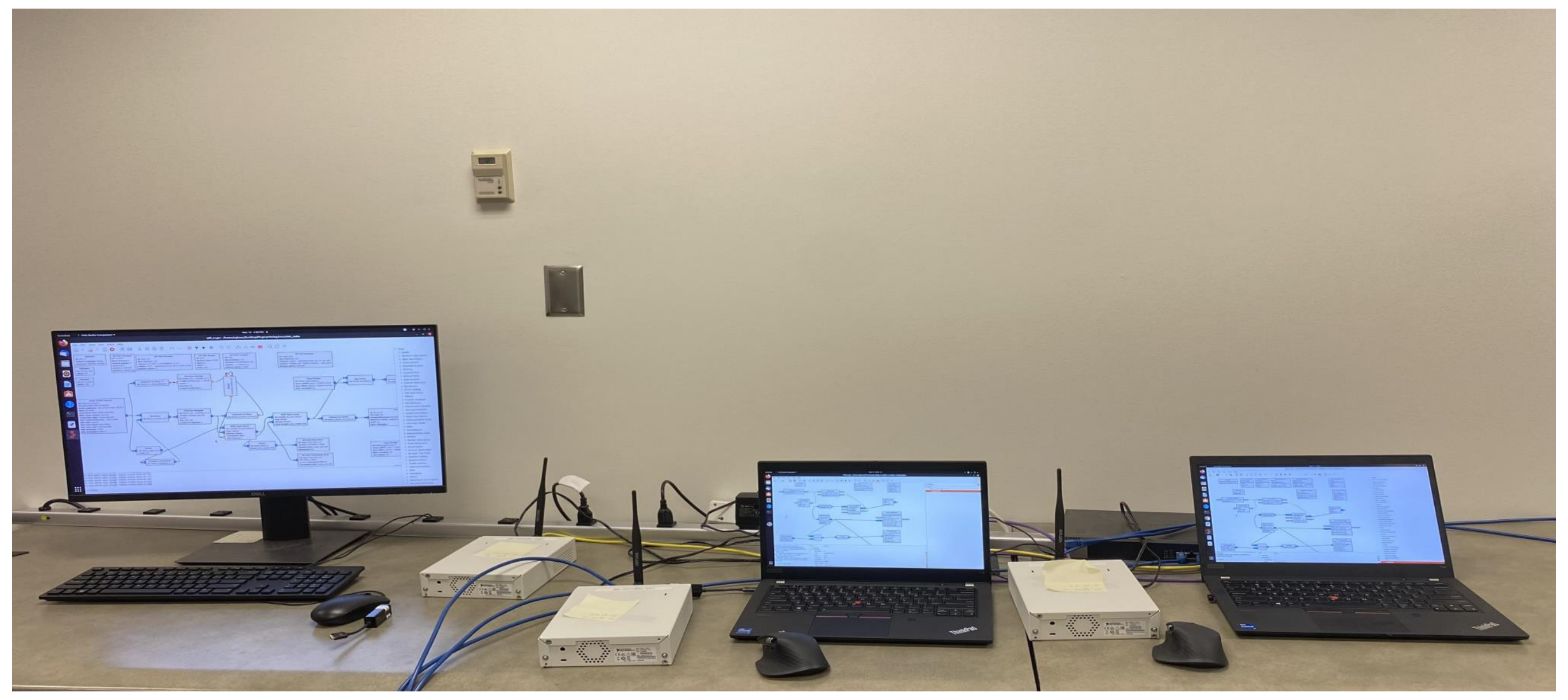

Figure 1: Testbed setup using USRP 2922 platform with 2 transmitters and 1 receiver, all running GNU Radio.

Preliminary Experiment

- The study used a USRP 2922 testbed and VERT2450 Antenna with two transmitters and a receiver to collect radio frequency data. The data was obtained through Wi-Fi transmissions with BPSK 1/2 modulation at a center frequency of 2.45 GHz, a 2 MHz bandwidth, and a 2 MHz sampling rate. Open-source GNU Radio code was used for data collection.
- The receiver was positioned one foot away from the transmitter, and both remained static. Data were collected for two days in a lab environment with three transmissions per day, each separated by one minute and a 15-second break.
- I/Q samples were taken at four points on the receiver side to capture the data: before FFT (frequency domain), after FFT (time domain), after equalizer (equalized), and metadata. Each transmitter broadcasted signals that were recorded for one minute, resulting in approximately 1,667 I/Q traces collected for each transmission from each transmitter.
- The experiment's training and testing phases used I/Q traces collected on Day 1 and Day 2, respectively, while Day 3 was used to evaluate the classifier's performance. To gather data, 5,000 I/Q traces were extracted from each transmitter.
- The collected I/Q traces were randomly split into three sets for training, validation, and testing purposes, with 64%, 16%, and 20% of the traces being allocated to each set, respectively, for each experiment conducted.

Preliminary Results

Table I

After FFT (frequency domain) L = 288 and w =64

	Same-day			Cross-day		
	modReLU	CReLU	ZReLU	modReLU	CReLU	ZReLU
S=0.5 L	0.49	0.5	0.49	0.45	0.48	0.44
S=2L	0.5	0.51	0.48	0.4	0.49	0.45
S=L	0.45	0.54	0.48	0.41	0.44	0.44

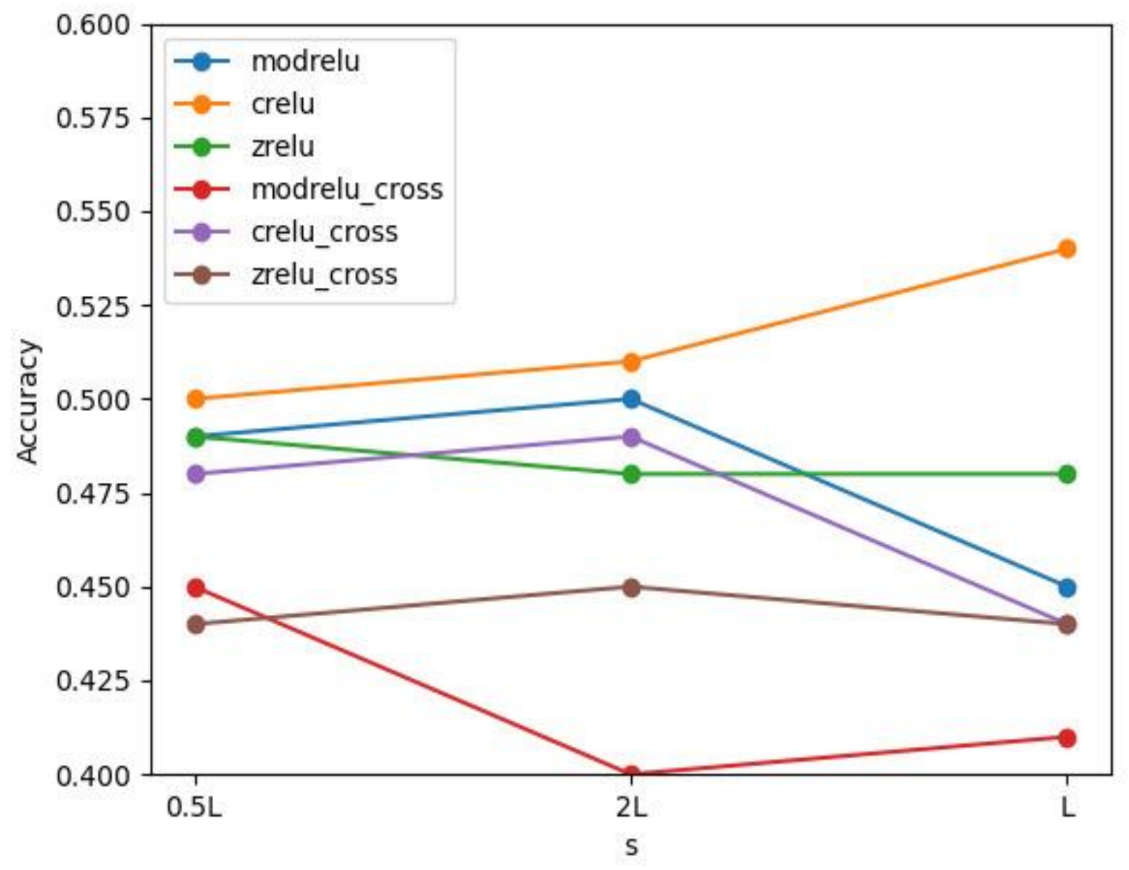

Figure 2: After FFT (frequency domain) L = 288 and w =64, same-day and cross-day accuracy for different complex activation functions

Table II

After FFT (frequency domain) s = 288 and L=288

	Same-day			Cross-day		
	modReLU	CReLU	ZReLU	modReLU	CReLU	ZReLU
w=64	0.51	0.51	0.51	0.44	0.52	0.4
w=128	0.5	0.55	0.54	0.41	0.5	0.44
w=256	0.52	0.54	0.47	0.41	0.54	0.44

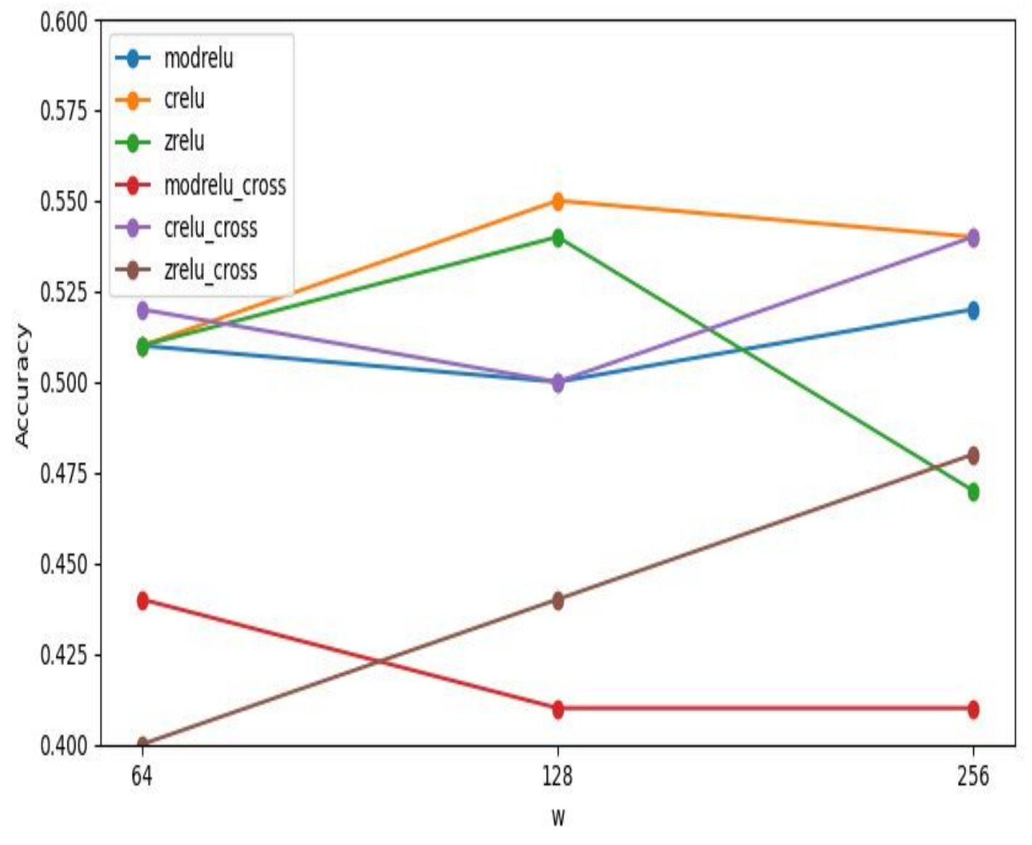

Figure 3: After FFT (frequency domain) s = 288 and L = 288, same-day and cross-day accuracy for different complex activation functions

Key Contributions

1. Direct use of complex numbers: Rather than converting complex numbers from two distinct arrays of floats, we passed complex numbers directly into the model.
2. Collection of a comprehensive dataset: We gathered a comprehensive dataset for Wi-Fi data at various locations, including before and after the Fast Fourier Transform (FFT), equalized, and with metadata.
3. Investigation of various complex activation functions and network architectures: Deep neural network architectures with various amounts of layers were investigated.
4. Extensive hyperparameter tuning: To enhance the efficacy of the model, hyperparameter tuning was carried out by adjusting variables like stride, length, and window size.

Future Works

1. We plan to collect data from a greater number of USRP devices, including both static and moving devices to capture a wider range of environmental data.
2. We aim to enhance our approach by utilizing more advanced techniques, such as the triplet network for training and testing, as well as the Generative adversarial network model.

References

1. Li, Haipeng, et al. "RadioNet: Robust deep-learning based radio fingerprinting." 2022 IEEE Conference on Communications and Network Security (CNS). IEEE, 2022.
2. Cole, Elizabeth K., et al. "Analysis of deep complex-valued convolutional neural networks for MRI reconstruction." arXiv preprint arXiv:2004.01738 (2020).
3. Cekic, Metehan, Soorya Gopalakrishnan, and Upamanyu Madhow. "Robust wireless fingerprinting: Generalizing across space and time." arXiv preprint arXiv:2002.10791 (2020).

ACKNOWLEDGEMENTS

This work was supported by the Nebraska Public Power District through the Nebraska Center for Energy Sciences Research and SWAN Lab at the University of Nebraska-Lincoln.

NEBRASKA CENTER FOR ENERGY SCIENCES RESEARCH