SARP: Spatial Agnostic Radio Fingerprinting with Pseudo-Labeling

Fahmida Afrin*, Haipeng Li†, Boyang Wang† and Nirnimesh Ghose*
*School of Computing, University of Nebraska-Lincoln, USA
†Department of Electrical and Computer Engineering, University of Cincinnati, USA
Email: fafrin2@huskers.unl.edu, li2hp@ucmail.uc.edu, boyang.wang@uc.edu, nghose@unl.edu

***Abstract*—Deep-learning radio fingerprinting is not robust against spatial variations, where a neural network trained on location A does not perform well over RF signals from location B. We promote the robustness of deep-learning radio fingerprinting against spatial variations by synergizing Complex-Valued Neural Networks (CVNNs) and pseudo-labeling. Compared to existing solutions, we leverage pseudo-labeling to fine-tune a CVNN without needing labeled RF signals from a new location. We collect large-scale real-world datasets across different locations. We conduct comprehensive evaluations of these datasets with multiple complex-valued activation functions. Our experimental results significantly improve the accuracy of radio fingerprinting when training data and test data are from two different locations (e.g., increasing accuracy from 40.0% to 62.8%).**

***Index Terms*—Radio fingerprinting, Complex-valued neural network, and Pseudo-labeling.**

I. Introduction

Radio fingerprinting [1], [2] identifies wireless transmitters by analyzing Radio Frequency (RF) signals on the receiver side. As RF signals present minor but unique shifts caused by hardware variations from manufacturing, it is feasible for the receiver to distinguish transmitters. Radio fingerprinting is an approach for physical-layer authentication, which is a critical and complementary component for enhancing wireless security, especially when traditional cryptography-based authentications are difficult to apply. Research studies show that radio fingerprinting can authenticate individual devices (such as smartphones and Internet of Things) and mission-critical targets (such as Unmanned Aerial Vehicles).

The majority of studies in radio fingerprinting leverage real-valued Convolution Neural Networks (CNN), in which complex values in RF signals are separated into two channels (i.e., one for real components and one for imaginary components) as inputs for a neural network. Despite promising results, we argue that RF signals are complex values in nature, and separating real and imaginary parts may lead to suboptimal performance in radio fingerprinting, especially when facing spatial variations. While there are a few studies [3]–[6] investigating Complex-Valued Neural Networks in the context of radio fingerprinting, the observations reported are limited to simulated RF signals and/or RF signals without spatial variations. *In other words, how and to what degree Complex-Valued Neural Networks can tackle spatial variations in radio fingerprinting remains mainly unknown.*

In this paper, by passing complex values directly into neural networks, our method keeps inter-correlations between real and imaginary components of I/Q (In-phase and Quadrature) samples in RF signals. Our main contributions and observations are:

- We have addressed the spatial variations of wireless channels by performing fine-tuning-based device fingerprinting without requiring labeled data from a new location.
- We conduct comprehensive evaluations of our datasets. We have the following key observations. (1) First, Concatenated Rectified Linear Units (CReLU) offers the best performance in radio fingerprinting compared to two other complex-valued activation functions (Complex Cardioid and Cart Leaky ReLU). (2) Second, our method can improve the accuracy of radio fingerprinting when training RF signals and testing RF signals from different locations. Specifically, a baseline CVNN without pseudo-labeling can only achieve 40.0% accuracy, while our method can derive an accuracy of 62.8% in identifying ten transmitters from a new location.

II. SARP: Spatial Agnostic Radio Fingerprinting with Pseudo-Labelling

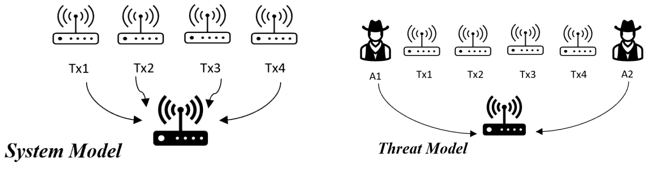
Radio fingerprinting often faces substantial performance degradation, given spatial variations. Specifically, a classifier trained with I/Q traces from one location would have difficulty identifying the same group of transmitters at a different location. A straightforward approach is to retrain the classifier with labeled I/Q traces from a new location.

A. System Model

Our implementation of spatial agnostic radio fingerprinting utilizes a system model consisting of multiple transmitters and one receiver, as depicted in Fig. 1. The radio fingerprinting process encompasses two main phases: the training phase and the authentication phase (also known as the test phase).

During the *training phase*, each transmitter in our system transmits Radio Frequency (RF) signals to the receiver. After obtaining the *training data*, the receiver transmits it to a server responsible for training a classifier. The server leverages training data to build a model to classify and distinguish RF signals from different transmitters.

In the *authentication phase*, when the receiver receives RF signals, it queries the trained classifier to predict the identity of



(a) (b)

Fig. 1: (a) System model of radio fingerprinting. (b) During the authentication phase, A1 or A2 are regarded as potential attackers who could impersonate Tx1, Tx2, Tx3, or Tx4.

the transmitter agnostic to the training location. The classifier determines which transmitter is transmitting at a given time by comparing the received RF signals with the learned patterns from the training phase.

B. Threat Model

We posit trust among all participants during the training phase but consider potential attackers in the *authentication phase*. In our threat model, attackers could be transmitters involved in the training phase (e.g., Tx3 in Fig. 1 impersonating Tx1 or Tx2) or transmitters absent during training (e.g., A1 or A2 in Fig. 1 impersonating Tx1, Tx2, or Tx3).

C. Our Main Idea

We address the problem by integrating Complex-Valued Neural Networks (CVNNs) and pseudo-labeling. Specifically, we first enhance the performance of radio fingerprinting by utilizing CVNNs. Second, we leverage pseudo-labeling to generate pseudo-labels for unlabeled RF signals from a new location and then utilize these pseudo-labeled RF signals to fine-tune the CVNN.

However, using these pseudo-labeled RF signals alone in the fine-tuning process would not help improve the performance of the CVNN over RF signals from a new location as the trained CVNN does not derive high accuracy over RF signals from a new location (i.e., these pseudo labels are less accurate) in the first place. To mitigate this issue, we adopt two improvements, including (1) applying a threshold to select pseudo-labeled RF signals with scores higher than the threshold only and (2) using a combination of labeled RF signals from the training data and the selected pseudo-labeled RF signals from a new location, in the fine-tuning process.

D. Fine-Tuning

Transfer learning with fine-tuning consists of three steps, including training, fine-tuning, and testing.

Training: The first step involves training a classifier with extensive source data to learn feature representations. This training phase lets the classifier obtain fundamental knowledge from the source domain.

Fine-Tuning: In this step, the knowledge acquired during training is transferred to the classification task in the target domain. Specifically, the last few layers (e.g., 1 or 2 layers) of a classifier are fine-tuned (i.e., updated) using a small amount of data from the target domain, referred to as *target training data*. The remaining layers of this classifier remain the same (i.e., frozen) during this process.

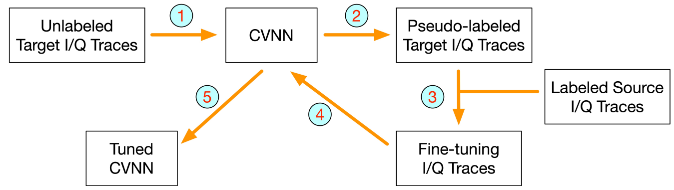


Fig. 2: Detailed steps of our fine-tuning with pseudo-labeling over unlabeled target I/Q traces in one iteration.

Testing (Authentication): The fine-tuned classifier is evaluated on *target test data* from the target domain using metrics such as accuracy. *In spatial agnostic radio fingerprinting, we consider I/Q traces acquired from one location as data of the source domain and I/Q traces captured from another location as data of the target domain.*

E. SARP: Our Proposed Methods

We designed our method to mitigate spatial variations by fine-tuning a classifier using pseudo-labeling on unlabeled target data from a new location. A CVNN is first trained using labeled source data. For fine-tuning, we leverage semi-supervised pseudo-labeling as presented in Fig. 2.

More specifically, during the *fine-tuning* step, (1) A set of unlabeled I/Q traces from the target dataset is first passed to the trained CVNN to assign confidence scores and labels, which are referred to as *pseudo-labels*; (2) Next, given a pre-defined threshold θ, where $\theta \in (0, 1)$, pseudo-labeled I/Q traces with confidence scores higher than pre-defined threshold are identified; (3) These identified pseudo-labeled I/Q traces are mixed with labeled source data to form fine-tuning traces; (4) The fine-tuning traces are passed to fine-tune the last few layers of the trained CVNN; (5) A fined-tuned CVNN is derived. The above five steps are repeated for multiple iterations with the same set of unlabeled I/Q traces from the target dataset and labeled source data until the confidence scores over these unlabeled I/Q traces from the target dataset do not further improve (i.e. no unlabeled I/Q traces can be further included in the fine-tuning traces).

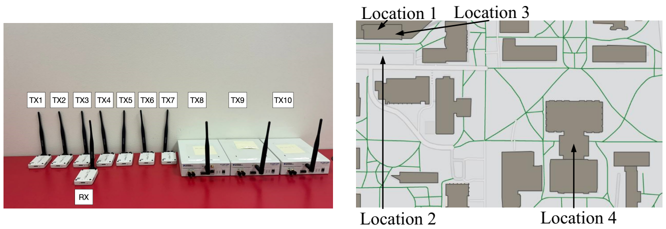
III. Dataset and Evaluation

We assess the performance of SARP in scenarios where source and target data are from different locations.

A. Our Testbed and Dataset

Our experimental setup, depicted in Fig. 3(a), comprises ten transmitters (three USRP 2922 devices and seven USRP B200-mini-i devices) as transmitters and one receiver (USRP B205-mini-i). To facilitate WiFi transmissions using BPSK 1/2 modulation, we utilized the open-source GNU Radio code from [7]. A center frequency of 2.45 GHz, a bandwidth of 2.5 MHz, and a sampling rate of 5 MHz were employed, resulting in the recording of I/Q samples.

We collect I/Q data at four locations: I/Q_Loc1 (Location 1), I/Q_Loc2 (Location 2), I/Q_Loc3 (Location 3), and I/Q_Loc4



(a) (b)

Fig. 3: (a) Testbed: One receiver and ten transmitters, (b) four locations for data collection: one outdoor (Location 2) and three indoors (Location 1, 3, and 4).

(Location 4), as shown in Fig. 3(b). All the devices were stationary during the data collection, with the transmitter and receiver positioned around 3 feet apart. Each area underwent three transmission sessions, each lasting 30 seconds, with a 30-second gap between two transmissions. Despite packet loss, approximately 4.50×10^8 I/Q samples were successfully collected in each transmission from the frequency domain.

B. Architectures of Complex-Valued Neural Network

We construct our neural network using the CVNN framework, illustrated in Fig. 4. In particular, we make use of the `cvnn.layers` module instead of `tf.keras.layers` in TensorFlow [8].

C. Experimental Settings

We run our experiments on a server equipped with a Tesla V100-PCIE-32GB GPU, supporting cuDNN version 8.6.0. We train a CVNN model for 100 epochs with a learning rate 0.1. We implement early stopping with a patience setting of 10 epochs. Given a dataset with raw I/Q samples, sliding windows are applied to extract I/Q traces as inputs for a neural network. From the raw I/Q samples of a dataset, we randomly selected 1,000 I/Q traces from each transmitter, resulting in a total of 10,000 I/Q traces given 10 transmitters. We use 72% for training, 8% for validation, and 20% for testing when we evaluate a CVNN.

D. Experiments

Experiment A: Baseline Performance of CVNN. We first examine the baseline performance of CVNN in radio fingerprinting without applying pseudo-labeling. A same-location scenario indicates that the training and test I/Q traces are from the same location (e.g., I/Q_Loc1). A cross-location scenario suggests that the training and test I/Q traces are from two locations (e.g., I/Q_Loc1 and I/Q_Loc2). In addition, we compare the performance of three complex-valued activation functions: Complex Cardioid, CReLU and Cart Leaky ReLU. We kept the trace length at $L = 288$ and the stride of sliding window as $s = \{288, 576\}$.

Finding A.1: As shown in Table I, we find that *CReLU* outperforms the other two activation functions in both same-location scenarios and cross-location scenarios. As a result, we will focus on CReLU in the remaining experiments.

Finding A.2: CVNN can achieve higher accuracy in the same-location scenarios but lower accuracy in the cross-location scenarios (as shown in Table I). When we measure the performance with device rank in Fig. 5 (a) and (b), our observation shows that (1) the higher accuracy from the same-location scenario can distinguish the transmitters correctly; (2) the lower accuracy from the cross-location scenario fails to distinguish the transmitters correctly.

Experiment B: Performance of CVNN with Pseudo-Labeling (cross-location). In this experiment, we evaluate the performance of our design with I/Q_Loc1 as the source dataset and I/Q_Loc2 as the target dataset. We explore multiple values during fine-tuning of $N = \{100, 200, 400, 800\}$ for unlabeled traces per transmitter in the target data. For fine-tuning, once we obtain the pseudo-labels of unlabeled target data, we mix those with 20% of the source data i.e. (the test source data) to fine-tune the last three layers of a CVNN.

Finding B.1: From Table II, we observe that the fine-tuning with pseudo-labeling improve the accuracy of cross-location scenarios significantly compared to baseline. For instance, the target data can achieve 62.8% accuracy. Due to these substantial increases on accuracy, fine-tuning with pseudo-labeling is able to distinguish transmitters correctly as we measure device rank as shown in Fig. 6.

Finding B.2: When we increase N, the number of I/Q traces per transmitter during fine-tuning, the accuracy increases. This is expected as a greater number of unlabeled I/Q traces from the target dataset can improve the effectiveness of fine-tuning.

Finding B.3: We run fine-tuning with unlabeled target data and testing with target data multiple times by choosing a different but fixed threshold value each time. We record the accuracy and training time, and fix $N = 800$ each time. As shown in Fig. 7, we observe choosing $\theta = 0.7$ derives a higher accuracy than the ones from other threshold values we examined.

Experiment C: Timing information. In this experiment, we report the time of our designs in training and fine-tuning. ***Finding C.1: Training Time***: We initially trained the networks using 7,200 I/Q traces with 288 samples/trace (which is equivalent to I/Q samples collected over 1.4 milliseconds.). During the fine-tuning phase, we utilized I/Q data collected between 20μs (200 I/Q traces with 288 samples/trace) and 160μs (800 I/Q traces with 288 samples/trace). In Table III, we show that fine-tuning with pseudo-labeling improves the cross-location performance.

IV. Conclusion

We propose several enhancements to the state-of-the-art radio fingerprinting. Firstly, we improve the model's accuracy by treating the entire complex structure of the I/Q data as a single input to a CVNN. Secondly, we leverage and utilize pseudo-labeling to tackle the domain shifts between training and test data due to the variations introduced by the wireless environment involving different locations. Our work provided concrete evidence that CVNN can outperform CNN in radio fingerprinting.

Acknowledgments

We thank the anonymous reviewers for their insightful comments. This work was supported by NSF Grants CNS-2225160

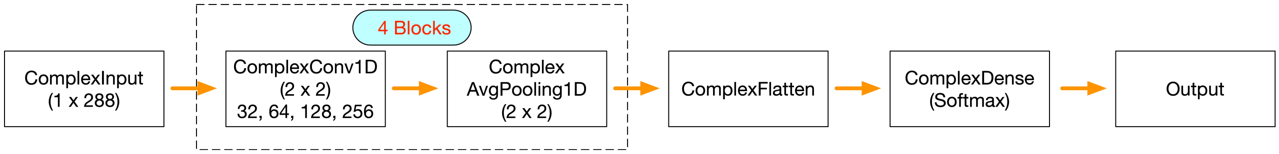


Fig. 4: Architecture of CVNN.

TABLE I: Accuracy of Baseline CVNN (trace length $L = 288$)

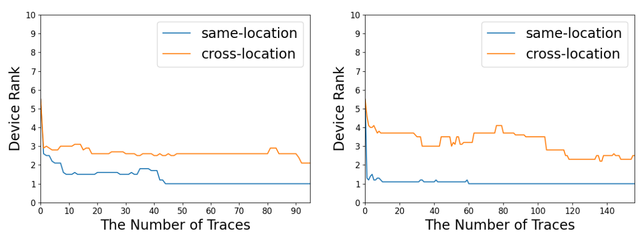
Stride s	Train: I/Q_Loc1; Test: I/Q_Loc1			Train: I/Q_Loc1, Test: I/Q_Loc2		
	Complex Cardioid	CReLU	Cart Leaky ReLU	Complex Cardioid	CReLU	Cart Leaky ReLU
288	68.28 ± 0.13	**78.45 ± 0.36**	72.30 ± 3.01	30.56 ± 2.33	**40.01 ± 1.25**	31.23 ± 2.70
576	65.33 ± 2.11	**71.23 ± 0.55**	68.55 ± 0.12	37.97 ± 1.90	**42.55 ± 6.37**	38.31 ± 2.08

TABLE II: Accuracy of Our Method (trace length $L = 288$, stride $s = 288$, train with 7,200 I/Q traces, fine-tune with $10 \times N$ pseudo-labeled traces, test with 2,000 traces)

Source	Target	Baseline (No pseudo-labeling)	N=100	N=200	N=400	N=800
I/Q_Loc1	I/Q_Loc2	40.01	51.20 ± 1.18	55.15 ± 0.56	58.23 ± 2.23	62.81 ± 0.20
I/Q_Loc3	I/Q_Loc4	38.35	51.44 ± 0.12	52.51 ± 7.54	56.31 ± 2.14	59.10 ± 3.48

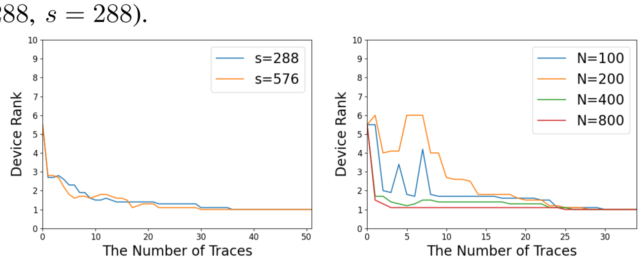
TABLE III: Performance Comparison: without and with Pseudo-Labeling (trace length $L = 288$, stride $s = 288$, train: I/Q_Loc1, test: I/Q_Loc2.)

Traces Per Device	CVNN Without Pseudo-Labeling (CReLU)		CVNN with Pseudo-labeling (θ=0.7)	
	Accuracy	Training Time (sec)	Accuracy	Fine-Tuning Time (sec)
1000	40.01 ± 1.25	120	62.81 ± 0.20	187
2000	48.05 ± 3.01	248	62.25 ± 0.39	295
4000	51.74 ± 2.12	457	61.21 ± 0.81	346
8000	51.52 ± 0.01	893	61.84 ± 0.30	535



(a) Train: I/Q_Loc1, test: I/Q_Loc2 (b) Train: I/Q_Loc3, test: I/Q_Loc4

Fig. 5: Average device rank of Baseline CVNN (CReLU, $L = 288$, $s = 288$).



(a) Train: I/Q_Loc1, test: I/Q_Loc2 (b) Train: I/Q_Loc1, test: I/Q_Loc2

Fig. 6: Average device rank of SARP (CReLU, $L = 288$).

and CNS-2225161 and the Nebraska Public Power District (NPPD) through NCESR at the University of Nebraska-Lincoln (cycle 17).

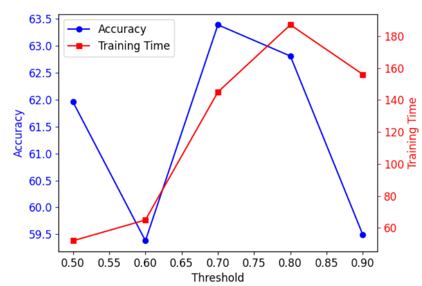


Fig. 7: Accuracy and Training Time (sec) for various threshold (θ) values used in pseudo-labeling ($L = 288$, $s = 288$).

References

[1] H. Li, K. Gupta, C. Wang, N. Ghose, and B. Wang, "RadioNet: Robust deep-learning based radio fingerprinting," in *Proc. of IEEE CNS*. IEEE, 2022, pp. 190–198.

[2] A. Al-Shawabka, F. Restuccia, S. D'Oro, T. Jian, B. C. Rendon, N. Soltani, J. Dy, S. Ioannidis, K. Chowdhury, and T. Melodia, "Exposing the fingerprint: Dissecting the impact of the wireless channel on radio fingerprinting," in *Proc. of INFOCOM*. IEEE, 2020, pp. 646–655.

[3] J. Chen, W.-K. Wong, B. Hamdaoui, A. Elmaghbub, K. Sivanesan, R. Dorrance, and L. L. Yang, "An analysis of complex-valued CNNs for RF data-driven wireless device classification," in *Proc. of IEEE ICC*. IEEE, 2022, pp. 4318–4323.

[4] S. Gopalakrishnan, M. Cekic, and U. Madhow, "Robust wireless fingerprinting via complex-valued neural networks," in *Proc. of IEEE GLOBECOM*. IEEE, 2019, pp. 1–6.

[5] Z. Liang, M. Tao, J. Xie, X. Yang, and L. Wang, "A radio signal recognition approach based on complex-valued cnn and self-attention mechanism," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 3, pp. 1358–1373, 2022.

[6] J. Stankowicz, J. Robinson, J. M. Carmack, and S. Kuzdeba, "Complex neural networks for radio frequency fingerprinting," in *Proc. of IEEE WNYISPW*. IEEE, 2019, pp. 1–5.

[7] B. Bloessl, "IEEE 802.11 a/g/p transceiver," https://github.com/bastibl/gr-IEEE802-11, [Online; accessed 22-Aug-2022].

[8] J. A. Barrachina, "Negu93/cvnn: Complex-valued neural networks," Nov. 2022. [Online]. Available: https://doi.org/10.5281/zenodo.7303587