

Deep Learning-Based RFI Detection and Mitigation for SMAP Using Convolutional Neural Networks

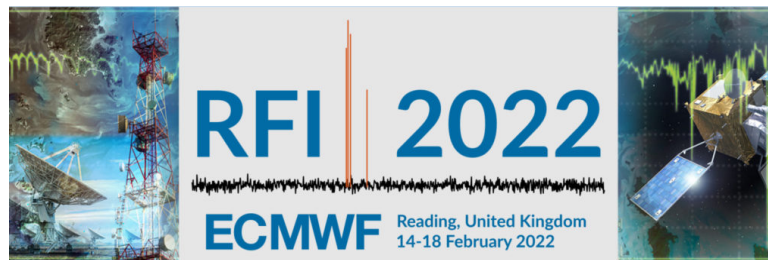
Ahmed Manavi Alam, Ali Gurbuz, Mehmet Kurum

Mississippi State University

Information Processing and Sensing Laboratory (IMPRESS)

Date: 02/15/2022

RFI 2022 Workshop



- Motivation
- Data Products
- Algorithm Flowchart
- Data Statistics
- Data Preparation
- Deep Learning Architecture
- Results
- Discussion and Future Work
- Acknowledgment

- Active Wireless Technologies Expansion
- Jeopardizes Success of a Mission
- Interference Mitigation Techniques
- A Data-Driven Approach

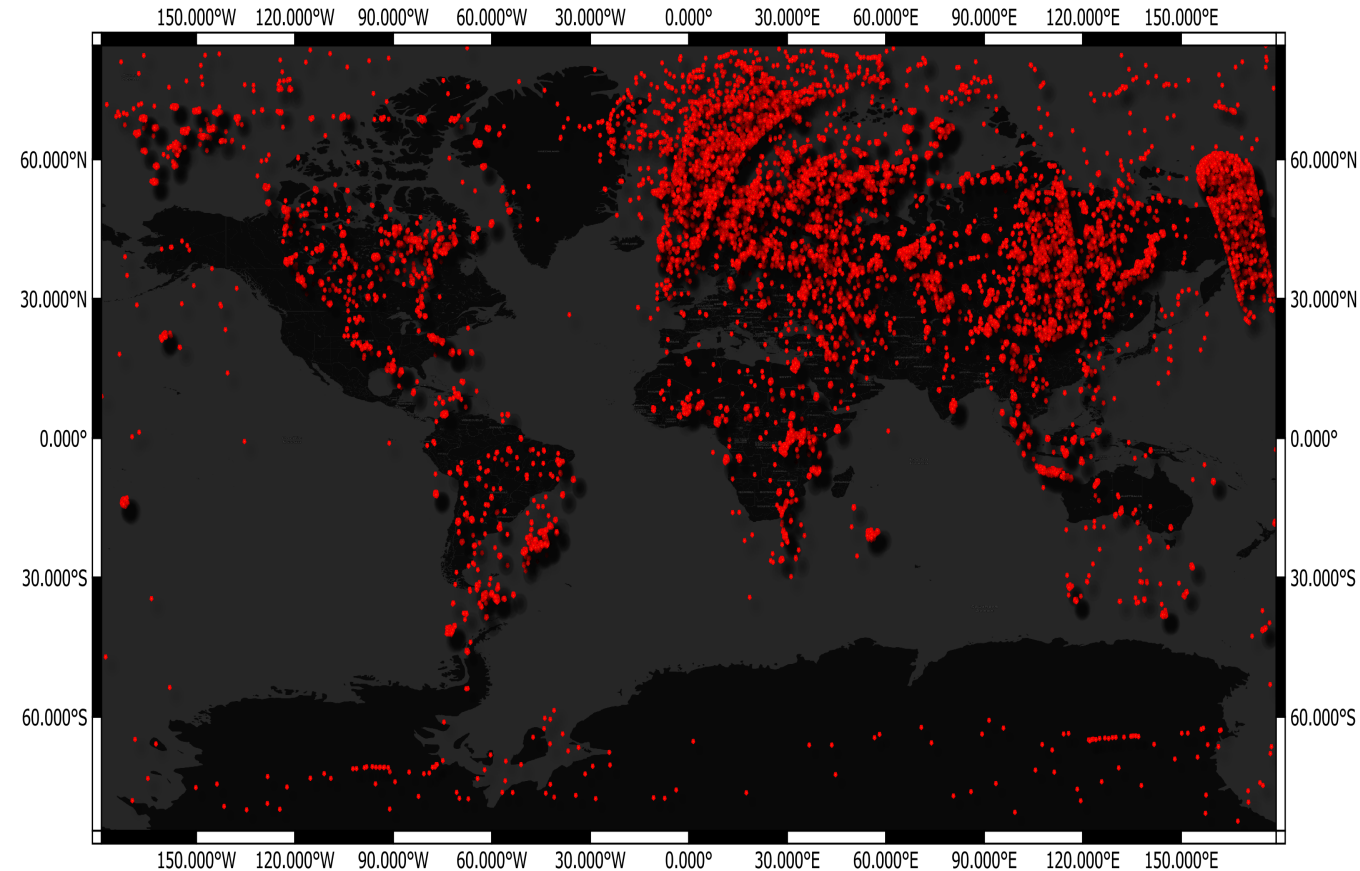


Figure: Global RFI cases from June 1 to June 4, 2017

Level 1A data products – Antenna Counts

<https://nsidc.org/data/SPL1AP/versions/2>

Antenna Moments (779×1931×16×4)

1st Raw Moment - $M_1 = \frac{1}{N} \sum_{i=1}^N X_1$

2nd Raw Moment - $M_3 = \frac{1}{N} \sum_{i=1}^N X_1^3$

3rd Raw Moment - $M_3 = \frac{1}{N} \sum_{i=1}^N X_1^3$

4th Raw Moment - $M_4 = \frac{1}{N} \sum_{i=1}^N X_1^4$

Stokes Parameters (779×1931×16)

3rd Stokes Parameters

4th Stokes Parameters

Level 1B data products – Brightness Temperatures

<https://nsidc.org/data/SPL1BTB/versions/5>

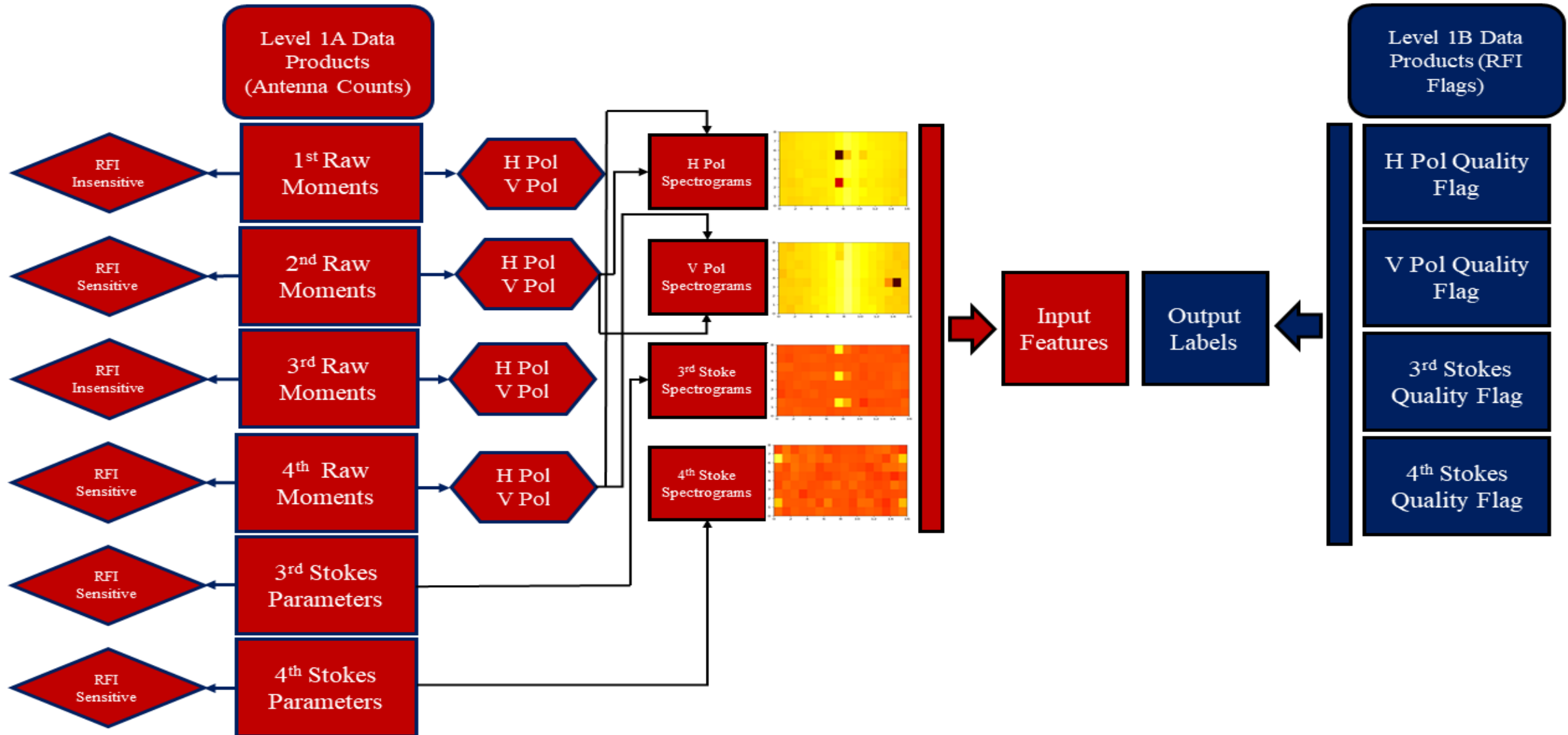
Quality Flags (779×241)

Horizontal Polarization

Vertical Polarization

3rd Stokes Parameters

4th Stokes Parameters



Antenna Counts Domain		Spectrogram Images	Observed Footprints	Observed Antenna Scans
Vertical	2 nd Raw	190,000	1,100,000	170,000
	4 th Raw	190,000	1,100,000	170,000
Horizontal	2 nd Raw	202,000	1,200,000	180,000
	4 th Raw	202,000	1,200,000	180,000
3 rd Stokes		71,000	900,000	150,000
4 th Stokes		66,000	800,000	140,000

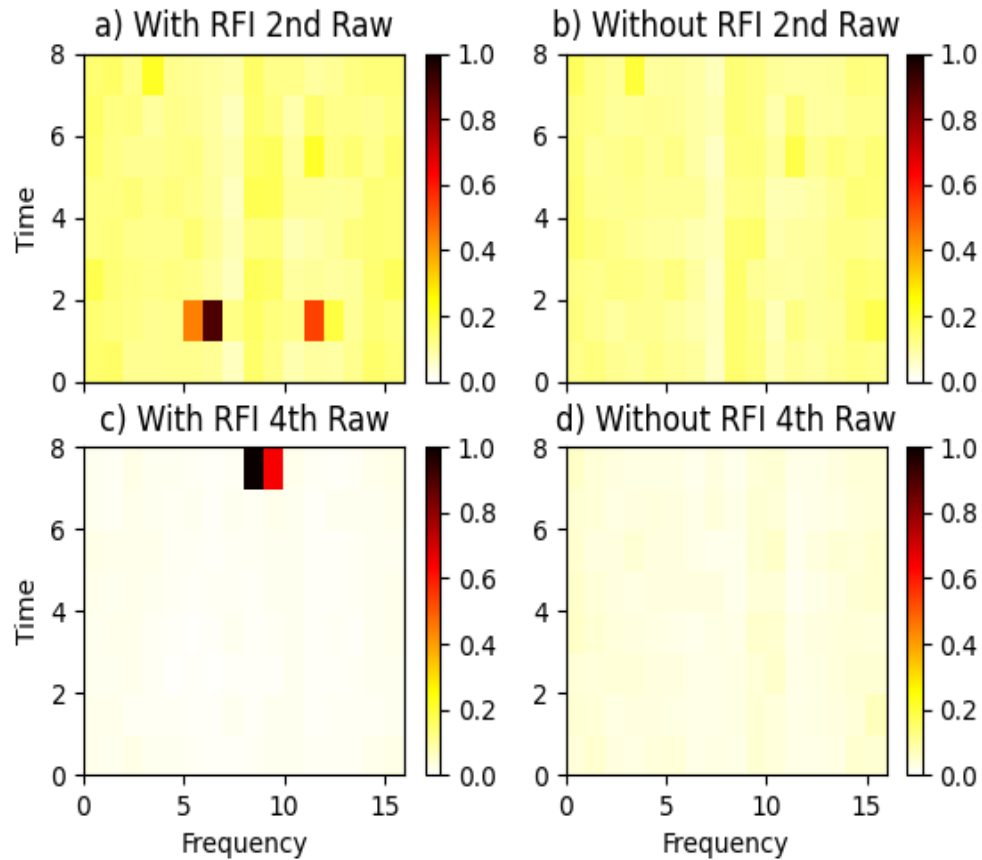


Figure: Spectrogram images from vertical polarization (Sensitive to RFI)

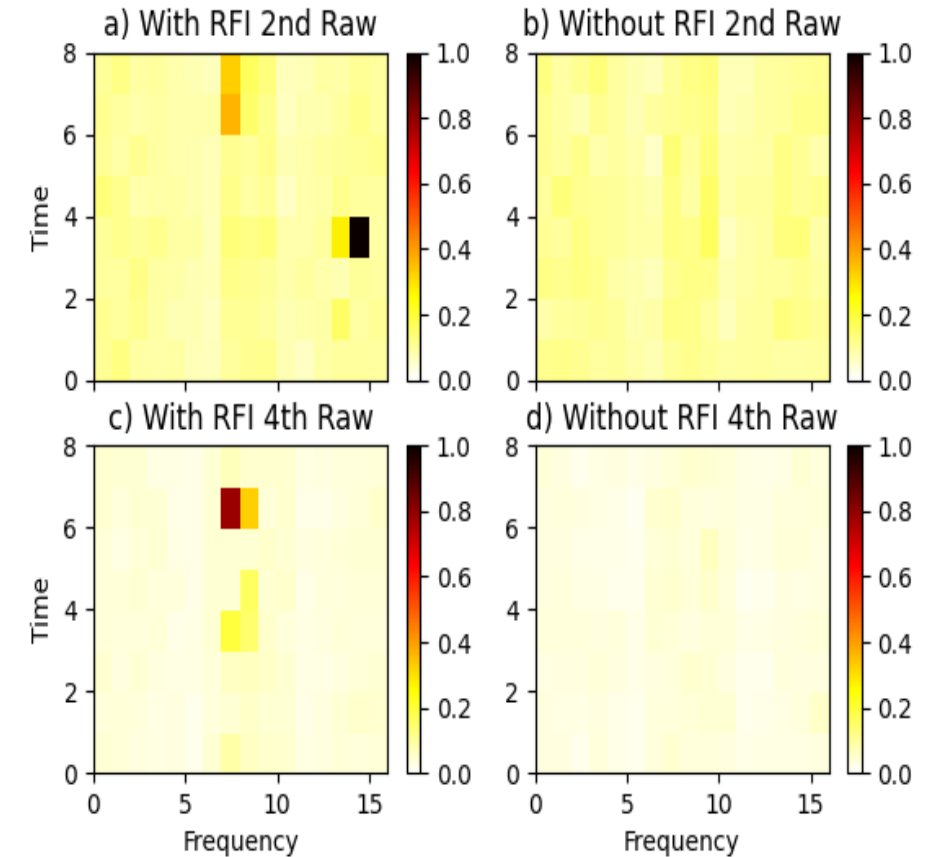


Figure: Spectrogram images from horizontal polarization (Sensitive to RFI)

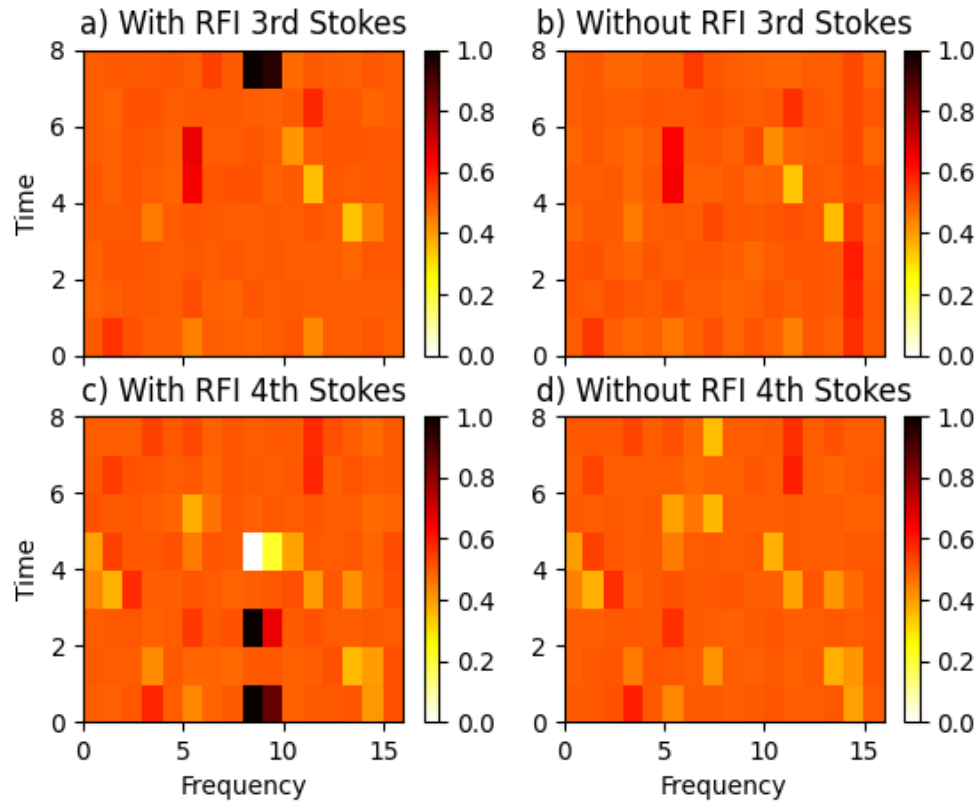


Figure: Spectrogram images from 3rd stokes and 4th stokes parameters (Sensitive to RFI)

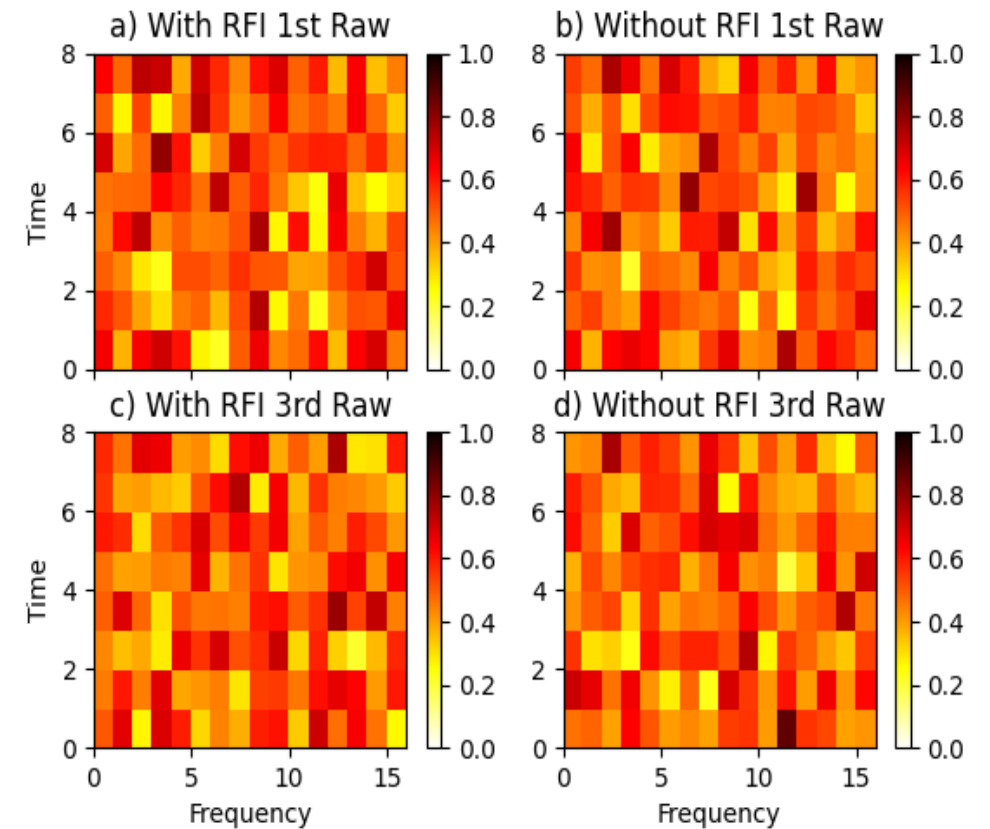


Figure: Spectrogram images from vertical polarization (Insensitive to RFI)

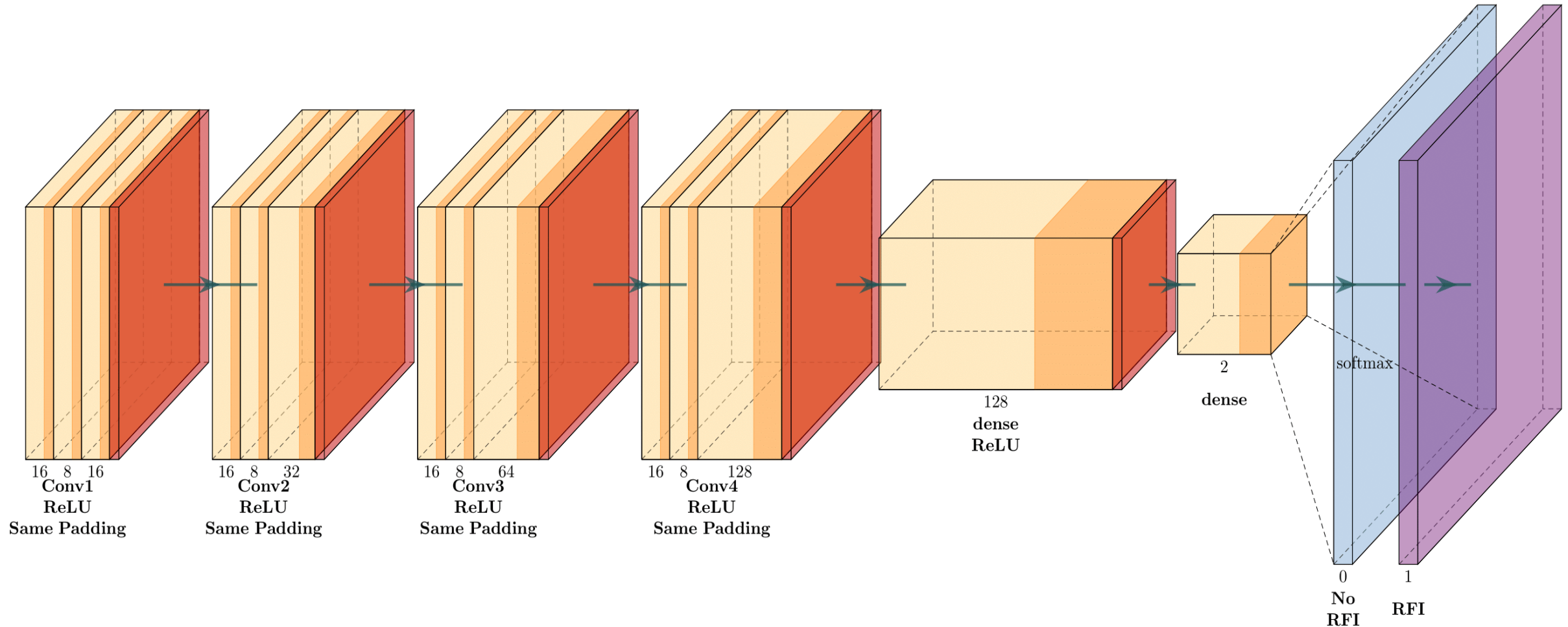


Figure: Deep learning architecture with 4 convolutional layers followed by a fully connected layer of 128 neurons

Hyper-Parameter	Details
Input Size	26×8
No. of Samples	1 million
Normalization	Max-Min
Padding	Same
Activation Function	ReLU
Optimizer	Adam
Loss Function	Binary Cross-Entropy
Kernel Size	2×2
Cross Validation	Train-Test Split & 5-fold
Epochs	60

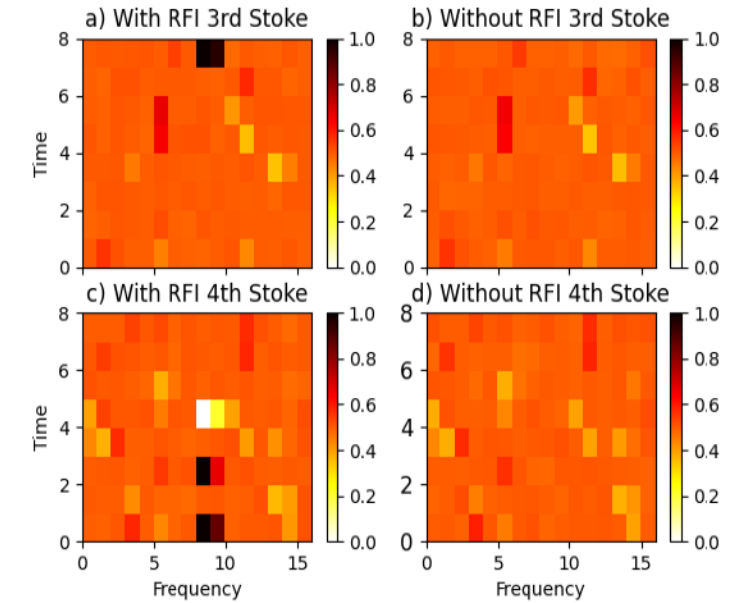
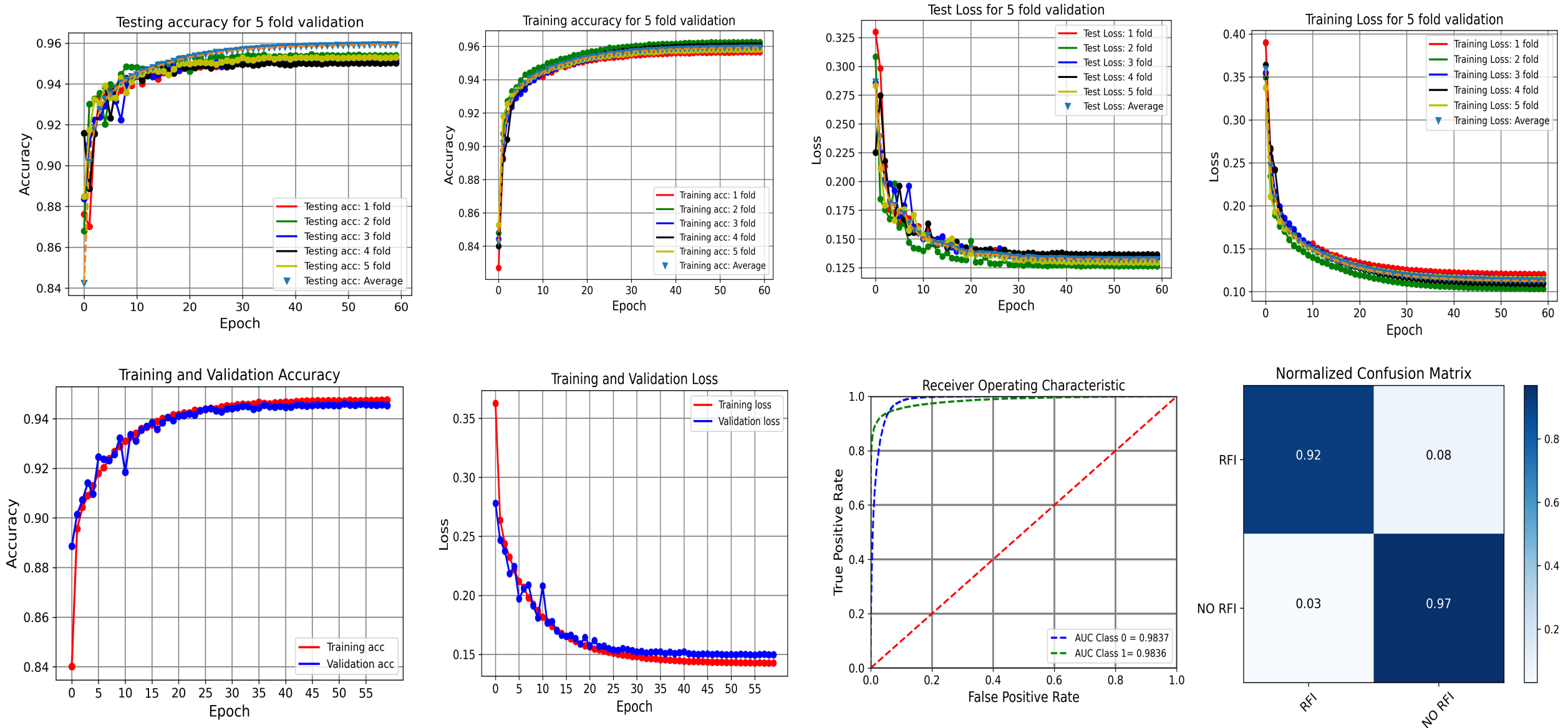
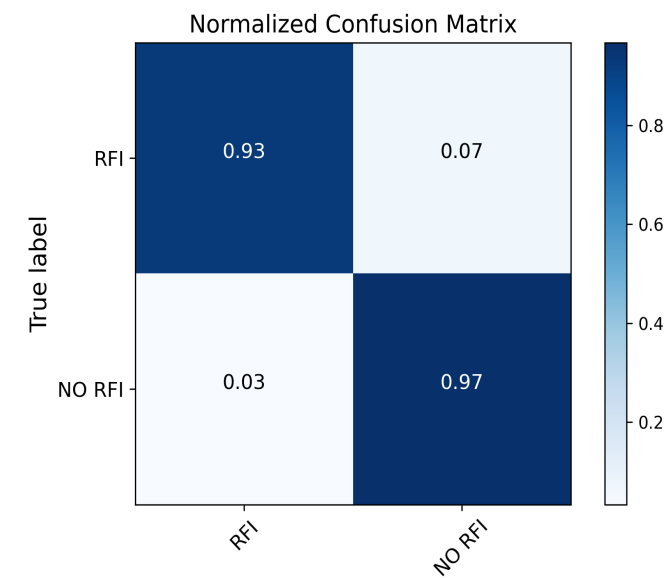
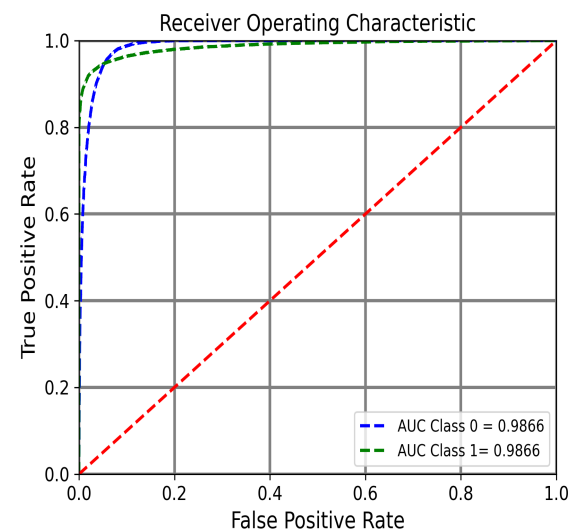
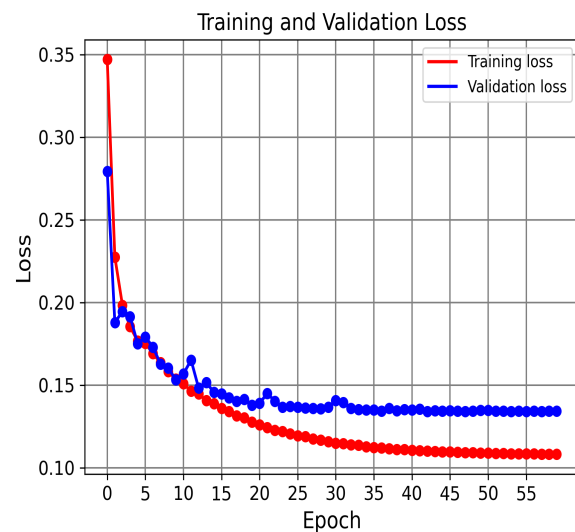
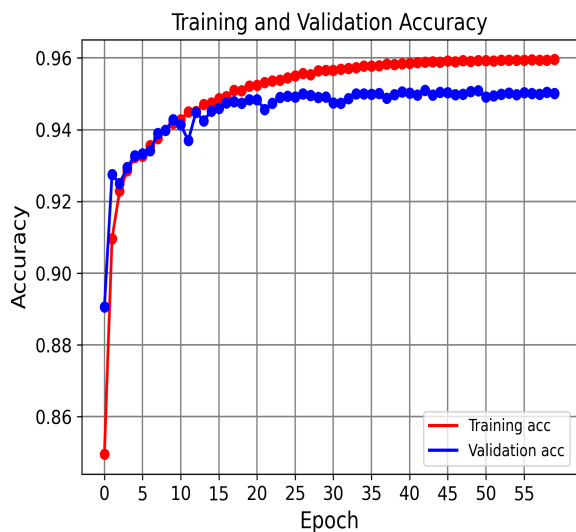
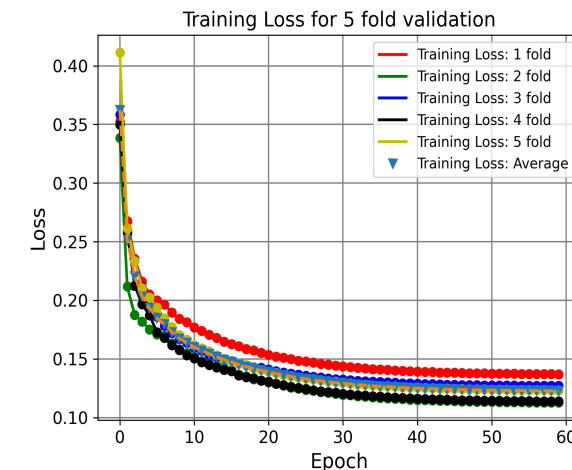
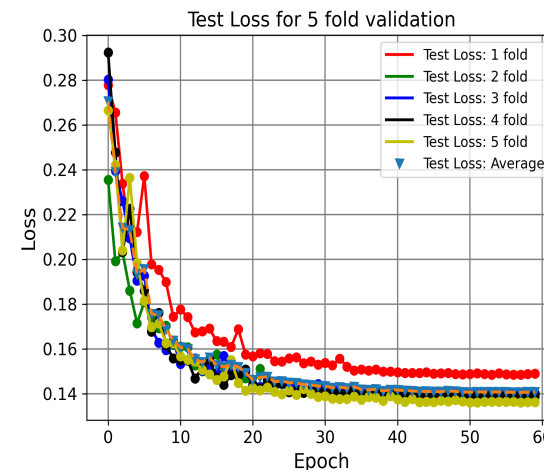
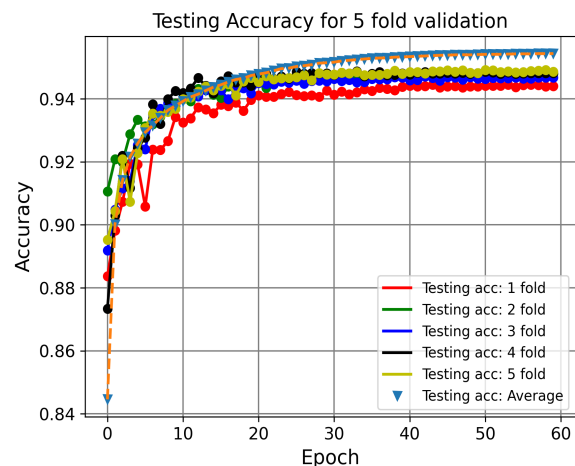
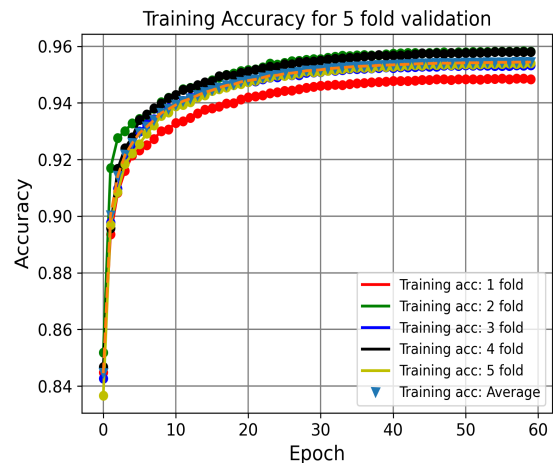
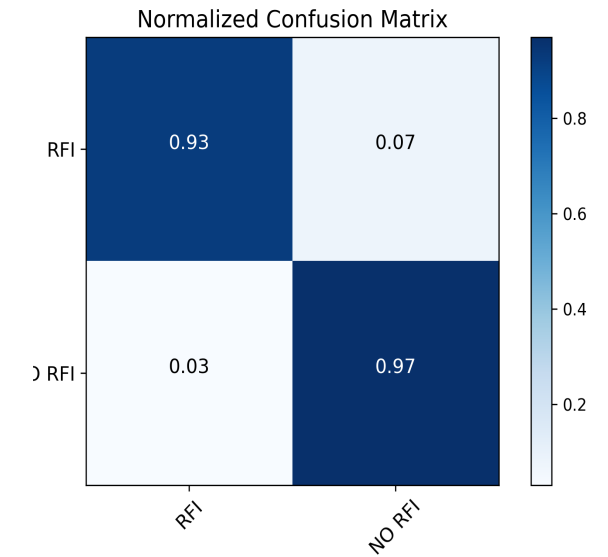
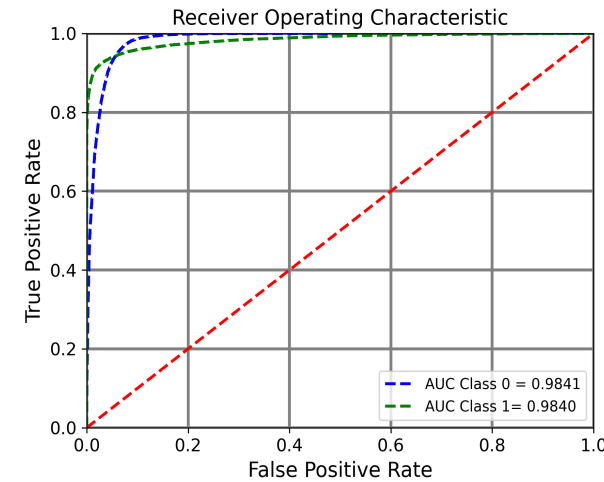
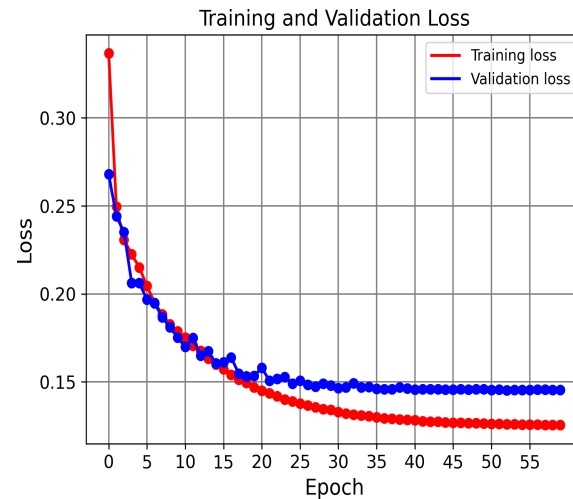
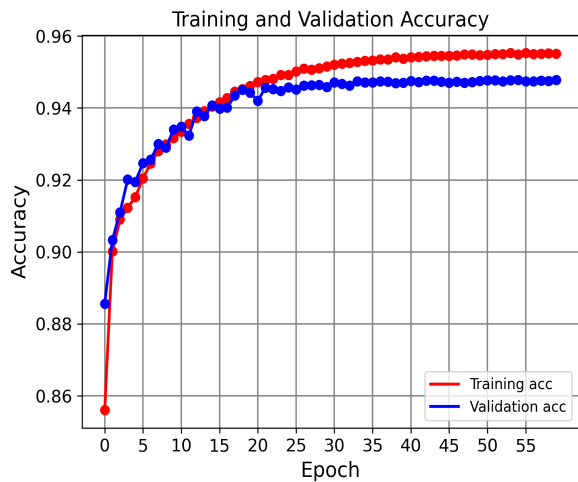
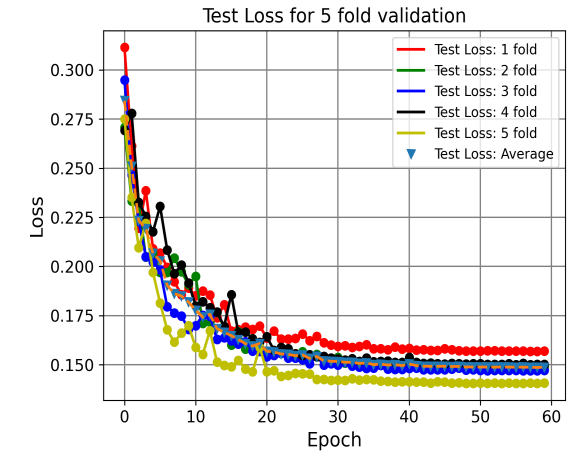
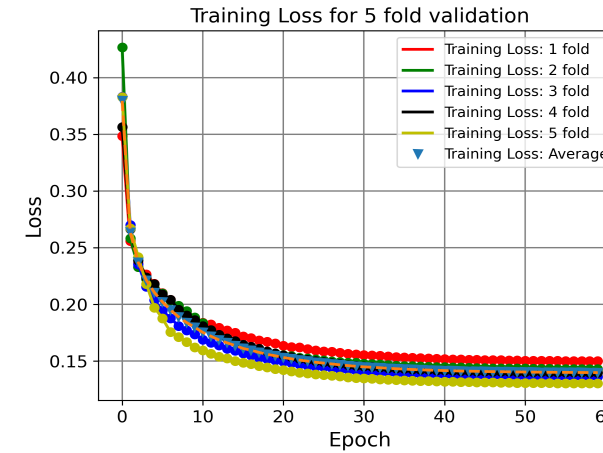
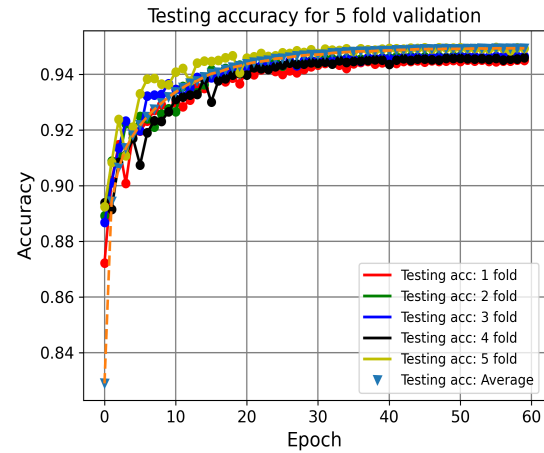
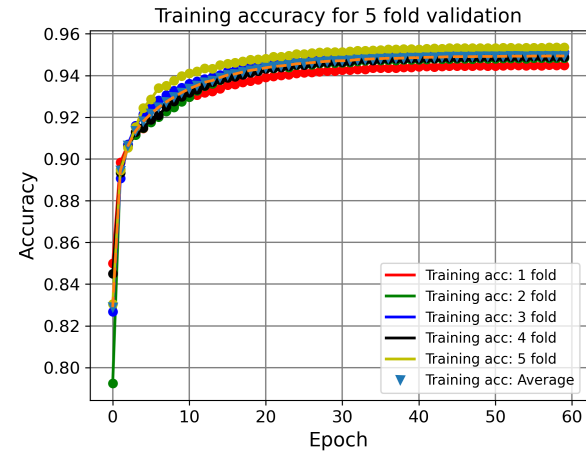


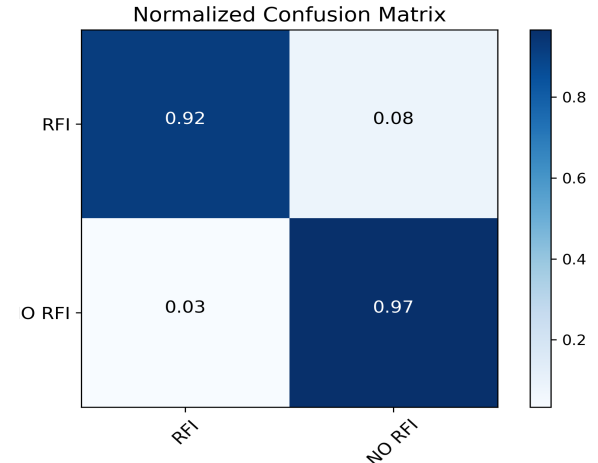
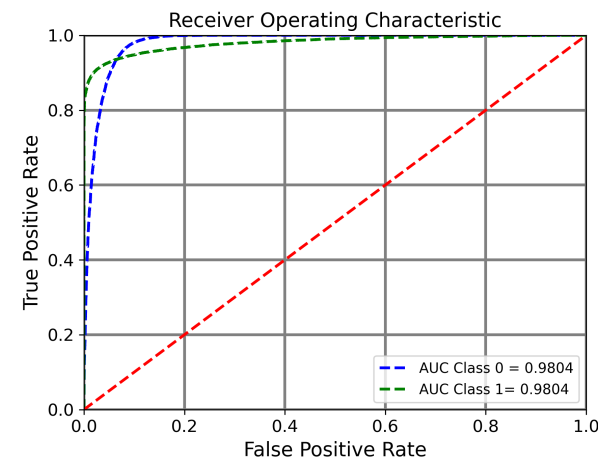
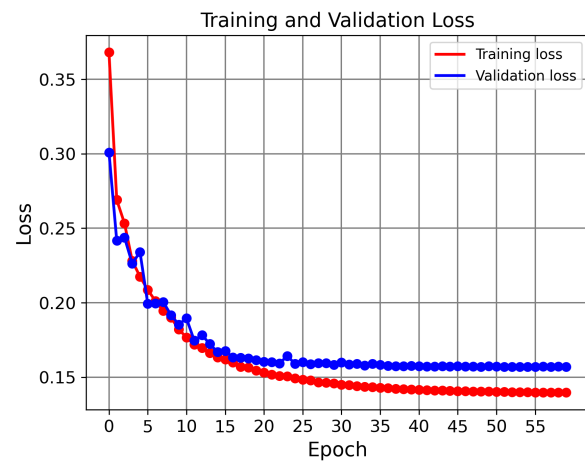
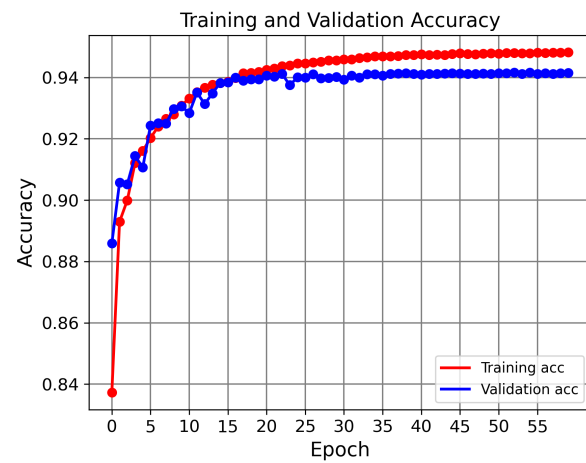
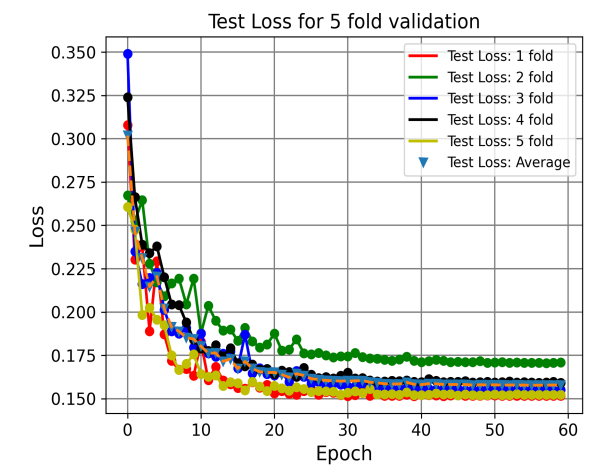
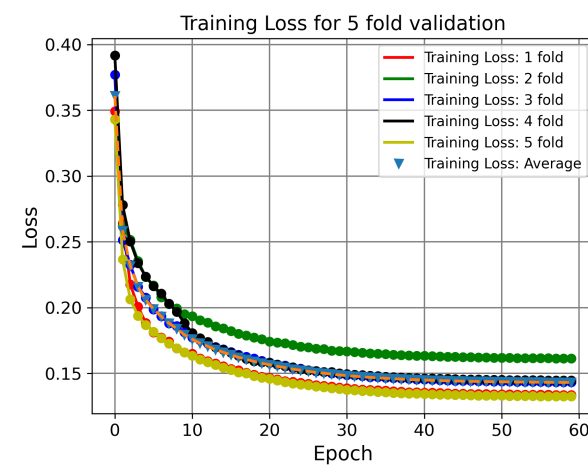
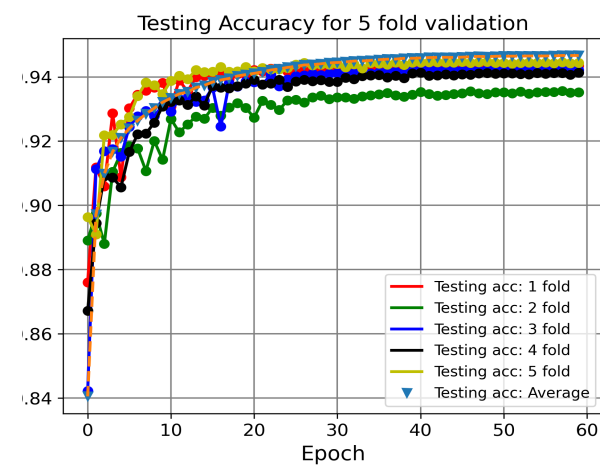
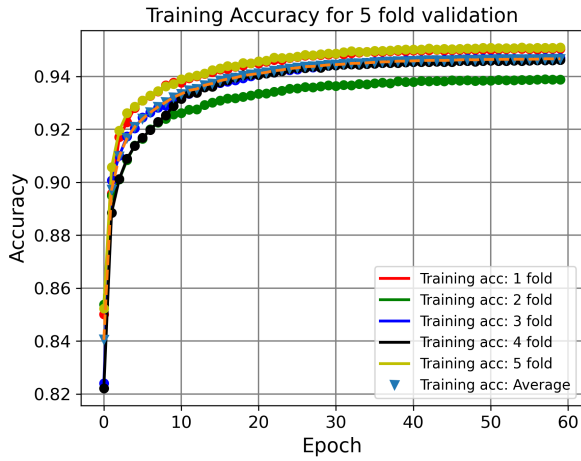
Figure: “Same” padding in convolutional layer helps to extract the features from spectrograms

Antenna Counts Domain		Accuracy	Precision
Vertical	2 nd Raw	95.1%	95.2%
	4 th Raw	94%	93.8%
Horizontal	2 nd Raw	94.3%	94.7%
	4 th Raw	94.6%	94.7%
Thirds Stokes		93.6%	94%
Fourth Stokes		93.5%	93.7%

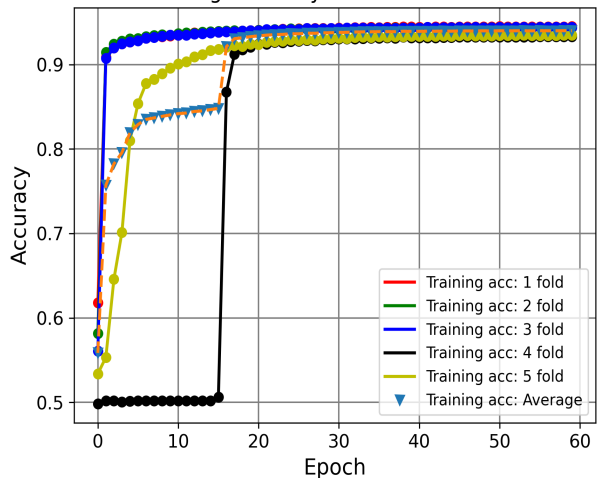




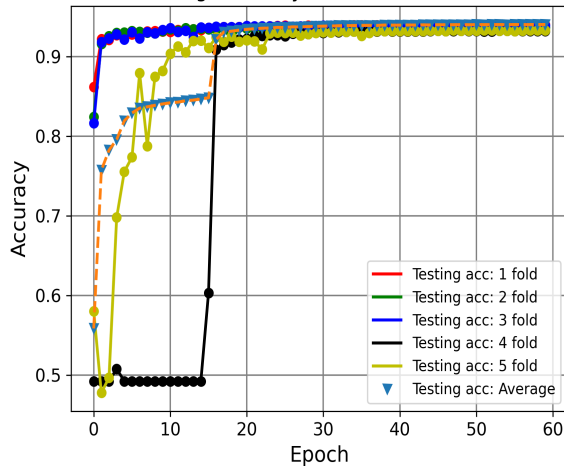




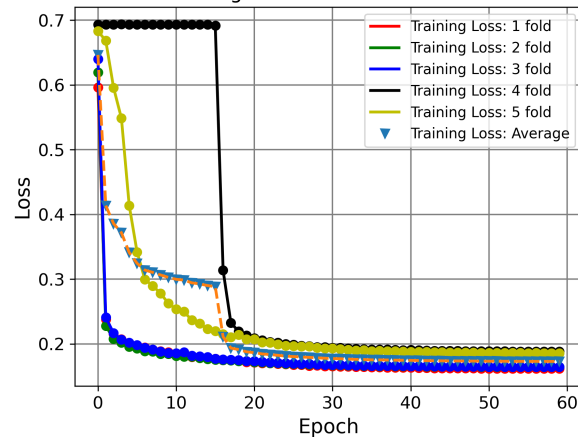
Training Accuracy for 5 fold validation



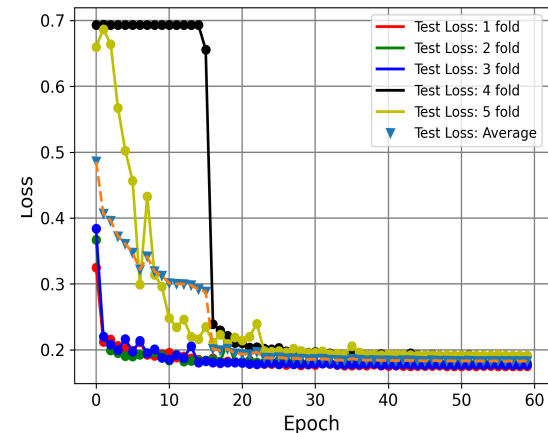
Testing Accuracy for 5 fold validation



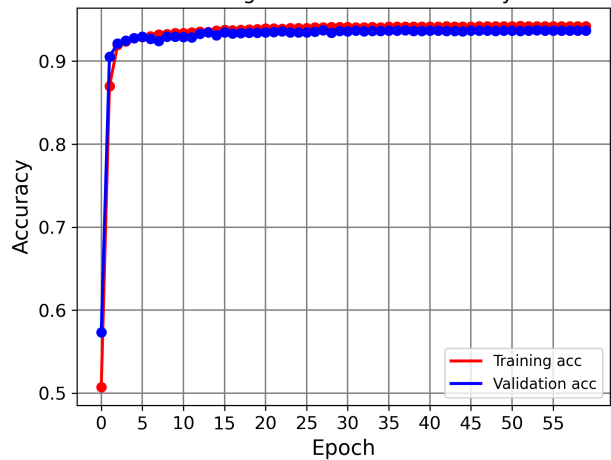
Training Loss for 5 fold validation



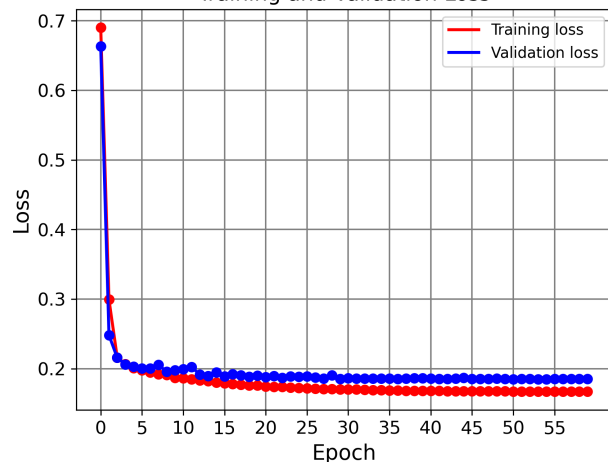
Test Loss for 5 fold validation



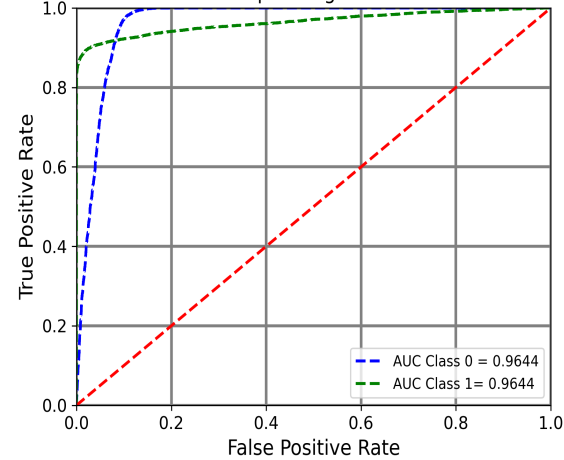
Training and Validation Accuracy



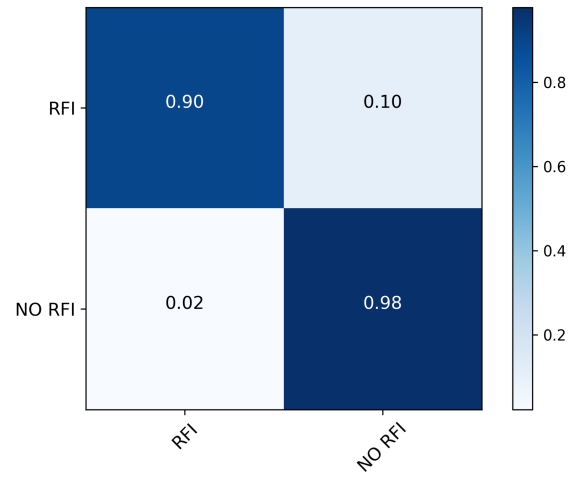
Training and Validation Loss

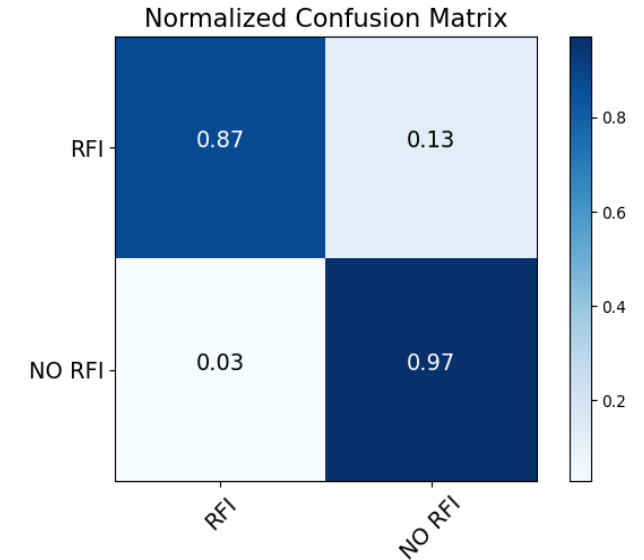
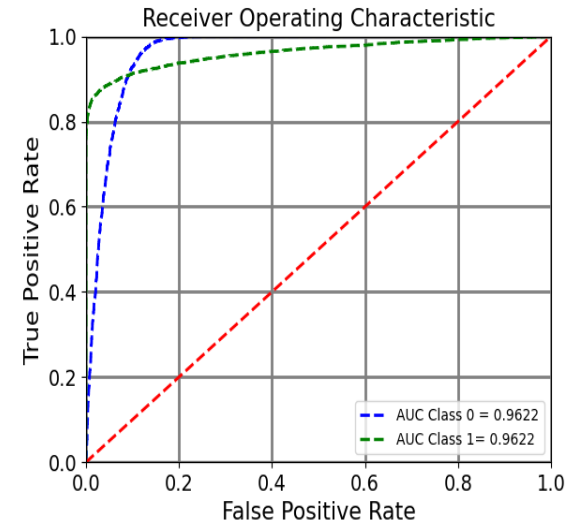
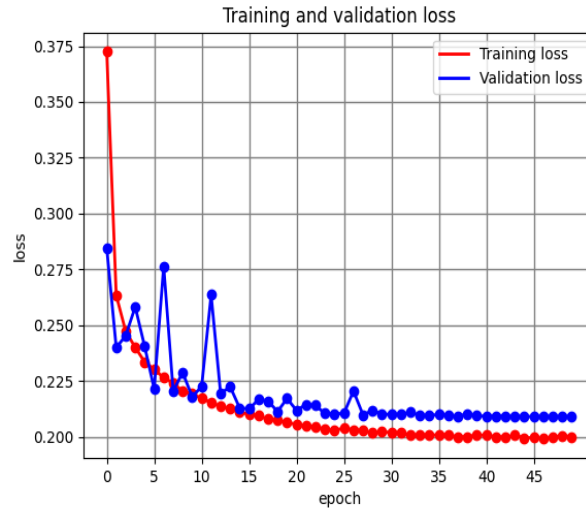
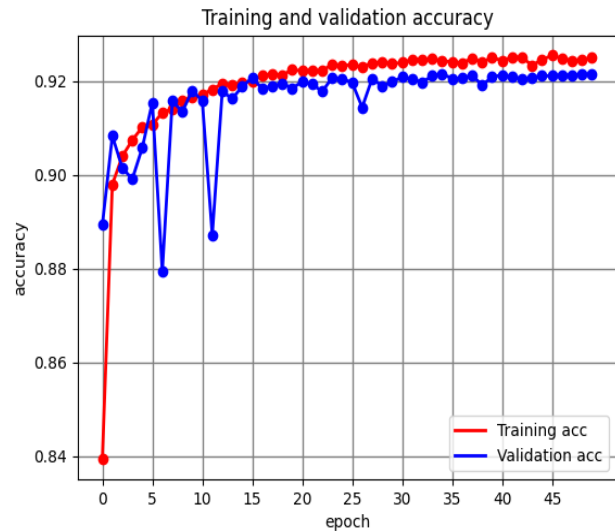
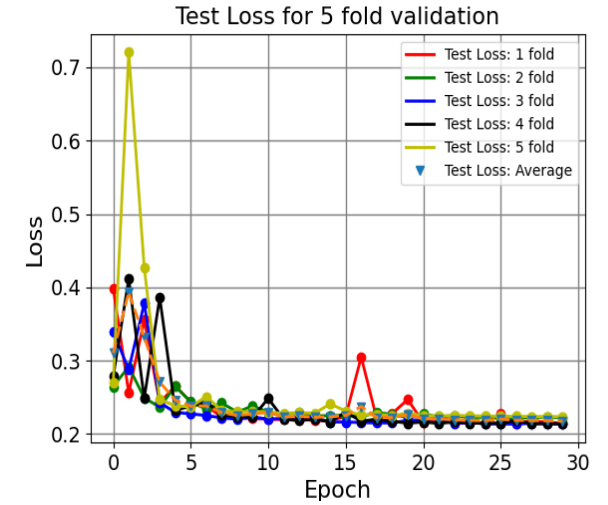
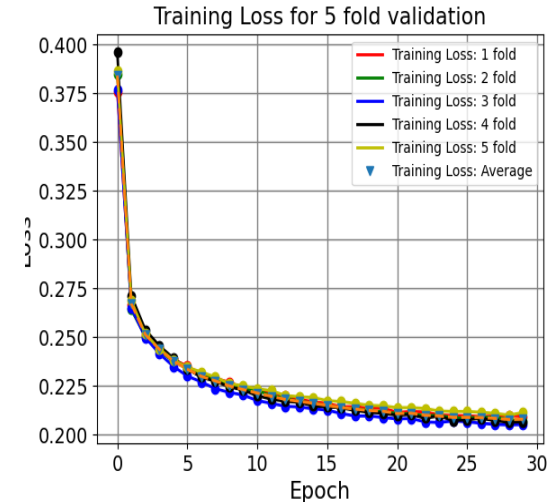
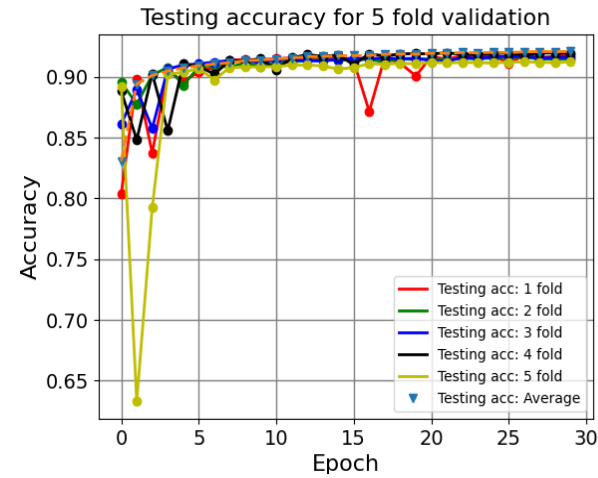
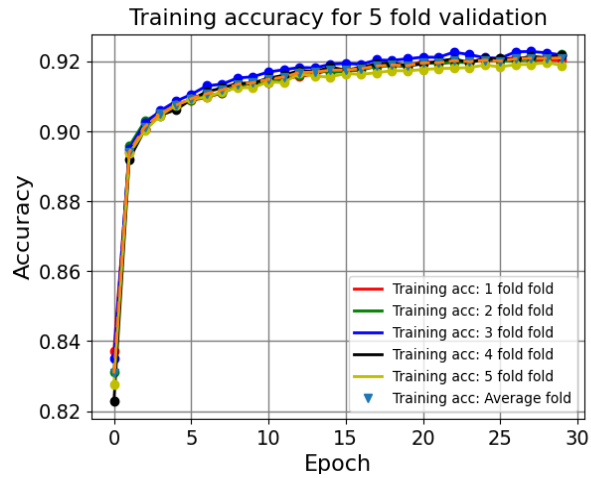


Receiver Operating Characteristic



Normalized Confusion Matrix





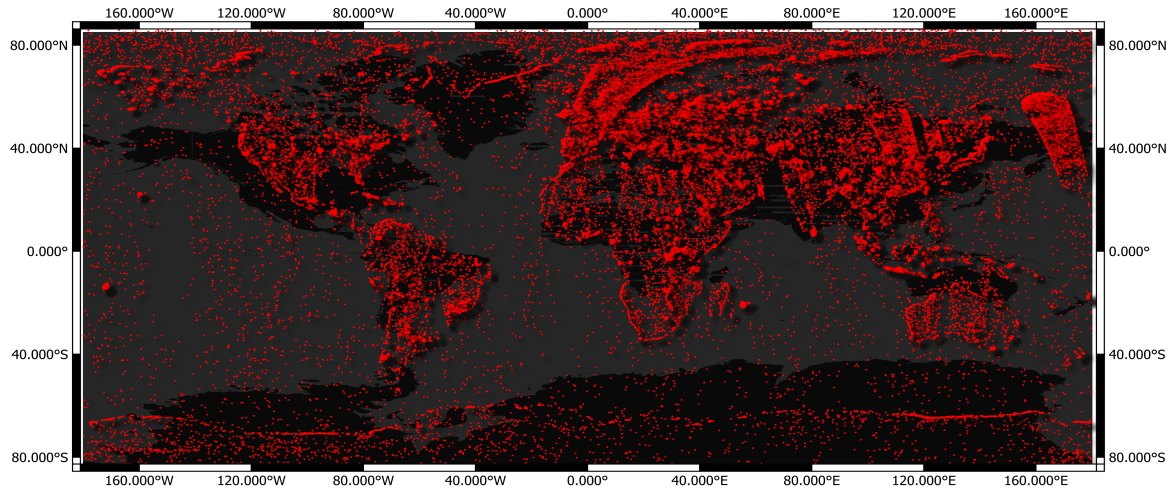


Figure: RFI prediction with deep learning vertical polarization

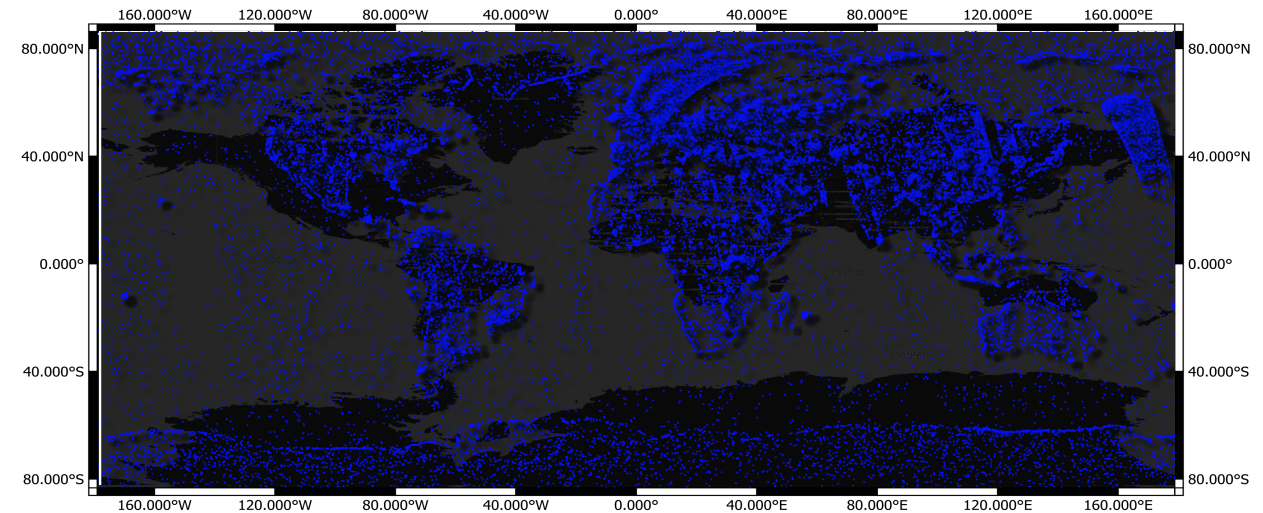


Figure: RFI prediction with SMAP quality flags

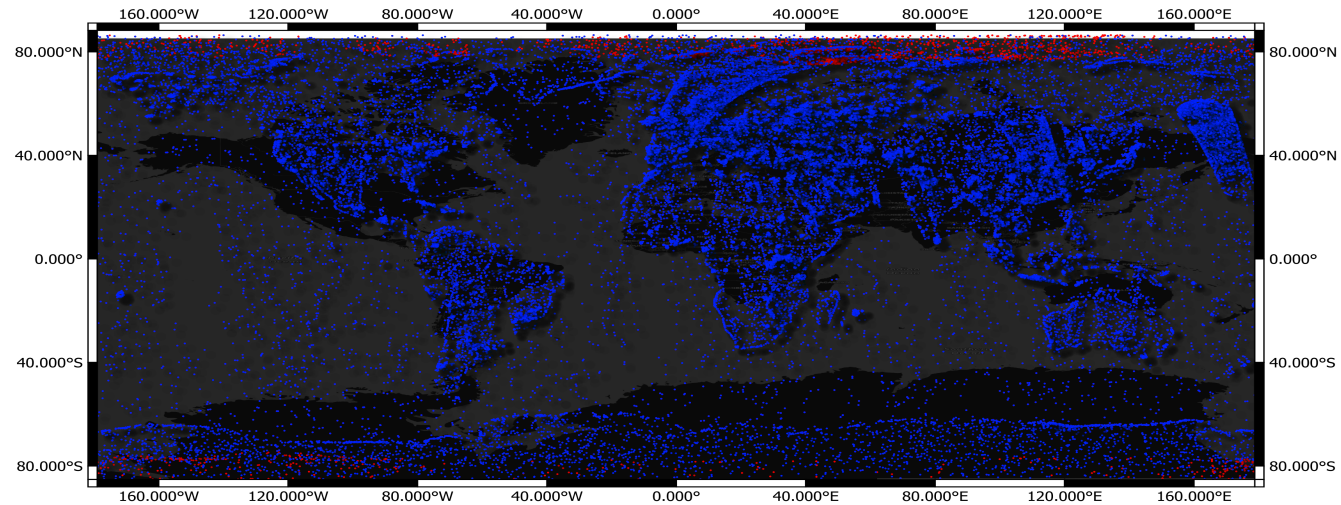


Figure: A comparison between SMAP and deep learning

- *In-situ* RFI contaminated sample collection is very challenging. To label the dataset for this study, SMAP's level 1B data quality flags have been used. So, there is less control over the dataset
- Future work aims to incorporate synthetically generated RFI into SMAP data products to have greater control over the dataset
- This will be helpful to distinguish lower and higher level RFI scenarios as well as single and multiple sources RFI

- This work was supported by National Science Foundation under Grant no. 2030291 and 204771

Thank You