ISG ISRS National Symposium 2023

GeoDiscover: Unravelling India's Spatial Frontier

Exploring Geospatial Ecosystem, Trend, and Innovations

SYMBIOSIS

Complex Valued U-Net for Segmentation of SAR Images

Saumya Vilas Roy, IIST

HariPriya S NRSC, ISRO
Deepak Mishra, IIST

Introduction

- Many real word process cannot be modeled efficiently by real values belonging to R such as SAR images, Signals, biological process belong in the complex domain.
- Processing of such data is mostly done by casting them into real domain and then using RVNN (Real Valued Neural Network).
- But it causes loss of information and inhibits us from finding meaningful relations so using the CVNN (Complex Valued Neural Network) would help us to preserve such information and find a much concrete solution.

Semantic Segmentation using U Net

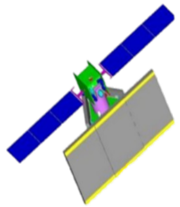
- The process of assigning a class label to each pixel of the image to indicate what is represented by the particular pixel.
- Semantic segmentation can be achieved by U-Net, SegNet, PSPNet.
- The idea is to supplement a network with a downsampling part where blocks are pooled together to reach a smaller size
- Later , upsampling is done where those pooling layers are replaced by unpooling operation to map the pixels to their respective labels.
- U-Net is a fully convolution networkand allows skip connection to propagate the information that might be dropped during the down-sampling layers

Complex activation functions

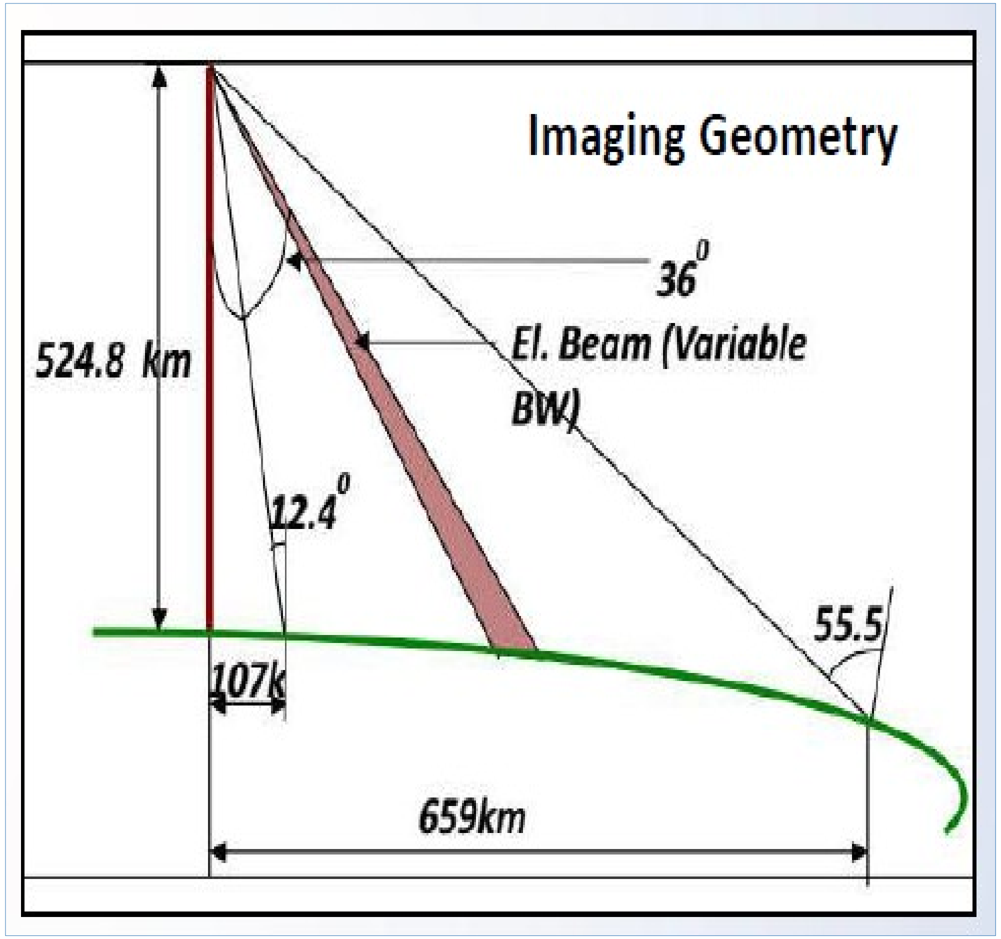
- The activation function in should be piece wise smooth to allow computation of the gradients, to create a complex domain activation function the real domain activation function is mostly extended into complex domain in either of these ways.
- Type-A: $\sigma A (z) = \sigma Re (Re(z) + i\sigma Im (Im(z))$,
- Type-B: $\sigma B (z) = \sigma r (|z|) \exp(i\sigma\phi (arg(z))$,
- Here σA , σB , σr , $\sigma\phi \in R$ and most popular activation functions like sigmoid, hyperbolic tangent (tanh) and Rectified Linear Unit (ReLU) and softmax is extendable this way.

EOS-04- C-Band SAR

EOS-04 (RISAT-1A) is Follow -on Mission of RISAT-1



- Launched on 14-February 2022
- Frequency: C-band (5.4 GHz).
- Imaging Modes: Stripmap, ScanSAR and Sliding- Spotlight -(FRS-1, FRS-2, MRS, CRS and HRS)
- Polarizations: Single, Dual, Compact (CP) & Full (FP)
- Swath Coverage: 10 Km to 223 Km
- Spatial Resolutions: 1m to 50m
- Data availability: 23rd March 2022 onwards

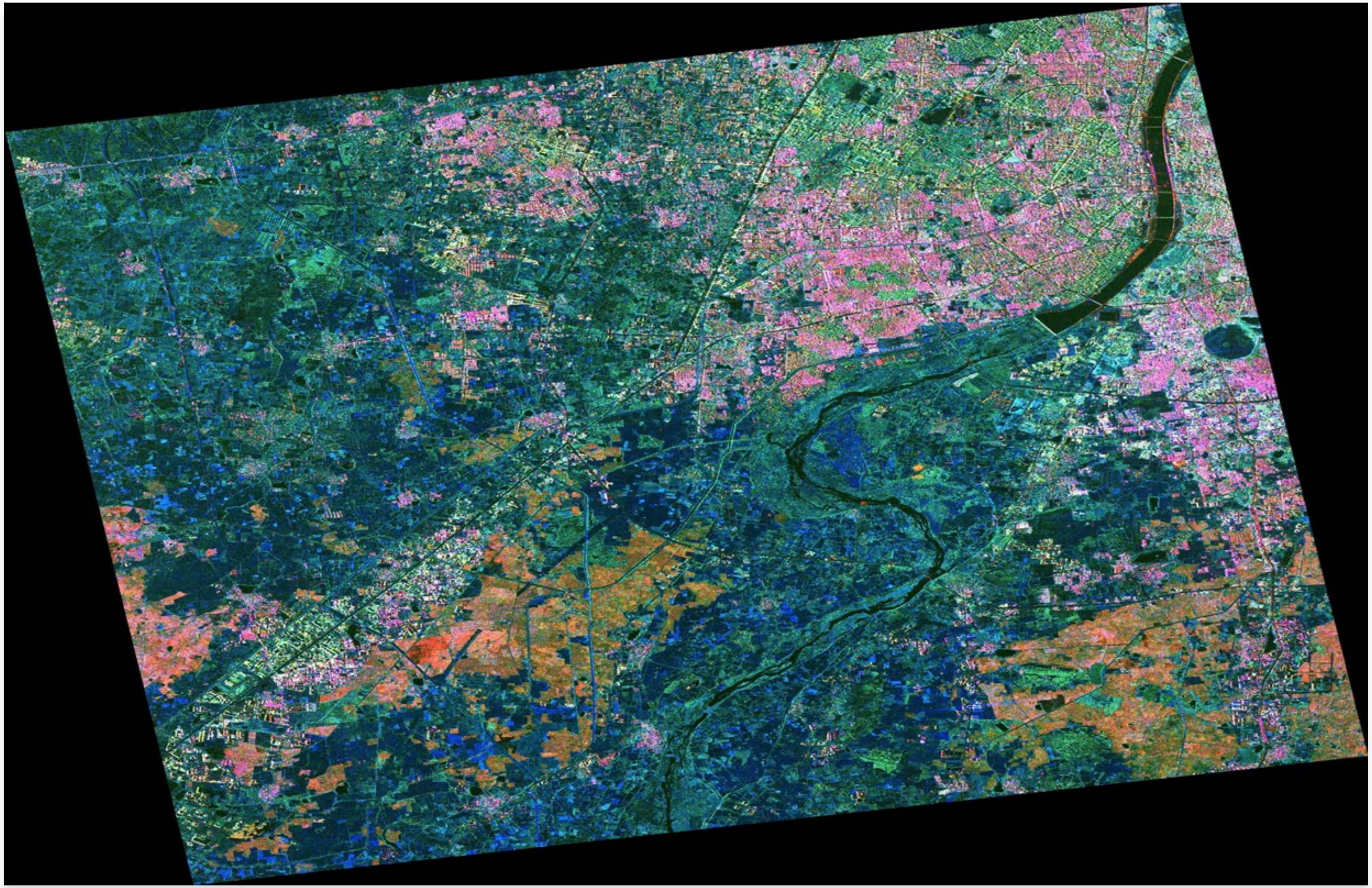


Product Specifications

MODES	FRS-1 (Fine Resolution Stripmap-1)	FRS-2 (Fine Resolution Stripmap-2)	MRS (Medium Resolution ScanSAR)	CRS (Coarse Resolution ScanSAR)
Swath (km)	25 #20	25 #20	160 #115	223 #168
Azimuth Resolution(m)	3	3	33	50
Slant range resolution(m)	2	4	8	8
Ground range resolution(m)	11-2.8	22-5.6	45-22	135-45
Worst Sigma Naught (dB)	≤- 18	≤-19	≤- 18	≤-18
Off-Nadir (km)	100 – 650 #100-400	100-650 #100-400	100 - 650 #100-400	100 - 650 #100-400
Look Angle (deg)	11-49 #11-37	11-49 #11-37	11-49 #11-37	11-49 #11-37
[illegible]	Single, Dual,	Single,	Single, Dual,	Single, Dual,

represents specifications for Full Polarimetric mode.

Parameter	EOS-04
Geo location accuracy (RMSE)	50 m
Radiometric Resolution (SLC)	3.1 dB
PSLR	-17 dB
Relative Radiometric accuracy	1dB
Absolute Radiometric accuracy	±1dB



Date of Imaging : 3rd April 2022, Full Quad Pol, Imaging Mode:FRS-1,Region:Ahmedabad, India

Red- Even Bounce
Green - Volume Scattering
Blue – Odd Bounce

Process Implementation:

- ❖ Full Pol EOS04 data in Single Look Complex Domain is ingested using the SARPOLTool v2.1 .
- ❖ The Scattering Matrix Elements with the real and imaginary components are combined to generate the T matrix- Coherency Matrix.
- ❖ For the first cut analysis ,the Pauli basis is used to assign the scattering type for each resolution cell based on the dominant polarimetric scattering mechanism.

Implementation Technique

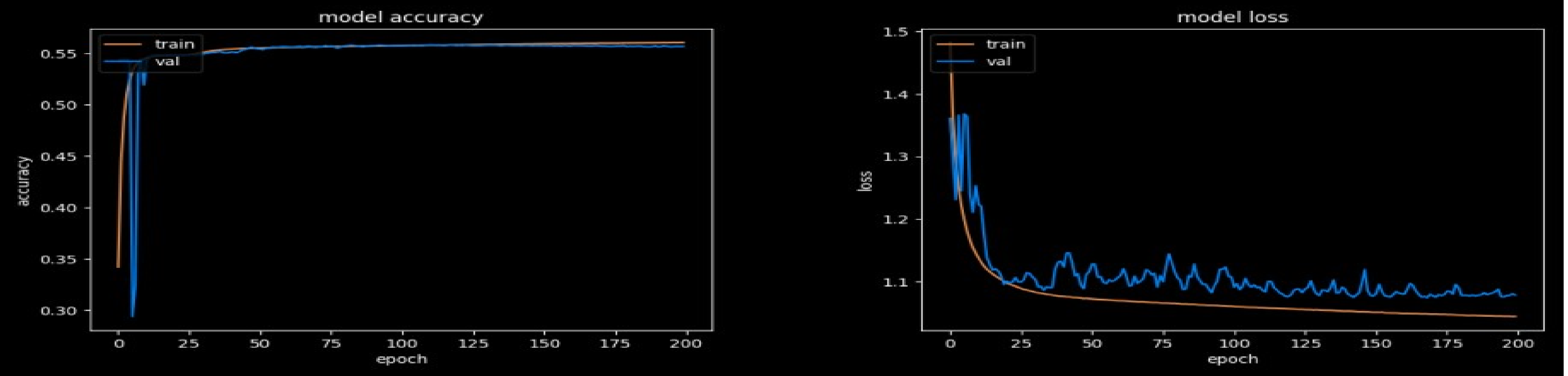
- The U-net architecture follows the conventional structure with encoder and decoder also it had a skip connection from every last conv2D layer in each step of the encoder to the relevant concatenation layer in the decoder to allow passage of information that might have been lost due to the kernels.
- The output filters in each layer or the encoder [32, 64, 128, 256] decoder has the same output filters in just reverse order [256, 128, 64, 32] and is referred to as ϕ CVNN.
- Another encoder [64, 128, 256, 512] and decoder [512, 256, 128, 64] were used as ρ CVNN in order to check for overfitting .Both models were trained with variable dropout layer rates.
- For the labeling of the data to generate the mask, RGB matrix values are used to the assign the labels
- The scattering values present in the matrix are averaged an thresholded as follows for 4 class segmentation:

Implementation Technique

To check the effectiveness of CVNN in segmentation several different methods and optimization were used to find the optimal model parameters and theresholding technique. The following methods were tried:

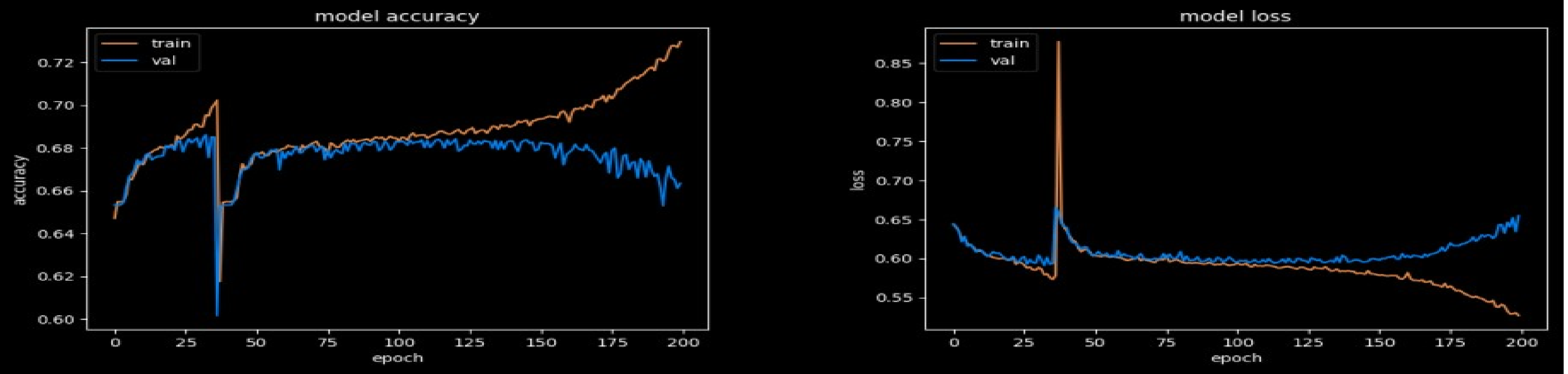
1. ϕ CVNN for segmentation on the 4 class mask.
2. ϕ CVNN for segmentation on the 2 class mask.
3. ϕ CVNN for segmentation with drop out rate 0.1 on the 2 class mask.
4. ϕ CVNN for segmentation with drop out rate 0.2 on the 2 class mask.
5. ϕ CVNN for segmentation with drop out rate 0.5 on the 2 class mask.
6. ρ CVNN for segmentation on the 2 class mask.
7. ρ CVNN for segmentation with drop out rate 0.1 on the 2 cla mask.

Plot Trends-CVNN with no dropouts



a. Model validation and training accuracy b. Model Loss and training accuracy

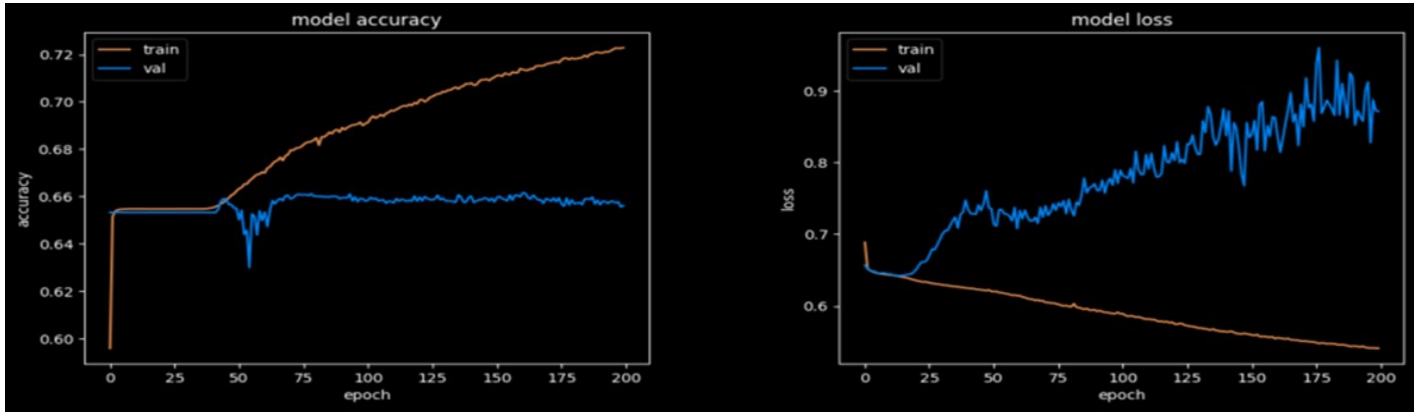
Figure 5.1: ϕ CVNN with no dropouts on 4 class segmentation



a. Model validation and training accuracy b. Model Loss and training accuracy

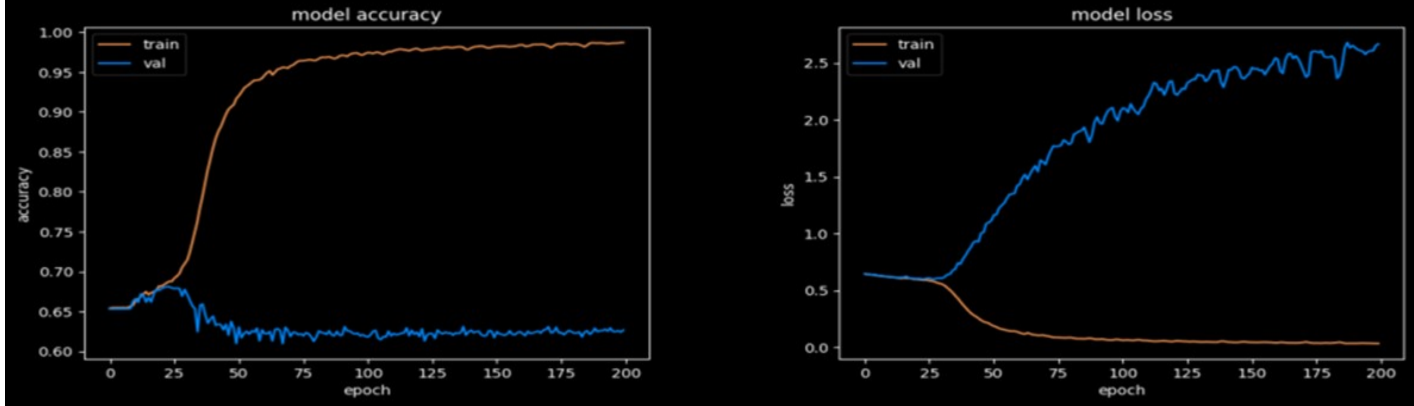
Figure 5.2: ϕ CVNN with no dropouts on 2 class segmentation

Plot Trends-CVNN with dropouts



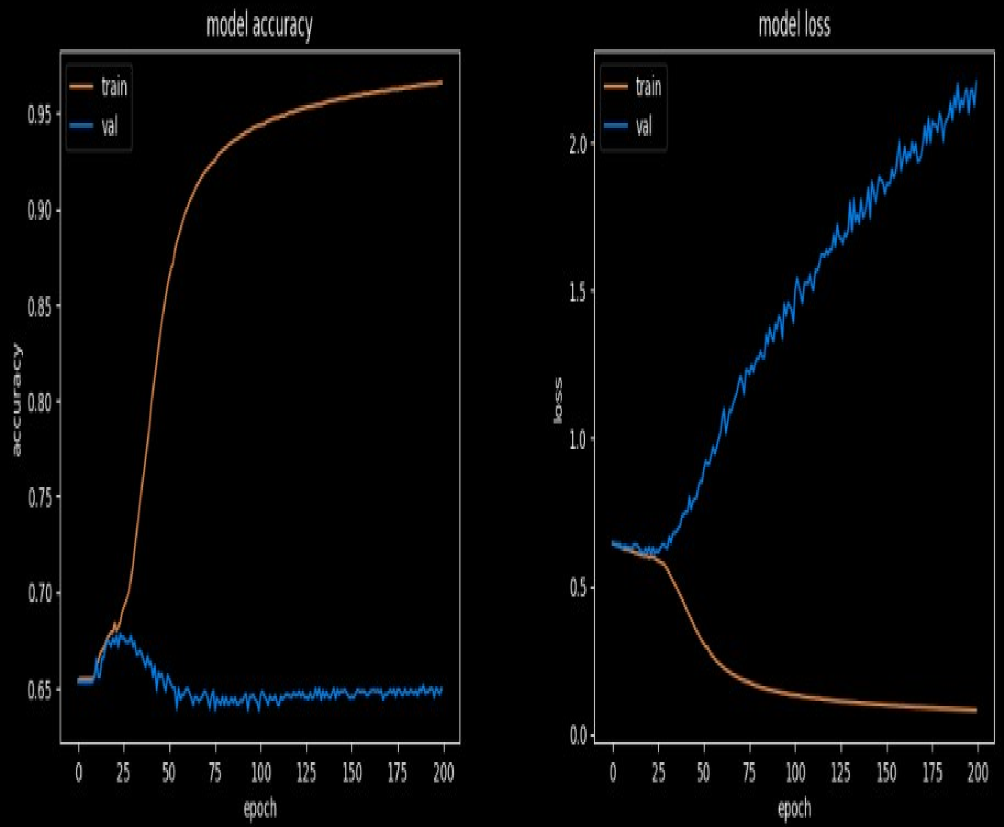
a. Model validation and training accuracy b. Model Loss and training accuracy

Figure 5.5: ϕ CVNN with dropout rate 0.5 on 2 class segmentation



a. Model validation and training accuracy b. Model Loss and training accuracy

Figure 5.6: ρ CVNN with no dropouts on 2 class segmentation



a. Model validation and training accuracy b. Model Loss and training accuracy

Figure 5.7: ρ CVNN with dropout rate 0.1 on 2 class segmentation

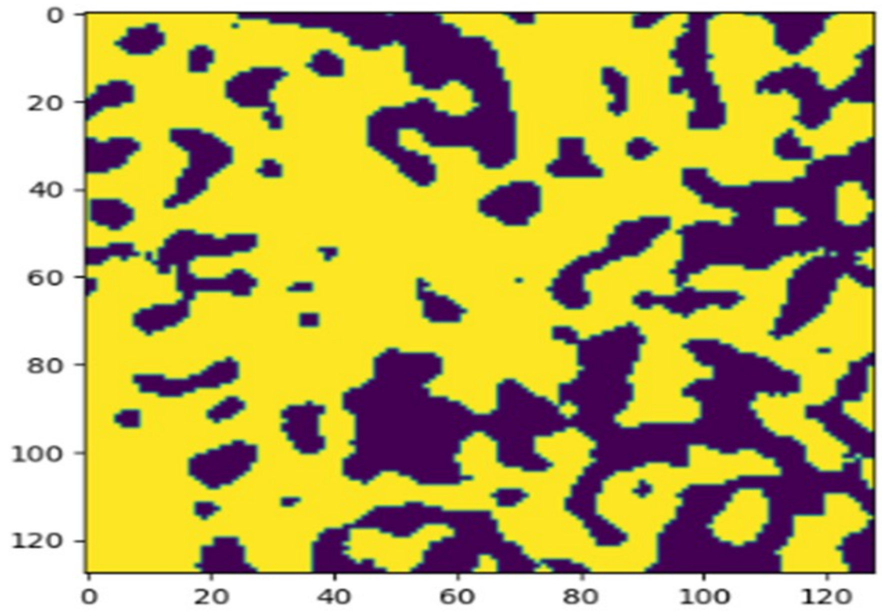
Results

Table 5.1: The testing results of the models

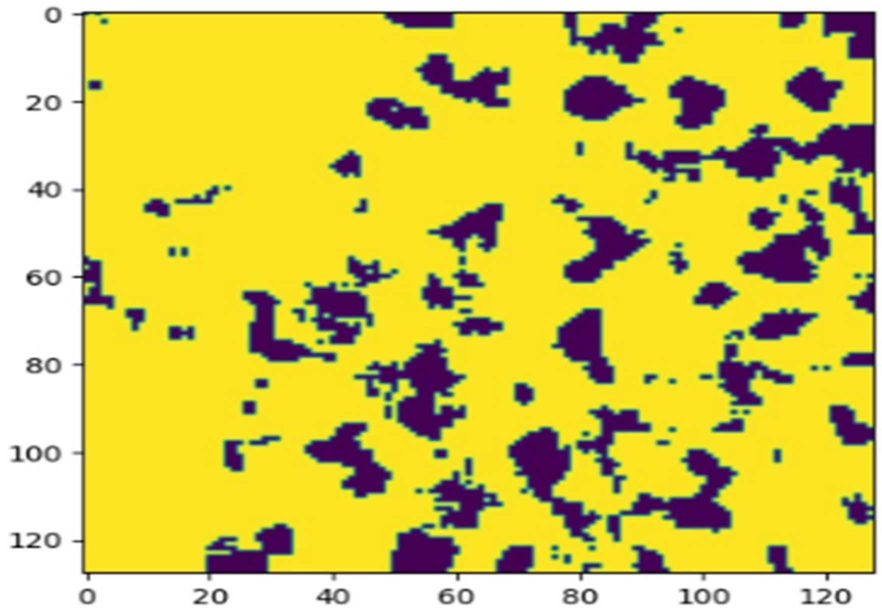
S. no.	Model	Dropout Val.	Output Class	Testing Loss	Testing Accuracy	Testing Precision
1	ϕCVNN	0	4	1.1034	0.5507	0.6499
2	ϕCVNN	0	2	0.6458	0.6710	0.6710
3	ϕCVNN	0.1	2	1.1405	0.6468	0.6468
4	ϕCVNN	0.2	2	0.9265	0.6519	0.6519
5	ϕCVNN	0.5	2	0.8619	0.6605	0.6605
6	ρCVNN	0	2	2.5442	0.6390	0.6390
7	ρCVNN	0.1	2	2.1523	0.6590	0.6590

Observations

- As we can observe due to high class variation due to unbalanced labeling the model is unable to reach a optimum solution as one label is very dominant.
- ϕ CVNN model goes to the most optimal solution after which it starts to overfit after 60 epochs
- The same model ϕ CVNN was trained with varying dropout rates show that dropouts are mostly ineffective in FCNN.
- ρ CVNN architecture was considered and were to test overfitting capabilities of highy com plex CVNN models and as we can see from trends from the validation loss and accuracy against the training loss and accuracy we can see the model is not learning anything after 25 epochs, even in adding dropout didn’t help much and model kept overfitting.



a. Target 2 class mask patch.



b. Predicted 2 class mask patch.

- The model ϕCVNN produces the highest accuracy with no dropouts.
- The dropout layers seem ineffective for CVNN U-Net.
- Increasing the number of filters in the U-Net leads to increases overfitting.
- Also the CVNN U-Net is very sensitive to the input mask used during the traning.

Summary

- As we can concur from the trends and results, the performance of CVNN for segmentation is highly depended on the labeling techniques used.
- Complex domain processing is a challenging problem itself for smooth training and to reach the optimal solution.
- As for data such as PolSAR or signals in complex domain it is very important avenue to keep the correlated information of phase and magnitude which is mostly lost if it’s adapted into Real domain R More research on optimization techniques for CVNN may allow us to create much optimal solutions for Complex data.

Future Scope

- Development of regularization techniques in complex domain as Ridge(L2) and LASSO(L1) regularizations won’t transform well in complex domain lack of bounds for complex values
- As the model is very sensitive to hyper-parameter tuning and further multiple models with varying learning rate or using variable learning rate might help us to optimize the results further, addition of more data for training might also help the model to improve further.
- Use of Ensemble methods Phase smoothing of the data in a neighborhood can be done to reduce variations can be employed.

References

- Xiangfang Li, Lijun Qian, Joshua Bassey et al-A Survey of Complex-Valued Neural Networks
- A. Hirose and S. Yoshida- Comparison of complex- and real-valued

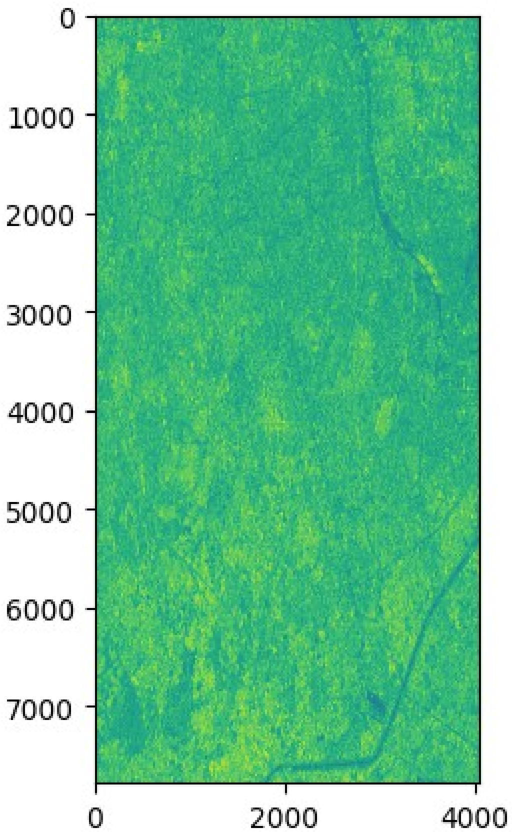
feedforward neural networks in their generalization ability- in Neural

Information Processing - 18th International Conference, ICONIP.

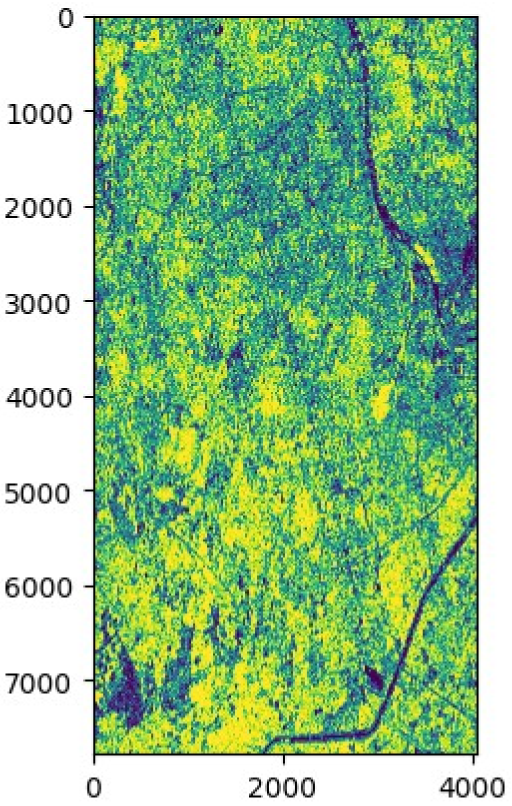
- N. Benvenuto and F. Piazza, “On the complex backpropagation algorithm,” IEEE Transactions on Signal Processidoi =

10.1109/IJCNN.2006.246722ng, vol. 40, pp. 967-969,

Thank you …



4-Class Thresholding



2-Class Thresholding.

- After thresholding with 4 class mask generation this way 54.85% were assigned to label 3, which caused very big class imbalance and led to very poor segmentation results
- To counter it the problem was converted from a 4 class problem to 2 class problem:
- Label 1 if B is >greater than .5
- Label 0 for the rest
- Training was done with these masks given as the labels, there were further pre-prossesing done by using Median filer on these masks to smooth them out.-