

**Development of machine learning algorithm for improved classification of
natural and man-made targets using PolSAR data based decomposition
models**

A Thesis submitted to the

UPES

For the Award of

Doctor of Philosophy

in

Computer Science and Engineering

By

Rajat Garg

August 2025

Supervisor

Dr. Anil Kumar

External Supervisor

Dr. Shashi Kumar

Dr. Manish Prateek



School of Computer Science (SOCS)

UPES

Dehradun- 248007: Uttarakhand, INDIA

**Development of machine learning algorithm for improved classification of
natural and man-made targets using PolSAR data based decomposition
models**

A Thesis submitted to the

UPES

For the Award of

Doctor of Philosophy

in

Computer Science and Engineering

By

Rajat Garg

(SAP ID: 500072290)

August 2025

Supervisor

Dr. Anil Kumar

Associate Professor

School of Computer Science

UPES

External Supervisor

Dr. Shashi Kumar

Scientist 'SF'

IIRS, ISRO

Dr. Manish Prateek

Professor and Pro Vice Chancellor

DBS Global University



School of Computer Science (SOCS)

UPES

Dehradun - 248007: Uttarakhand, INDIA

Declaration

I hereby declare that the thesis entitled **“Development of machine learning algorithm for improved classification of natural and man-made targets using PolSAR data based decomposition models”** for the award of Degree of Doctor of Philosophy and submitted in the School of Computer Science, UPES, is an authentic record of my own work under the supervision of Dr. Anil Kumar (Supervisor), Associate Professor, SoCS, UPES, Dr. Shashi Kumar (Ext. Co-Supervisor), Scientist ‘SF’, IIRS, ISRO and Dr. Manish Prateek (Ext. Co-Supervisor) Professor and Pro VC, DBS Global University.

The matter presented in this thesis has not been submitted by me for the award of any degree of this or any other Institute/University.



RAJAT GARG

School of Computer Science

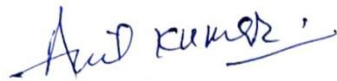
UPES, Bidholi, Dehradun

Uttarakhand, INDIA

Date: 19/08/25

CERTIFICATE

I certify that **Rajat Garg** has prepared his thesis entitled "**Development of machine learning algorithm for improved classification of natural and man-made targets using PolSAR data based decomposition models**", for the award of PhD degree of the UPES, under my guidance. He has carried out the work at the School of Computer Science, UPES.



Internal Supervisor

Dr. Anil Kumar

Associate Professor, Cluster Head: Artificial Intelligence

SoCS, UPES, Dehradun

Uttarakhand, India

Date: 06/08/2025

भारत सरकार
अंतरिक्ष विभाग
भारतीय सुदूर संवेदन संस्थान
4, कालीदास मार्ग, पो. बाक्स सं. 135
देहरादून- 248 001, भारत
दूरभाष : +91-135-2524399
फैक्स : +91-135-2741987, 2748041



Government of India
Department of Space
Indian Institute of Remote Sensing
4, Kalidas Road, P.B. No. 135,
Dehradun - 248 001, India
Telephone : +91-135-2524399
Fax : +91-135-2741987, 2748041

CERTIFICATE

I certify that **Rajat Garg** has prepared his thesis entitled “**Development of machine learning algorithm for improved classification of natural and man-made targets using PolSAR data based decomposition models**”, for the award of PhD degree of the UPES, under my guidance. He has carried out the work at the School of Computer Science, UPES.

Shashi Kumar
05/08/2025

Co-Supervisor
Dr. Shashi Kumar
Scientist 'SF'
IIRS, ISRO
Dehradun, Uttarakhand, India
Date:

भारतीय अंतरिक्ष अनुसंधान संगठन Indian Space Research Organisation



DBS GLOBAL
UNIVERSITY

CERTIFICATE

I certify that **Rajat Garg** has prepared his thesis entitled “**Development of machine learning algorithm for improved classification of natural and man-made targets using PolSAR data based decomposition models**”, for the award of PhD degree of the UPES under my guidance. He has carried out the work at the School of Computer Science, UPES.


Co-Supervisor 05/08/2025

Dr. Manish Prateek
Professor & Pro Vice Chancellor
DBS Global University, Dehradun
Uttarakhand, INDIA
Date:

Established vide The Uttarakhand Private University (Amendment) Act 2024, (Section 1(2) of Act No.- 05, 2024)
UGC Recognized under Section 2(f)

 122-MI, Selaqui, Dehradun (UK) INDIA

 0135-6141300/301

 info@dgu.ac.in

 www.dgu.ac

ABSTRACT

The world has been witnessing rapid growth and technology advancement in recent years and if left unchecked, the side effects could easily engulf the positive impact that is expected from the new innovation and development. Recently, we have been seeing catastrophic changes in the climate throughout the world which leads to extreme weather conditions and increased natural calamities like earthquakes, landslides, tsunamis, floods, and droughts in various parts of the world. This necessitates the regular monitoring of earth's surface and thus accurate land use land cover classification becomes very important.

With the advent of machine learning and deep learning algorithms, there has been a lot of improvement in the efficiency with which image processing applications can be handled. No doubt, the remote sensing community has quickly adopted the technology advancement for better performance. There has been a growing use of deep neural network based models for image classification task due to their high efficiency in handling image related applications. The main problem with these models is that they require a very large amount of data for effective training and thereby good predictions from them. Also, these models have an inherent black-box nature due to which they lack explainability as how they have reached a particular solution.

There has been a lot of improvement in the land use land cover classification through model based decomposition techniques but one challenge still being faced by the remote sensing community while handling SAR data is the oriented urban targets that cause volume type of scattering and therefore get misclassified as vegetation instead of urban targets. This work focuses on reducing this misclassification by leveraging the power of deep learning models through transfer learning method. The output from intermediary layers are presented in the current work to reduce the black-box nature of the deep learning model used.

The current work uses data from multiple sensors and locations to carry out the LULC study. The UAVSAR data which is openly available has been used for the analysis of Houston region in USA. The multi-frequency PolSAR data from spaceborne sensors like ALOS-2/PALSAR-2, TerraSAR-X, and RISAT-1 and multispectral optical data from Sentinel-2 were used for analysing the parts of Mumbai and nearby areas in Maharashtra, India. This work has

mainly three sections. In the first section, a comparative analysis of some popular machine learning models like Support Vector Machine, K Nearest Neighbor, and Random Forest is carried out for the PolSAR and Optical image segmentation task where it was found that Random Forest model performs the best with a mIOU of 0.6153 for ALOS2 data, 0.5815 for TerraSAR-X data, and 0.4031 for RISAT-1 data. The second section focuses on utilizing a pre-trained DeeplabV3+ model for PolSAR image segmentation task through transfer learning and minimizing the misclassification of oriented urban targets. The results obtained are compared with various other machine learning and deep learning models like Random Forest, K Nearest Neighbor, Support Vector Machine, Feature Pyramid Network, U-Net. The DeeplabV3+ model outperforms all the others with an overall accuracy of 85.65%.

ACKNOWLEDGEMENT

Man is a social being and that is why all activities of the man are completed with the help of God's grace, kind blessings of the elders, and of course good blessings of the younger. I have availed the co-operation of many personalities in this research work.

I would like to convey my unfathomed gratitude to my supervisor **Dr. Anil Kumar** and Co-Supervisors **Dr. Shashi Kumar** and **Dr. Manish Prateek** for their invaluable guidance and support to pursue this research work. From finding a suitable title for the research to penning down the thesis, a constant unreserved help and guidance was present. It was because of this help and guidance, I was able to finish my work and to overcome the obstacles that I encountered during the research work. Their dedication and sincerity had always inspired in me to reach to a higher level of thinking. The suggestions, interpretations, patronizing affection, and painstaking supervision would always be cherished in my memory. Their knowledge and clarity of thoughts have helped me to complete the research on the dot. I remember the times, when I was under pressure and lost my patience but they have always encouraged, energized me by sharing their past experiences and philosophical knowledge of life. I really enjoyed working with them.

I would also like to extend my gratitude to Space Applications Centre (SAC), ISRO, Ahmedabad (India), for the financial support under the NASA-ISRO Synthetic Aperture Radar (NISAR) project titled 'Implementation of Evolutionary Computing Algorithm for Polarimetric SAR data processing and classification'.

I am thankful to my seniors for their never-ending encouragement and assisting in departmental work. Friends have remained my life-long encourager, this would have been rendered incomplete without their love and support. I would also like to acknowledge colleagues associated with other NISAR projects for their help during various stages of my research.

Finally, blessings from my parents and elders, love from brothers and sisters who always stood by my side, supported, and motivated me despite the great difficulties have helped in moulding the thesis in this shape, which would have otherwise remained incomplete.

RAJAT GARG

Table of Contents

Declaration.....	iii
Certificate.....	iv
Certificate.....	v
Certificate.....	vi
Abstract.....	vii
Acknowledgement.....	ix
Abbreviations.....	xi
List of Figures.....	xiii
List of Tables.....	xvi
List of Appendices.....	xvii
Chapter 1: Introduction.....	01
1.1. Background and Motivation.....	01
1.2. Objectives.....	08
1.3. Methodology.....	09
1.4. Chapter Scheme.....	10
Chapter 2: Literature Review.....	16
2.1. PolSAR data.....	16
2.2. Pre-Processing of PolSAR data.....	19
2.2.1. Radiometric Calibration.....	19
2.2.2. Multilooking.....	20
2.2.3. Speckle Filtering.....	21
2.2.4. Terrain Correction.....	22
2.2.5. Coherency or Covariance Matrix Generation.....	23
2.3. PolSAR Decomposition methods.....	24
2.4. Machine Learning Models.....	29
2.5. Deep Learning Models.....	30
Chapter 3: Preparation of Dataset for Semantic Segmentation.....	33
3.1. Mumbai Region.....	33
3.2. Workflow of image segmentation task on Mumbai region dataset.....	35
3.2.1. Pre-processing.....	36
3.2.2. PolSAR Decomposition.....	36
3.2.3. Dataset Generation.....	39
3.2.4. Feature Extraction.....	40
3.3. Houston Region.....	40
3.4. Workflow of image segmentation using Houston region dataset.....	42
Chapter 4: Implementation of Machine Learning for semantic segmentation of multi-frequency SAR and Optical multispectral data.....	46
4.1. Machine Learning.....	46
4.1.1. Random Forest.....	46

4.1.2. K-Nearest Neighbor.....	47
4.1.3. Support Vector Machine.....	48
4.2. Deep Learning.....	50
4.2.1. Multilayer Perceptron.....	50
4.2.2. DeepLabV3+.....	51
Chapter 5: Results and Discussion.....	55
5.1. Mumbai Region.....	55
5.1.1. Performance evaluation metrics.....	55
5.1.2. Performance evaluation of image segmentation into four classes.....	56
5.1.3. Performance evaluation of water class.....	63
5.2. Houston Region.....	64
Chapter 6: Conclusion.....	69
References.....	73
List of Publications.....	89
Appendix.....	90

ABBREVIATIONS

ALOS	Advanced Land Observation Satellite
ASPP	Atrous Spatial Pyramid Pooling
CA	Class-wise accuracy
CAGR	Compound Annual Growth Rate
CC	Correlation Coefficient
CRF	Conditional Random Field
CGLS	Copernicus Global Land Service
DBNs	Deep Belief Networks
DCN	Data Communication Network
DCNN	Deep Convolutional Neural Network
DL	Deep Learning
DNN	Deep Neural Network
ERS-1	European Remote Sensing Satellite-1
ESA	European Space Agency
EWS	Extra Wide swath
FBD	Fine Beam Dual
FBP	Fine Beam Polarimetric
FCC	Fine Resolution Stripmap
FCN	Fully Convolutional Networks
FPN	Feature Pyramid Network
FRS1	Fine Resolution Stripmap mode 1
HH	Horizontal Horizontal
HV	Horizontal Vertical
IOU	Intersection over Union
IOU _F	Intersection over Union for Forest class
IOU _G	Intersection over Union for Ground class
IOU _U	Intersection over Union for Urban class
IOU _W	Intersection over Union for Water class
ISRO	Indian Remote Sensing Organization

JAXA	Japan Aerospace Exploration Agency
K	Kappa value or coefficient
KNN	K-Nearest Neighbour
LULC	Land Use Land Cover
MCGM	Municipal Corporation of Greater Mumbai
MCSM	Multi-Criteria Decision Making
mIOU	Mean Intersection over Union
MLP	Multilayer Perceptron
MMR	Mumbai Metropolitan Region
NASA	National Aeronautics and Space Administration
PALSAR	Phased Array type L-band Synthetic Aperture Radar
PASCAL VOC	Pascal Visual Object Classes
PolSAR	Polarimetric Synthetic Aperture Radar
PROBA-V	Project for On-Board Autonomy-Vegetation
RADARSAT-2	Radar Satellite-2
RBF	Radial Basis Function
RESNET	Residual Networks
RF,	Random Forest
RBF	Radial Basis Function
RGB	Red, Green, Blue
RISAT 2	Radar Imaging Satellite-2
RISAT-1	Radar Imaging Satellite-1
SAE	Stacked Auto Encoders
SAR	Synthetic Aperture Radar
ScanSAR	Scanning Synthetic Aperture Radar
S4R	Singh Four Component Decomposition
SDP	System Design Process
SLC	Single Look Complex
SVM	Support Vector Machine
SEASAT	Sea Satellite
TerraSAR-X	Terra Synthetic Aperture Radar-X
UAVSAR	Uninhabited Aerial Vehicle Synthetic Aperture Radar

VH	Vertical-Horizontal polarization
VV	Vertical-Vertical polarization
Y4R	Yamaguchi Four Component Decomposition

List of Figures

Figure 1.1:	Working of SAR sensor.....	02
Figure 1.2:	A trend of research using deep learning in remote sensing.....	06
Figure 1.3:	Proposed Research Methodology.....	09
Figure 2.1:	SLC σ° image corresponding to the VH band of Sentinel-1 data for Rohtak, India.....	20
Figure 2.2:	MLC σ° image corresponding to the VH band of Sentinel-1 data for Rohtak, India with a 4:1 ratio of number of range looks and azimuth looks.....	20
Figure 2.3:	σ° image corresponding to the VH band of Sentinel-1 data for Rohtak, India after applying the refined lee filter.....	21
Figure 2.4:	σ° image corresponding to the VH band of Sentinel-1 data for Rohtak, India after performing terrain correction.....	22
Figure 3.1:	Location of the study area (a) Land cover map of India with a resolution of 100 m derived from PROBA-V data for the year 2019 (b) Land cover map of Maharashtra with a light yellow color polygon representing the Mumbai Metropolitan Region (MMR) (c) Freeman-Durden decomposition-based color composite image of ALOS-2 PALSAR-2 data for Mumbai.....	34
Figure 3.2:	The flow chart of the methodology adopted for image segmentation using machine learning algorithms (RF= Random Forest, KNN = K-Nearest Neighbor, SVM = Support Vector Machine).....	35
Figure 3.3:	Representing the different decomposition results of (a) Entropy based decomposition, (b) Huynen Decomposition, and (c) Barnes decomposition.....	36
Figure 3.4:	Representing optical images that are based on true colour composites, false colour composites, and the results of Barnes decomposition of various areas in Mumbai. (a) Patch1's (Varaladevi Talao, Bhiwandi, Maharashtra) true-color composite based optical image, (b) false-color composite based optical image, and (c) Barnes decomposition based false-color composite SAR image, and (d) Patch 2's (Taloje River, Navi Mumabai, Maharashtra) true-color composite based optical image, (e) false-color composite based optical image, and (f) Barnes decomposition based false-color composite SAR image respectively.....	39
Figure 3.5:	Displaying the input image and related processed and annotated images of Mumbai's Vihar Lake, near Powai. (a) Sentinel-2 true-color composite image patch and (b) its corresponding label; (c) SAR image patch and (d) its corresponding label.....	40
Figure 3.6:	Represents the research area's position as compiled by using ArcGIS 10.6. (a) A map of Texas's border (available at: https://gis.txdot.opendata.arcgis.com/). (b) the PolSARPro v6.0 opensource remote sensing tool-generated G4U decomposition based false-color composite imagery of UAVSAR data over a portion of Houston, Texas (USA).....	41
Figure 3.7:	Represents the work flow of the methodology adopted for image segmentation using traditional machine learning algorithms and DeepLabv3+ model.....	43
Figure 3.8:	Represents the ground truth image, false-color composite image, and labeled image of the patches of Houston City, Texas(USA) (a) Google Earth image, (b) The false-color composite image, and (c) Ground truth label for Patch1, (d) Google Earth image, (e) The false-color composite image of G4U decomposition, and (f) Ground truth label for Patch2.....	44
Figure 4.1	Random Forest Classifier.....	47
Figure 4.2:	KNN Classifier.....	48
Figure 4.3:	Support Vector Machine Classifier.....	49
Figure 4.4:	Encoder-Decoder based architecture of DeepLabv3+.....	51

Figure 4.5:	In DeepLabv3+, the Xception architecture has been changed to include: (1) more layers, (2) depthwise separable convolutions with striding in place of all max pooling layers, and (3) additional batch normalisation and ReLU added after each 3×3 depthwise convolution.....	52
Figure 4.6:	Depicts the normal and atrous convolution. (a) a normal convolution with kernel size = 3, stride = 1, and padding is used to retain the input image size ; (b) an atrous convolution with an additional parameter called dilation rate that deletes every alternate row and column.....	53
Figure 4.7:	Represents the feature maps extracted from 1st hidden layer of Block1, Block7 and Block14.....	54
Figure 5.1:	Results of image segmentation for the Varaladevi Talao, Bhiwandi, Maharashtra patch 1. (a) Sentinel-2 true color composite image (where Red → B4, Green → B3, and Blue → B2), and (b) RF classifier, (c) KNN classifier, and (d) SVM classifier, respectively..	56
Figure 5.2:	Results of image segmentation for test patch 1. (a) Sentinel-2 false-color composite imagery (where Red → B8, Green → B4, Blue → B3), and (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively.....	56
Figure 5.3:	The results of image segmentation on patch 1 utilising (a) Barnes decomposition based false colour composite RGB image (where Red → Even bounce scattering, Green → Volume scattering, and Blue→Surface bounce scattering), (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for ALOS-2 data.....	57
Figure 5.4:	Results of patch 2's image segmentation detecting the area Taloje River, Navi Mumabai, Maharashtra (a) Sentinel-2 true colour composite image (where Red → B4, Green → B3, Blue→B2), and (b) RF classifier, (c) KNN classifier, and (d) SVM classifier results for the related image segmentation.....	58
Figure 5.5:	Results of image segmentation on Patch 2 where (a) Sentinel-2 false color composite based image (where Red → B8, Green → B4, Blue → B3) for test patch 2 and corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively.....	58
Figure 5.6:	Image segmentation results for patch 2. (a) Barnes decomposition based false color composite RGB image (where Red → Even bounce scattering, Green → Volume scattering, and Blue → Surface bounce scattering), and the corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for ALOS-2 data.....	59
Figure 5.7:	mIOU comparison for SAR (ALOS-2/PALSAR-2) and Optical (Sentinel-2) datasets for four classes (Water, Urban, Forest, Ground) for Patch1 and Patch 2 respectively...	60
Figure 5.8:	IOU comparison for SAR (ALOS-2/PALSAR-2) and Optical (Sentinel-2) datasets for water class for Patch 1 and Patch 2 respectively.....	61
Figure 5.9:	(a) Barnes decomposition based false color composite RGB image of selected patch (Vihar and Powai lake, Mumbai) where Red → Even bounce scattering, Green → Volume scattering, and Blue → Surface bounce scattering and the corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for TerraSAR-X data.....	61
Figure 5.10:	(a) 3-Component compact-pol decomposition based false color composite RGB image of selected patch (Vihar and Powai lake, Mumbai) where Red → Even bounce scattering, Green → Volume scattering, and Blue → Surface bounce scattering and the corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for RISAT-1 data.....	62
Figure 5.11:	(a) Barnes decomposition based false color composite RGB image of selected patch (Vihar and Powai lake, Mumbai) where Red → Even bounce scattering, Green → Volume scattering, and Blue → Surface bounce scattering and the corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for ALOS-2 data.....	62
Figure 5.12:	Represents the ground truth of Houston City, Texas (patch 1) and respective segmentation results of various classifiers (a) Google Earth image of test patch 1 (b)	

	False color composite image for G4U decomposition on UAVSAR data of a part of Houston City, Texas (USA) and the corresponding image segmentation results for (c) RF classifier, (d) KNN classifier, (e) SVM classifier, (f) FPN, (g) UNET, (h) Deeplab V3+ respectively.....	65
Figure 5.13:	(a) represents the ground truth through the Google Earth image of test patch 2 from Houston City, Texas (USA) (b) False color composite of UAVSAR image of the same test patch and the corresponding image segmentation results for (c) RF classifier, (d) KNN classifier, (e) SVM classifier, (f) FPN, (g) UNET, and (h) DeepLabv3+ respectively.....	66
Figure 5.14:	Demonstrates the performance metrics. (a) and (b) show the pixel accuracy for patches 1 and 2, respectively; and (c) shows the overall pixel accuracy for both patches for RF, KNN, SVM, FPN, UNET and DeepLabv3+ algorithms.....	67

List of Tables

Table 1.	Historical and current satellites carrying SAR sensor of different frequencies.....	16
Table 2.	Details of multi-frequency SAR datasets used in current work.....	34
Table 3.	Details of UAVSAR data and instrument.....	41
Table 4.	Results from a performance assessment of image segmentation using mIOU for optical (Sentinel-2) and SAR (ALOS-2) data utilizing both true-color composite (OpticalTCC) and false-color composite (OpticalFCC) images.....	60
Table 5.	The findings of performance testing for water class using IOU and image segmentation using mIOU for TerraSAR-X, RISAT-1, and ALOS-2 data.....	63
Table 6.	Results of performance assessment for the Water class utilizing IOU for optical (Sentinel-2) and SAR (ALOS-2) data using both true-color composite (OpticalTCC) and false-color composite (OpticalFCC) imagery.....	63
Table 7.	The results of performance evaluation of traditional machine learning algorithms and DeepLabv3+ algorithm.....	66

List of Appendices

Appendix 1.....	90
-----------------	----

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

Over the last 60 years of its presence, the utility of high-resolution microwave images configured by Synthetic Aperture Radar (SAR) ranges from military exploration to commercial all-weather surveillance (Jansing, 2021). In 1951, at Goodyear, Arizona, Carl Wiley proposed a concept based on the observation that objects within a radar beam move at varying velocities relative to the antenna. He theorized that this difference in speed would result in each item producing a distinct Doppler shift. Therefore, to create a comprehensive image, an accurate frequency analysis of the radar reflections is required. He named this technique Doppler Beam Sharpening (Jansing, 2021). As the year 1957 was coming to an end, airborne synthetic aperture radar was producing remarkable findings.

Up until the 1970s, SAR was only used in military missions. In 1978, the first earth-viewing SAR satellite targeting civilian applications named SEASAT was launched on the Seasat spacecraft (Way & Smith, 1991). The launch of SEASAT brought with it a sense of realization among the scientific community that the SAR can have several applications in the geoscience and engineering disciplines. After SEASAT, ERS-1 was the second spaceborne satellite carrying SAR sensor launched in 1991 (Ouchi, 2013).

Synthetic Aperture Radar (SAR) employs an intricate process to construct a remarkably large virtual antenna by utilizing the comprehensive Doppler history derived from the radar echoes that are generated as a result of the forward velocity of the spacecraft in motion. It is essential to note that despite the antenna's physical dimensions being relatively compact in size, the imagery that is produced as a result of this sophisticated technology exhibits an impressive level of azimuth resolution, making it highly effective for detailed observations and analyses (Fingas & Brown, 2014). The operational efficacy of SAR is contingent upon a meticulously integrated array of advanced onboard navigational and control systems, which encompass various high-precision inertial navigation devices and Doppler systems that collectively ensure an exceptional degree of location accuracy and precision (Hensley & Tighe, 2008). Moreover, SAR functions as a side-looking sensor, which is characterized by its ability to coherently capture signals; this operational capability signifies that it can detect not only the amplitude

but also the phase of the signals through a process known as quadrature demodulation. In order to produce a comprehensive two-dimensional representation of the Earth's surface that is reflective of its characteristics in terms of reflectivity, the side-looking sensor is strategically employed to methodically scan the landscape along a predetermined orbital or flight trajectory (Bamler, 2000). The intricate geometry and operational principles underlying the SAR sensor are depicted in figure 1.1, which serves to visually explain the complex interactions involved in its functioning and design.

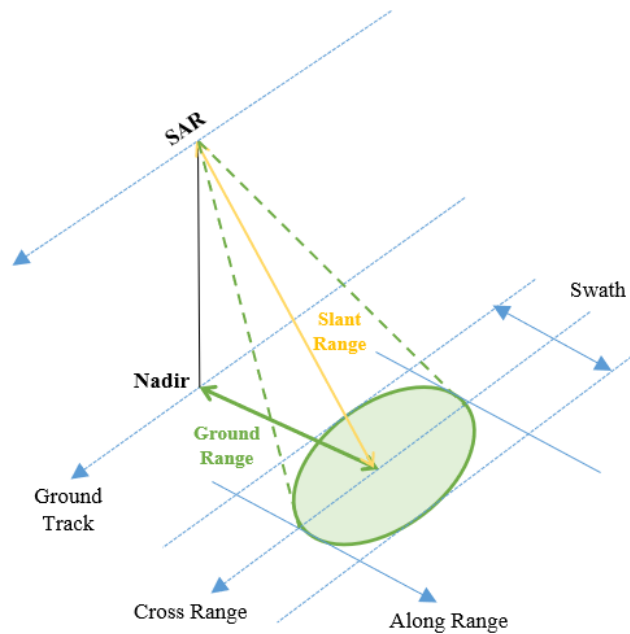


Figure 1.1: Working of SAR sensor

SAR imagery serves as a vital tool for quantifying the radar backscatter corresponding to specific terrestrial locations (Karachristos et al., 2024). The presence of bright regions indicates a substantial reflection of radar energy back toward the sensor, while darker areas denote minimal energy return to the radar apparatus (Joshi & Das, 2013). The backscatter characteristics for a designated area at a specific wavelength are influenced by various factors including the electrical properties of the target, the physical dimensions of scatterers within the targeted region, and the moisture content, with drier targets exhibiting a darker appearance and wetter entities appearing more luminous (Barrett et al., 2009). Additionally, the aspects of wavelength, polarization, and observation angles of the SAR pulses are pivotal in defining the intensity of backscatter (Joshi & Das, 2013; Barrett et al., 2009).

Polarization constitutes a fundamental attribute of electromagnetic waves. It can be characterized as the orientation and uniformity of the electric or magnetic field component situated in a plane orthogonal to the wave's trajectory (López-Martínez & Pottier, 2021). When observed along the trajectory of the wave, the electric field vector of a fully polarized wave arranges itself in a systematic configuration typically described as an ellipse, known as the polarization ellipse. The parameters defining a polarization ellipse include the wave's amplitude, orientation, and ellipticity (López-Martínez & Pottier, 2021). The magnitude of the electric field vector and its rotation rate function as indicators of the wave's amplitude and frequency, respectively. Polarization can be classified into linear, circular, and elliptical categories according to their orientation and ellipticity (López-Martínez & Pottier, 2021).

A synthetic aperture radar (SAR) system illuminates continuous sections of the Earth's surface along one lateral aspect of its flight trajectory (Lacomme et al., 2001). The radar sensor captures the radar signal that is reflected back towards it from the Earth's surface. The radar antenna quantifies the intensity and delay of the return signals. SAR systems are configured with both horizontal (H) and vertical (V) orientation channels, denoted as HH for horizontal transmission and horizontal reception, and HV for horizontal transmission and vertical reception, respectively (Karachristos et al., 2024). In this context, HH is classified as a co-polarized or like-polarized channel, whereas HV is designated as a cross-polarized channel (López-Martínez & Pottier, 2021). Such SAR systems can exhibit three levels of polarization complexity:

- Single polarized: HH, VV, HV, or VH
- Dual-polarized signals encompass (HH and VV), (HH and HV), or (HV and VV)
- Full or quad polarization incorporates all four channels (HH, VV, HV, VH)

The full or quad polarimetric mode is preferred for a thorough investigation of the backscatter since it maintains the whole vector character of the electromagnetic radiations. In addition to measuring the phase difference between various polarisation channels, a fully polarimetric SAR system also helps in the extraction of the polarimetric information (Kumar & Sharma, 2023). The complete vector character of the electromagnetic waves is addressed by SAR polarimetry. Based on the reaction of the target surface in various polarisation states, information about the surface may be extracted. A polarised radar wave's polarisation changes based on the precise properties of the earth's surface when it comes into contact with it. This

comprises the geometry, form, reflectivity, orientation, and geophysical characteristics like moisture content and surface roughness (Kumar & Sharma, 2023). As a result, the polarisation of the backscattered radar wave differs from the incident wave's polarization. Because of this, earth features respond differently in various polarisation channels, and distinct earth features may be recognized from the radar image based on this reaction. Features such as ocean waves, ice, soil moisture, geological features, vegetation, and man-made structures are better detected in the PolSAR imagery as compared to images captured by the optical sensors (Verma, 2012).

The scattering matrix serves as a repository for the polarimetric data pertaining to the target surface. Although the scattering matrix offers immediate understanding regarding pure or isolated scatterers, an examination utilizing the covariance or coherency matrix is imperative for distributed scatterers. These matrices, which embody the second-order statistics derived from the scattering matrix, encompass intricate information concerning the targets.

For humans, interpreting a radar image is more difficult than interpreting an optical image, and this task becomes even more difficult for a machine. By dividing the received signal into a total of distinct scattering processes expressing particular polarimetric characteristics, the PolSAR decomposition is used to characterize and categorize scattering from natural vegetation, man-made buildings, and landscape objects. To build an accurate picture of the chosen region of interest, PolSAR data must be properly divided into scattering power matrices. The better our decomposition result, the easier it is to recognize different data characteristics. Also, since decomposition is also a classification, with good decomposition results, we can obtain highly accurate segmentation of a PolSAR image. Now, the major steps in processing level 1 SAR data are coherency/covariance matrix generation, speckle filtering, and decomposition. The covariance/coherency matrix is conceptualized as an aggregation of various scattering modalities through a model-based decomposition methodology. A first-order Bragg's surface scatterer serves to represent surface scattering phenomena, whereas a dihedral corner reflector is indicative of double-bounce scattering effects. Volume scattering is delineated by an ensemble of randomly oriented dipolar entities (Freeman & Durden, 1998). In the creation of a pseudo-color RGB image, it is customary to assign Red to double-bounce, Blue to surface, and Green to volume scattering matrices.

PolSAR data classification can be assisted by polarimetric decomposition (Qi et al., 2012). (Huynen, 1970) articulated the target decomposition predicated upon Chandrasekhar's

theoretical framework regarding the scattering of light by miniscule anisotropic particles (Chandrasekhar, 1960). Since then, specialists have offered several target decompositions, which have been thoroughly examined in the literature review section.

Deep learning paradigms, in conjunction with diverse machine learning methodologies, have undergone extraordinary and transformative advancements in the domains of geoinformatics and remote sensing technologies over the recent years, thereby substantially augmenting our capacity to scrutinize spatial data and derive significant insights from it. Specifically, machine learning algorithms that function within the framework of supervised learning are engineered to leverage labeled datasets as their fundamental input, whereupon they engage in a process of elucidating latent patterns intrinsic within the data during the training phase, and subsequently apply the knowledge acquired through this training to formulate insightful predictions when confronted with novel, previously unobserved datasets during the testing phase. The rising environmental issues, encompassing urgent matters like deforestation, pollution of aquatic resources, and decline in soil quality, may be suitably addressed through the utilization and evaluation of land use land cover (LULC) maps, as demonstrated by the findings of (Liang et al., 2012). Furthermore, the constrained availability of remote sensing data constitutes a significant impediment that obstructs the generation of exceptionally precise land cover maps, which are indispensable for addressing and resolving these urgent and pressing environmental challenges that we presently encounter. For optimal performance, data-driven machine learning models demand a huge quantity of data, but there is practically no benchmark labeled dataset available for model training in the case of microwave remote sensing application involving semantic segmentation task. Also, the remote sensing data of most spaceborne sensors is very costly and thus cannot be accessed for academic purposes. The task of manual data annotation is very cumbersome and time-consuming which necessitates some other way around. Artificial neural networks predicated deep learning frameworks which are otherwise an element of machine learning, endeavor to emulate the operations of the human brain. To take the advantage of knowledge gathered from the model training in one task as a beginning point for another activity, a transfer learning strategy can be brought into use. This work suggests employing transfer learning to get a good outcome from minimal labeled data for training and testing the deep learning model implemented. DeepLabV3+ encoder-decoder architecture is based on a modified Xception model with atrous convolutions (Chen et al., 2018).

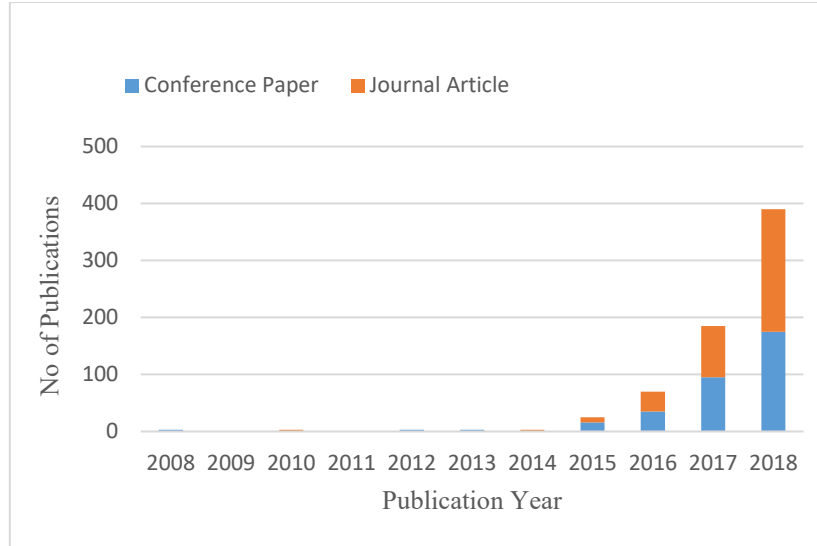


Figure 1.2. A trend of research using deep learning in remote sensing (Ma et al., 2019).

Understanding the scattering mechanisms that give birth to patterns in the various polarisation parameters is a fundamental challenge in evaluating polarimetric SAR data. One of the major problems being faced today in decomposition is the modeling of highly oriented urban targets with respect to radar illumination direction. An oriented urban target looks more like vegetation than the urban area which leads to misclassification of such targets into vegetation or they might not get classified into any category. This problem will be addressed with the help of ML methods. It is widely recognized that the emergence of machine learning and deep learning has unveiled a multitude of prospects in nearly every domain and the remote sensing domain is still exploring this current advancement in the computing world. An increasing interest in this interdisciplinary field among the research community can be observed from the bar chart presented in figure 1.2 (Ma et al., 2019). With the help of improved decomposition and using a suitable deep learning algorithm for image segmentation, the target will be to develop a python code capable of processing level 1 fully polarimetric SAR data of RISAT-1, TerraSAR-X, and UAVSAR sensors which will include pre-processing of the dataset, decomposition and image segmentation.

Biomass estimation (Kumar et al., 2012), agricultural monitoring (Liu et al., 2019), forest fire detection (Hosseini & Lim, 2023), mineral prospecting (Sabrian et al., 2017), vegetation cover mapping (Gazzea et al., 2023), ice dynamics modeling (Gomez et al., 2019), oil spill detection (Chaudhary & Kumar, 2021), and defence monitoring (Gao et al., 2023) are all possible uses of SAR based remote sensing. The applications of SAR technology in the above-mentioned uses are yet to be thoroughly investigated since there is still time till SAR

technology becomes fully affordable for smaller-scale applications. All of the aforementioned factors are projected to boost the market. Featuring benefits such as all-weather functionality, continuous functioning regardless of time, and immunity to atmospheric conditions, responsiveness to dielectric properties, sensitivity to artificial objects, sensitivity to target structure (with polarimetry), precise measurements of distance (with interferometry), and subsurface penetration, the market for SAR can sprout massively in the future.

This exponentially growing market working round the clock will generate yottabytes of data in the near future which will require advanced computational power and usage of recent technologies like big data and machine learning for data handling and processing. The SAR market is expected to grow at a CAGR of 11.6% as per a study conducted by (Mordor Intelligence, 2022). With the miniaturization of satellites, many micro and nano satellites are coming up recently which cannot afford a large antenna and are bound to use SAR technology. So, the requirement of SAR data processing is here to stay with a plethora of opportunities cropping up in near future.

Just like the tremendous market potential of SAR, the research aspect is equally fascinating. After more than 50 years of research into the target decomposition, there has been a lot of improvement in PolSAR decomposition results. The misclassification of urban areas into the forest has been reduced by a huge amount but most of the improvement has occurred in areas consisting of perfectly aligned buildings while areas consisting of oriented urban targets, still need a lot of improvement. This is one major issue that is being faced today that will be targeted in this research work. Solving this problem would greatly improve the image segmentation results and therefore improved LULC mapping. The dataset used in this work would be unique in itself and comprise of multi-frequency PolSAR data from multiple locations like Houston, Texas, USA, and Mumbai, India. The PolSAR data from different airborne and spaceborne sensors like UAVSAR, TerraSAR-X, and RISAT-1 would be used for the implementation of the image segmentation models. It is a known fact that data is the foundation of all the megatrends occurring in the new world order. However, in the context of microwave remote sensing, the existence of a reference annotated dataset is exceedingly uncommon, and furthermore, the PolSAR data is not readily obtainable in substantial quantities, which constitutes a fundamental requirement for the implementation of data-driven machine learning algorithms. Consequently, the dataset for the training of the machine learning and deep learning models in this study would be developed in-house.

1.2 Objectives

The principal aim of this scholarly investigation is the implementation of a machine learning algorithm for improved classification using PolSAR data based decomposition models.

Sub Objectives:

- i. Devising an improved machine learning classification model for the identification of LULC using PolSAR data.
- ii. Implementation of a deep learning algorithm for image segmentation.
- iii. Performance evaluation of the devised model on PolSAR data in comparison with existing state-of-the-art machine learning models.

1.3 Methodology

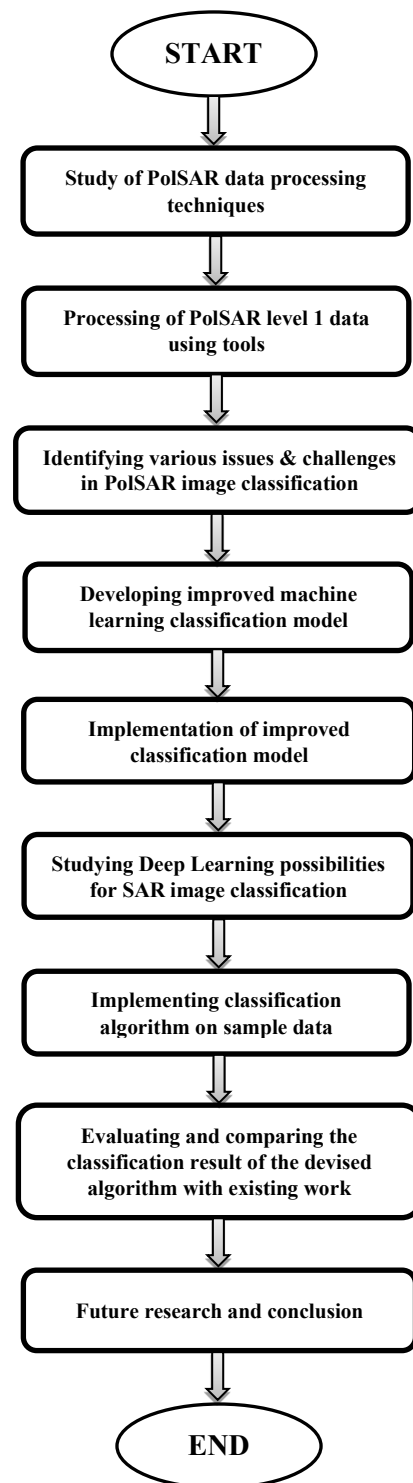


Figure 1.3 Proposed Research Methodology

1.4 Chapter Scheme

This doctoral dissertation is segmented into six sections. The initial section serves as the introductory segment that establishes the foundation by furnishing a comprehensive overview of Synthetic Aperture Radar (SAR) and Polarimetric SAR (PolSAR) technologies. It delineates their progression from their nascent applications in military contexts to their contemporary functions in civilian and commercial spheres, such as urban development, environmental surveillance, and disaster management. The section underscores SAR's distinctive capacity to acquire high-resolution imagery under all meteorological conditions, irrespective of the time of day, rendering it exceedingly advantageous for remote sensing endeavors. The key focus of this research is the classification and segmentation of PolSAR data, which provides information not only about the surface of the Earth but also its physical properties. The challenge, however, lies in the complexity of interpreting this data due to various scattering mechanisms and inherent noise, such as speckle, which can obscure important details. To mitigate these challenges, the present study investigates the utilization of deep learning architectures - specifically, models grounded in convolutional neural networks (CNNs) - for the purpose of semantic segmentation. Such models, particularly those that integrate sophisticated methodologies such as atrous convolution and multi-scale feature extraction, are exceptionally equipped to navigate the intricacies inherent in PolSAR data. DeepLabv3+, recognized as a cutting-edge model for semantic segmentation, is posited as the principal framework for this inquiry due to its proficiency in capturing intricate details and yielding precise segmentations within complex urban landscapes.

The first chapter concludes by acknowledging the rapid technological advancements in SAR and PolSAR imaging and their expanding use in civilian and scientific domains. While traditional decomposition methods provided initial tools for analyzing PolSAR data, they fall short in handling the intricate scattering mechanisms found in mixed land cover, involving urban and natural environments. These limitations, along with the inherent noise present in SAR data, necessitate more robust solutions. Deep learning, and specifically semantic segmentation models like DeepLabv3+, have emerged as powerful tools in this context. These models leverage the massive amounts of data generated by PolSAR sensors to learn complex spatial and spectral relationships, making them ideal for tasks such as land use classification and land cover assessment. The conclusion emphasizes that this research will demonstrate how these models can improve the accuracy of PolSAR segmentation, helping to unlock new

insights in fields like urban development, environmental conservation, and disaster management.

The subsequent chapter presents a comprehensive examination of the scholarly literature pertaining to the analysis of PolSAR data, encompassing both conventional and modern techniques. The literature on PolSAR data analysis is reviewed extensively, starting with traditional approaches and moving toward more contemporary methods. Early work in this field largely relied on target decomposition techniques like the Freeman-Durden and Yamaguchi decompositions. These methods helped researchers break down complex PolSAR signals into simpler scattering components—such as surface, volume, and double-bounce scattering—making the data more interpretable. However, these techniques often struggled to accurately classify urban environments due to the complex scattering mechanisms present in cities, where buildings and other structures reflect signals in ways that mimic natural features like vegetation. The review then transitions into machine learning methods, which have shown promise in improving classification performance. Methods such as Support Vector Machines (SVMs), Random Forests (RFs), and k-Nearest Neighbors (KNN) have been utilized on PolSAR data with differing levels of efficacy. Nonetheless, these methodologies continue to necessitate the manual development of features and face challenges arising from the high dimensionality and intricacy of PolSAR data, particularly when addressing extensive datasets or highly variable environments. In recent years, deep learning has surfaced as a more effective strategy for remote sensing applications, including PolSAR segmentation. Deep learning architectures, notably convolutional neural networks (CNNs), possess the capability to autonomously extract complex feature representations from the input data, thereby removing the requirement for manual feature extraction. This attribute renders deep learning models especially adept for PolSAR data, which encompasses detailed spatial and spectral information. Transfer learning, a key concept in deep learning, is also introduced as a solution to the problem of limited labeled data. By leveraging pre-trained models, researchers can fine-tune these models for specific tasks, such as PolSAR segmentation, even with limited training data.

The chapter concludes by acknowledging the contributions of traditional methods to the field of PolSAR data analysis. Techniques like target decomposition have provided valuable insights into the scattering mechanisms present in PolSAR images, allowing for basic classification of natural and man-made targets. However, these methods have clear limitations, particularly in complex environments with mixed land cover regions. Machine learning

methods marked an improvement by allowing more sophisticated pattern recognition, but they too fall short when faced with large, complex datasets. These methods require handcrafted features and struggle to generalize well across different types of data. This is where deep learning models excel. Through their capacity to autonomously acquire hierarchical feature representations, deep learning architectures, particularly CNNs, offer a more robust solution for PolSAR segmentation tasks. Models like DeepLabv3+ not only handle the complexity of PolSAR data but also provide better scalability and adaptability, making them a more practical choice for large-scale remote sensing applications. The chapter highlights the shift in the interest of the research community toward deep learning methods and justifies the use of models like DeepLabv3+ for this study. These models, with their advanced architectures and capabilities, represent the next step in improving PolSAR segmentation accuracy, particularly in complex urban and natural environments.

Chapter 3 offers a comprehensive exposition of the datasets employed within the research and the preprocessing methodologies implemented upon them. The datasets consist of multi-frequency PolSAR data from several sources, including ALOS-2, TerraSAR-X, RISAT-1, and Sentinel-2 satellites, with data collected from diverse geographic regions such as Mumbai, India, and Houston, USA. These locations were selected due to their varying land cover types, which present a range of challenges for segmentation tasks. In Mumbai, for instance, the mix of urban structures, water bodies, and vegetation creates a complex environment, while Houston's urban sprawl and natural features offer a different set of challenges. PolSAR data is complex and often noisy, requiring significant preprocessing before it can be fed into a deep learning model. The chapter discusses several key preprocessing techniques, including radiometric calibration, which adjusts for sensor noise, and multilooking, which improves image quality by averaging multiple looks at the same scene. Speckle filtering, another critical step, reduces the granular noise inherent in SAR images, making the data easier to interpret. Terrain correction is also applied to account for distortions caused by the Earth's surface geometry. Once the data is preprocessed, it is split into smaller patches and augmented to create a larger, more diverse dataset. Augmentation techniques such as rotation, flipping, and scaling are applied to the dataset to simulate different viewing angles and conditions. This aspect holds significant relevance for the training of deep learning architectures, as it contributes to enhancing the model's capacity for generalization when confronted with novel, previously unobserved data.

The preprocessing steps outlined in this chapter are essential for ensuring that the data is clean, well-structured, and suitable for deep learning applications. Without these steps, the noise and distortions present in the raw PolSAR data would significantly hinder the performance of the segmentation models. Speckle filtering, in particular, is crucial for reducing the noise that can obscure important features in the data. The application of radiometric calibration and terrain correction ensures that the data accurately reflects the true physical characteristics of the scene, further enhancing the quality of the segmentation results. In conjunction with preprocessing, dataset augmentation assumes a pivotal function in enhancing the efficacy of deep learning models. Through the artificial expansion of both the size and diversity of the dataset, augmentation facilitates the model's capacity to achieve improved generalization., which is particularly important in remote sensing tasks where labeled data is often limited. This chapter emphasizes that careful attention to preprocessing and dataset generation is key to achieving high accuracy in PolSAR segmentation tasks.

Chapter 4 discusses the methodology and model architecture used in this research. The discussion commences by delineating the conventional machine learning paradigms that function as reference points for evaluative comparison. Models such as Random Forest (RF) and Support Vector Machine (SVM) are employed to create a foundational performance standard for the segmentation endeavor. These models, while useful in simpler classification tasks, are limited in their ability to handle the complexity of PolSAR data, especially in terms of capturing fine details and multi-scale features. The central focus of this chapter is the DeepLabv3+ model, an advanced semantic segmentation model that employs an encoder-decoder architecture. The encoder within the DeepLabv3+ architecture is tasked with the extraction of advanced features from the input data through the utilization of an extensive convolutional neural network, whereas the decoder enhances the segmentation output by guaranteeing the precise delineation of object boundaries. A key feature of DeepLabv3+ is its use of atrous (or dilated) convolution, which allows the model to capture multi-scale contextual information without losing spatial resolution. This is particularly important in PolSAR segmentation, where the ability to capture both fine details and large-scale patterns is crucial. Transfer learning is another critical component of the methodology. Utilizing a model that has been pre-trained on a substantial dataset like ImageNet, the study makes use of the existing knowledge embedded in the model and adjusts it for the precise task of PolSAR segmentation. This method effectively diminishes the quantity of labeled data needed for training, presenting

a feasible solution for remote sensing assignments, particularly in cases where labeled datasets are generally limited.

This chapter demonstrates that DeepLabv3+ is a highly effective model for semantic segmentation, particularly when applied to complex datasets like PolSAR. Traditional machine learning models like RF and SVM, while useful as benchmarks, lack the sophistication needed to handle the intricate patterns and features present in PolSAR data. In contrast, DeepLabv3+ excels in capturing multi-scale information, making it well-suited for tasks that require both fine detail and broad context. The model's performance is improved through the utilization of transfer learning, enabling it to generalize effectively despite having limited training data. This is a crucial advantage in remote sensing, where obtaining large, labeled datasets can be difficult and expensive. Overall, the methodology outlined in this chapter provides a robust framework for applying deep learning to PolSAR segmentation, offering a significant improvement over traditional methods.

In Chapter 5, the findings of the research are elucidated and scrutinized. The efficacy of the DeepLabv3+ model is juxtaposed with conventional machine learning models such as Random Forest (RF) and Support Vector Machine (SVM) through the application of various metrics, encompassing mIOU, accuracy, and precision. The datasets from Mumbai and Houston are used to evaluate the models, with a focus on their ability to accurately segment urban areas, natural landscapes, and mixed land cover. The results clearly show that DeepLabv3+ outperforms the traditional models, particularly in complex environments like urban areas. In Mumbai, for example, traditional models struggled to distinguish between buildings and vegetation, often misclassifying urban areas as natural features. DeepLabv3+, on the other hand, was able to accurately segment these regions, thanks to its ability to capture multi-scale features and refine object boundaries. The chapter also discusses some of the challenges encountered during the research, such as the limited availability of labeled data. However, the use of transfer learning helped mitigate this issue by allowing the model to leverage pre-trained knowledge, significantly improving its performance despite the data constraints.

The findings presented in chapter five demonstrate the clear superiority of deep learning models like DeepLabv3+ for PolSAR data segmentation. While traditional machine learning models like RF and SVM can be useful in simpler tasks, they are not well-suited for handling

the complexity of PolSAR data, especially in urban environments. DeepLabv3+, with its ability to capture fine details and multi-scale information, provides a much more accurate and reliable solution for segmentation tasks. The employment of performance indicators, specifically mean Intersection over Union (mIOU) and accuracy, underscores the efficacy of the model in differentiating between various land cover categories, especially in complex environments characterized by the presence of multiple scattering mechanisms. The chapter concludes that DeepLabv3+ is the most effective tool for PolSAR segmentation, offering significant improvements over traditional methods.

In chapter 6, the key findings of the research are summarized, with a focus on the advantages of using deep learning models for PolSAR data segmentation. The chapter emphasizes the significant improvements in segmentation accuracy and precision achieved by DeepLabv3+, particularly in complex environments like urban areas. The investigation further emphasizes the pragmatic advantages associated with the implementation of transfer learning, facilitating efficient model training even in scenarios where labeled data is scarce. Additionally, the chapter addresses the constraints of the research, including the limited availability of publicly accessible labeled PolSAR datasets, and proposes avenues for future scholarly inquiry. One potential avenue is the integration of optical data with PolSAR data, which could further improve classification accuracy by providing additional information about the surface characteristics.

The dissertation culminates by reasserting the efficacy of deep learning architectures, specifically encoder-decoder frameworks such as DeepLabv3+, in the segmentation of PolSAR data. These architectures, due to their capacity to assimilate intricate feature representations and effectively capture information across multiple scales, provide a substantial advancement compared to conventional machine learning techniques. Through the implementation of methodologies such as transfer learning, deep learning models can be proficiently trained even in scenarios with constrained labeled data, rendering them a pragmatic and scalable alternative for real-world remote sensing applications. The chapter also emphasizes the need for continued research in this field, particularly in the development of benchmark datasets that can be used to train and evaluate deep learning models. By improving the availability and quality of labeled PolSAR datasets, future research can further enhance the performance of these models, unlocking new possibilities for applications in urban planning, environmental monitoring, and disaster response.

CHAPTER 2

LITERATURE REVIEW

2.1 PolSAR data

PolSAR data is available for download at online portals of multiple space agencies namely ESA, NASA, JAXA, and ISRO. Most of the PolSAR data captured by these agencies is not open source but the data from spaceborne sensor Sentinel 1 and airborne sensor UAVSAR is freely available on the respective portals of ESA and NASA. The researchers can download the Dual-Pol, C-Band Sentinel-1 data from the ESA portal (scihub.copernicus.eu) and Full-Pol, L-Band UAVSAR data from the NASA's portal (uavsar.jpl.nasa.gov).

Table 1. Historical and current satellites carrying SAR sensor of different frequencies (Solberg et al., 2008) (Al-Ruzouq et al., 2020).

Mission	ALOS/ PALSAR	TerraSAR-X	RADARSAT-2	RISAT-1	Sentinel-1	ALOS-2/ PALSAR-2
Country	Japan	Germany	Canada	India	ESA	Japan
Launch Date	24 Jan 2006	15 June 2007	14 Dec 2007	26 Apr 2012	3 Apr 2014	24 May 2014
Frequency [GHz]	1.27	9.65	5.405	5.35	5.405	1.236
Band	L-Band	X-Band	C-Band	C-Band	C-Band	L-Band
Polarization	Dual, Full	Full	Full	Dual	Dual	Full
Mode of Acquisition	Fine Beam Dual (FBD), Fine Beam Polarimetric (FBP)	Stripmap, Spotlight, ScanSAR	Spotlight, ScanSAR, Single Beam	Fine Resolution Stripmap (FRS1)	Stripmap, Interferometric wide swath, Extra-wide swath.	Stripmap, Spotlight, ScanSAR

The table above lists some of the PolSAR based satellites which operate on different bands targeting different applications of remote sensing. The ALOS/PALSAR is a Japanese origin satellite which was launched on 24 January 2006 (Shimada et al., 2014). It operates on 1.27 GHz frequency (L-Band) having multiple Polarization capabilities and multiple modes of data acquisition. The ALOS/PALSAR satellite captures data with 10 m resolution and covers a wide variety of applications like soil moisture estimation, biomass estimation, agriculture monitoring.

Launched on June 15, 2007, the TerraSAR-X represents a satellite engineered by the German space agency. Its principal objective is to furnish high-resolution, multi-mode X-Band SAR data applicable across a myriad of fields. These fields include disaster monitoring, hydrological and climatological analysis, geological investigations, environmental change observation, oceanographic exploration, and cartographic generation. This satellite functions at a frequency of 9.65 GHz, employing Full Polarization and a Stripmap mode of data acquisition. The dimensions of its antennas are 4.8 meters in length and 0.7 meters in width, oriented towards the predetermined nominal look direction. The comprehensive performance incidence angle range for spotlight modes extends from 20° to 55° (with an accessible range of 15° to 60°). The azimuth resolution in stripmap mode achieves 3.3 meters (6.6 meters for dual polarization), whereas the ground range resolution spans from 1.70 meters to 3.49 meters. In Spotlight mode, the azimuth resolution is recorded at 1.1 meters (single polarization) and 2.2 meters (dual polarization), while the ground range resolution ranges from 1.48 meters to 3.49 meters and from 0.74 meters to 1.77 meters (under conditions of 300 MHz bandwidth and reduced swath extent in range) (Roth et al., 2005).

The Canadian satellite RADARSAT-2, with a mass of 2300 kg, was launched on December 14, 2007, and its capability to function in the Selective Dual Polarization (SDP) mode significantly enhances its capacity to deliver information pertinent to applications addressing targets such as transparent vegetation volumes with diverse structural characteristics or land surfaces exhibiting varying levels of roughness (Livingstone et al., 2005). It operates at a frequency of 5.405 GHz, utilizing full polarization and multiple modes of acquisition, including Spotlight, ScanSAR, and Single Beam. With an antenna length measuring 15 meters and an antenna height of 1.5 meters, the imaging modes prove to be highly beneficial for military applications, encompassing situational awareness and change detection (Abreu et al.).

The RISAT-1 is an Indian Radar Imaging Satellite with a mass of 1858 kg out of which SAR payload contributes to 950kg. The launch of the satellite on 26 April 2012 marked a new era in the remote sensing programme of ISRO. With a frequency of 5.35GHz, it supports multiple resolution and swath requirements (Valarmathi et al., 2013). Dual polarisation operation is supported for both stripmap and scanSAR modes. A quad polarisation stripmap mode is now offered to enable for extra resource categorisation. All of these modes allow for resolutions of 3 to 50 m and swaths of 25 to 223 km (Valarmathi et al., 2013). It holds numerous

applications like flood mapping, agriculture monitoring and identification of soil moisture, vegetation, forestry, geology and man-made object identification.

The Sentinel-1 is the project of European Space Agency with a mass of 2300kg launched on 3 April 2014 with frequency of 5.405 GHz and Polarization Selectable between HH+HV and VV+VH. Interferometric Wide-swath mode in conjunction with the Wave mode adheres to the secure and dependable service, lowers operational costs, and also satisfies future requests by creating a consistent long-term archive ("Sentinel-1 - Mission Summary - Sentinel Online"). It supports mapping to provide humanitarian help in emergency situations, monitoring marine ecosystems, tracking sea-ice areas and the polar environment, tracking risks associated with land surface mobility, and mapping of land surfaces, including those used for agriculture, water, and soil. The roll steering mode of the satellite is a crucial component. Altitude fluctuation is offset by the constant roll maneuver that is performed around the orbit (similar to yaw steering in azimuth).

The ALOS-2/PALSAR-2 is a Japanese radar imaging satellite launched on 24 May 2014 with the frequency of 1.236 GHz. It operates on stripmap mode of acquisition with full polarization. The mission of the project revolves around four-fold objectives i.e. environmental monitoring (Hydrosphere, Cryosphere, Biosphere and Geosphere); disaster monitoring; natural resources (Agriculture and Ocean); technology development (Remote Sensing). It boasts about four new techniques- Chirp Modulation (Azimuth Phase Coding); Dual Receive Antenna System; New Data Compression and High Power Efficiency Device (Kankaku et al., 2013).

The PolSAR data is mostly available as GRD/SLC product on the website of the space agencies and it is available as covariance matrix also in case of UAVSAR airborne sensor. The GRD/SLC product undergoes various pre-processing steps before it is broken into distinct scattering types for classification.

A Synthetic Aperture Radar (SAR) image is composed of a two-dimensional grid of pixels organized in rows and columns. Each pixel represents a specific area on the Earth's surface, with its size determined by the SAR system's specifications (Lee & Pottier, 2017). Within each SAR resolution cell, pixels contain complex numbers that encode both amplitude and phase information based on the reflectance of individual scatterers. The radar backscattering coefficient σ_0 , also known as surface reflectivity, is influenced by both radar

system parameters (such as polarization, frequency, and incidence angle of electromagnetic waves) and surface characteristics (including local incidence angle, topography, roughness, moisture, and dielectric properties of the medium) (Lee & Pottier, 2017).

SAR imaging utilizes a side-looking sensor that illuminates perpendicular to the flight path. The cross-track dimension in SAR imagery is determined by measuring the time linked to the direct distance (slant range) from the radar to the surface point. This method results in inherent geometrical distortions in SAR images due to the difference between slant range and ground range. The two primary sources of specific geometrical distortions are layover and foreshortening (Lee & Pottier, 2017). To address these distortions and eliminate other noise in SAR data, pre-processing steps are necessary, which will be discussed in the subsequent section.

2.2 Pre-Processing of PolSAR data

It has been articulated in the preceding section that the PolSAR data necessitates a series of pre-processing steps prior to its categorization into discrete scattering types utilizing decomposition models; this section will elaborate on several critical pre-processing techniques employed in the treatment of PolSAR data.

2.2.1 Radiometric Calibration

The Synthetic Aperture Radar (SAR) instrument endeavors to provide sufficient information regarding the characteristics of the target through the electromagnetic interaction of the emitted signal with the target, thereby facilitating the extraction of geophysical data from the resultant images. Multi-parameter imaging, which varies parameters such as the radar wavelength, polarization, and imaging geometry to create a full description of the target, is the most effective method for target characterization (Freeman & Curlander, 1989). In order to interpret the data, it is necessary to relate the amplitude and phase of the generated picture pixels to certain target properties (such as complex reflectivity, radar cross-section and sigma) while taking into account all of the data system's components. Different sources of radiometric calibration errors are atmospheric propagation, sensor platform, antenna, radar sensor, signal processor, image processor and geophysical processor. The majority of the end-to-end SAR system's deterministic or slowly changing random errors may potentially be addressed while the SAR image is being formed (Freeman & Curlander, 1989).

2.2.2 Multilooking

The process of multilooking Synthetic Aperture Radar (SAR) data can be accomplished through the averaging of range and/or azimuth resolution. This methodological approach results in an enhancement of the radiometric resolution; however, it concurrently leads to a deterioration of spatial resolution. The resultant multilooked imagery exhibits reduced noise levels and pixels that approximate a square shape.

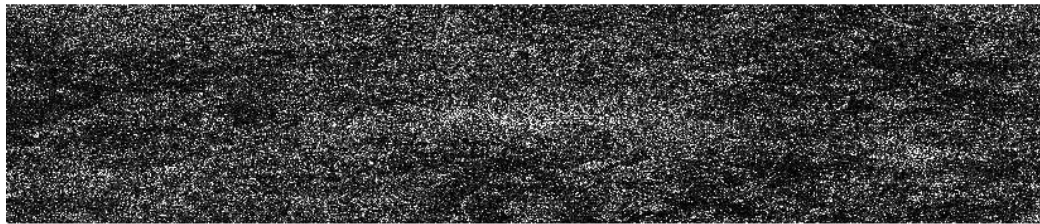


Figure 2.1 SLC σ° image corresponding to the VH band of Sentinel-1 data for Rohtak, India.

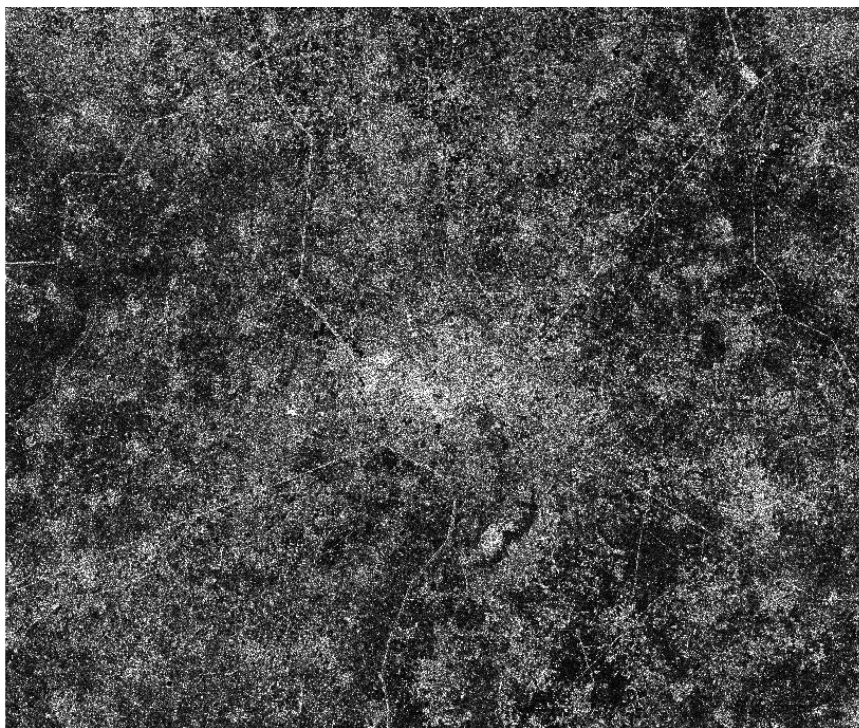


Figure 2.2 MLC σ° image corresponding to the VH band of Sentinel-1 data for Rohtak, India with a 4:1 ratio of number of range looks and azimuth looks.

A representative illustration of Sentinel-1 Synthetic Aperture Radar (SAR) data prior to and subsequent to the Multilooking process is depicted in figures 2.1 and 2.2, respectively. The TOPSAR Sentinel-1 dataset underwent processing utilizing the SNAP toolbox and has not undergone terrain correction. In the process of multilooking the input Single Look Complex (SLC) image and generating square pixels, the parameter for range looks was established at 4,

while the parameter for azimuth looks was determined to be 1. Consequently, the resultant mean Ground Range (GR) square pixel dimension was calculated to be 14.328 meters.

2.2.3 Speckle Filtering

Speckle constitutes a granular noise phenomenon that manifests in Synthetic Aperture Radar (SAR) data as a consequence of the interference resulting from waves that are reflected off multiple elementary scatterers (Lee et al., 1994). The existence of speckle within SAR imagery significantly diminishes the clarity and interpretability of the target features observable in the data. The principal concept underpinning speckle filtering involves the application of filtering processes to adjacent pixels that exhibit analogous scattering characteristics to the central pixel (Lee & Pottier, 2017). There exists a multitude of methodologies for the attenuation of speckle noise; however, these strategies can be broadly categorized into three distinct classifications.

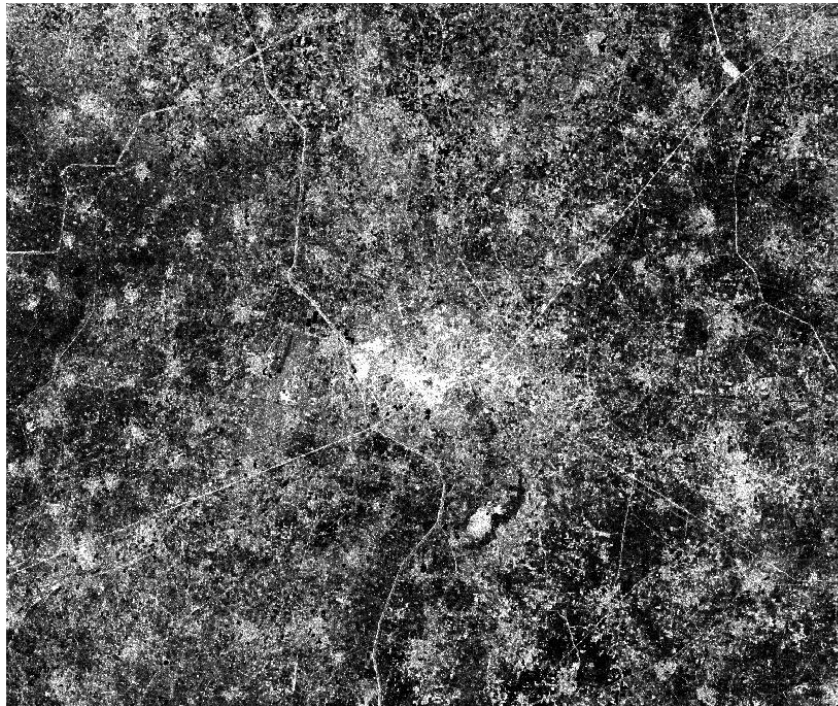


Figure 2.3 σ^0 image corresponding to the VH band of Sentinel-1 data for Rohtak, India after applying the refined lee filter.

In the first category, multiple frames taken from a section of the available azimuth spectral bandwidth are averaged incoherently to enhance the appearance of a SAR picture while the noise variance is decreased (Lee & Jurkevich, 1989). The processing of the multi-look SAR is an illustration of this. In the second category, the filtering techniques are used to remove the

speckle in the spatial domain after an image has been formed (Lee & Jurkevich, 1989). The third category consists of methods for merging SAR imagery with various polarization states and/or frequency bands for speckle reduction.

The present work deals with first and second category of techniques for reducing the speckle noise present in the data. Figure 2.3 shows an output image of the same Sentinel-1 imagery shown in figure 2.2 after the application of refined lee filter for suppressing the speckle noise. The filtered output is not terrain corrected. The earliest version of the refined lee filter was developed more than 3 decades back considering the limited availability of computational power during those times. Recently, this refined filter operates in much larger window (9×9 or higher) for improved speckle reduction (Lee & Pottier, 2017).

2.2.4 Terrain Correction

Images obtained using synthetic aperture radar (SAR) show radiometric image distortions brought on by terrain undulations. An imaging model using the digital elevation model (DEM) is built up to rebuild the imaging geometry and to correct the SAR image's geometry. The terrain corrected image is shown in figure 2.4.

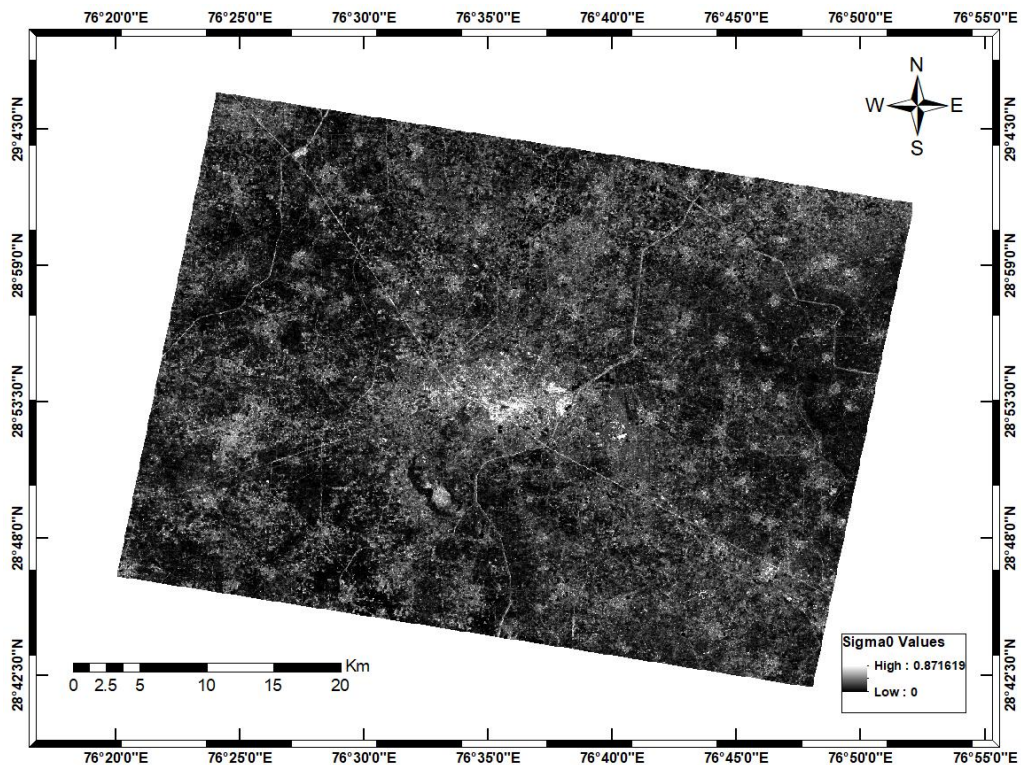


Figure 2.4 σ^0 image corresponding to the VH band of Sentinel-1 data for Rohtak, India after performing terrain correction.

2.2.5 Coherency or Covariance Matrix Generation

When a radar beam with a certain polarisation interacts with the target surface, its polarisation state changes. After this contact, the wave that was reradiated from the target surface will have a response in both the horizontal and vertical polarisations (Hensley & Tighe, 2008). The backscattered wave carrying these responses in each polarisation channel is recorded using the 2*2 scattering matrix represented as

$$[S] = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \quad (2.1)$$

where each component of the matrix reflects the target's backscatter response in a certain polarisation channel. The off-diagonal elements of the matrix are indicative of the cross-polarization information, signifying that the transmitted and received electromagnetic waves possess orthogonal polarizations, whereas the diagonal elements of the matrix correspond to the co-polarization information, indicating that the transmitted and received waves exhibit identical polarization states. The characteristics pertaining to a pure target that manifests a specific scattering mechanism are encapsulated by the scattering matrix. Nonetheless, the properties of the terrestrial environment are frequently more intricate or heterogeneous, presenting a spectrum of scattering responses. In such instances, the insights obtained from the scattering matrix prove insufficient for a comprehensive characterization of the surface's physical attributes. Consequently, the covariance and coherency matrices, which are comprised of second-order statistics derived from the scattering matrix, are employed. These matrices are formulated from the scattering matrix by utilizing the vectorized representation derived from both the Lexicographic and Pauli bases.

Assuming the reciprocity condition ($S_{HV} = S_{VH}$) for the monostatic case, the lexicographic vector form is given by

$$k_L = \begin{bmatrix} S_{HH} \\ \sqrt{2}S_{HV} \\ S_{VV} \end{bmatrix} \quad (2.2)$$

From the lexicographic vector, the coherency matrix is computed as $[C] = k_L k_L^\dagger$ and expressed below

$$[C] = \begin{pmatrix} S_{HH}S_{HH}^* & \sqrt{2}S_{HH}S_{HV}^* & S_{HH}S_{VV}^* \\ \sqrt{2}S_{HV}S_{HH}^* & 2S_{HV}S_{HV}^* & \sqrt{2}S_{HV}S_{VV}^* \\ S_{VV}S_{HH}^* & \sqrt{2}S_{VV}S_{HV}^* & S_{VV}S_{VV}^* \end{pmatrix} \quad (2.3)$$

where $\dagger$ represents the complex conjugate and transpose, $*$ represents the complex conjugate.

$$\text{Similarly, the coherency matrix is obtained as } \langle [T] \rangle = \langle k_p k_p^\dagger \rangle. \quad (2.4)$$

where $\langle \rangle$ represents the mean calculated across the entire dataset, and denotes the Pauli vector provided by

$$k_p = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ 2S_{HV} \end{bmatrix} \quad (2.5)$$

The coherency matrix is a widely used representation of scattering matrix properties of the targets. The coherency matrix obtained from equations (2.4) and (2.5) is given by

$$\langle [T] \rangle = \frac{1}{2} \begin{bmatrix} \langle |S_{HH} + S_{VV}|^2 \rangle & \langle (S_{HH} + S_{VV})(S_{HH} - S_{VV})^* \rangle & 2\langle (S_{HH} + S_{VV})S_{HV}^* \rangle \\ \langle (S_{HH} - S_{VV})(S_{HH} + S_{VV})^* \rangle & \langle |S_{HH} - S_{VV}|^2 \rangle & 2\langle (S_{HH} - S_{VV})S_{HV}^* \rangle \\ 2\langle S_{HV}(S_{HH} + S_{VV})^* \rangle & \langle S_{HV}(S_{HH} + S_{VV})^* \rangle & 4\langle |S_{HV}|^2 \rangle \end{bmatrix} \quad (2.6)$$

2.1 PolSAR decomposition methods

A comprehensive multivariate statistical characterization is crucial for numerous objects of interest in radar remote sensing, due to the interplay between coherent speckle noise and the stochastic vectors arising from scattering effects in both surface and volume contexts. To effectively classify or invert scattering data, it is pertinent to cultivate the concept of an "average" or "dominant" scattering mechanism for such entities (Lee & Pottier, 2017). The notion of a "distributed target," possessing its own inherent structure, emerges from this averaging process, distinguishing it from a "stationary target" or "pure single target." Such an interpretation is intended to be elucidated by "target decomposition theorems," which rely on plausible physical constraints, such as the average target's invariance with respect to alterations in the wave polarization basis (Lee & Pottier, 2017).

The coherency or covariance matrix is used to decompose the radar backscatter into multiple scattering types. By dividing the received signal into a total of distinct scattering processes expressing particular polarimetric characteristics, PolSAR decomposition is used to characterize and categorize scattering from natural and man-made targets (Cloude & Pottier, 1996; Lee & Pottier, 2017). The target decomposition approaches are broadly categorised into the following groups: Coherent decomposition (Cameron & Leung, 1990; Krogager, 1990) and non-coherent decompositions (Huynen, 1970; Barnes, 1988; Freeman & Durden, 1998; Yamaguchi et al., 2005; Singh et al., 2013).

The deterministic assumption behind coherent decompositions (Cameron & Leung, 1990; Krogager, 1990) is that every resolution cell belongs to a single, prevailing scattering mechanism (Karachristos et al., 2024). As a result, these methods try to distinguish this fundamental scattering behavior and relate it to a basic scatterer with an elementary geometric structure. This involves presenting the backscattering matrix S , which is usually measured in the HV basis, as a coherent aggregation of basis matrices, all in alignment with the fundamental scattering technique. The first coherent decomposition was proposed by Pauli. Furthermore, it is advantageous to incorporate the Pauli basis when the objective is to derive the physical significance of the scattering matrix. The following expressions can be used to express any arbitrary 2×2 scattering matrix, assuming the Pauli decomposition basis (López-Martínez & Pottier, 2021):

$$S = \begin{bmatrix} a + b & c - jd \\ c + jd & a - b \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + c \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix} \quad (2.7)$$

On the other hand, non-coherent decomposition techniques (Huynen, 1970; Freeman & Durden, 1998; Yamaguchi et al., 2005) are based on the idea that the target diffuses over several neighboring PolSAR cells due to the impact of speckle noise present in SAR images. Since there isn't a dominating scattering mechanism in each resolution cell in this instance, second-order statistics must be used to retrieve information. The idea of covariance and coherency matrix may be used to identify non-coherent targets, which have intrinsic stochastic behavior (Karachristos et al., 2024). The second order coherency and covariance matrices are then obtained from the first order scattering matrix as discussed in the previous section of this chapter. There are mainly three groups into which the non-coherent decompositions fall. These three types of decompositions are as follows: the first comprises decompositions that utilize

the dichotomy of the Kennaugh matrix; the second kind utilizes eigenvector analysis; and the third type is known as model-based decompositions.

The first formal target decomposition was given by Huynen and it falls under the category of Target Dichotomy decompositions. Huynen proposed a singular plus N-target residue theorem that delineates an arbitrary Mueller matrix $\langle [M_A] \rangle = M_s + M_N$ (Cloude & Pottier, 1996). The research conducted by Huynen (Huynen, 1970) has established the foundational basis for the roll invariant Barnes decomposition (Barnes, 1988). In the context of vector spaces, the fundamental characteristic that TN is roll invariant signifies that the vector space it engenders is orthogonal to the vector space produced by the pure target T0. Moreover, this orthogonality persists even when the line of sight undergoes rotation. The contributions of Huynen and Barnes have been elaborated upon extensively in Chapter 3.

The decomposition based on Eigen-vectors (Cloude & Pottier, 1997) entails the computation of the eigenvectors and eigenvalues of the 3×3 covariance C_3 and coherency T_3 matrices to establish a diagonal configuration as delineated by equation 2.7.

$$T_3 = U_P \Sigma_P U_P^{-1} \text{ and } C_3 = U_C \Sigma_C U_C^{-1} \quad (2.8)$$

where Σ_P and Σ_C represents the 3×3 diagonal matrices with non-negative real elements and U_C and U_P are the 3×3 unitary matrices of the orthogonal eigenvectors of covariance C_3 and coherency T_3 matrices, respectively.

It is imperative to highlight that the eigenvalues associated with covariance C_3 and coherency T_3 matrices are identical, and the eigenvectors are interconnected with $U_P = U_{3(L \rightarrow P)} U_C$.

An eigenvector of the 3×3 Hermitian averaged coherency T_3 matrix can be determined to obtain a set of three uncorrelated targets. Consequently, a fundamental statistical model can be formulated by decomposing T_3 into three distinct targets, $\{T_{0i}\}_{i=1,3}$, each corresponding to a deterministic scattering mechanism related to a singular analogous scattering matrix. The unitary eigenvector u_i signifies the scattering type, while the eigenvalue λ_i elucidates the contribution arising from the deterministic scattering mechanism. This decomposition can be succinctly expressed (Lee & Pottier, 2017):

$$T_3 = \sum_{i=1}^3 \lambda_i u_i u_i^{*T} = T_{01} + T_{02} + T_{03} \quad (2.9)$$

Utilizing a methodology to identify the dominant scattering mechanism through the extraction of the largest eigenvalue (λ_1), Cloude pioneered the investigation of such eigenvector-based decomposition (Cloude, 1985). In this context, the recovered coherency matrix (T_{01}) possesses a rank of 1, has an associated scattering matrix (S), and can be articulated as the outer product of a singular target vector k_1 with

$$T_{01} = \lambda_1 u_1 \cdot u_1^{*T} = k_1 \cdot k_1^{*T} \quad (2.10)$$

The extent of the corresponding scattering matrix is characterized by the sole non-zero eigenvalue λ_1 , which equates to the square of the Frobenius norm of the target vector k_1 (Cloude, 1985). This target vector can be expressed as:

$$k_1 = \sqrt{\lambda_1} u_1 = \frac{e^{j\varphi}}{\sqrt{2A_0}} \begin{bmatrix} 2A_0 \\ C + jD \\ H - jG \end{bmatrix} = e^{j\varphi} \begin{bmatrix} \sqrt{2A_0} \\ \sqrt{B_0 + Be^{+j\arctan(D/C)}} \\ \sqrt{B_0 - Be^{-j\arctan(G/H)}} \end{bmatrix} \quad (2.11)$$

Interestingly, the moduli of this target vector's three components equal those of the three Huynen target generators (Lee & Pottier, 2017).

The Model-based decomposition methodologies (Freeman & Durden, 1998; Yamaguchi et al., 2005; Yamaguchi et al., 2011; Singh et al., 2013; Singh & Yamaguchi, 2018; Bhattacharya et al., 2015) have recently been extensively employed and investigated for numerous applications. The inaugural model-based decomposition was formulated by Freeman and Durden in 1998, incorporating principles of radar physics alongside mathematical constructs (Freeman & Durden, 1998). A three-component scattering model was suggested which used 5 of the 9 covariance matrix elements with the assumption of reflection symmetry for naturally occurring targets. It had a major drawback of over-estimating the volume scattering and under-estimating the even-bounce scattering which gives an inaccurate estimation of the urban area especially.

In order to address the issue of underestimating even-bounce scattering, Yamaguchi et al. (2005) proposed a four-component scattering model in 2005, designating the fourth type of scattering as Helix scattering. This innovative model incorporated six elements from the covariance matrix. Although this decomposition yielded more accurate estimations for urban regions, significant enhancements were still necessary.

After Yamaguchi's model, a new model called MCSM was given by (Zhang et al., 2008). This model introduced one more scattering type called wire scattering. This model works well for some areas but fails for others. It solves the problem of under-estimation of even-bounce scattering but as a side effect, it distorts the vegetation area classifying vegetation as urban area. After this, (Yamaguchi et al., 2011) gave a decomposition model based on a coherency matrix with a small modification. The quantity of independent variables associated with $[T]$ was diminished from nine to eight via the rotational manipulation of the coherency matrix, thereby enhancing the ratio of information retrieved. This small modification of matrix rotation further improved the decomposition results and the problem of under-estimation of urban areas got resolved to a great extent.

A universal four-component decomposition with unitary rotation called S4R was proposed by (Singh et al., 2013), like the Y4R model, S4R integrates coherency matrix rotation, making $ReT23 = 0$ and reducing the depolarized contribution from $T33$. It does, however, differentiate volume scattering associated with vegetation and oriented-dihedral configurations by employing the extended volume scattering model for dihedral structures to improve the differentiation of oriented-urban areas from vegetative cover. Nonetheless, the challenge of identifying scattering from structures within densely populated urban environments remains unresolved, and the present study contributes to addressing this issue.

The PolInSAR-based decompositions (Chen et al., 2014; M E & Kumar, 2020; Shafai & Kumar, 2020; Ramya & Kumar, 2021) are highly effective in accurately characterizing man-made structures, making them a valuable tool in applications like urban mapping, infrastructure monitoring, and disaster management. These methodologies depend on the integration of polarimetric and interferometric Synthetic Aperture Radar (SAR) data, thereby facilitating the acquisition of comprehensive insights regarding both the scattering processes and the elevation or configuration of terrestrial entities. When an InSAR pair of datasets is available, PolInSAR can provide highly accurate insights into land cover classification by effectively differentiating between urban structures, vegetation, and other natural features. However, one significant drawback is that PolInSAR requires planned acquisitions of interferometric data, which involves capturing SAR images from slightly different vantage points, often during consecutive satellite passes. This requirement for a coordinated acquisition schedule limits the availability of the necessary datasets, particularly in regions where satellite tasking is not easily controllable or where frequent revisit times are not possible. As a result, the reliance on

interferometric pairs restricts the widespread application of these otherwise powerful techniques.

2.4 Machine Learning Models

The process of teaching a machine to make autonomous, intelligent judgments through supervised or unsupervised learning is included in machine learning. Using remote sensing data, supervised classification is utilized to produce maps of land cover. Numerous comparison studies for various classification algorithms using optical and SAR data, as well as the fusion of the two types of data, have been conducted (Rawat et al., 2021). SAR imaging was used in comparative research by (Camargo et al., 2019) to classify the land cover of the Brazilian Savana Forest. The findings revealed that SVM, RF, and MLP machine learning algorithms perform better than Naive Bayes and J48, with SVM having the greatest overall accuracy. In contrast, a comparative research analysis pertaining to the Mediterranean Forest region conducted by (Lapini et al., 2020) demonstrated that the Random Forest classifier exhibits superior performance relative to the Support Vector Machine classifier, with the disparate volume of data across classes likely serving as a contributing factor. An evaluation of Land-Use Land-Cover classification utilizing machine learning classifiers on Landsat data indicated that Random Forest represents the most effective methodology for land cover classification, achieving an accuracy rate of 89% (Belgiu & Drăguț, 2016; Talukdar et al., 2020). Furthermore, another comparative investigation of land-cover classification employing optical and Synthetic Aperture Radar (SAR) data revealed that the accuracy of the Landsat-8 classified map was recorded at 91.34%, whereas the accuracy of the SAR classified output was noted to be 93.15% (Ali et al., 2018). Utilizing multitemporal data (Yu et al., 2017) and applying speckle filtering before classification (Gašparović & Dobrinić, 2020) both enhance classification performance. An upscaling method used to classify the land cover in mangrove ecosystems has a 65.5% overall accuracy (Bihamta Toosi et al., 2020). Another work illustrates that by combining multiple polarimetric decomposition outputs across L and S bands significantly enhances land-cover class separability and classification accuracy using Support Vector Machine model achieving classification accuracy of 93.55% (Verma et al., 2023). Another promising framework for the segmentation of PolSAR imagery based on Unsupervised learning approach was introduced using multi-view similarity and clustering over superpixels achieving mIOU of 0.8259 and F1-score of 0.8995 respectively (Li et al.,

2024). This work offers a great alternative to supervised deep-learning models, particularly important in remote sensing applications with limited availability of annotated datasets.

2.5 Deep Learning Models

Deep learning architectures, especially convolutional neural networks (CNNs), possess the capability to autonomously derive feature representations from the input data, thereby obviating the necessity for manual feature extraction. This characteristic presents a notable benefit when analyzing PolSAR data, as it enables the model to assimilate both intricate details and extensive patterns. For example, in urban areas, CNNs can learn to distinguish between buildings, roads, and vegetation based on their unique scattering signatures, improving the accuracy of the segmentation.

The models of Deep Neural Networks (DNN) encompass various categories such as deep Boltzmann machines (Salakhutdinov & Hinton, 2009), Deep Belief Networks (DBNs) (Hinton et al., 2006), Convolutional Neural Networks (LeCun et al., 2010; Chen et al., 2014), and Stacked Auto Encoders (SAEs) (Ma et al., 2015). The capability of DNN models to acquire highly abstract feature sets, which can augment discriminative features while suppressing insignificant variations in the input data, has been evidenced across numerous applications (LeCun et al., 2015). Consequently, the DNN framework propels advancements in the forefront of numerous computer vision and remote sensing domains (Chun Liu et al., 2015; Zhang et al., 2016). In a preceding study, (Lv et al., 2015) established a DBN specifically for synthetic aperture radar (SAR)-based land cover classification within urban environments. Following that, (Liu et al., 2016) created a Wishart DBN to first classify and use local spatial information to enhance the outcomes of the classification. A DNN was used by (Gong et al., 2016) to create a change detection map directly from two shots, showing an increase in detection performance over other conventional approaches. A Deep Convolutional Network (DCN) for target detection in the SAR data was presented by (Chen et al., 2016) which consisted of sparsely connected layers. Another study (Hashemi et al., 2024) underscores promising avenues for leveraging deep neural networks and sensor fusion to enhance segmentation performance in land cover mapping based on PolSAR data. Multi-sensor fusion, data augmentation, pre-processing of data like speckle noise filtering are some of the strategies that can be used to improve the PolSAR segmentation performance when using small amount of labelled data (Shehzad et al., 2025).

A sophisticated deep learning architecture, referred to as Deeplab, was introduced by Chen et al. (2016), which adeptly amalgamated the attributes of Deep Convolutional Neural Networks (DCNN) alongside fully-connected Conditional Random Fields (CRF) to achieve superior performance on the PASCAL VOC dataset. This model not only surpassed the performance of the second best model developed by Mostajabi et al. (2014) by a substantial margin of 7.2%, but it also exhibited enhanced computational efficiency due to the integration of the atrous algorithm within its framework. The Deeplab model underwent considerable enhancement, leading to the introduction of Deeplabv2 (Chen et al., 2018) in 2016. In addition to its efficacy in localizing object boundaries, this model effectively combined the outputs of DCNN with fully-connected CRFs and introduced atrous spatial pyramid pooling (ASPP) to facilitate effective segmentation of objects across multiple scales. To effectively capture objects and contextual visual information at varying levels, ASPP investigates an input convolutional feature layer utilizing filters with diverse sampling rates and effective fields-of-view (Chen et al., 2018). The Deeplabv3, as proposed by Chen et al. (2017), demonstrated enhanced performance relative to its predecessor Deeplab models without the need for the incorporation of dense CRFs. This work proposed that by applying atrous convolution in parallel it is possible to capture multi-scale context with the help of various atrous rates. Also, the ASPP proposed in earlier version of Deeplab is augmented with image level features that gives further boost to the performance by encoding the global perspective (Chen et al., 2017). An encoder-decoder network using Xception as its backbone and layers for convolution, pooling, and filtering are used in the DeepLabv3+ (Chen et al., 2018) model. In this model, the Xception backbone network's features are retrieved, and atrous convolution is employed in the last stage of the backbone to regulate the size of the feature map. Because it outperforms other designs like RESNET-101 for test datasets like ImageNet (Russakovsky et al., 2015), Xception has been selected as the backbone in this case. Altogether, using PolSAR image data, the DeepLabv3+ model can extract intricate textural patterns and polarimetric relevant features.

Researchers within the domain of remote sensing have exerted considerable effort to assimilate land use and land cover classification through the application of deep learning techniques, subsequent to the notable advancements in deep learning methodologies observed in the discipline of computer vision. Both users and the remote sensing community are afforded a more profound understanding of the current landscape by consulting land use and land cover maps corresponding to the area of interest. This facilitates the monitoring of temporal variations in forest cover, surface water bodies, agricultural systems, and biomass

quantification, thereby mitigating the economic and human impacts resulting from floods and other natural calamities. The continuous collection of extensive datasets by multiple satellite sensors engenders a repository of big data. Consequently, there exists a pressing necessity for sophisticated deep learning models that are both computationally efficient and cost-effective to extract pertinent information from this extensive dataset and to categorize the data in accordance with specific requirements. The deep learning models are highly scalable and can be applied to large datasets without requiring significant modifications. This is particularly important for PolSAR data, which is often collected at large scales and contains millions of data points. Deep learning models can handle this complexity and scale, making them a practical solution for PolSAR segmentation tasks.

In recent times, an extensive array of applications utilizes methodologies rooted in machine learning and deep learning techniques. This is due to the fact that these algorithms are relatively general and excel in applications involving data analytics (Yang & Shami, 2020). Only after adjusting a machine learning or deep learning model's hyperparameters can it reach its full potential (Elshaw et al., 2019). No single solution works for all subjects since each one is distinct in its own right. For various problems or datasets, the same method may have various hyperparameter optimal values. By applying the same amount of optimization to various machine learning models, hyperparameter optimization also aids in meaningful comparative analysis.

Hyperparameter optimization is a complex problem that involves a large search space and evolutionary algorithms are well suitable for such problems (Yao et al., 2018). One of the most frequently used evolutionary algorithms for hyperparameter optimization is the Genetic Algorithm (Lessmann et al., 2005).

CHAPTER 3

Dataset Generation and Augmentation to support Semantic Segmentation on PolSAR images

3.1 Mumbai Region

Being a world city, Mumbai has grown into a sprawling urban city with diverse features ranging from urban buildings to natural targets which makes it an ideal location for the case study. Mumbai boasts of holding the title of “Financial Capital of India” due to its unfathomable influence on the Indian economy. The Municipal Corporation of Greater Mumbai (MCGM) estimated that as of the middle of 2016, Mumbai city comprised of 1,26,89,644 residents. The Mumbai Metropolitan Region was established to build a large metropolitan cluster that encompasses over 4355 square kilometers and is made up of 7 Municipal Corporations and 13 Municipal Councils. Mumbai city has 227 municipal wards altogether. Due to its proximity to the sea, the city's soil is primarily sandy, and it is categorized as a Seismic Zone III region (Brihanmumbai Municipal Corporation, 2021). The efficacy of machine learning algorithms in the categorization of land use and land cover was investigated utilizing PolSAR data obtained from the Advanced Land Observing Satellite-2 (ALOS-2) Phased Array type L-band Synthetic Aperture Radar (PALSAR-2), TerraSAR-X, Radar Imaging Satellite-1 (RISAT-1), and Sentinel-2 satellites belonging to the copernicus programme satellite constellation administered by the European Space Agency. Sentinel-2 Level-1C data utilized in this study was obtained on March 29, 2016. (descending pass). Figure 5 depicts the location of the research region. Figure 3.1(a) illustrates a representation of India's terrestrial coverage, which was developed utilizing the Global Land Cover dataset at a spatial resolution of 100 m, sourced from the Copernicus Global Land Service (CGLS) (Buchhorn et al., 2021). The PROBA-V dataset, derived from near real-time epoch 2019 global land cover, encompasses 23 distinct classifications, one of which is the category denoted as "no input data available" (Buchhorn et al., 2021). The class "no input data available" was eliminated from the land cover map of India, leaving a total of 20 distinct classes, as shown in figure 3.1(a). Figure 3.1 displays the state of Maharashtra's boundary in black (a).

Figure 3.1(b) depicts the state of Maharashtra's land cover map with the Mumbai Metropolitan Region (MMR) highlighted in a light yellow polygon. A portion of the Mumbai Metropolitan Region (MMR) was scanned using ALOS-2 PALSAR-2 full-pol SAR data. Using

Freeman-Durden decomposition-based output, figure 3.1(c) displays a false-color composite image of the research region. Figure 3.1(c) highlights in green, the targets that cause volume scattering. These scatters, which are the sources of strong backscattering, are shown in red as double-bounce or even-bounce scatters. Figure 3.1(c) shows the surface scattering, also known as odd-bounce scattering, as blue. Table 2 lists the specifications of the ALOS-2/PALSAR-2, TerraSAR-X, and RISAT-1 data.

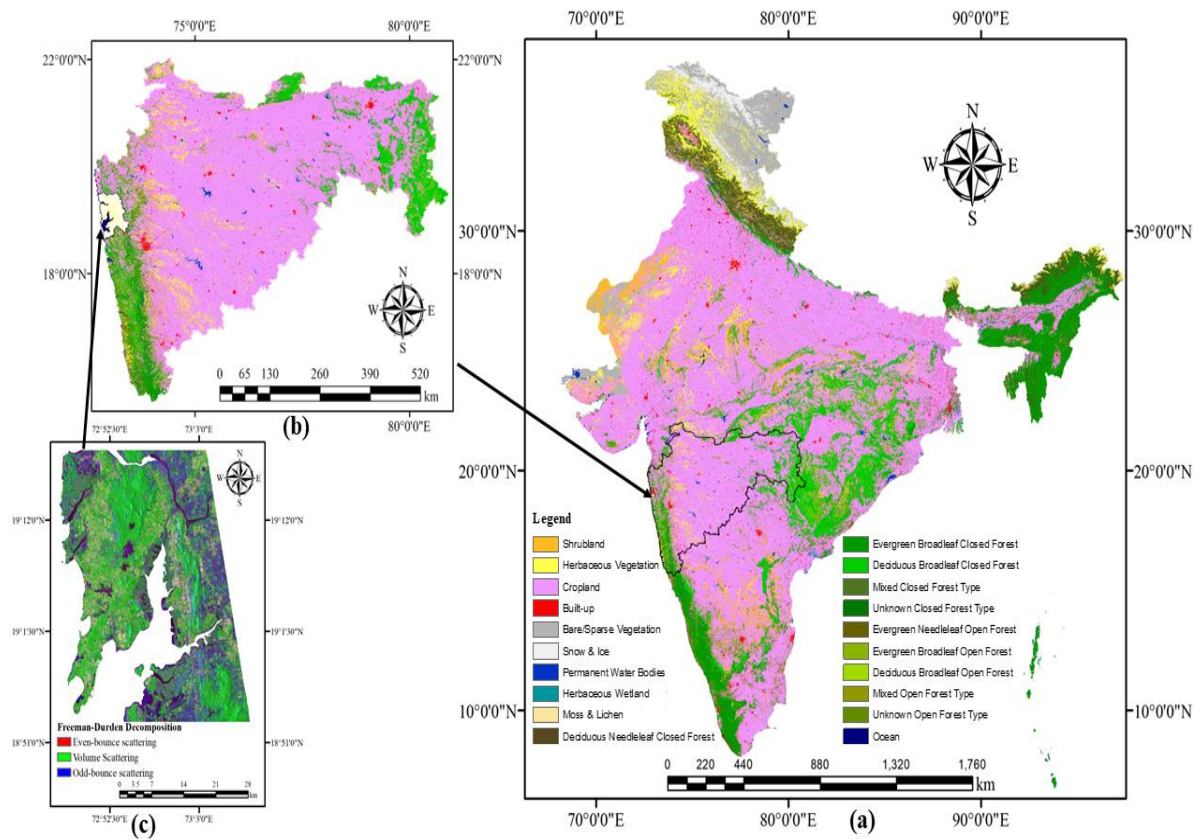


Figure 3.1 Location of the study area (a) Land cover map of India with a resolution of 100 m derived from PROBA-V data for the year 2019 (b) Land cover map of Maharashtra with a light yellow color polygon representing the Mumbai Metropolitan Region (MMR) (c) Freeman-Durden decomposition-based color composite image of ALOS-2 PALSAR-2 data for Mumbai (Garg et al., 2021).

Table 2. Details of multi-frequency SAR datasets used in current work

SAR system parameters	ALOS-2/PALSAR-2	TerraSAR-X	RISAT-1
Platform	Spaceborne	Spaceborne	Spaceborne
Satellite	Advanced Land Observing Satellite-2 (ALOS-2)	TerraSAR-X	Radar Imaging Satellite (RISAT-1)
Sensor	Phased Array type L-band Synthetic Aperture Radar (PALSAR-2)	SAR	SAR

Date of Acquisition	14 April 2016	11 March 2015	14 November 2012
Pass	Ascending	Descending	Ascending
Altitude	628 km	514 km	536 km
Data Type	Single Look Complex (SLC) Fully Polarimetric	Complex Fully Polarimetric	Single Look Complex (SLC)
Polarisation	Quad-pol (HH+HV+VH+VV)	Quad-pol (HH+HV+VH+VV)	Dual-pol (HH+HV)
Mode of Acquisition	Stripmap mode	Stripmap mode	FRS1
Frequency	1.236 GHz	9.65 GHz	5.35 GHz
Look direction	Right	Right	Left

3.2 Workflow of image segmentation task on Mumbai region dataset

The workflow adopted for the image segmentation of the multi-frequency PolSAR data and the optical multi-spectral data is presented in the figure 3.2. The steps include pre-processing of the SLC product and PolSAR decomposition to get the pseudo-color image in case of SAR data and RGB image generation in case of optical data. Once, a colored image is obtained, the next steps are common for both SAR and Optical data which include dataset

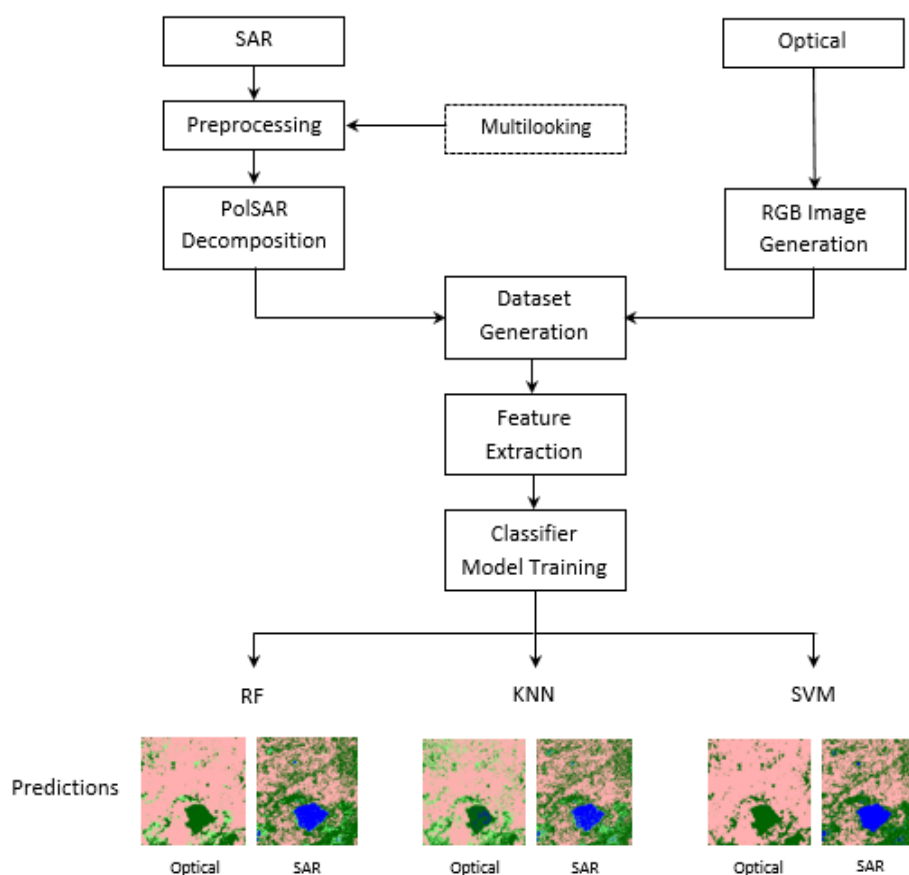


Figure 3.2 The flow chart of the methodology adopted for image segmentation using machine learning algorithms (RF= Random Forest, KNN = K-Nearest Neighbor, SVM = Support Vector Machine).

generation, feature extraction, and model training. The steps involved are discussed in detail ahead.

3.2.1 Pre-Processing

Radiometric calibration and multi-looking on SAR data are often carried out during pre-processing. But, only multi-looking is carried out because the accessible dataset utilized in this work was already calibrated. Since SAR is a side-looking sensor, the near-range objects appear compressed in comparison to the far-range ones due to slant-range distortion. As a result, the size of the acquired image changes from close to distant. Square pixels are produced by a method called multi-looking. The image statistics determine the number of looks. The Alos-2 SAR data is multi-looked to generate square pixels of 5.78m in size. No pre-processing technique is implemented on the TerraSAR-X, RISAT-1, and Sentinel-2 datasets.

3.2.2 PolSAR decomposition

The PolSAR decomposition is employed to characterize and categorize the scattering from both natural and artificial targets by dividing the received signal into a total of distinct scattering processes that each display certain polarimetric properties (Cloude & Pottier, 1996)

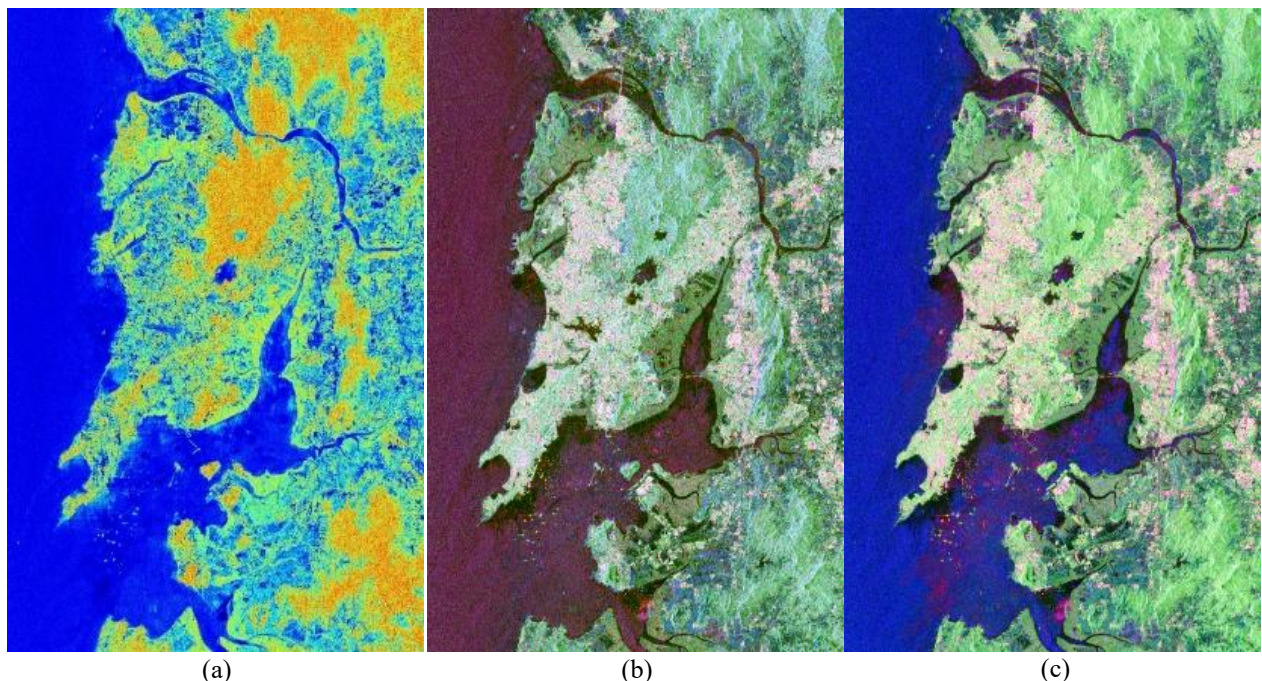


Figure 3.3 Representing the different decomposition results as presented by (Garg et al., 2021) of (a) Entropy based decomposition [32], (b) Huynen decomposition [11], and (c) Barnes decomposition [38].

(Lee & Pottier, 2017). In this study, quad-pol L-band ALOS-2/PALSAR-2 data and X-band TerraSAR-X data are divided into several scattering power matrices using a roll-invariant

Barnes decomposition to create a false-color composite RGB image. Huynen's work (Huynen, 1970) has been the main ground for the Barnes decomposition. The results of Huynen decomposition and Barnes decomposition are shown in figure 3.3b and 3.3c respectively.

Huynen gave a single plus N-target residue theorem which expresses an arbitrary Meuller matrix $[M_A]$ as expressed in equation 1 (Cloude & Pottier, 1996).

$$\langle [M_A] \rangle = M_S + M_N = \begin{bmatrix} A_0 + B_0^S & C & H & F^S \\ C & A_0 + B^S & E^S & G \\ H & E^S & A_0 - B^S & D \\ F^S & G & D & B_0^S - A_0 \end{bmatrix} + \begin{bmatrix} B_0^N & 0 & 0 & F^N \\ 0 & B^N & E^N & 0 \\ 0 & E^N & -B^N & 0 \\ F^N & 0 & 0 & B_0^N \end{bmatrix} \quad (1)$$

The variables A_0 , C , H , G , and D are held constant, while the remaining variables are selected in such a manner that the stipulation requiring the $[M_S]$ to possess an equivalent rank one coherency matrix is fulfilled (Cloude & Pottier, 1996). The Huynen decomposition, when articulated in terms of the coherency matrix $[T]$, can be expressed through equation 2.

$$[T] = \begin{bmatrix} 2A_0 & C - jD & H + jG \\ C + jD & B_0 + B & E + jF \\ H - jG & E - jF & B_0 - B \end{bmatrix} = [T_0] + [T_N] \quad (2)$$

where $[T_0]$ = Pure Target, $[T_N]$ = N-Target

For T_N to be roll-invariant, equation 3 should be satisfied.

$$T_N(\theta)q = 0 \Rightarrow U(\theta)T_N U(\theta)^{-1}q = 0 \quad (3)$$

where $U(\theta)$ is a 3*3 rotation matrix, q is the Eigenvector of $U(\theta)^{-1}$ as shown in equation 4 (Maghsoudi, 2012).

$$so, q_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, q_2 = \begin{bmatrix} 0 \\ 1 \\ j \end{bmatrix}, q_3 = \begin{bmatrix} 0 \\ j \\ 1 \end{bmatrix} \quad (4)$$

The normalized target vector k_0 associated with each of the above eigenvectors can be determined through equation 5 (Maghsoudi, 2012):

$$Tq = T_0q + T_Nq = T_0q = k_0 k_0^T * q \quad \left. \vphantom{\begin{matrix} Tq = T_0q + T_Nq = T_0q = k_0 k_0^T * q \end{matrix}} \right\} \quad 37$$

$$\Rightarrow K_0 = \frac{Tq}{k_0^T * q} = \frac{Tq}{\sqrt{q^T * qT}} \quad (5)$$

$$q^T * qT = q^T * k_0 k_0^T * q = |k_0^T * q|^2$$

Substituting equations 2 and equation 4 in equation 5, we can obtain the three normalized target vectors as below (Maghsoudi, 2012):

$$k_{01} = \frac{\frac{\langle [T] \rangle_{q_1}}{\sqrt{q_1^T * \langle [T] \rangle_{q_1}}}}{\frac{1}{\sqrt{\langle 2A_0 \rangle}}} \begin{bmatrix} \langle 2A_0 \rangle \\ \langle C \rangle + j \langle D \rangle \\ \langle H \rangle - j \langle G \rangle \end{bmatrix} \quad (6)$$

$$k_{02} = \frac{\frac{\langle [T] \rangle_{q_2}}{\sqrt{q_2^T * \langle [T] \rangle_{q_2}}}}{\frac{1}{\sqrt{2(\langle B_0 \rangle - \langle F \rangle)}}} \begin{bmatrix} \langle C \rangle - \langle G \rangle + j \langle H \rangle - j \langle D \rangle \\ \langle B_0 \rangle + \langle B \rangle - \langle F \rangle + j \langle E \rangle \\ \langle E \rangle + j \langle B_0 \rangle - j \langle B \rangle - j \langle F \rangle \end{bmatrix} \quad (7)$$

$$k_{03} = \frac{\frac{\langle [T] \rangle_{q_3}}{\sqrt{q_3^T * \langle [T] \rangle_{q_3}}}}{\frac{1}{\sqrt{2(\langle B_0 \rangle + \langle F \rangle)}}} \begin{bmatrix} \langle H \rangle + \langle D \rangle + j \langle C \rangle + j \langle G \rangle \\ \langle E \rangle + j \langle B_0 \rangle + j \langle B \rangle + j \langle F \rangle \\ \langle B_0 \rangle - \langle B \rangle + \langle F \rangle + j \langle E \rangle \end{bmatrix} \quad (8)$$

where A_0 , B , B_0 , C , D , E , F , G , and H are the same parameters as used in Hyunen's Muller Matrix mentioned in equation 1.

In case of Barnes decomposition (Barnes, 1988), 3 single target vectors k_{01} , k_{02} , and k_{03} are calculated using equation 6, equation 7, and equation 8. This decomposition has been chosen in preference to other roll invariant parameter-based decomposition techniques because, as shown in figure 3.4, it provides a more accurate assessment of urban and water targets. Figures 3.5b and 3.5d display the Barnes decomposition findings for test patches.

However, a three-component compact decomposition (Charbonneau et al., 2010) based on the initial stokes component (S_0), the relative phase angle (δ), and the degree of polarization (m) as described in equation 9 has been utilized to examine the compact-pol RISAT-1 data.

$$\begin{bmatrix} V_R \\ V_G \\ V_B \end{bmatrix} = \begin{bmatrix} \sqrt{S_0 m \frac{(1 - \sin \delta)}{2}} \\ \sqrt{S_0 (1 - m)} \\ \sqrt{S_0 m \frac{(1 + \sin \delta)}{2}} \end{bmatrix}, \quad \delta(deg) = \tan^{-1} \left(\frac{S_3}{S_2} \right), \quad m = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \quad (9)$$

where, S_0 , S_1 , S_2 , and S_3 are Stokes components, δ is the relative phase angle and m is the degree of polarization.

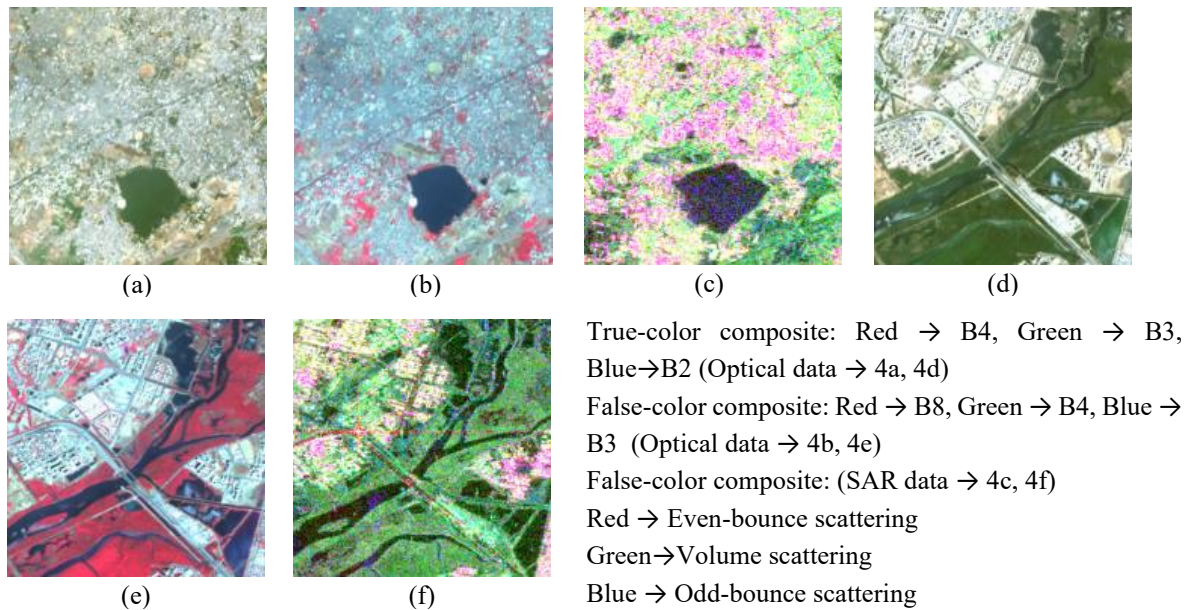


Figure 3.4 Representing optical images that are based on true colour composites, false colour composites, and the results of Barnes decomposition of various areas in Mumbai as presented by (Garg et al., 2021). (a) Patch1's (Varaladevi Talao, Bhiwandi, Maharashtra) true-color composite based optical image, (b) false-color composite based optical image, and (c) Barnes decomposition based false-color composite SAR image, and (d) Patch 2's (Taloje River, Navi Mumbai, Maharashtra) true-color composite based optical image, (e) false-color composite based optical image, and (f) Barnes decomposition based false-color composite SAR image respectively.

3.2.3 Dataset Generation

The target decomposition of the ALOS-2 SAR data produces a false-color composite image that is divided into smaller patches, each 500×500 in size, for comparative evaluation. 10 patches are annotated among these to provide a train and test dataset using GIMP 2.8 tool. Similarly, for Sentinel-2 optical data, false color composite and true color composite RGB images are created and then smaller patches of 500×500 pixels are cut. For ensuring spatial and temporal consistency, google earth imagery from the same period was used for generating accurate annotations.

Figure 3.4 displays the two test patches used to evaluate the models, and figure 3.5 displays an example image and label from the training dataset for both optical and SAR data. Eighty percent of the entire dataset is employed for the purpose of training, whereas the residual twenty percent is designated for testing and validation.

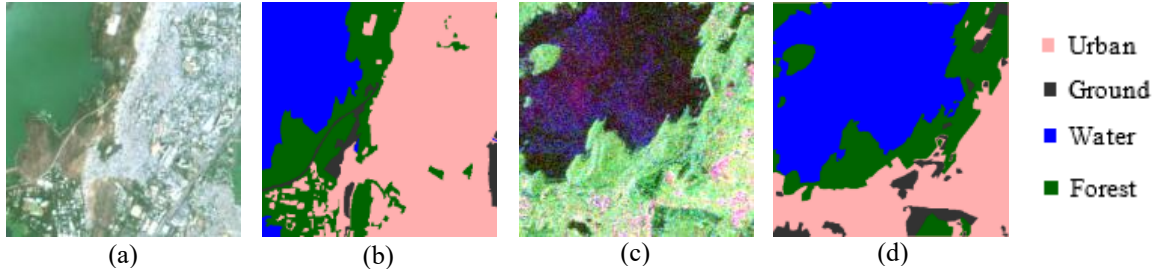


Figure 3.5 Displaying the input image and related processed and annotated images of Mumbai's Vihar Lake, near Powai as presented by (Garg et al., 2021). (a) Sentinel-2 true-color composite image patch and (b) its corresponding label; (c) SAR image patch and (d) its corresponding label.

3.2.4 Feature Extraction

Following annotation and dataset creation, feature extraction is a vital stage in putting any machine learning algorithm into practice since the effectiveness of the chosen features governs how well the model performs. In this study, in addition to the three primary RGB bands, feature vectors for the classifier's training are produced using the Gabor, Scharr, Gaussian, and Median filters. These characteristics are created to separate the textual data from the picture for thorough classifier model training. The sixteen separate Gabor features were produced by adjusting the Gabor filter's settings. For both the optical and the SAR datasets, the ML models are trained using a total of 23 features.

3.3 Houston Region

Another dataset employed in this research was procured from NASA-JPL's publicly accessible platform (Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR), 2021). This dataset comprises L-band fully polarimetric data derived from the 2 m-resolution UAVSAR aerial sensor. The target area of Houston City, illustrated in figure 3.6, was selected due to its representation of a variety of land cover characteristics. Located in the state of Texas, Houston constitutes a significant urban center. The predominant soil types in this region are prairie soils, which are primarily clayey and loamy in composition. The efficacy of machine learning techniques for semantic segmentation applications was assessed utilizing the L-band Uninhabited Aerial Vehicle Synthetic Aperture Radar data. The UAVSAR fully polarimetric SAR sensor scanned a portion of the Houston metropolitan area.

A false-color composite image of the research area's G4U decomposition-based output is shown in figure 3.6b. The elements that contribute to volume scattering are highlighted by the color green. The surface scattering, or odd-bounce scattering, is depicted in blue, while

double bounce or even-bounce scatters, which lead to strong backscatter, are indicated in red. Table 3 lists the specifications of the UAVSAR equipment and the data utilized.

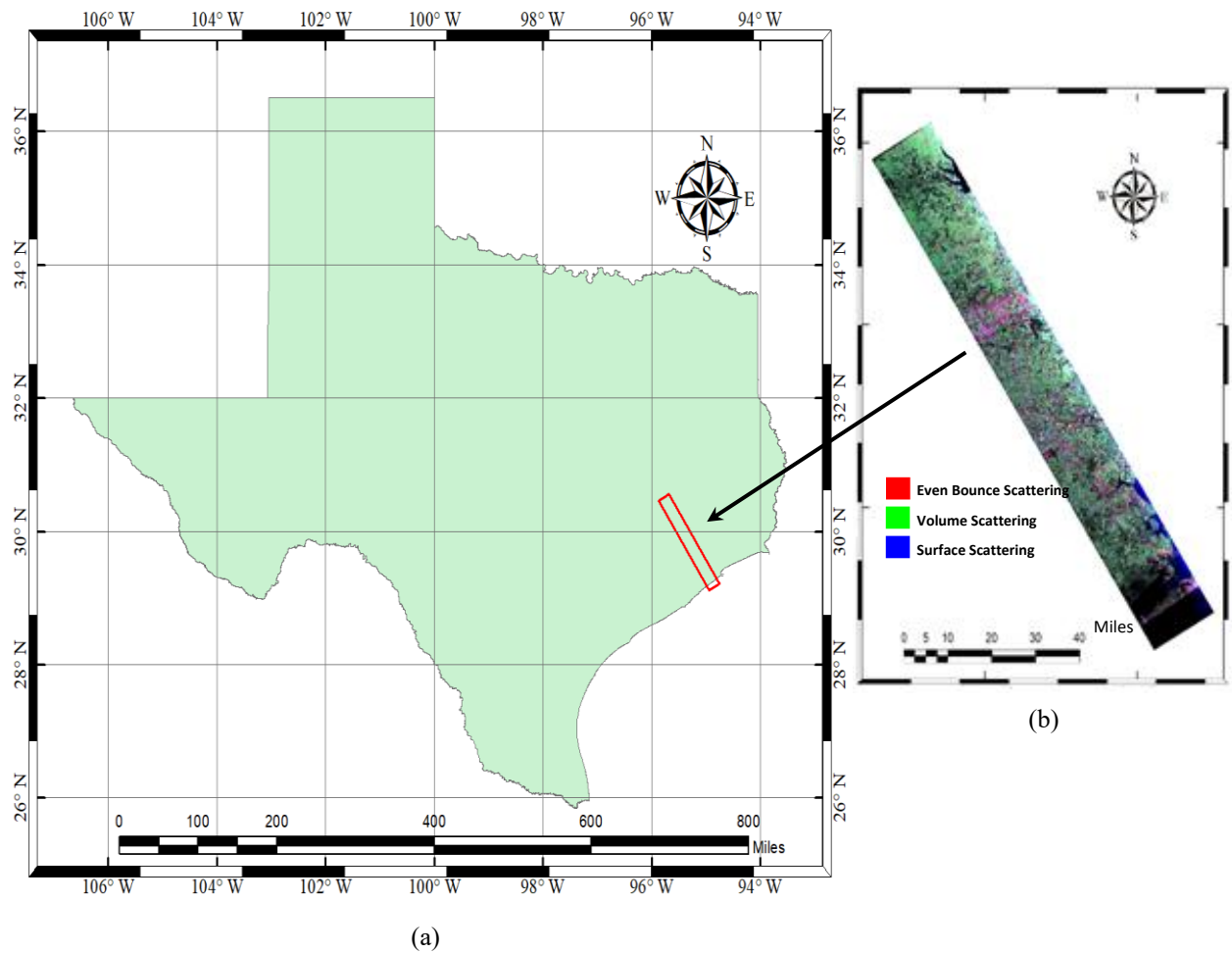


Figure 3.6 Represents the research area's position as compiled by using ArcGIS 10.6. (a) A map of Texas's border (available at: <https://gis-txdot.opendata.arcgis.com/>). (b) the PolSARPro v6.0 opensource remote sensing tool-generated G4U decomposition based false-color composite imagery of UAVSAR data over a portion of Houston, Texas (USA) (Garg et al., 2021b).

Table 3. Details of UAVSAR data and instrument

SAR system parameters	Details
Platform	Airborne
Flight	Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR)
Date of Acquisition	2 Sep 2017
Acquisition Mode	PolSAR
Frequency	1.26 GHz
Wavelength	23.79 cm
Bandwidth	80 MHz

Pulse Duration	5-50 μ s
Transmit Power	3.1 kW
Look angle range	25° - 65°
Data Type	Multi Look Complex (MLC)
Polarisation	Fully Polarimetric Quad-pol (HH+HV+VH+VV)
Swath width	16 km
Range resolution (Meters)	4.99 m (MLC product)
Azimuth resolution (Meters)	7.2 m (MLC product)

3.4 Workflow of image segmentation using Houston region data

The Ortho-Rectified Quad-Pol data from UAVSAR airborne sensor is available as the covariance matrix (C3). This matrix is then transformed into a coherency matrix (T3). Speckle noise generally affects the SAR dataset, which causes issues with object recognition and classification. Different specialized filters that were created primarily to reduce the speckle effect from polarimetric SAR data are utilized to reduce speckle from PolSAR data. One of the most popular and effective speckle filters, the Lee filter (Lee, 1980) is used in this study for speckle filtering of the airborne sensor data. After speckle-filtering, PolSAR decomposition is used to build a false-color composite image, which is then divided into smaller patches to form a training dataset for the ML models.

One of the earliest model-based decompositions that was well accepted was proposed by (Freeman & Durden, 1998). The SAR backscatter was divided by this model into three distinct forms of scattering: odd-bounce scattering, double-bounce scattering, and volume scattering. Later, (Yamaguchi et al., 2005) suggested an enhanced model with the addition of Helix scattering as a new form of scattering. The general 4-component decomposition method with a unitary transformation that utilizes the full polarimetric information in decomposition was proposed by (Singh et al., 2013) based on the Yamaguchi decomposition. The T_{13} component of the coherency matrix was not utilized in earlier models, but this approach permitted its use. Additionally, it had an expanded volume scattering model that distinguishes between dipole and dihedral volume scattering generated by the cross-polarized HV component. Compared to the current model-based decompositions, this resulted in improved double-bounce scattering from urban targets. This is the rationale for the usage of the G4U decomposition approach in the current work.

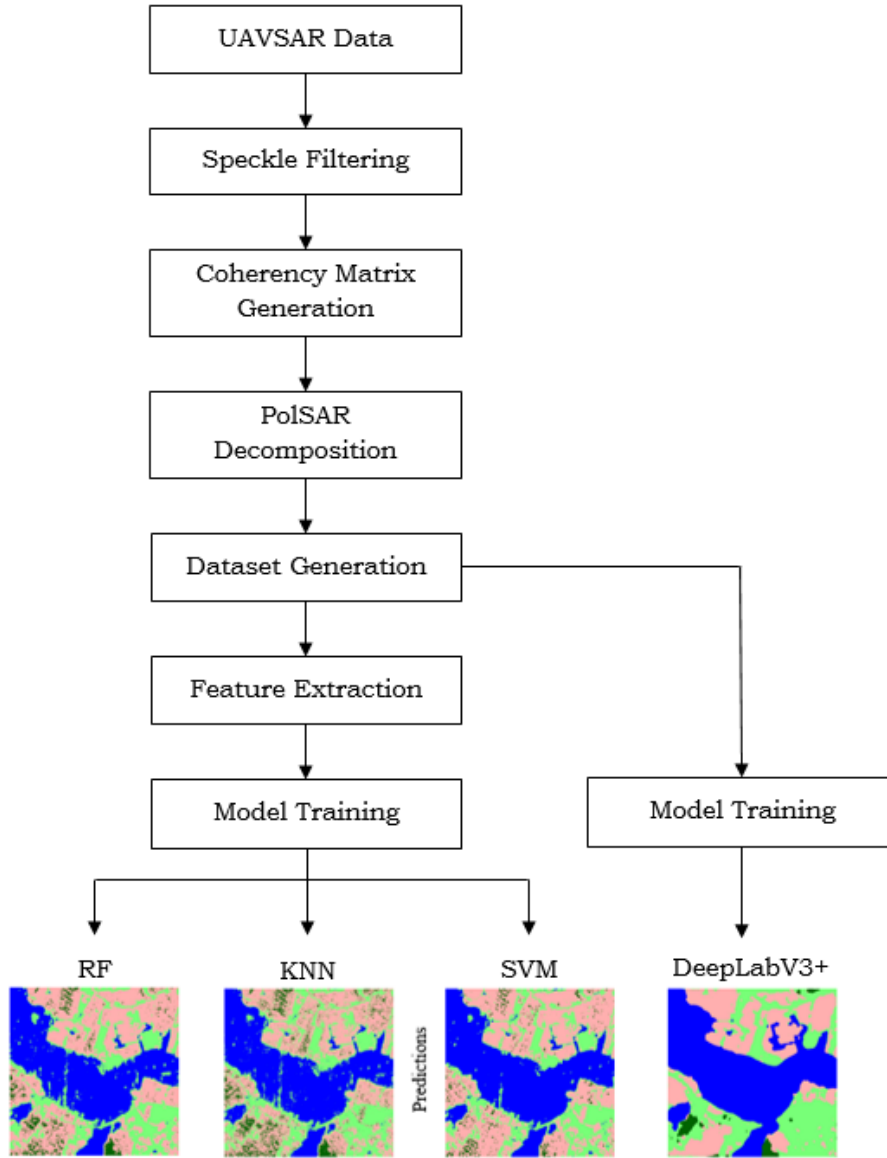


Figure 3.7 Represents the work flow of the methodology adopted for image segmentation using traditional machine learning algorithms and DeepLabv3+ model (Garg et al., 2021b).

Four component decomposition is expressed in equation (17)

$$\langle [T] \rangle = P_s[T]_s + P_d[T]_d + P_v[T]_v + P_c[T]_{\text{helix}} \quad (17)$$

where P_s , P_d , P_v , and P_c are scattering powers to be determined. $[T]_s$, $[T]_d$, $[T]_v$, and $[T]_{\text{helix}}$ are expansion matrices corresponding to surface, double-bounce, and helix scattering respectively.

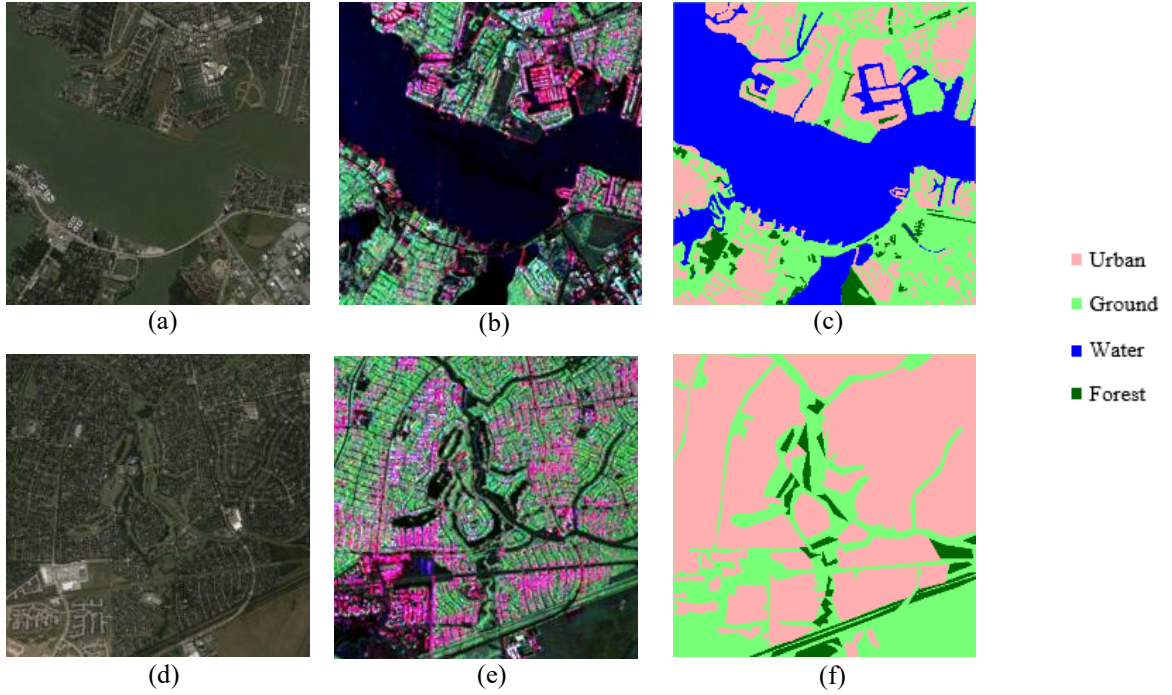


Figure 3.8 Represents the ground truth image, false-color composite image, and labeled image of the patches of Houston City, Texas(USA) as presented in (Garg et al., 2021b) (a) Google Earth image, (b) The false-color composite image, and (c) Ground truth label for Patch1, (d) Google Earth image, (e) The false-color composite image of G4U decomposition, and (f) Ground truth label for Patch2.

With the help of the Gimp 2.8 tool, the false-color composite image is fragmented into smaller segments of 500×500 pixels, and 20 patches are annotated to form a dataset for the DeepLabv3+ model and machine learning algorithms. For ensuring spatial and temporal consistency, google earth imagery from the same period was used for generating accurate annotations. Two of these 20 patches were maintained for testing and validation, while the other 18 were utilized for training. For the deployment of machine learning models, feature extraction is carried out manually after generating the necessary dataset using the Gabor, Gaussian, Canny edge detector, and Median filters. However, in the case of DeepLabv3+, feature extraction is carried out automatically by the Xception-based encoder. Figure 3.7 presents the whole procedure used for the job of semantic segmentation. Figure 3.8 illustrates the Google image, the decomposed false-color composite, and the associated annotated label for the test patches, respectively.

The Gabor filter is represented by equation (18) which was originally given by (Gabor, 1946).

$$g(x, y, T, \emptyset) = h_x(x; T, \emptyset) \cdot h_y(y; \emptyset) = \left\{ \exp \left| \left(\frac{x_\emptyset^2}{2\sigma_x^2} \right) \right| \cos \left| \left(\frac{2\pi x_\emptyset}{T} \right) \right| \cdot \left\{ \exp \left| \left(-\frac{y_\emptyset^2}{2\sigma_y^2} \right) \right| \right\} \right\} \quad (18)$$

where h_x – Bandpass filter, h_y – Gaussian filter

The Gaussian filter (Deng & Cahill, 1993) is represented by equation (19)

$$G(x, y) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-(x^2 + y^2)/2\sigma^2) \quad (19)$$

where σ^2 is the variance of the Gaussian filter, and l ($-l \leq x, y \leq l$) is the size of the filter kernel

The imagery is initially subjected to Gaussian filtering in the implementation of the Canny edge detector (Dong et al., 2013; Canny, 1986), after which the magnitude and direction of the gradient are determined using the below equation (20)

$$G(i, j) = \sqrt{G_x^2(i, j) + G_y^2(i, j)} \quad , \quad \theta(i, j) = \arctan \frac{G_x(x, y)}{G_y(x, y)} \quad (20)$$

where $G_x(i, j)$ and $G_y(i, j)$ are the partial-derivative at point (i, j) in x-direction and y-direction respectively.

CHAPTER 4

IMPLEMENTATION OF MACHINE LEARNING FOR SEMANTIC SEGMENTATION OF MULTIFREQUENCY SAR AND OPTICAL MULTISPECTRAL DATA

4.1 Machine Learning

The technique of making a computer capable of making its own informed judgments is known as machine learning. While there exists a plethora of both supervised and unsupervised algorithms tailored for a multitude of specific applications across various domains, it is noteworthy to mention that amongst these, the Random Forest, Support Vector Machine, and K-Nearest Neighbor algorithms stand out as the most prevalently utilized methodologies for addressing classification challenges within the realm of machine learning.

4.1.1 Random Forest

Breiman (Breiman, 2001) made the initial suggestion for a random forest classifier in 2001. It is based on Breiman's own Bagging predictors (Breiman, 1996) proposal. A random forest comprises a multitude of "n" decision trees, and it derives the final output of the model through a process of majority voting, thereby adhering to the same underlying principles that govern the operation of a singular decision tree. Because each tree functions independently of the others, the algorithm may be executed in parallel. The Random Forest classifier requires a substantial training sample despite being extremely resistant to outliers and noise. Assuming a dataset $\{r_1, r_2, \dots, r_n\} \in \mathbb{R}^{L \times n}$, where L denotes the total number of features present in the dataset and n signifies the total number of samples collected. While z contributes the connection between r_i and r_j , where $z \in Z = \{1, -1\}$, r_i indicates the position of sample i within the multidimensional space $\mathbb{R}^{L \times n}$. When r_i and r_j are members of the same class, $z = 1$, and when they are members of different classes, $z = -1$. The ensemble for the m^{th} tree is $f_p(r) = f(r, \theta_m)$, and a tree $f(r, \theta_m)$ is constructed in conjunction with the training dataset and the stochastic random vector θ_m , which encapsulates the various random components inherent to the structure of the tree. Furthermore, the collection of $\{\theta_m\}$ is characterized by its independence and identical distribution (Breiman, 2001).

The estimated probability that is utilized for predicting the class z associated with a given sample is delineated in equation (10),

$$P(z|r) = \frac{1}{M} \sum_{m=0}^M P_m(z|r) \quad (10)$$

where $P_m(z|r)$ represents the estimated density function of the class labels that corresponds to the m^{th} decision tree within the ensemble, and M denotes the total number of trees that comprise the random forest.

The decision function (Dong et al., 2015) of forest is defined in Eq. (11),

$$C(r) = \arg \max_{j \in Z} P(j|r) \quad (11)$$

The margin function (Dong et al., 2015) is defined in Eq. (12),

$$ml(r,z) = P(z|r) - \max_{j \in Z, j \neq z} P_m(j|r) \quad (12)$$

where $j \neq z$ and if $ml(r,z) > 0$, we can obtain a correct result.

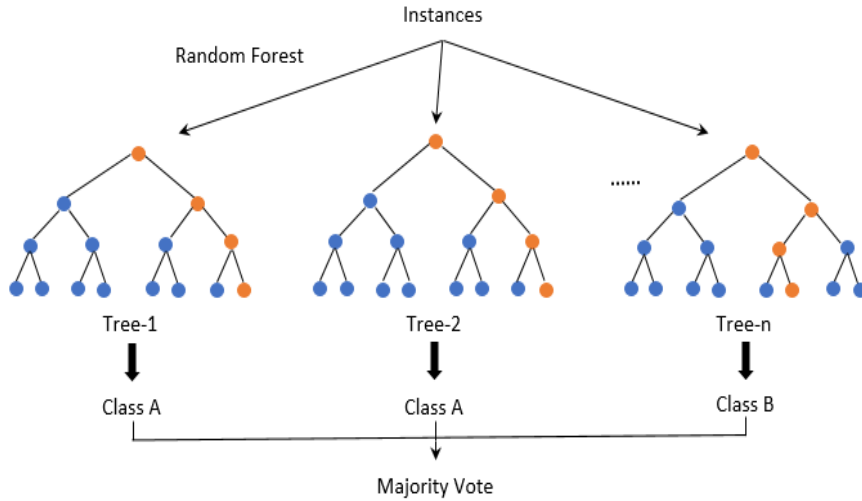


Figure 4.1 Random Forest Classifier

4.1.2 K-Nearest Neighbor

The nonparametric classifier K-Nearest Neighbor has proven to be an effective machine learning technique (Cover and Hart, 1967). In accordance with the principles underlying the Nearest Neighbor rule, if one considers a particular set comprising n pairings designated as $(x_1, \theta_1), \dots, (x_n, \theta_n)$ it is essential to recognize that the x_i 's take values in a metric space designated as X , wherein a specific metric d has been established, while the θ_i 's values correspond to elements within the set $\{1, 2, \dots, M\}$. Every element within the θ set is

conceptualized as an index that denotes the specific category to which the i^{th} individual is affiliated, and concomitantly, each corresponding x_i element represents the outcome of a series of measurements that have been conducted on that individual. In the context of introducing a new pair characterized as (x, θ) , where x is the measurable quantity and θ necessitates estimation utilizing data extracted from a collection of meticulously identified locations. Then $x' \in \{x_1, x_2, \dots, x_n\}$ is recognized as the nearest neighbor to x if it satisfies the equation (13),

$$\min d(x_i, x) = d(x_n', x); \quad i = 1, 2, \dots, n. \quad (13)$$

Based on the stipulations of the Nearest Neighbor rule, it follows that x is categorized under the same category θ_n' as its nearest neighbor, x_n' . If $\theta_n' \neq \theta$ holds, a mistake happens. The Nearest Neighbor rule solely uses the nearest neighbor categorization and ignores the $n-1$ remaining classifications θ_i .

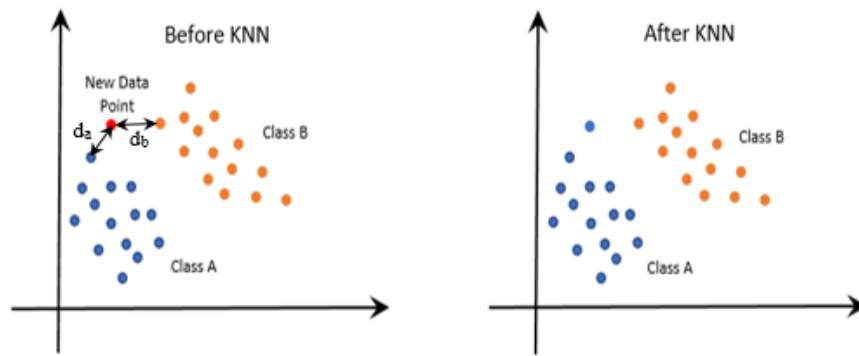


Figure 4.2 KNN Classifier

The KNN method operates on the presumption that members belonging to the same class are gathered nearby, as shown in figure 4.2. While there are several ways to determine how far away the target point is from its k neighbors, Euclidean distance is the one that is most frequently employed. The target point is given the class for which the greatest number of neighbors are there after calculating the distance from k neighbors, which is known as the majority vote. Since KNN must determine the distance from each point in its vicinity, it is referred to as a lazy learner and requires extra time to train when compared to other machine learning algorithms (Guo et al., 2003).

4.1.3 Support Vector Machine

When utilized as a linear SVM, the vector machine classifier performs as a parametric classifier, however advanced SVMs used to classify complicated non-linear data are non-

parametric. The calculation is not carried out in the high-dimensional feature space where the complex dataset may be projected as linear, using kernels like the radial basis function (RBF). All required calculations are carried out directly in the input space when kernels are used (Hearst et al., 1998). The objective is to ascertain an optimal hyper-plane that maximizes the separation from the data points corresponding to the two distinct classes depicted in figure 4.3. Support Vector Machines (SVMs) are notably prominent within the domain of remote sensing, as they exhibit the capability to effectively manage limited training datasets and typically yield superior classification accuracy compared to conventional methods (Mountrakis et al., 2011).

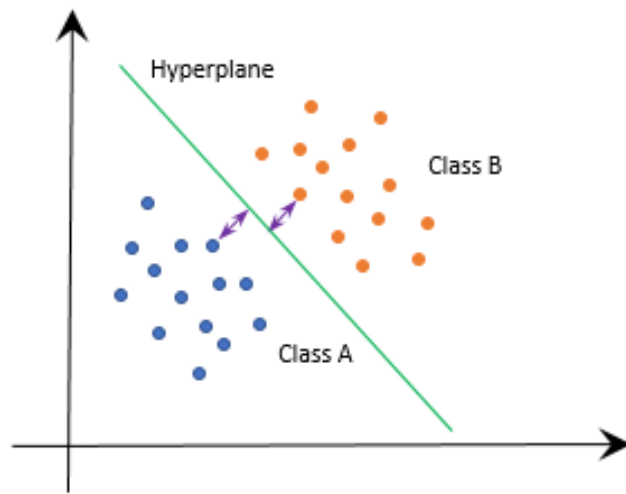


Figure 4.3 Support Vector Machine Classifier

Three different kernel types, represented by equations (14), (15), and (16), are principally utilized with non-parametric SVMs: polynomial kernel, RBF kernel, and sigmoid kernel.

Polynomial kernel (k),

$$k(x,y) = (x \cdot y)^d \quad (14)$$

is associated with a mapping function Φ in the vector space that encompasses all products of precisely d dimensions within R^N

RBF kernel (k),

$$k(x,y) = \exp(-\|x-y\|^2/(2\sigma^2)) \quad (15)$$

where $1/2\sigma^2 = \gamma$ and γ is designated as the free parameter

Sigmoid kernel (k),

$$k(x,y) = \tanh(\kappa(x,y) + \theta) \quad (16)$$

such that κ represents the gain parameter and θ denotes the offset constant.

4.2 Deep Learning

Deep learning offers an alternate point of view towards feature learning, where powerful, abstract, and invariant elements are advanced from start to finish, progressively, from crude information (example- image pixels) to semantic labels, which is a critical benefit compared to the past cutting edge algorithms (Nogueira et al., 2017; LeCun et al., 2015). In deep learning, there are different approaches to implementing semantic segmentation. A basic deep neural network algorithm is called a Multi-Layer Perceptron (MLP). Fully Convolutional Networks (FCN) serve as the foundation for popular deep learning approaches. Contrary to traditional CNN, which uses fixed fully connected networks, FCN simply comprises convolution and pooling layers, allowing it to handle inputs of any size (Shelhamer et al., 2017). Encoder-decoder networks make up the architecture of a Semantic Segmentation system in general. A classification network like VGG/RESNET/Xception can be an encoder network. By upsampling, a decoder network projects the encoder network's low-resolution output into a high-resolution pixel space. The job of semantic segmentation has been effectively accomplished using an encoder-decoder network (Ronneberger et al., 2015; Badrinarayanan, 2017).

4.2.1 Multilayer Perceptron

A multilayer perceptron (MLP) constitutes a fully-connected feed-forward neural architecture, signifying that each neuron within a particular layer is interconnected with every neuron in the subsequent layers; however, there exists no reciprocal interaction among the layers, and the input traverses unidirectionally towards the output (Del Frate et al., 2007). Each node's outputs are weighted units, then a nonlinear activation function is applied to them to identify the data that cannot be linearly separated (Pacifci et al., 2007). Mathematically, the input activation a^l is used to derive the output activation a^{l+1} at layer $l + 1$ as (Zhang et al., 2018):

$$a^{(l+1)} = \sigma(w^{(l)}a^{(l)} + b^{(l)}) \quad (21)$$

where l corresponds to a specific layer, $w^{(l)}$ represents the weight at layer l , $b^{(l)}$ represents the bias at layer l and σ denotes the activation function.

For an n-layer MLP network, the first input layer is $a^{(1)} = x$ while the last output layer is (Del Frate et al., 2007):

$$h_{w,b}(x) = a^{(m)} \quad (22)$$

The weight (w) and bias (b) in equation 22, are learned by supervised learning with an objective to minimize the loss function (Zhang et al., 2018):

$$J(W, b; x, y) = \frac{1}{2} \|h_{w,b}(x) - y\|^2 \quad (23)$$

4.2.2 DeepLabv3+

The DeepLabv3+ framework constitutes an advancement of the DeepLabv3 architecture. Figure 4.4 depicts the comprehensive encoder-decoder framework of DeepLabv3+.

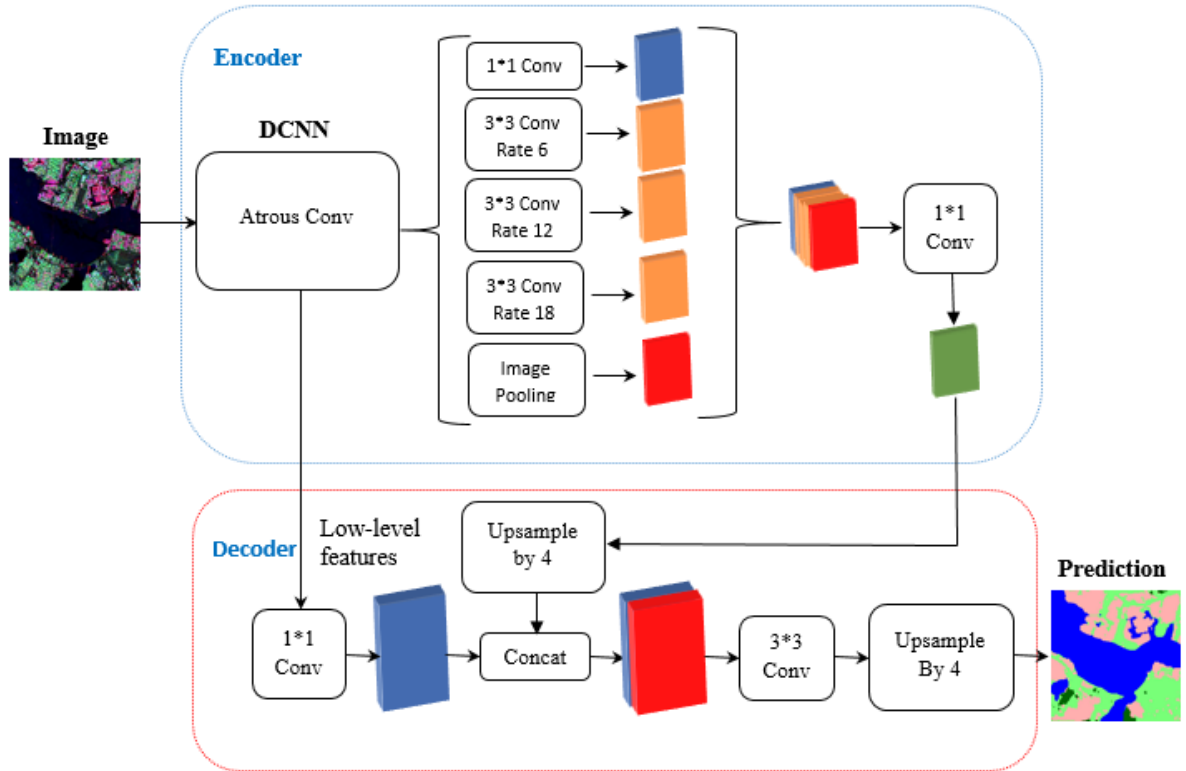


Figure 4.4 Encoder-Decoder based architecture of DeepLabv3+ (Chen et al., 2018).

The decoder component is responsible for refining the segmentation results at the delineation of objects, whereas the encoder component's function is to encode the multi-scale contextual information through the application of atrous convolutions.

Figure 4.5 depicts the updated Xception model used by the DeepLabv3+ as an encoder. The initial iterations of Deep Convolutional Neural Networks (DCNNs) exhibited a limitation whereby the spatial resolution of the output feature maps was substantially diminished as a result of the continual application of max-pooling and striding at each layer of these architectures (Chen et al., 2016).

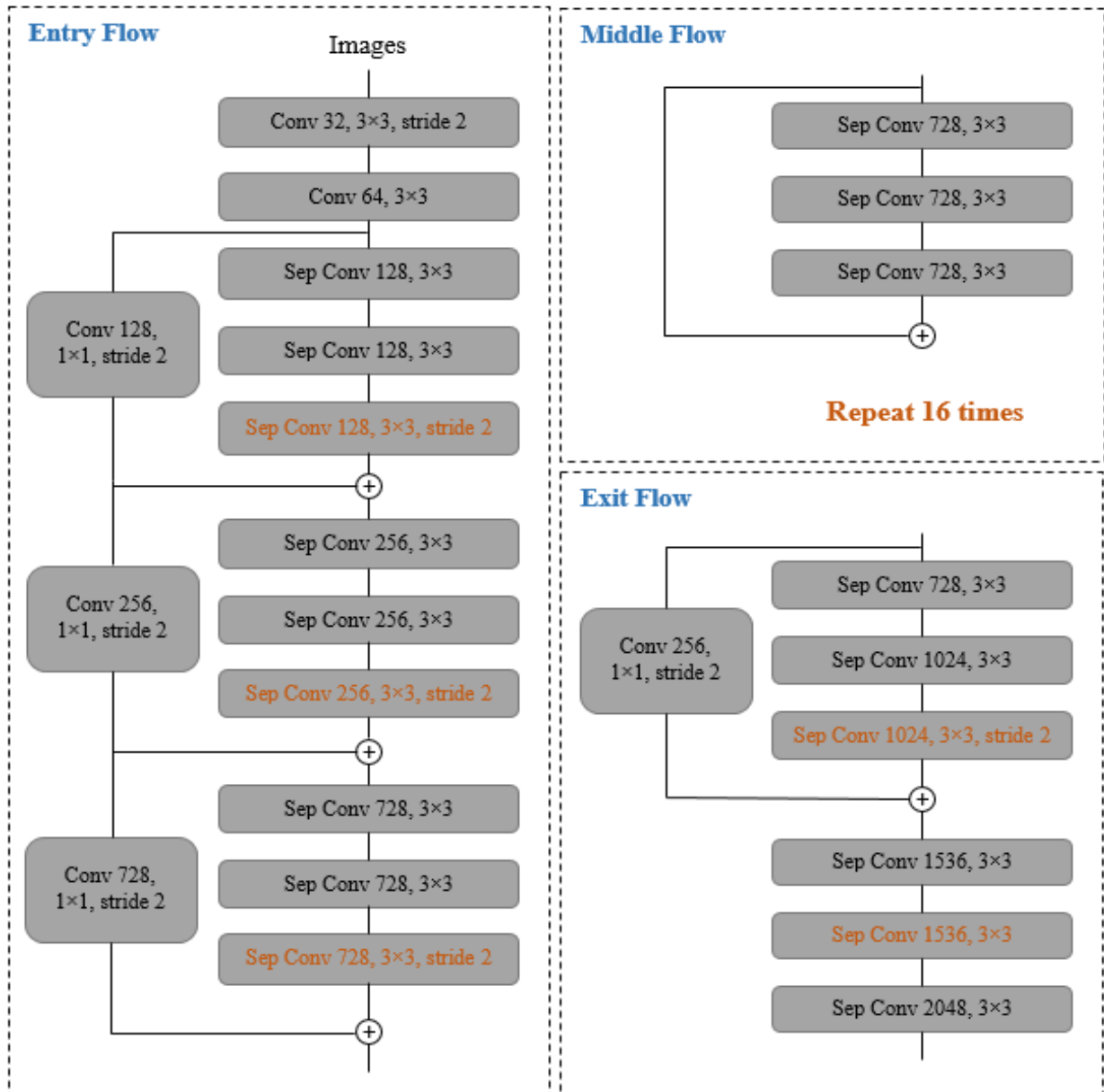


Figure 4.5 In the DeepLabv3+ framework, the Xception architecture has undergone modifications that encompass: (1) an increased number of layers, (2) the implementation of depthwise separable convolutions with striding substituting all max pooling layers, and (3) the incorporation of supplementary batch normalization and ReLU activations subsequent to each 3×3 depthwise convolution (Chen et al., 2018).

Atrous convolution, which was first created for computing undecimated wavelet transform (Holschneider et al., 1990), is employed to solve this issue. Any desired resolution can be used to compute responses using it (Chen et al., 2016).

Kernel size, stride, and padding are the three parameters that may be used to construct a convolution operation. However, an atrous convolution, sometimes referred to as a dilated convolution, contains a fourth parameter called dilation rate (r), which is the distance between the kernel values. This particular instance of atrous convolution is comparable to regular convolution if $r = 1$. For a two-dimensional signal, the atrous convolution is executed over the input feature map x in accordance with equation (24) for each point i on the feature map y and the convolution filter w (Gong et al., 2016).

$$y[i] = \sum_k x[i + r \cdot k]w[k] \quad (24)$$

where r = atrous rate, signifies the sampling stride utilized for the input data.

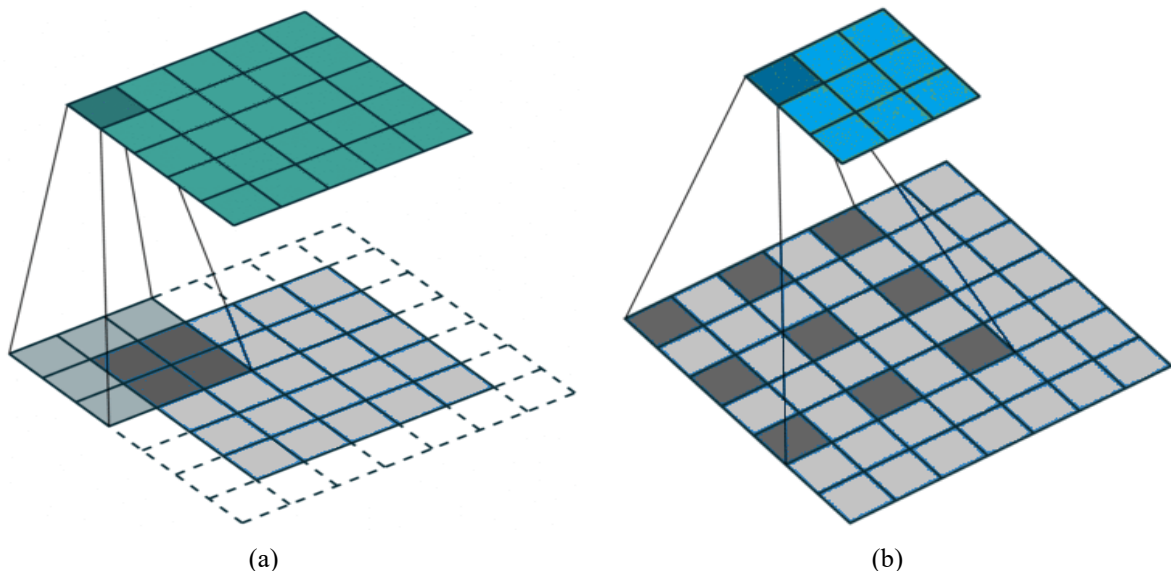


Figure 4.6 depicts the normal and atrous convolution. (a) a normal convolution with kernel size = 3, stride = 1, and padding is used to retain the input image size ; (b) an atrous convolution with an additional parameter called dilation rate that deletes every alternate row and column.

Wider fields of view are possible with atrous convolution with the same computing cost. Figure 4.6 shows a comparison between atrous convolution and normal convolution. Atrous Spatial Pyramid Pooling (ASPP), a technique that employs parallel atrous convolutions with diverse rates, was implemented in DeepLabv3 to facilitate the acquisition of contextual information across multiple scales. The resultant feature map contains essential semantic

details; however, the convolution operation involving striding tends to omit data related to object edges. To obtain denser feature maps, this challenge can be addressed by utilizing atrous convolution, albeit this approach may incur significant computational costs and inefficiencies. In the context of encoder-decoder architectures, no features are dilated within the encoder pathway, yet they gradually recover precise object boundaries within the decoder pathway, thus enhancing computational efficiency. DeepLabv3+ adds a straightforward yet efficient decoder module for recovering the sharp object boundaries to the DeepLabv3 model, combining the benefits of both approaches.

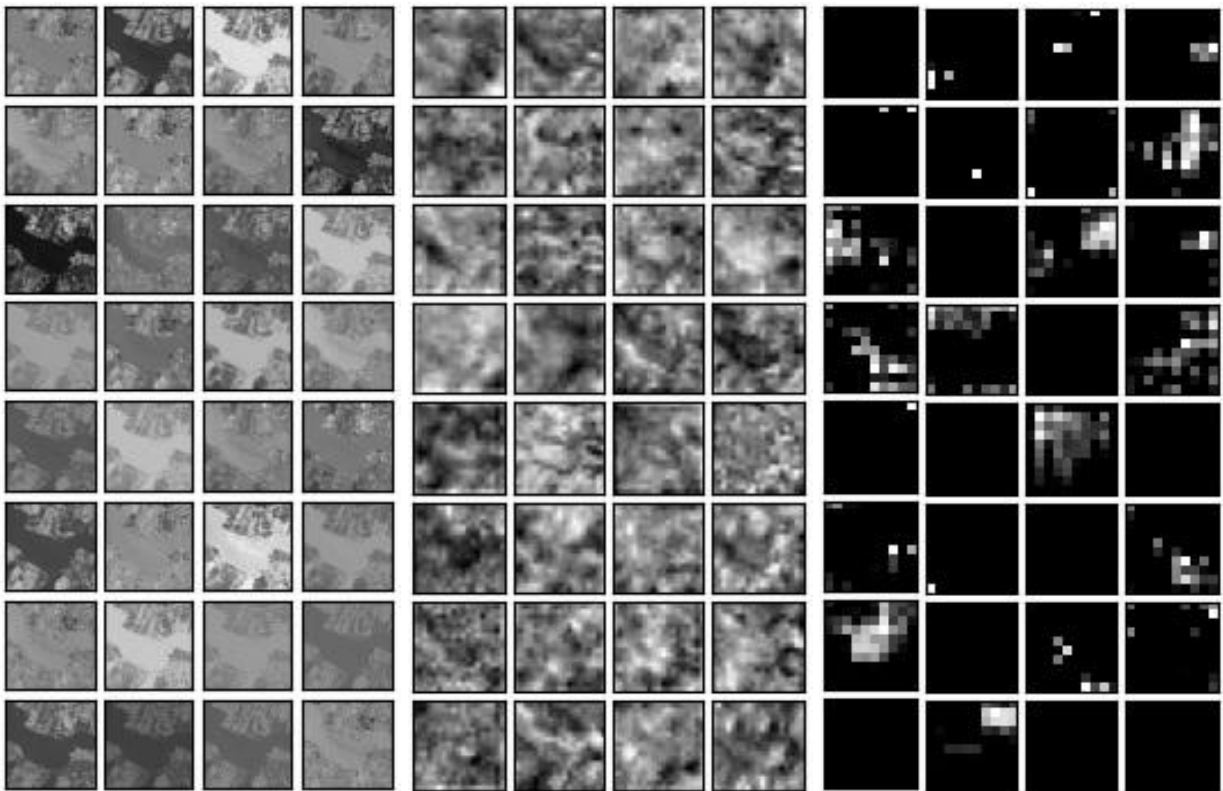


Figure 4.7 Represents the feature maps extracted from 1st hidden layer of block1, block7 and block14.

To understand the working of the encoder network of the Deeplabv3+, the feature maps corresponding to the 1st hidden layer of block1, block7, and block 14 are fetched and presented in figure 4.7. It can be noted that while the upper layers focus on finer details, the lower levels work on simpler, coarser aspects.

CHAPTER 5

RESULTS and DISCUSSION

5.1 Mumbai Region

For patch 1, the resulting land cover maps generated after semantic segmentation of true-color composite based optical image, false-color composite based optical image, and false-color composite based SAR image are illustrated in figures 5.1, 5.2, and 5.3, respectively. For Patch 2, the corresponding maps are presented in figures 5.4, 5.5, and 5.6. Furthermore, adopting a single patch for training and testing, figures 5.9, 5.10, and 5.11 demonstrate the land cover maps for the TerraSAR-X, RISAT-1, and ALOS-2 PALSAR-2 datasets, respectively. Tables 4, 5, and 6 present, correspondingly, the outcomes of the evaluation of the findings using the mIOU and IOU evaluation matrices.

5.1.1 Performance evaluation metrics

The mIOU has been employed to assess how well machine learning systems accomplish image segmentation tasks. It is one of the rigorous metrics that considers every pixel in the validation and assessment of image segmentation tasks. By averaging the IOU for all classes, mIOU is determined (Kumar & Prateek, 2020). In equations (25) and (26) respectively, the formulas for IOU and mIOU are mentioned.

$$IOU = \frac{true_positive}{true_positive + false_positive + false_negative} \quad (25)$$

$$mIOU = \frac{IOU_U + IOU_F + IOU_W + IOU_G}{4} \quad (26)$$

where,

IOU_U = Intersection over Union for Urban class

IOU_F = Intersection over Union for Forest class

IOU_W = Intersection over Union for Water class

IOU_G = Intersection over Union for Ground class

5.1.2 Performance evaluation of image segmentation into four classes

The mIOU for Sentinel-2 true-color composite based optical images ranged from 0.2323 to 0.4517, for false-color composite based optical images varied from 0.4130 to 0.5832, and for ALOS-2 PALSAR-2 data varied from 0.4033 to 0.5178 for the image segmentation task, as indicated by the bar chart in figure 28 and also reported in table 4. This was discovered during comparing the results of Sentinel-2 and ALOS-2 PALSAR-2 data. Figure 5.1 shows the segmentation task results for test patch 1 using the true-color composite optical image.

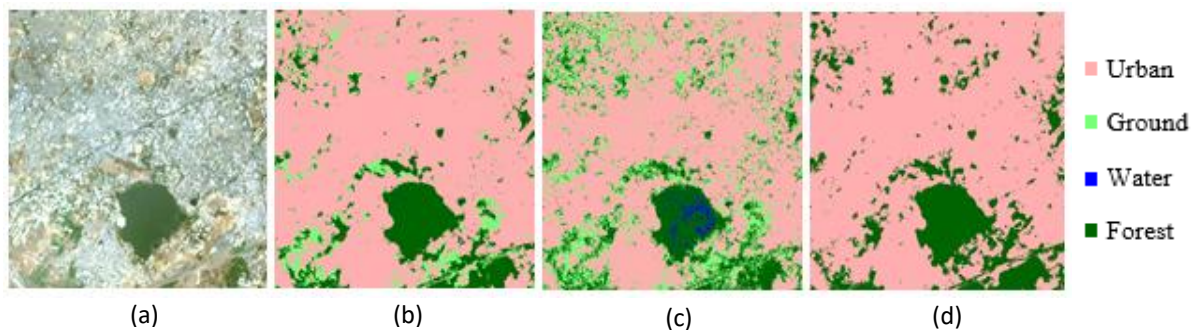


Figure 5.1 Results of image segmentation for the Varaladevi Talao, Bhiwandi, Maharashtra patch 1. (a) Sentinel-2 true color composite image (where Red $\rightarrow$ B4, Green $\rightarrow$ B3, and Blue $\rightarrow$ B2), and (b) RF classifier, (c) KNN classifier, and (d) SVM classifier, respectively (Garg et al., 2021).

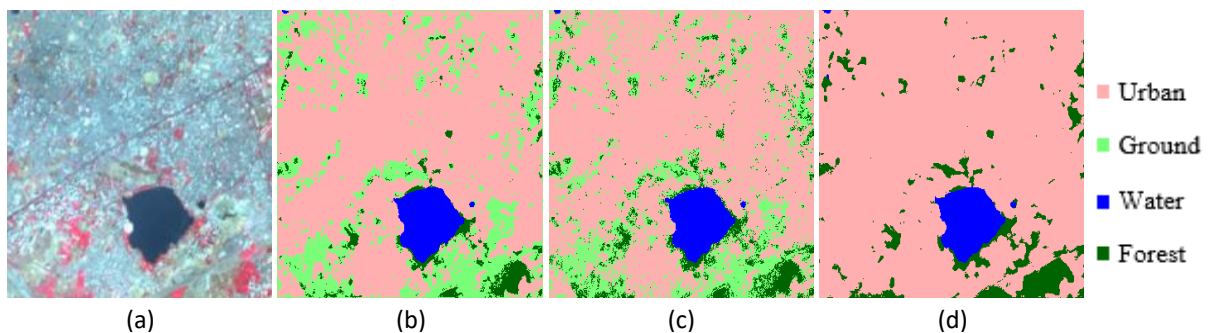


Figure 5.2 Results of image segmentation for test patch 1. (a) Sentinel-2 false-color composite imagery (where Red $\rightarrow$ B8, Green $\rightarrow$ B4, Blue $\rightarrow$ B3), and (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively (Garg et al., 2021).

Red, green, and blue colors are represented by the B4, B3, and B2 bands of the Sentinel-2 data in Figure 5.1(a), which is a true-color composite picture of patch 1 of optical data. Figures 5.1(b), 5.1(c), and 5.1(d), correspondingly, show the outcomes for the semantic segmentation task on patch1 using the RF classifier, the KNN classifier, and the SVM classifier. The SVM classifier failed to detect the water body or the ground class as a result of which it has the least mIOU while both RF and KNN classifiers recorded higher mIOU as reported in table 4 and bar chart in figure 5.7. It should be noted that both the RF and KNN

classifiers are able to identify the urban, forest, and ground classes with good accuracy but they fail to identify the water body. Figure 5.2 shows the outcomes of the segmentation job performed on the false color composite optical image. The RF classifier, followed by the KNN classifier and the SVM classifier, had the highest mIOU for patch 1 (0.4941), as seen in the bar chart in figure 5.7 and table 4. Red, green, and blue colors correspond to the B8, B4, and B3 bands of the Sentinel-2 data in figure 5.2(a), which shows the false-color composite picture of patch 1 of optical data. Figures 5.2(b), 5.2(c), and 5.2(d), correspondingly, show the outcomes for the semantic segmentation task on Patch1 using the RF classifier, the KNN classifier, and the SVM classifier. The segmentation findings make it evident that the water body being incorrectly classified as vegetation when using the true-color composite picture has been resolved by utilizing the false-color composite image. In addition, close examination and comparison of figures 5.1b and 5.2b and 5.1c and 5.2c reveals that, with the exception of the SVM classifier, which is unable to detect the ground class in both cases, the ground class is better identified in false color composite based image segmentation results than in true color composite based results.

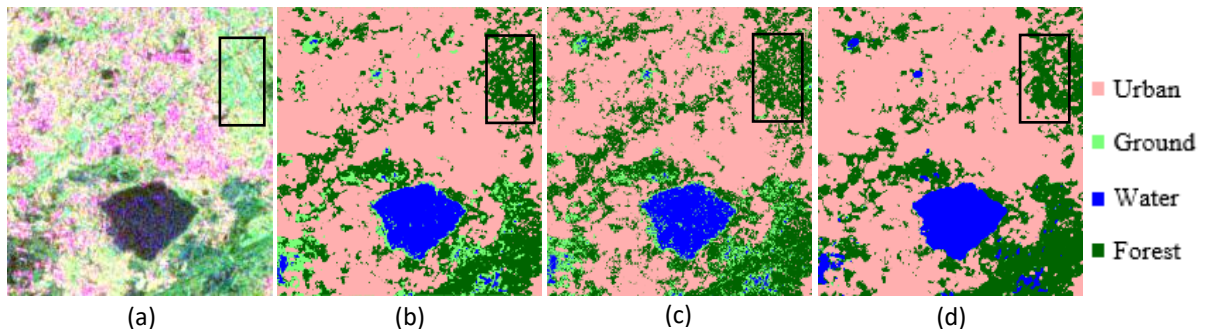


Figure 5.3 The results of image segmentation on Patch1 utilising (a) Barnes decomposition based false colour composite RGB image (where Red → Even bounce scattering, Green → Volume scattering, and Blue → Surface bounce scattering), (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for ALOS-2 data (Garg et al., 2021).

The false-color composite representation of test patch 1 derived from the ALOS-2 PALSAR-2 dataset is illustrated in Figure 5.3(a), wherein the hues red, green, and blue correspond to the phenomena of even bounce scattering, volume scattering, and surface bounce scattering, respectively. Figures 5.3(b), 5.3(c), and 5.3(d) sequentially depict the results of the semantic segmentation task employing the Random Forest (RF) classifier, the K-Nearest Neighbors (KNN) classifier, and the Support Vector Machine (SVM) classifier. Unlike the oriented urban targets denoted by the black rectangular area which detects urban targets

partially, the other classifiers were equally capable of detecting the urban region. According to table 4, the random forest classifier had the highest mIOU and was followed by the KNN and SVM classifiers in terms of classification accuracy for the remaining classes.

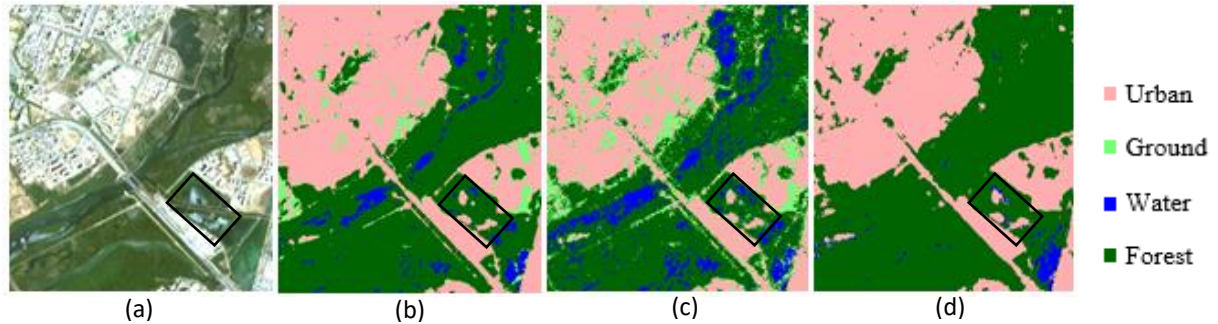


Figure 5.4 Results of patch 2's image segmentation detecting the area Taloje River, Navi Mumabai, Maharashtra (a) Sentinel-2 true colour composite image (where Red $\rightarrow$ B4, Green $\rightarrow$ B3, Blue $\rightarrow$ B2), and (b) RF classifier, (c) KNN classifier, and (d) SVM classifier results for the related image segmentation (Garg et al., 2021).

Figure 5.4 illustrates the segmentation task results for test patch 2 using the true color composite optical picture. The KNN classifier recorded the highest mIOU of 0.4517 in patch 2, while SVM recorded the lowest mIOU, as can be seen from the bar chart in figure 5.4 and table 4. Red, green, and blue colors correspond to the B4, B3, and B2 bands of the Sentinel-2 data in Figure 5.4(a), which shows the true color composite image of patch 2 of optical data. Figures 5.4(b), 5.4(c), and 5.4(d), respectively, show the outcomes for the semantic segmentation task using the RF classifier, KNN classifier, and SVM classifier. Only the KNN classifier was able to detect the water bodies present in the test patch to a certain degree. SVM fared badly in detecting water bodies and ground class as well, resulting in the least mIOU for SVM and the highest mIOU for KNN as reported in table 4. All the other classifiers could detect urban features rather well.

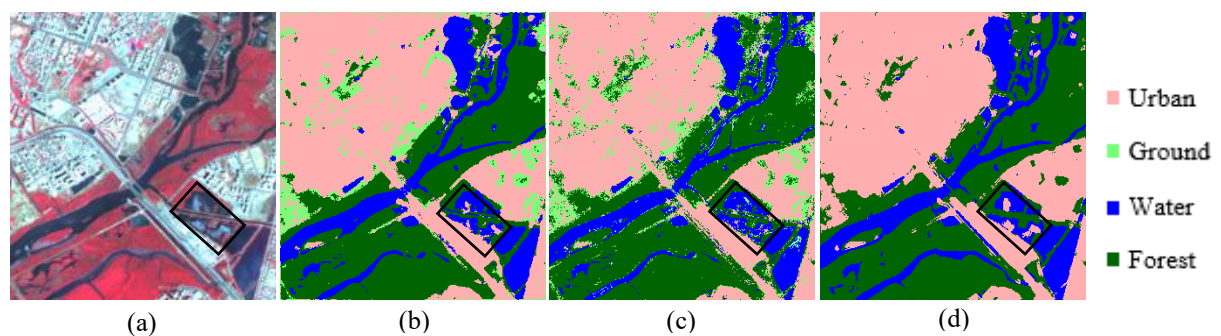


Figure 5.5 Results of image segmentation on Patch 2 where (a) Sentinel-2 false color composite based image (where Red $\rightarrow$ B8, Green $\rightarrow$ B4, Blue $\rightarrow$ B3) for test patch 2 and corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively (Garg et al., 2021).

Figure 5.5 demonstrates the segmentation task results for the false color composite optical image. The RF classifier, followed by the KNN classifier and the SVM classifier, had the highest mIOU for patch 2 (0.5832), as seen in the bar chart in figure 5.7 and table 4. Red, green, and blue colors correspond to the B8, B4, and B3 bands of the Sentinel-2 data in figure 5.5(a), which shows the false-color composite picture of patch 2 of optical data. Figures 5.5(b), 5.5(c), and 5.5(d) demonstrate, respectively, the outcomes for the semantic segmentation task on Patch2 using the RF classifier, the KNN classifier, and the SVM classifier.

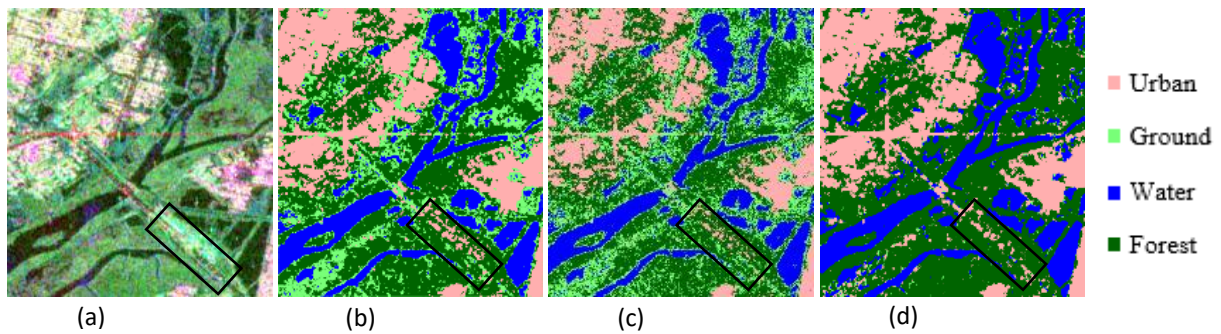


Figure 5.6 Image segmentation results for patch 2. (a) Barnes decomposition based false color composite RGB image (where Red $\rightarrow$ Even bounce scattering, Green $\rightarrow$ Volume scattering, and Blue $\rightarrow$ Surface bounce scattering), and the corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for ALOS-2 data (Garg et al., 2021).

The problem of misclassifying water features as vegetation has been overcome in patch 2 much like it was in patch 1 using the false-color composite image for optical data. Furthermore, employing the false-color composite picture for the segmentation job greatly reduces the impact of cloud cover, according to a thorough examination of the region enclosed by the black box in figures 5.4 and 5.5. When utilizing a true-color composite image, the clouds above the water bodies lead them to be mistakenly classified as urban targets (figure 5.4), yet the same location is recognized as a water body when using a false-color composite image (figure 5.5). As seen in figure 5.5d, only the SVM classifier is incorrectly identifying clouds as urban targets.

Figure 5.6(a) shows the false-color composite image of test patch 2 from the ALOS-2 PALSAR-2 data, and figures 5.6(b), 5.6(c), and 5.6(d) shows the equivalent results for the semantic segmentation task using the RF classifier, KNN classifier, and SVM classifier, respectively. It should be noticed that while the ground class could not be recognized by the SVM classifier, the water, urban, and forest classes were well detected by all three classifiers. Additionally, the area enclosed by the black rectangle serves as a visual cue that the KNN

classifier underestimates the urban area relative to the random forest classifier. As seen in table 4, this led to the highest mIOU observed by the RF classifier and equivalent mIOU by the KNN and SVM classifiers.

Table 4. Results from a performance assessment of image segmentation using mIOU for optical (Sentinel-2) and SAR (ALOS-2) data utilizing both true-color composite (OpticalTCC) and false-color composite (OpticalFCC) images (Garg et al., 2021)

S. No.	Model	Patch1			Patch2		
		Optical _{TCC} (mIOU)	Optical _{FCC} (mIOU)	SAR (mIOU)	Optical _{TCC} (mIOU)	Optical _{FCC} (mIOU)	SAR (mIOU)
1.	RF	0.2866	0.4941	0.4663	0.4156	0.5832	0.5178
2.	KNN	0.2862	0.4569	0.4384	0.4517	0.5611	0.4655
3.	SVM	0.2323	0.4130	0.4033	0.3451	0.5156	0.4600

Overall, from the data shown in figures 5.1 to 5.6 and the bar charts in figures 5.7 and 5.8, it can be concluded that while optical data can detect urban areas more accurately than SAR data, the latter is more effective at identifying water bodies and distinguishing them from the ground and road network. When using a true color composite image with optical data, it is very difficult to distinguish between the water class and the forest class. However, when using SAR data, both classes can be easily distinguished because there is little backscatter from the water body and high volume scattering from the vegetated forest area. Using the false-color composite image for the optical data, however, solves this issue with optical data.

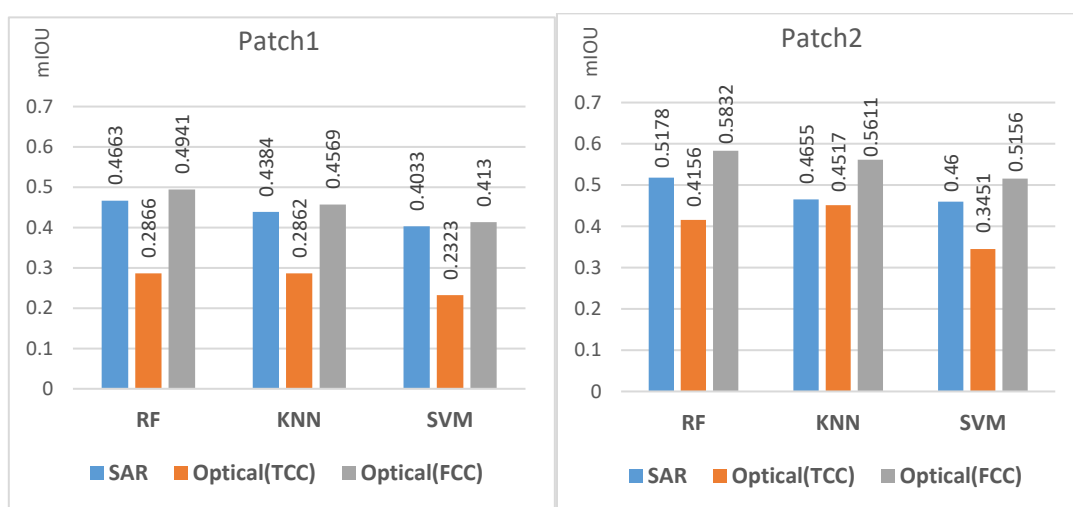


Figure 5.7 mIOU comparison for SAR (ALOS-2/PALSAR-2) and Optical (Sentinel-2) datasets for four classes (Water, Urban, Forest, Ground) for Patch1 and Patch 2 respectively (Garg et al., 2021).

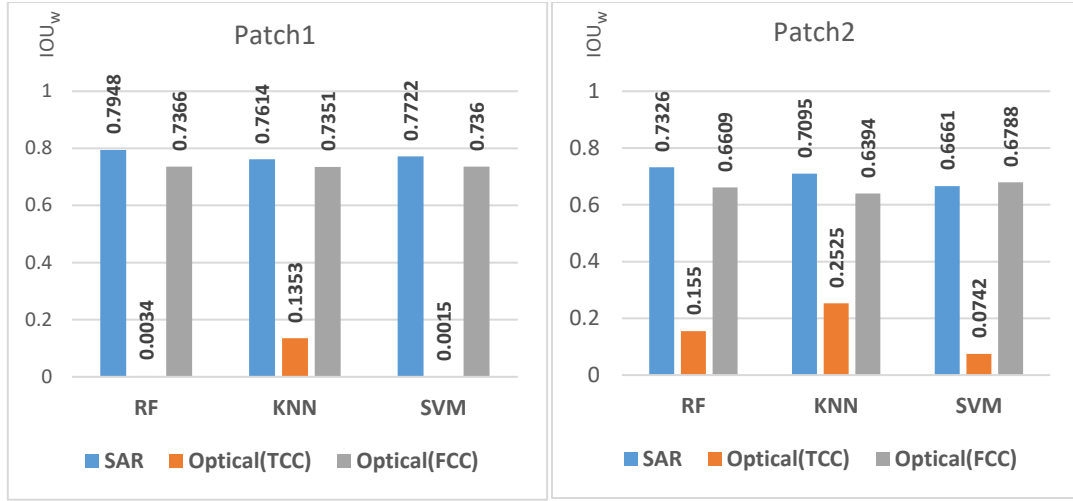


Figure 5.8 IOU comparison for SAR (ALOS-2/PALSAR-2) and Optical (Sentinel-2) datasets for water class for Patch 1 and Patch 2 respectively (Garg et al., 2021).

The segmentation results on the TerraSAR-X data are shown in figure 5.9 for the classification job using a single picture patch. As recorded in table 4, the RF classifier achieved the highest mIOU of 0.5815 while the SVM classifier registered the lowest mIOU. The false-color composite image created by Barnes decomposition is shown in figure 5.9(a), where the colors red, green, and blue represent the even bounce scattering, volume scattering, and surface or odd bounce scattering, respectively.

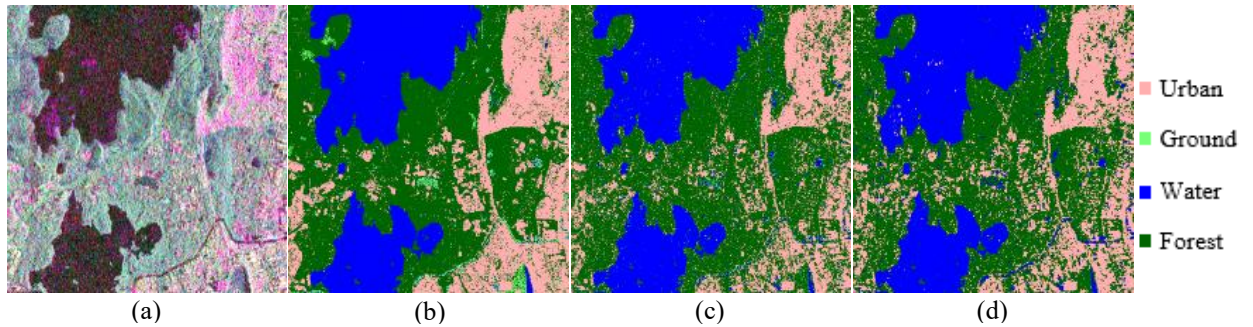


Figure 5.9 (a) Barnes decomposition based false color composite RGB image of selected patch (Vihar and Powai lake, Mumbai) where Red → Even bounce scattering, Green → Volume scattering, and Blue → Surface bounce scattering and the corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for TerraSAR-X data (Garg et al., 2021).

Figures 5.9(b), 5.9(c), and 5.9(d), respectively, show the outcomes for the semantic segmentation challenge using the RF classifier, KNN classifier, and SVM classifier. The segmentation findings show that, in contrast to the KNN and SVM classifiers, which

incorrectly categorized the ground targets as water bodies, the RF classifier was able to properly identify all of the classes.

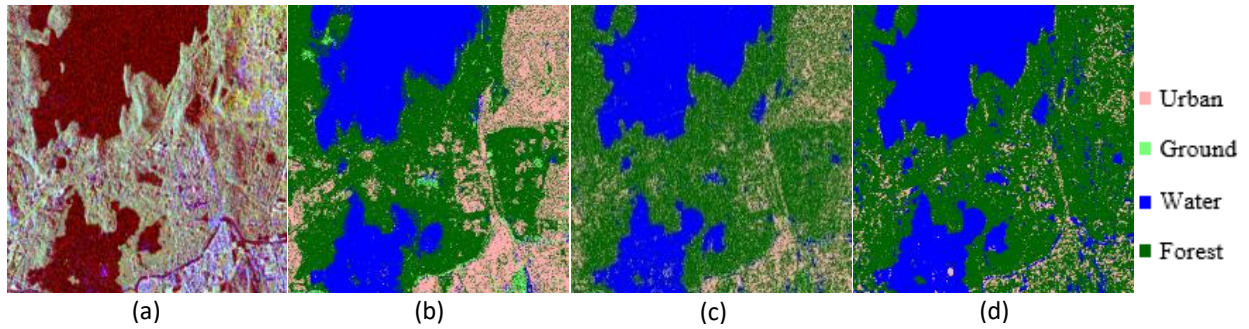


Figure 5.10 (a) 3-Component compact-pol decomposition based false color composite RGB image of selected patch (Vihar and Powai lake, Mumbai) where Red → Even bounce scattering, Green → Volume scattering, and Blue → Surface bounce scattering and the corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for RISAT-1 data (Garg et al., 2021).

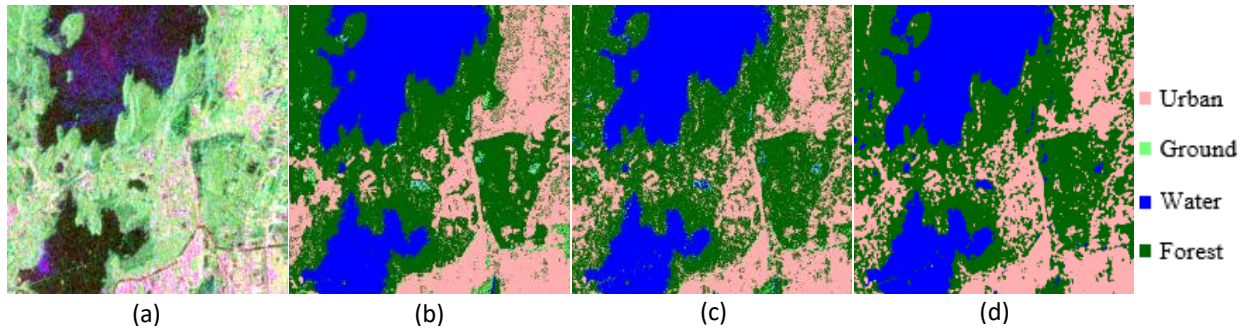


Figure 5.11 (a) Barnes decomposition based false color composite RGB image of selected patch (Vihar and Powai lake, Mumbai) where Red → Even bounce scattering, Green → Volume scattering, and Blue → Surface bounce scattering and the corresponding image segmentation results using (b) RF classifier, (c) KNN classifier, and (d) SVM classifier respectively for ALOS-2 data (Garg et al., 2021).

As delineated in table 5, the Random Forest (RF) classifier achieved the highest mean Intersection over Union (mIOU) of 0.4031 for the RISAT-1 dataset, whereas the Support Vector Machine (SVM) classifier exhibited the lowest mIOU. Figure 5.10(a) illustrates the false-color composite image generated through the three-component compact polar decomposition, with the colors red, green, and blue representing even bounce scattering, volume scattering, and surface or odd bounce scattering, respectively. Figures 5.10(b), 5.10(c), and 5.10(d) sequentially depict the results from the semantic segmentation task employing the RF classifier, K-Nearest Neighbors (KNN) classifier, and SVM classifier. As seen in figure 5.10, the KNN and SVM classifiers could hardly recognize the urban characteristics, whereas the RF classifier was able to accurately distinguish the urban targets. Additionally, the KNN and SVM classifiers performed poorly in comparison to the RF classifier because they

misclassified the ground targets as water bodies. The RF classifier recorded the greatest mIOU of 0.6153 for the ALOS-2 PALSAR-2 data, while the SVM classifier recorded the lowest mIOU, as recorded in table 5.

The composite image represented in false color, derived through the Barnes decomposition method, is illustrated in figure 5.11(a), where the colors red, green, and blue are indicative of even bounce scattering, diffused scattering, and odd bounce scattering, respectively. Figures 5.11(b), 5.11(c), and 5.11(d) present the results of the semantic segmentation task employing the RF classifier, KNN classifier, and SVM classifier, correspondingly. As seen in figure 5.11, the KNN and SVM classifiers performed poorly when compared to the RF classifier because they misclassified the ground targets as water bodies. The reason why RISAT-1 data performed poorly in the LULC classification task when compared to ALOS-2, PALSAR-2, and TerraSAR-X data is that the compact-pol data includes less information than the quad-pol data.

Table 5. The findings of performance testing for water class using IOU and image segmentation using mIOU for TerraSAR-X, RISAT-1, and ALOS-2 data

S. No.	Model	TerraSAR-X (Quad-Pol)		RISAT-1 (Compact-Pol)		ALOS-2 (Quad-Pol)	
		(mIOU)	(IOU _w)	(mIOU)	(IOU _w)	(mIOU)	(IOU _w)
1.	RF	0.5815	0.9273	0.4031	0.7478	0.6153	0.9428
2.	KNN	0.5546	0.9168	0.3435	0.7029	0.6011	0.9423
3.	SVM	0.5498	0.9055	0.3699	0.7178	0.5571	0.9323

5.1.3 Performance evaluation of water class

The IOU for the water class varied from 0.0034 to 0.2525 for the true-color composite based optical image, from 0.6394 to 0.7366 for the false-color composite based optical image, and from 0.6661 to 0.7948 for the ALOS-2 PALSAR-2 data, as recorded in table 6.

Table 6. Results of performance assessment for the Water class utilizing IOU for optical (Sentinel-2) and SAR (ALOS-2) data using both true-color composite (Optical_{TCC}) and false-color composite (Optical_{FCC}) imagery.

S. No.	Model	Patch1			Patch2		
		Optical _{TCC}	Optical _{FCC}	SAR	Optical _{TCC}	Optical _{FCC}	SAR
		IOU _w	IOU _w	IOU _w	IOU _w	IOU _w	IOU _w
1.	RF	0.0034	0.7366	0.7948	0.155	0.6609	0.7326

2.	KNN	0.1353	0.7351	0.7614	0.2525	0.6394	0.7095
3.	SVM	0.0015	0.7360	0.7722	0.0742	0.6788	0.6661

The KNN classifier successfully identified the water body in the true-color composite based optical picture with IOU values of 0.1353 and 0.2525 for patches 1 and 2, respectively, while the other two classifiers nearly missed it, particularly in the case of patch 1, as seen in figures 5.1 and 5.4. But employing the false-color composite image for the optical data based image segmentation work solves this misclassification issue, as seen in the segmentation results in figures 5.2 and 5.5, for patch 1 and patch 2, respectively. The highest IOU for the water class in patch 1 was 0.7948, while the highest IOU for the water class in patch 2 was 0.7326 in the case of the RF classifier, as seen in the bar chart in figure 5.8.

The findings in figure 5.9 show that the RF classifier performed the best for a single patch based classification job using TerraSAR-X data, with an IOU for the water class ranging from 0.9055 to 0.9273 as stated in table 6. The findings in figure 35 show that the RF classifier performed the best in the case of the RISAT-1 data, where the IOU for the water class ranged from 0.7029 to 0.7478. The IOU for the water class in the ALOS-2 PALSAR-2 data ranged from 0.9323 to 0.9428, with the RF classifier doing the best, as seen in the findings in figure 36.

5.2 Houston Region

Figures 5.12 and 5.13, which show the Google Earth imagery in figure 5.12a and figure 5.13a and the false-color composite images for the chosen patches in figure 5.12b and figure 5.13b, exhibit the outcomes of the semantic segmentation job for patch 1 and patch 2, respectively. When compared to conventional machine learning techniques, it can be shown that DeepLabv3+ produces exceptionally accurate and smooth land cover maps. Areas A and B in figure 5.12 correspond to a river that belongs to the water class, but all conventional methods have pointed to a tiny piece of ground there, however, DeepLabv3+ has accurately predicted the water class. Additionally, it is evident from the SAR imagery that region C has a tiny forest, but as shown in figures 5.12c, 5.12d, and 5.12e, the SVM method completely misses it while the RF and KNN techniques catch it sparingly. However as can be seen in figure 5.12f, 5.12g, and 5.12h that FPN, UNET and DeepLabv3+ has successfully captured the vegetation. But, both FPN and UNET fails to capture the minute details like water at ship dock and intricate boundaries that are captured by the DeepLabV3+.

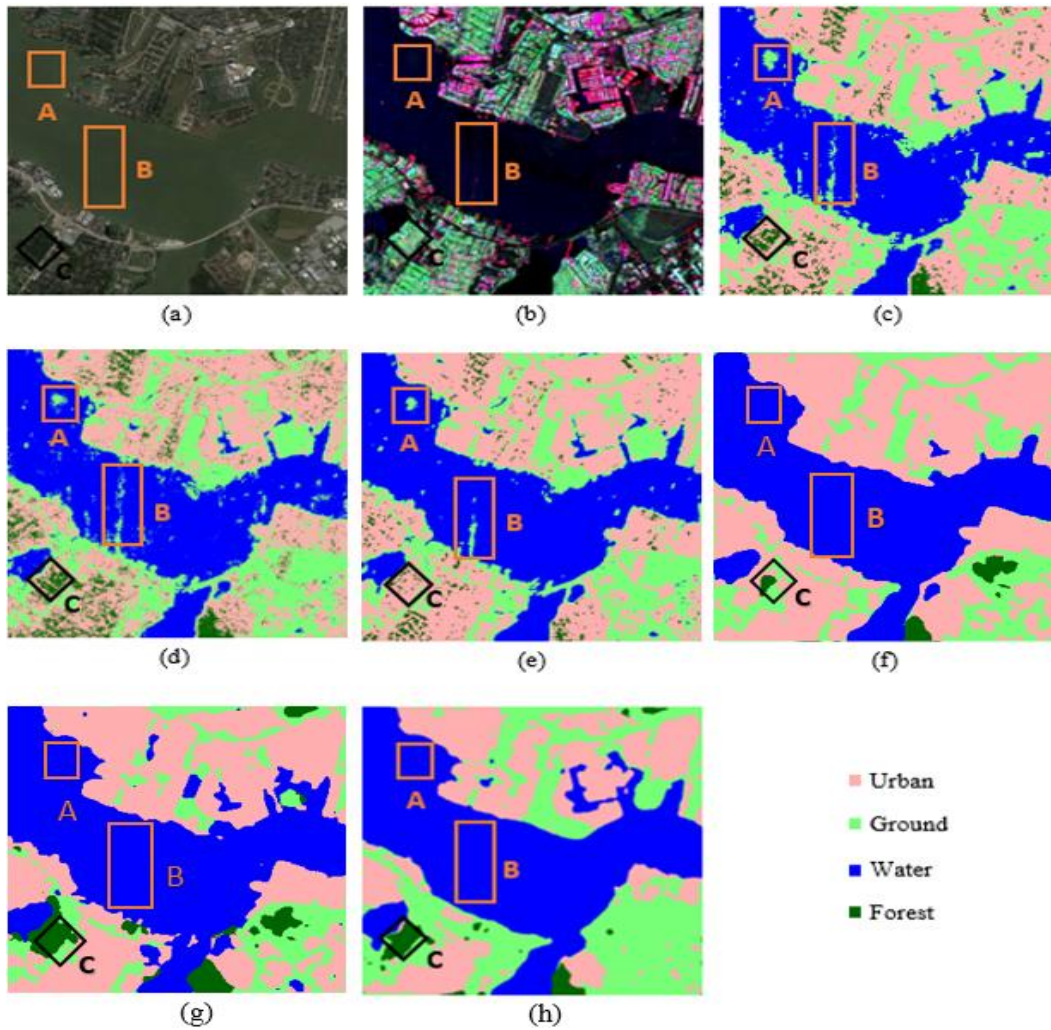


Figure 5.12 Represents the ground truth of Houston city, Texas (patch 1) and respective segmentation results of various classifiers (a) Google Earth image of test patch 1 (b) False color composite image for G4U decomposition on UAVSAR data of a part of Houston city, Texas (USA) and the corresponding image segmentation results for (c) RF classifier, (d) KNN classifier, (e) SVM classifier, (f) FPN, (g) UNET, (h) Deeplab V3+ respectively (Garg et al., 2021b).

As seen in figure 5.13, areas D and E correspond to orientated urban structures that the decomposed image incorrectly classifies as vegetation (green color). This has an impact on the prediction of conventional machine learning algorithms as well, and KNN is the worst affected. As a result, most of the orientated urban features are misclassified as vegetation. The oriented urban targets could be accurately mapped by the deep learning algorithms namely FPN, UNET, and DeepLabv3+ under the urban class with little misclassification, as shown in figure 5.13f, 5.13g, and 5.13h.

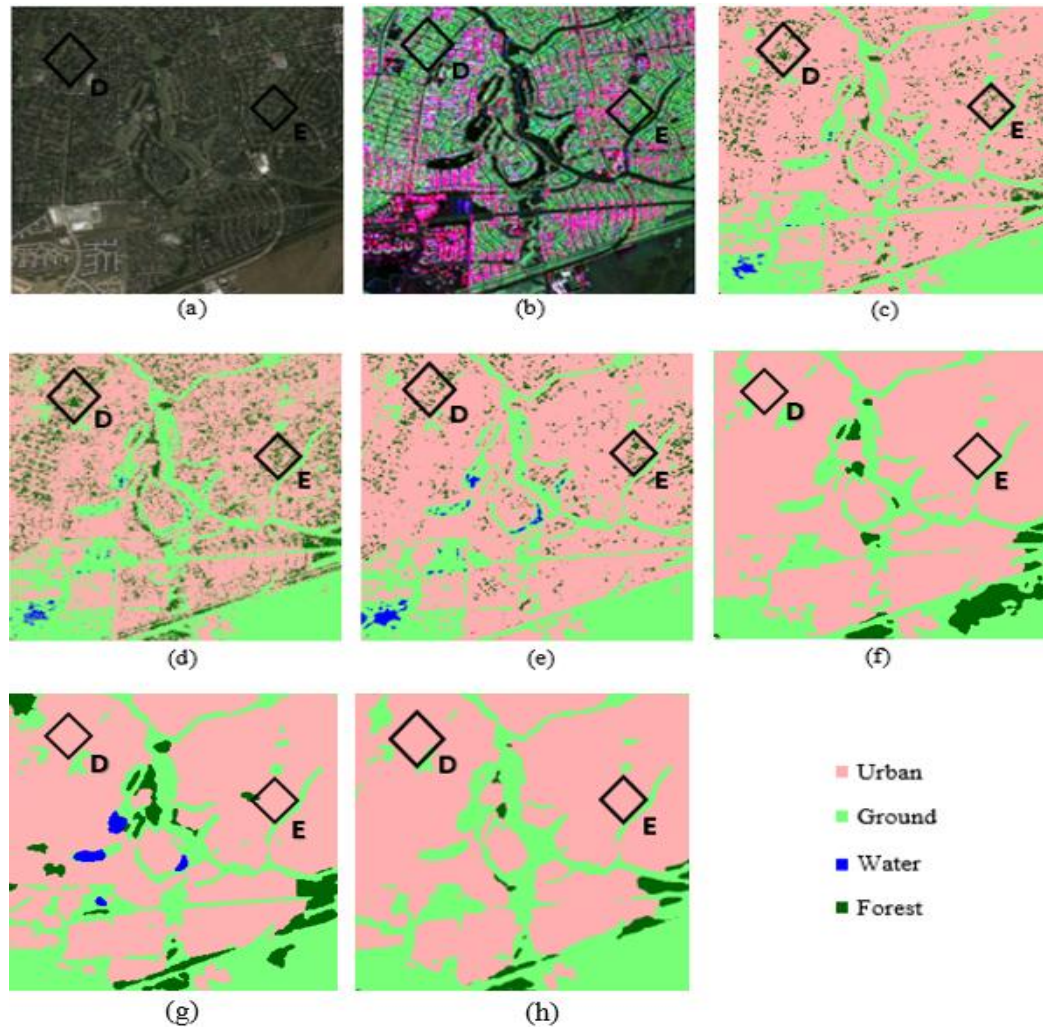


Figure 5.13 (a) represents the ground truth through the Google Earth image of test patch 2 from Houston City, Texas (USA) (b) False color composite of UAVSAR image of the same test patch and the corresponding image segmentation results for (c) RF classifier, (d) KNN classifier, (e) SVM classifier, (f) FPN, (g) UNET, and (h) DeepLabv3+ respectively (Garg et al., 2021b).

Table 7. The results of performance evaluation of traditional machine learning algorithms and DeepLabv3+ algorithm (Garg et al., 2021b)

Parameter	RF	KNN	SVM	FPN	UNET	DeepLabv3+
Pixel Accuracy _{Patch 1}	74.74%	78.68%	74.44%	75.28%	75.87%	83.51%
Pixel Accuracy _{Patch 2}	81.09%	74.31%	79.59%	84.62%	86.52%	87.78%
Overall Pixel Accuracy	77.92%	76.48%	77.02%	79.95%	81.19%	85.65%
Precision (Urban Class)	0.8958	0.7984	0.8977	0.9172	0.9149	0.9228
Training Time	4204 sec	41091 sec	2803 sec	—	—	13411 sec
Inference Time (per scene)	3.04 sec	864.94 sec	0.46 sec	—	—	0.31 sec

The image segmentation job is evaluated using a variety of performance assessment matrices, including Accuracy and Precision. In this work, the performance of the semantic segmentation job has been assessed using a pixel accuracy metric. Since one of the goals of

this study is to address the problem of oriented urban targets that are frequently misclassified as vegetation, the Precision metric for the urban class has also been presented. Table 6 summarises all of the outcomes based on these performance indicators.



Figure 5.14 Demonstrates the performance metrics. (a) and (b) show the pixel accuracy for patches 1 and 2, respectively; and (c) shows the overall pixel accuracy for both patches for RF, KNN, SVM, FPN, UNET and DeepLabv3+ algorithms.

The performance metrics for the implemented algorithms are depicted in figure 5.14. As seen in figure 5.14a for the test patch 1 or scene 1, the DeepLabv3+ algorithm achieved the highest pixel accuracy of 83.51%, while the KNN classifier recorded the highest pixel accuracy of 78.68% among machine learning algorithms. As reported in table 7, both Random forest and SVM classifiers recorded pixel accuracy below 75%. The DeepLabv3+ algorithm recorded the highest pixel accuracy of 87.78% for test patch 2, as shown in figure 5.14b, while the Random

Forest classifier achieved the best pixel accuracy of 81.09% among machine learning algorithms. Trailing behind were the SVM and KNN classifiers with pixel accuracy of 79.59% and 74.31%, respectively. It can be observed from figure 5.12g and figure 5.13g that the UNET algorithm is more sensitive to the water class. It has over-estimated the water class slightly in figure 5.12g and figure 5.13g.

The DeepLabv3+ algorithm, seen in figure 5.14c, achieved the highest overall pixel accuracy of 85.51%, whereas the machine learning algorithms Random Forest, KNN, and SVM recorded overall pixel accuracies of 77.92%, 76.48%, and 77.02%, respectively. The FPN model achieves an overall pixel accuracy of 79.95% and the UNET based model achieves an overall pixel accuracy of 81.21%. Because RF classifier fails to correctly map the river water in patch 1, the performance of RF is shown to be better in patch 2, which has more urban area than patch 1, which has a good contribution from all classes. Furthermore, the KNN algorithm's performance suffers in the case of patch 2, mostly as a result of its inability to recognize orientated urban objects accurately and mislabeling them as vegetation. As reported in table 6, DeepLabv3+ logged the highest precision of 0.9228, while all other machine learning algorithms logged a precision value of less than 0.90 for the urban class for the same region. This demonstrates that the oriented urban targets that are misclassified by the decomposition methods can be accurately classified by the DeepLabv3+ model.

CHAPTER 6

CONCLUSION

The progression of machine learning and deep learning algorithms for the purpose of land cover classification has faced a major obstacle: the absence of a universal, labeled benchmark dataset. If such a dataset were available, it would allow researchers to train and evaluate models in a more consistent manner, leading to more accurate land cover maps. Unfortunately, this gap has hindered progress, especially in the creation of highly precise maps that are essential for tackling today's pressing environmental challenges.

Land cover maps have become increasingly important as we confront issues like land degradation, uncontrolled urban expansion, deforestation, water pollution, and climate change. Accurate maps are indispensable for making informed decisions in urban planning, environmental management, and sustainable development. This study has shown that, despite the limitations of small datasets, deep learning models can be leveraged to produce accurate land cover maps when combined with transfer learning techniques.

One of the key findings of this research is the effective use of transfer learning to address the scarcity of labeled data in microwave remote sensing. In most cases, training a deep learning model from scratch requires large quantities of labeled data and significant computational power. However, transfer learning provides a way around these limitations by utilizing pre-trained models. This approach dramatically reduces the training time and computational resources needed while still delivering high levels of accuracy. In short, transfer learning proved to be a practical and efficient solution for land cover mapping in this study.

A comparative analysis of PolSAR and optical data from the Mumbai region revealed some interesting results. For instance, false-color composite images based on optical data delivered the best classification results overall, especially in distinguishing between different types of land cover, like vegetation and urban structures. However, when it came to specific features, such as detecting water bodies or distinguishing urban structures under less-than-ideal conditions, SAR data outperformed optical data.

The research further assessed the efficacy of various machine learning algorithms, revealing that the Random Forest (RF) classifier consistently surpassed alternative techniques such as k-Nearest Neighbors (KNN) and Support Vector Machines (SVM). KNN, known for being slow to train and perform inference, proved impractical for large-scale tasks, while SVM showed promise in detecting water bodies but fell short in more complex urban environments. These findings underscore the importance of choosing the right algorithm depending on the specific nature of the land cover to be classified.

One of the persistent challenges in using PolSAR data is the accurate classification of urban areas. Urban environments are incredibly complex, with buildings, roads, and other infrastructure often reflecting electromagnetic waves in ways that can be confused with natural features, such as vegetation. This results in the frequent misclassification of urban areas as natural regions, which can skew land cover maps.

In response to this challenge, this study employed transfer learning with the Deeplabv3+ model. The Deeplabv3+ model, known for its state-of-the-art segmentation performance, was able to overcome the issue of misclassifying urban areas. Its use of atrous convolution and an encoder-decoder architecture allowed it to capture fine details and multi-scale contextual information, which proved essential for distinguishing between urban and natural features. This finding is especially significant as it challenges the prevalent assumption that extensive datasets are essential for deep learning models to exceed the efficacy of conventional machine learning techniques. Despite the constraints of a smaller dataset, Deeplabv3+ successfully outperformed recognized algorithms such as Random Forest, KNN, and SVM.

The results from experiments on the Houston dataset, captured by a UAVSAR sensor, provided further evidence of Deeplabv3+'s effectiveness. The model achieved the highest pixel accuracy—83.51% for Patch 1 and 87.78% for Patch 2. These results highlight Deeplabv3+'s ability to handle the complexities of PolSAR data, particularly in urban settings. Among the machine learning models, Random Forest was the top performer, reaching an overall pixel accuracy of 77.92%.

Interestingly, the study also highlighted the drawbacks of the KNN classifier, particularly its inefficiency. KNN had the longest average inference time of more than 860

seconds per scene. These figures make KNN an impractical choice for large-scale applications in remote sensing, especially when compared to more efficient models like Random Forest and Deeplabv3+. Deeplabv3+, by contrast, excelled in detecting intricate urban features such as narrow streets and complex building patterns, delivering highly accurate results in a fraction of the time.

The findings from this study pave the way for further research and development in the field of PolSAR segmentation. One critical area for future exploration is the creation of a large, publicly accessible benchmark dataset. Such a dataset would enable researchers to train and evaluate different models more effectively, providing a standard against which various approaches can be measured.

Another promising avenue for future research is the fusion of optical and PolSAR data for semantic segmentation. Optical data is particularly effective for detecting vegetation and urban structures under normal conditions, while PolSAR data shines in less ideal conditions, such as cloudy weather or low light. By combining these two data types in physics inspired deep learning models, future studies could create even more accurate land cover maps. For example, optical data could help distinguish between urban structures and vegetation, while PolSAR could enhance the detection of water bodies and other features that are hard to spot with optical data alone.

Additionally, there is great potential in exploring other deep learning models beyond Deeplabv3+. Models such as U-Net, SegNet, and Mask R-CNN have shown promising results in other fields of image segmentation and could be applied to PolSAR data. Testing these models on datasets similar to those used in this study would provide new insights into their strengths and limitations, and potentially open up new opportunities for innovation in land cover mapping.

Another future focus could be on improving computational efficiency. While models like Deeplabv3+ offer exceptional accuracy, they can still be resource-intensive. Techniques such as model compression, distributed computing, or even developing lightweight models could make these advanced tools more accessible, particularly for regions or institutions with limited computational resources.

In conclusion, this study elucidates the efficacy and promise inherent in deep learning architectures, particularly Deeplabv3+, for the enhancement of PolSAR data segmentation. The utilization of transfer learning has effectively mitigated some of the most significant obstacles in PolSAR analysis, such as the erroneous classification of urban regions and the insufficiency of labeled training datasets. The results unequivocally demonstrate that while conventional machine learning algorithms, such as Random Forest and SVM, retain their relevance, deep learning frameworks offer a superior and more proficient approach for intricate segmentation challenges.

Moving forward, the fusion of PolSAR and optical data, the availability of standardized datasets, and the exploration of other deep learning models will likely drive the next wave of advancements in this field. These innovations could revolutionize land cover mapping, offering more detailed, accurate, and timely insights to support environmental monitoring, urban planning, and disaster management worldwide.

References

Abreu R. D. et. al., “RADARSAT-2 mode selection for maritime surveillance”, DRDC, Ottawa, Canada.

Ali, M.Z., Qazi, W., Aslam, N. (2018). A comparative study of ALOS-2 PALSAR and landsat-8 imagery for land cover classification using maximum likelihood classifier. *Egypt J. Remote Sens. Sp. Sci.* 21, S29–S35. <https://doi.org/10.1016/j.ejrs.2018.03.003>

Al-Ruzouq, R., Gibril, M. B. A., Shanableh, A., Kais, A., Hamed, O., Al-Mansoori, S., & Khalil, M. A. (2020). Sensors, Features, and Machine Learning for Oil Spill Detection and Monitoring: A Review. *Remote Sensing*, 12(20), 3338. <https://doi.org/10.3390/rs12203338>

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), 2481–2495. <https://doi.org/10.1109/tpami.2016.2644615>

Bamler R. (2000). Principles of Synthetic Aperture Radar. *Surveys in Geophysics*, 21, 147-157.

Barnes, R.M. (1988). In: Roll Invariant Decompositions for the Polarisation Covariance Matrix, Internal report, pp. 02173–10073

Barrett BW, Dwyer E, Whelan P., (2009). Soil Moisture Retrieval from Active Spaceborne Microwave Observations: An Evaluation of Current Techniques, *Remote Sensing*, 1(3), 210-242. <https://doi.org/10.3390/rs1030210>

Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>

Bhattacharya, A., Singh, G., Manickam, S., & Yamaguchi, Y. (2015). An Adaptive General Four-Component Scattering Power Decomposition With Unitary Transformation of Coherency Matrix (AG4U). *IEEE Geoscience and Remote Sensing Letters*, 12(10), 2110–2114. <https://doi.org/10.1109/lgrs.2015.2451369>

Bihanta Toosi, N., Soffianian, A. R., Fakheran, S., Pourmanafi, S., Ginzler, C., & T. Waser, L. (2020). Land Cover Classification in Mangrove Ecosystems Based on VHR Satellite Data and Machine Learning—An Upscaling Approach. *Remote Sensing*, 12(17), 2684. <https://doi.org/10.3390/rs12172684>

Breiman L. (2001). Random Forests. *Machine Learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>

Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123–140. <https://doi.org/10.1007/bf00058655>

Brihanmumbai Municipal Corporation. (2021). https://portal.mcgm.gov.in/irj/portal/anonymouse/qlvitalstatsreport?guest_user=english.

Buchhorn, M., Smets, B., Bertels, L., Roo, B. De, Lesiv, M., Tsendbazar, N.-E., Herold, M., Fritz, S., Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe [WWW Document]. Copernicus L. Monit. Serv. <https://doi.org/10.5281/ZENODO.3939050>.

Buchhorn, M., Smets, B., Bertels, L., Roo, B. De, Lesiv, M., Tsendbazar, N.-E., Li, L., Tarko, A., Copernicus Global Land Operations “Vegetation and Energy” “CGLOPS-1” I3.3. ed, Copernicus Global Land Operations. Copernicus Global Land Operations, Zenodo, Geneve, Switzerland. <https://doi.org/10.5281/zenodo.3938963>.

Camargo, F. F., Sano, E. E., Almeida, C. M., Mura, J. C., & Almeida, T. (2019). A Comparative Assessment of Machine-Learning Techniques for Land Use and Land Cover Classification of the Brazilian Tropical Savanna Using ALOS-2/PALSAR-2

Polarimetric Images. *Remote Sensing*, 11(13), 1600.
<https://doi.org/10.3390/rs11131600>

Cameron, W.L., Leung, L.K. (1990). Feature motivated polarization scattering matrix decomposition. In: *IEEE International Conference on Radar*. pp 549–557.
<https://doi.org/10.1109/RADAR.1990.201088>

Canny, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6), 679–698.
<https://doi.org/10.1109/tpami.1986.4767851>

Chandrasekhar S., (1960). Radiative Transfer, Dover Publications, New York.

Charbonneau, F. J., Brisco, B., Raney, R. K., McNairn, H., Liu, C., Vachon, P. W., Shang, J., DeAbreu, R., Champagne, C., Merzouki, A., Geldsetzer, T. (2010). Compact polarimetry overview and applications assessment. *Can. J. Remote Sens*, 36, S298–S315. <https://doi.org/10.5589/m10-062>

Chaudhary, V., & Kumar, S. (2021). Dark spot detection for characterization of marine surface slicks using UAVSAR quad-pol data. *Scientific Reports*, 11(1).
<https://doi.org/10.1038/s41598-021-88301-9>

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2016). Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. <https://doi.org/10.48550/arXiv.1412.7062>

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018, April 1). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834–848. <https://doi.org/10.1109/tpami.2017.2699184>

Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation.

Computer Vision – ECCV 2018, 833–851. https://doi.org/10.1007/978-3-030-01234-2_49

Chen, L.-C., Papandreou, G., Schroff, F., and Adam, H. (2017). Rethinking Atrous Convolution for Semantic Image Segmentation”. <https://doi.org/10.48550/arXiv.1706.05587>

Chen, S., Wang, H., Xu, F., & Jin, Y. Q. (2016). Target Classification Using the Deep Convolutional Networks for SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8), 4806–4817. <https://doi.org/10.1109/tgrs.2016.2551720>.

Chen, S.-W., Wang, X.-S., Li, Y.-Z., & Sato, M. (2014). Adaptive Model-Based Polarimetric Decomposition Using PolInSAR Coherence. *IEEE Transactions on Geoscience and Remote Sensing*, 52(3), 1705–1718. <https://doi.org/10.1109/tgrs.2013.2253780>

Chun Liu, C. L., Junjun Yin, J. Y., & Jian Yang, J. Y. (2015). Application of deep learning to polarimetric SAR classification. IET International Radar Conference 2015. <https://doi.org/10.1049/cp.2015.1182>

Cloude, S. (1985). Target decomposition theorems in radar scattering. *Electronics Letters*, 21(1), 22–24. <https://doi.org/10.1049/el:19850018>

Cloude, S., & Pottier, E. (1996). A review of target decomposition theorems in radar polarimetry. *IEEE Transactions on Geoscience and Remote Sensing*, 34(2), 498–518. <https://doi.org/10.1109/36.485127>

Cloude, S.R., Pottier, E. (1997). An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*. 35, 68–78. <https://doi.org/10.1109/36.551935>

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/tit.1967.1053964>.

Del Frate, F., Pacifici, F., Schiavon, G., & Solimini, C. (2007). Use of Neural Networks for Automatic Classification from High-Resolution Images. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4), 800–809. <https://doi.org/10.1109/tgrs.2007.892009>

Deng, G., & Cahill, L. (1993). An adaptive Gaussian filter for noise reduction and edge detection. *1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference*. <https://doi.org/10.1109/nssmic.1993.373563>

Dong, Y., Du, B. & Zhang, L. (2015). Target detection based on random forest metric learning”, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8, 1830–1838. <https://doi.org/10.1109/JSTARS.2015.2416255>

Dong, Y., Li, M., Li, J. (2013). Image retrieval based on improved Canny edge detection algorithm. *Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)*. <https://doi.org/10.1109/mec.2013.6885296J>

Elshaw R., Maher, M., Sakr, S. (2019). Automated machine learning: State-of-the-art and open challenges. <http://arxiv.org/abs/1906.02287>

Fingas, M., & Brown, C. E. (2014). Oil Spill Remote Sensing. *Handbook of Oil Spill Science and Technology*, 311–356. <https://doi.org/10.1002/9781118989982.ch12>

Freeman, A. & Curlander, J. C. (1989). Radiometric Correction and Calibration of SAR Images. *Photogrammetric Engineering and Remote Sensing*, 55(9).

Freeman, A., & Durden, S. (1998). A three-component scattering model for Polarimetric SAR data”, *IEEE Transactions on Geoscience and Remote Sensing*, 36(3), 963–973. <https://doi.org/10.1109/36.673687>

Gabor D. (1946). Theory of communication, *Journal of the Institution of Electrical Engineers*, 93, 429-441.

Gao, W., Liu, Y., Zeng, Y., Liu, Q., & Li, Q. (2023). SAR Image Ship Target Detection Adversarial Attack and Defence Generalization Research. *Sensors*, 23(4), 2266. <https://doi.org/10.3390/s23042266>

Garg, R., Kumar, A., Prateek, M., Pandey, K., & Kumar, S. (2021). Land cover classification of spaceborne multifrequency SAR and optical multispectral data using machine learning. *Advances in Space Research*, 69(4), 1726–1742. <https://doi.org/10.1016/j.asr.2021.06.028>

Garg, R., Kumar, A., Bansal, N., Prateek, M., & Kumar, S. (2021b). Semantic segmentation of PolSAR image data using advanced deep learning model. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-94422-y>

Garg, R., Singh, K., Singh, I., Kumar, A., Kumar, S. and Tiwari, R. (2023) "Semantic Segmentation of PolSAR Data Using CNN Based Deep Learning Models," *2023 11th International Conference on Intelligent Systems and Embedded Design (ISED)*, Dehradun, India, pp. 1-4, <https://doi.org/10.1109/ISED59382.2023.10444583>.

Gašparović, M., & Dobrinić, D. (2020, June 17). Comparative Assessment of Machine Learning Methods for Urban Vegetation Mapping Using Multitemporal Sentinel-1 Imagery. *Remote Sensing*, 12(12), 1952. <https://doi.org/10.3390/rs12121952>

Gazzea, M., Solheim, A., & Arghandeh, R. (2023). High-resolution mapping of forest structure from integrated SAR and optical images using an enhanced U-net method. *Science of Remote Sensing*, 8, 100093. <https://doi.org/10.1016/j.srs.2023.100093>

Gogna, A., & Tayal, A. (2013, December). Metaheuristics: review and application. *Journal of Experimental & Theoretical Artificial Intelligence*, 25(4), 503–526. <https://doi.org/10.1080/0952813x.2013.782347>

- Gomez, R., Arigony-Neto, J., De Santis, A., Vijay, S., Jaña, R., & Rivera, A. (2019). Ice dynamics of union glacier from SAR offset tracking. *Global and Planetary Change*, 174, 1–15. <https://doi.org/10.1016/j.gloplacha.2018.12.012>
- Gong, M., Zhao, J., Liu, J., Miao, Q., & Jiao, L. (2016, January). Change Detection in Synthetic Aperture Radar Images Based on Deep Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 27(1), 125–138. <https://doi.org/10.1109/tnnls.2015.2435783>
- Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). KNN Model-Based Approach in Classification. *On the Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*, 986–996. https://doi.org/10.1007/978-3-540-39964-3_62
- Hashemi, M. G., Jalilvand, E., Alemohammad, H., Tan, P., & Das, N. N. (2024). Review of synthetic aperture radar with deep learning in agricultural applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 218, 20–49. <https://doi.org/10.1016/j.isprsjprs.2024.08.018>
- Hearst, M., Dumais, S., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and Their Applications*, 13(4), 18–28. <https://doi.org/10.1109/5254.708428>
- Hensley S., and Tighe L. (2008). Interferometric Synthetic Aperture Radar (IfSAR).
- Hinton, G. E., Osindero, S., & Teh, Y. W. (2006, July). A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, 18(7), 1527–1554. <https://doi.org/10.1162/neco.2006.18.7.1527>
- Holschneider, M., Kronland-Martinet, R., Morlet, J., & Tchamitchian, P. (1990). A Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform. *Inverse Problems and Theoretical Imaging*, 286–297. https://doi.org/10.1007/978-3-642-75988-8_28

Hosseini, M., & Lim, S. (2023). Burned area detection using Sentinel-1 SAR data: A case study of Kangaroo Island, South Australia. *Applied Geography*, 151, 102854. <https://doi.org/10.1016/j.apgeog.2022.102854>

Huynen, J.R., (1970). Phenomenological Theory of Radar Targets Ph.D. Dissertation. Technical University, Delft, Netherlands.

Itano, F., de Abreu de Sousa, M. A., & Del-Moral-Hernandez, E. (2018). Extending MLP ANN hyper-parameters Optimization by using Genetic Algorithm. *2018 International Joint Conference on Neural Networks (IJCNN)*. <https://doi.org/10.1109/ijcnn.2018.8489520>

Jansing, E. D. (2021). Introduction to Synthetic Aperture Radar: Concepts and Practice. McGraw Hill Professional.

Joshi A. K., and Das S. K. (2013). Use of SAR data in exploration of uranium mineralization. *RISAT Utilization Programme Compilation of Pre-RISAT Results*, SAC, Ahmedabad.

Kankaku, Y., Suzuki, S., & Osawa, Y. (2013). ALOS-2 mission and development status. *2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*. <https://doi.org/10.1109/igarss.2013.6723302>

Karachristos, K., Koukiou, G., & Anastassopoulos, V. (2024). A Review on PolSAR Decompositions for Feature Extraction. *Journal of Imaging*, 10(4), 75. <https://doi.org/10.3390/jimaging10040075>

Krogager, E. (1990). New decomposition of the radar target scattering matrix. *Electron. Lett.* 26, 1525-1527(2). <https://doi.org/10.1049/el:19900979>

Kumar, S., Pandey, U., Kushwaha, S. P., Chatterjee, R. S., & Bijker, W. (2012). Aboveground biomass estimation of tropical forest from Envisat advanced synthetic aperture radar data using modeling approach. *Journal of Applied Remote Sensing*, 6(1), 063588. <https://doi.org/10.1117/1.jrs.6.063588>

Kumar, A., & Prateek, M. (2020). Localization of Nuclei in Breast Cancer Using Whole Slide Imaging System Supported by Morphological Features and Shape Formulas. *Cancer Management and Research*, Volume 12, 4573–4583. <https://doi.org/10.2147/cmar.s248166>

Kumar, S., & Sharma, A. (2023). Synthetic Aperture Radar Remote Sensing. In CRC Press eBooks (pp. 1–12). <https://doi.org/10.1201/9781003204466-1>

Lacomme, P., Hardange, J. P., Marchais, J. C., & Normant, E. (2001). Synthetic Aperture Radar. In Elsevier eBooks (pp. 233–264). <https://doi.org/10.1016/b978-189112113-5.50017-6>

Lapini, A., Pettinato, S., Santi, E., Paloscia, S., Fontanelli, G., & Garzelli, A. (2020). Comparison of Machine Learning Methods Applied to SAR Images for Forest Classification in Mediterranean Areas. *Remote Sensing*, 12(3), 369. <https://doi.org/10.3390/rs12030369>

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>

LeCun, Y., Kavukcuoglu, K., & Farabet, C. (2010). Convolutional networks and applications in vision. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*. <https://doi.org/10.1109/iscas.2010.5537907>

Lee J. S. (1980). Digital Image Enhancement and Noise Filtering by Use of Local Statistics. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-2(2), 165–168.

Lee, J. S. and Jurkevich, I. (1989). Segmentation of SAR images. *IEEE Trans. Geoscience and Remote Sensing*, 27 (6), 674–680.

Lee, J. S., & Pottier, E. (2017). Polarimetric Radar Imaging. In CRC Press eBooks. <https://doi.org/10.1201/9781420054989>

- Lee, J. S., Jurkevich, L., Dewaele, P., Wambacq, P., & Oosterlinck, A. (1994). Speckle filtering of synthetic aperture radar images: A review. *Remote Sensing Reviews*, 8(4), 313–340. <https://doi.org/10.1080/02757259409532206>
- Li, M., Zou, H., Dong, Z., Qin, X., Liu, S., & Zhang, Y. (2024). Unsupervised semantic segmentation of POLSAR images based on multiview similarity. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, 5317–5331. <https://doi.org/10.1109/jstars.2024.3365664>
- Liu, C. A., Chen, Z. X., Shao, Y., Chen, J. S., Hasi, T., & Pan, H. Z. (2019). Research advances of SAR remote sensing for agriculture applications: A review. *Journal of Integrative Agriculture*, 18(3), 506–525. [https://doi.org/10.1016/s2095-3119\(18\)62016-7](https://doi.org/10.1016/s2095-3119(18)62016-7)
- Liu, F., Jiao, L., Hou, B., & Yang, S. (2016). POL-SAR Image Classification Based on Wishart DBN and Local Spatial Information. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6), 3292–3308. <https://doi.org/10.1109/tgrs.2016.2514504>
- Livingstone, C.E., Sikaneta, I., Gierull, C., Chiu, S., Beaulne, P. (2005) RADARSAT-2 System and Mode Description. In *Integration of Space-Based Assets within Full Spectrum Operations* (pp. 15-1 – 15-22). Meeting Proceedings RTO-MP-SCI-150, Paper 15. Neuilly-sur-Seine, France: RTO. Available from: <http://www.rto.nato.int/abstracts.asp>
- López-Martínez, C., & Pottier, E. (2021). Basic Principles of SAR Polarimetry. In *Remote sensing and digital image processing* (pp. 1–58). https://doi.org/10.1007/978-3-030-56504-6_1
- Lv, Q., Dou, Y., Niu, X., Xu, J., Xu, J., & Xia, F. (2015). Urban Land Use and Land Cover Classification Using Remotely Sensed SAR Data through Deep Belief Networks. *Journal of Sensors*, 2015, 1–10. <https://doi.org/10.1155/2015/538063>

M E, B. P., & Kumar, S. (2020). PolInSAR decorrelation-based decomposition modelling of spaceborne multifrequency SAR data. *International Journal of Remote Sensing*, 42(4), 1398–1419. <https://doi.org/10.1080/01431161.2020.1829155>

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 166–177. <https://doi.org/10.1016/j.isprsjprs.2019.04.015>

Ma, X., Geng, J., & Wang, H. (2015). Hyperspectral image classification via contextual deep learning. *EURASIP Journal on Image and Video Processing*, 2015(1). <https://doi.org/10.1186/s13640-015-0071-8>

Maghsoudi, Y. (2012). Analysis of Radarsat-2 Full Polarimetric Data for Forest Mapping”, University of Calgary (Canada). <https://doi.org/10.11575/PRISM/4817>

Mordor Intelligence, “Synthetic Aperture Radar Market – Growth, Trends, COVID-19 Impact, and Forecasts (2022 - 2027)”, 09 August 2022. <https://www.mordorintelligence.com/industry-reports/synthetic-aperture-radar-market>

Mostajabi, M., Yadollahpour, P., and Shakhnarovich, G. (2014). Feedforward semantic segmentation with zoom-out features. arXiv:1412.0774, 2014.

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247–259. <https://doi.org/10.1016/j.isprsjprs.2010.11.001>

Nogueira, K., Penatti, O. A., & dos Santos, J. A. (2017). Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognition*, 61, 539–556. <https://doi.org/10.1016/j.patcog.2016.07.001>

Ouchi, K. (2013). Recent Trend and Advance of Synthetic Aperture Radar with Selected Topics. *Remote Sensing*, 5(2), 716–807. <https://doi.org/10.3390/rs5020716>

Pacifici, F., Chini, M., & Emery, W. J. (2009). A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing of Environment*, 113(6), 1276–1292. <https://doi.org/10.1016/j.rse.2009.02.014>

Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. *JMLR* 12, 2825-2830.

Qi, Z., Yeh, A.G-O., Li, X., Lin, Z. (2012) “A novel algorithm for land use and land cover classification using RADARSAT-2 polarimetric SAR data”, *Remote Sens. Environ.*, 118, 21–39. <https://doi.org/10.1016/j.rse.2011.11.001>.

Ramya, M., & Kumar, S. (2021). PolInSAR coherence-based decomposition modeling for scattering characterization: A case study in Uttarakhand, India. *Science of Remote Sensing*, 3, 100020. <https://doi.org/10.1016/j.srs.2021.100020>

Rawat, A., Kumar, A., Upadhyay, P., & Kumar, S. (2021). Deep learning-based models for temporal satellite data processing: Classification of paddy transplanted fields. *Ecological Informatics*, 61, 101214. <https://doi.org/10.1016/j.ecoinf.2021.101214>

Ronneberger, O., Fischer, P. & Brox, T. (2015). U-Net: convolutional networks for biomedical image segmentation BT. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. 234–241. <https://doi.org/10.48550/arXiv.1505.04597>

Roth, Achim & Hoffmann, Jörn & Esch, Thomas. (2005). TerraSAR-X: How can high resolution SAR data support the observation of urban areas?, *In: ISPRS Proceedings. ISPRS WG VIII/1. ISPRS Joint Conference URBAN2005 and URS2005*, Tempe, AZ (USA)

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet

Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3), 211–252. <https://doi.org/10.1007/s11263-015-0816-y>

S. Lessmann, R. Stahlbock, S.F. Crone. (2005). Optimizing hyperparameters of support vector machines by genetic algorithms, *Proc. 2005 Int. Conf. Artif. Intell. ICAI'05*.

Sabrian, P. G., Saepuloh, A., Syafrizal, N., & Hede, A. N. H. (2017). Identification of Altered Minerals Based on Synthetic Aperture Radar (SAR) For Mineral Exploration in a Tropical Area. *IOP Conference Series. Earth and Environmental Science*, 71, 012021. <https://doi.org/10.1088/1755-1315/71/1/012021>

Salakhutdinov, R. & Hinton, G. (2009). Deep Boltzmann Machines. Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics, in *Proceedings of Machine Learning Research*, 5:448-455. Available from <https://proceedings.mlr.press/v5/salakhutdinov09a.html>.

Sentinel-1 - Mission Summary - Sentinel Online - Sentinel Online, Sentinel Online. <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/overview/mission-summary>

Shafai, S. S., & Kumar, S. (2020). PolInSAR Coherence and Entropy-Based Hybrid Decomposition Model. *Earth and Space Science*, 7(10). <https://doi.org/10.1029/2020ea001279>

Shelhamer, E., Long, J., & Darrell, T. (2017). Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 640–651. <https://doi.org/10.1109/tpami.2016.2572683>

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., & Lucas, R. (2014). New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sensing of Environment*, 155, 13–31. <https://doi.org/10.1016/j.rse.2014.04.014>

Shunlin Liang, Xiaowen Li, Jindi Wang, (2012). Land Cover and Land use Changes, *Advanced Remote Sensing*, 703-772. <https://doi.org/10.1016/B978-0-12-385954-9.00024-1>

Singh, G., & Yamaguchi, Y. (2018). Model-Based Six-Component Scattering Matrix Power Decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, 56(10), 5687–5704. <https://doi.org/10.1109/tgrs.2018.2824322>

Singh, G., Yamaguchi, Y., & Park, S. E. (2013). General Four-Component Scattering Power Decomposition with Unitary Transformation of Coherency Matrix. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5), 3014–3022. <https://doi.org/10.1109/tgrs.2012.2212446>

Solberg R., Malnes E., Amlien J., et. al. (2008). State of the art for tropical forest monitoring by remote sensing, Norwegian Computing Center. <https://nr.no/en/publication/957998/>

Shahzad, M., Singh, A., & Saha, S. (2025). Synthetic aperture radar image analysis in era of deep learning. In Elsevier eBooks (pp. 125–140). <https://doi.org/10.1016/b978-0-44-326484-9.00014-2>

Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y. A., & Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. *Remote Sensing*, 12(7), 1135. <https://doi.org/10.3390/rs12071135>

Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR). (2021). <https://uavsar.jpl.nasa.gov/cgi-bin/data.pl>

Valarmathi, N., Tyagi, R. N., Kamath, S. M., Reddy, B. T., VenkataRamana, M., Srinivasan, V. V., Dutta, C., Veena, N., Venketesh, K., Raveendranath, G. N., Babu, G. R. C., Prasad, K. S., Badagandi, R. R., Natarajan, P., Sudhakar, S., Subhalakshmi, J., Rao, S., & Reddy, M. K. (2013). RISAT-1 spacecraft configuration: architecture, technology and performance. *Current Science*, 104(4), 462–471. <http://www.jstor.org/stable/24089747>

Verma, S., Kumar, S., & Dsouza, H. (2023). Polarimetric Decomposition and Machine Learning-Based Classification of L- and S-Band Airborne SAR (LS-ASAR) data. *Earth and Space Science*, 10(6). <https://doi.org/10.1029/2022ea002796>

Verma R. (2012). Polarimetric decomposition based on general characterization of scattering from urban areas and multiple component scattering model. [Universiteit Twente]. https://library.itc.utwente.nl/papers_2012/msc/gfm/verma.pdf

Way, J., & Smith, E. (1991). The evolution of synthetic aperture radar systems and their progression to the EOS SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 29(6), 962–985. <https://doi.org/10.1109/36.101374>

Xueyun Chen, Shiming Xiang, Cheng-Lin Liu, & Chun-Hong Pan. (2014). Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks. *IEEE Geoscience and Remote Sensing Letters*, 11(10), 1797–1801. <https://doi.org/10.1109/lgrs.2014.2309695>

Yamaguchi, Y., Moriyama, T., Ishido, M., & Yamada, H. (2005). Four-component scattering model for polarimetric SAR image decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, 43(8), 1699–1706. <https://doi.org/10.1109/tgrs.2005.852084>

Yamaguchi, Y., Sato, A., Boerner, W. M., Sato, R., & Yamada, H. (2011). Four-Component Scattering Power Decomposition with Rotation of Coherency Matrix. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6), 2251–2258. <https://doi.org/10.1109/tgrs.2010.2099124>

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415, 295–316. <https://doi.org/10.1016/j.neucom.2020.07.061>

Yao, Q., et al. (2018). Taking Human out of Learning Applications: A Survey on Automated Machine Learning. <http://arxiv.org/abs/1810.13306>.

Yu, Y., Li, M., & Fu, Y. (2017). Forest type identification by random forest classification combined with SPOT and multitemporal SAR data. *Journal of Forestry Research*, 29(5), 1407–1414. <https://doi.org/10.1007/s11676-017-0530-4>

Zhang, C. (2018). Deep learning for land cover and land use classification. Lancaster University. <https://doi.org/10.17635/lancaster/thesis/428>

Zhang, L., Zhang, L., & Du, B. (2016). Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geoscience and Remote Sensing Magazine*, 4(2), 22–40. <https://doi.org/10.1109/mgrs.2016.2540798>

Zhang, L., Zou, B., Cai, H., & Zhang, Y. (2008). Multiple-Component Scattering Model for Polarimetric SAR Image Decomposition. *IEEE Geoscience and Remote Sensing Letters*, 5(4), 603–607. <https://doi.org/10.1109/lgrs.2008.2000795>

List of Publications

Name: Rajat Garg

ORCID ID: <https://orcid.org/0000-0001-7000-1795>

1. Garg, R., Kumar, A., Bansal, N., Prateek, M., & Kumar, S. (2021). Semantic segmentation of PolSAR image data using advanced deep learning model. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-94422-y>
2. Garg, R., Kumar, A., Prateek, M., Pandey, K., & Kumar, S. (2022). Land cover classification of spaceborne multifrequency SAR and optical multispectral data using machine learning. *Advances in Space Research*, 69(4), 1726–1742. <https://doi.org/10.1016/j.asr.2021.06.028>
3. Rajat Garg, Kartikay Singh, Isha Singh, et. al. (2023). “Semantic Segmentation of PolSAR Data Using CNN Based Deep Learning Models”, 2023 11th International Conference on Intelligent Systems and Embedded Design (ISED), Dehradun, India. 1-4. <https://doi.org/10.1109/ISED59382.2023.10444583>
4. Anil Kumar, Rajat Garg, Susheela Dahiya, Manish Prateek, Shashi Kumar. (2021). “Semantic Segmentation of L&S Band SAR Data after Tuning the Hyper Parameters in Machine Learning Models”, IEEE InGARSS. <https://doi.org/10.1109/InGARSS51564.2021.9792137>
5. Kumar, A., Garg, R., & Kumar, S. (2023a). Implementation of Machine learning classification models on Multifrequency Band SAR dataset. In CRC Press eBooks (pp. 89–105). <https://doi.org/10.1201/9781003204466-4>
6. Kumar, A., Garg, R., & Kumar, S. (2023b). Implementation of Neural Network-Based Classification models on Multifrequency Band SAR dataset. In CRC Press eBooks (pp. 107–121). <https://doi.org/10.1201/9781003204466-5>
7. Rajat Garg, Anil Kumar, and Shashi Kumar. (2020). “Implementation of Semantic Segmentation of UAVSAR Data”, The Proceedings of National Seminar on ‘Recent Advances in Geospatial Technology & Applications’, IIRS Dehradun, India

Appendix-1

1. Xception Model Summary

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	(None, 299, 299, 3)	0	
block1_conv1 (Conv2D)	(None, 149, 149, 32)	864	input_1[0][0]
block1_conv1_bn (BatchNormalization)	(None, 149, 149, 32)	128	block1_conv1[0][0]
block1_conv1_act (Activation)	(None, 149, 149, 32)	0	block1_conv1_bn[0][0]
block1_conv2 (Conv2D)	(None, 147, 147, 64)	18432	block1_conv1_act[0][0]
block1_conv2_bn (BatchNormalization)	(None, 147, 147, 64)	256	block1_conv2[0][0]
block1_conv2_act (Activation)	(None, 147, 147, 64)	0	block1_conv2_bn[0][0]
block2_sepconv1 (SeparableConv2)	(None, 147, 147, 128)	8768	block1_conv2_act[0][0]
block2_sepconv1_bn (BatchNormalization)	(None, 147, 147, 128)	512	block2_sepconv1[0][0]
block2_sepconv2_act (Activation)	(None, 147, 147, 128)	0	block2_sepconv1_bn[0][0]
block2_sepconv2 (SeparableConv2)	(None, 147, 147, 128)	17536	block2_sepconv2_act[0][0]
block2_sepconv2_bn (BatchNormalization)	(None, 147, 147, 128)	512	block2_sepconv2[0][0]
conv2d (Conv2D)	(None, 74, 74, 128)	8192	block1_conv2_act[0][0]
block2_pool (MaxPooling2D)	(None, 74, 74, 128)	0	block2_sepconv2_bn[0][0]
batch_normalization (BatchNormalization)	(None, 74, 74, 128)	512	conv2d[0][0]
add (Add)	(None, 74, 74, 128)	0	block2_pool[0][0] batch_normalization[0][0]
block3_sepconv1_act (Activation)	(None, 74, 74, 128)	0	add[0][0]
block3_sepconv1 (SeparableConv2)	(None, 74, 74, 256)	33920	block3_sepconv1_act[0][0]
block3_sepconv1_bn (BatchNormalization)	(None, 74, 74, 256)	1024	block3_sepconv1[0][0]
block3_sepconv2_act (Activation)	(None, 74, 74, 256)	0	block3_sepconv1_bn[0][0]
block3_sepconv2 (SeparableConv2)	(None, 74, 74, 256)	67840	block3_sepconv2_act[0][0]
block3_sepconv2_bn (BatchNormalization)	(None, 74, 74, 256)	1024	block3_sepconv2[0][0]
conv2d_1 (Conv2D)	(None, 37, 37, 256)	32768	add[0][0]
block3_pool (MaxPooling2D)	(None, 37, 37, 256)	0	block3_sepconv2_bn[0][0]
batch_normalization_1 (BatchNormalization)	(None, 37, 37, 256)	1024	conv2d_1[0][0]
add_1 (Add)	(None, 37, 37, 256)	0	block3_pool[0][0] batch_normalization_1[0][0]
block4_sepconv1_act (Activation)	(None, 37, 37, 256)	0	add_1[0][0]
block4_sepconv1 (SeparableConv2)	(None, 37, 37, 728)	188672	block4_sepconv1_act[0][0]
block4_sepconv1_bn (BatchNormalization)	(None, 37, 37, 728)	2912	block4_sepconv1[0][0]
block4_sepconv2_act (Activation)	(None, 37, 37, 728)	0	block4_sepconv1_bn[0][0]
block4_sepconv2 (SeparableConv2)	(None, 37, 37, 728)	536536	block4_sepconv2_act[0][0]
block4_sepconv2_bn (BatchNormalization)	(None, 37, 37, 728)	2912	block4_sepconv2[0][0]
conv2d_2 (Conv2D)	(None, 19, 19, 728)	186368	add_1[0][0]
block4_pool (MaxPooling2D)	(None, 19, 19, 728)	0	block4_sepconv2_bn[0][0]
batch_normalization_2 (BatchNormalization)	(None, 19, 19, 728)	2912	conv2d_2[0][0]
add_2 (Add)	(None, 19, 19, 728)	0	block4_pool[0][0] batch_normalization_2[0][0]
block5_sepconv1_act (Activation)	(None, 19, 19, 728)	0	add_2[0][0]
block5_sepconv1 (SeparableConv2)	(None, 19, 19, 728)	536536	block5_sepconv1_act[0][0]

block5_sepconv1_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block5_sepconv1[0][0]
block5_sepconv2_act (Activation)	(None, 19, 19, 728)	0	block5_sepconv1_bn[0][0]
block5_sepconv2 (SeparableConv2)	(None, 19, 19, 728)	536536	block5_sepconv2_act[0][0]
block5_sepconv2_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block5_sepconv2[0][0]
block5_sepconv3_act (Activation)	(None, 19, 19, 728)	0	block5_sepconv2_bn[0][0]
block5_sepconv3 (SeparableConv2)	(None, 19, 19, 728)	536536	block5_sepconv3_act[0][0]
block5_sepconv3_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block5_sepconv3[0][0]
add_3 (Add)	(None, 19, 19, 728)	0	block5_sepconv3_bn[0][0]
			add_2[0][0]
block6_sepconv1_act (Activation)	(None, 19, 19, 728)	0	add_3[0][0]
block6_sepconv1 (SeparableConv2)	(None, 19, 19, 728)	536536	block6_sepconv1_act[0][0]
block6_sepconv1_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block6_sepconv1[0][0]
block6_sepconv2_act (Activation)	(None, 19, 19, 728)	0	block6_sepconv1_bn[0][0]
block6_sepconv2 (SeparableConv2)	(None, 19, 19, 728)	536536	block6_sepconv2_act[0][0]
block6_sepconv2_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block6_sepconv2[0][0]
block6_sepconv3_act (Activation)	(None, 19, 19, 728)	0	block6_sepconv2_bn[0][0]
block6_sepconv3 (SeparableConv2)	(None, 19, 19, 728)	536536	block6_sepconv3_act[0][0]
block6_sepconv3_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block6_sepconv3[0][0]
add_4 (Add)	(None, 19, 19, 728)	0	block6_sepconv3_bn[0][0]
			add_3[0][0]
block7_sepconv1_act (Activation)	(None, 19, 19, 728)	0	add_4[0][0]
block7_sepconv1 (SeparableConv2)	(None, 19, 19, 728)	536536	block7_sepconv1_act[0][0]
block7_sepconv1_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block7_sepconv1[0][0]
block7_sepconv2_act (Activation)	(None, 19, 19, 728)	0	block7_sepconv1_bn[0][0]
block7_sepconv2 (SeparableConv2)	(None, 19, 19, 728)	536536	block7_sepconv2_act[0][0]
block7_sepconv2_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block7_sepconv2[0][0]
block7_sepconv3_act (Activation)	(None, 19, 19, 728)	0	block7_sepconv2_bn[0][0]
block7_sepconv3 (SeparableConv2)	(None, 19, 19, 728)	536536	block7_sepconv3_act[0][0]
block7_sepconv3_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block7_sepconv3[0][0]
add_5 (Add)	(None, 19, 19, 728)	0	block7_sepconv3_bn[0][0]
			add_4[0][0]
block8_sepconv1_act (Activation)	(None, 19, 19, 728)	0	add_5[0][0]
block8_sepconv1 (SeparableConv2)	(None, 19, 19, 728)	536536	block8_sepconv1_act[0][0]
block8_sepconv1_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block8_sepconv1[0][0]
block8_sepconv2_act (Activation)	(None, 19, 19, 728)	0	block8_sepconv1_bn[0][0]
block8_sepconv2 (SeparableConv2)	(None, 19, 19, 728)	536536	block8_sepconv2_act[0][0]
block8_sepconv2_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block8_sepconv2[0][0]
block8_sepconv3_act (Activation)	(None, 19, 19, 728)	0	block8_sepconv2_bn[0][0]
block8_sepconv3 (SeparableConv2)	(None, 19, 19, 728)	536536	block8_sepconv3_act[0][0]
block8_sepconv3_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block8_sepconv3[0][0]
add_6 (Add)	(None, 19, 19, 728)	0	block8_sepconv3_bn[0][0]
			add_5[0][0]
block9_sepconv1_act (Activation)	(None, 19, 19, 728)	0	add_6[0][0]
block9_sepconv1 (SeparableConv2)	(None, 19, 19, 728)	536536	block9_sepconv1_act[0][0]
block9_sepconv1_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block9_sepconv1[0][0]
block9_sepconv2_act (Activation)	(None, 19, 19, 728)	0	block9_sepconv1_bn[0][0]

block9_sepconv2 (SeparableConv2)	(None, 19, 19, 728)	536536	block9_sepconv2_act[0][0]
block9_sepconv2_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block9_sepconv2[0][0]
block9_sepconv3_act (Activation)	(None, 19, 19, 728)	0	block9_sepconv2_bn[0][0]
block9_sepconv3 (SeparableConv2)	(None, 19, 19, 728)	536536	block9_sepconv3_act[0][0]
block9_sepconv3_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block9_sepconv3[0][0]
add_7 (Add)	(None, 19, 19, 728)	0	block9_sepconv3_bn[0][0]
			add_6[0][0]
block10_sepconv1_act (Activation)	(None, 19, 19, 728)	0	add_7[0][0]
block10_sepconv1 (SeparableConv2)	(None, 19, 19, 728)	536536	block10_sepconv1_act[0][0]
block10_sepconv1_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block10_sepconv1[0][0]
block10_sepconv2_act (Activation)	(None, 19, 19, 728)	0	block10_sepconv1_bn[0][0]
block10_sepconv2 (SeparableConv2)	(None, 19, 19, 728)	536536	block10_sepconv2_act[0][0]
block10_sepconv2_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block10_sepconv2[0][0]
block10_sepconv3_act (Activation)	(None, 19, 19, 728)	0	block10_sepconv2_bn[0][0]
block10_sepconv3 (SeparableConv2)	(None, 19, 19, 728)	536536	block10_sepconv3_act[0][0]
block10_sepconv3_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block10_sepconv3[0][0]
add_8 (Add)	(None, 19, 19, 728)	0	block10_sepconv3_bn[0][0]
			add_7[0][0]
block11_sepconv1_act (Activation)	(None, 19, 19, 728)	0	add_8[0][0]
block11_sepconv1 (SeparableConv2)	(None, 19, 19, 728)	536536	block11_sepconv1_act[0][0]
block11_sepconv1_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block11_sepconv1[0][0]
block11_sepconv2_act (Activation)	(None, 19, 19, 728)	0	block11_sepconv1_bn[0][0]
block11_sepconv2 (SeparableConv2)	(None, 19, 19, 728)	536536	block11_sepconv2_act[0][0]
block11_sepconv2_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block11_sepconv2[0][0]
block11_sepconv3_act (Activation)	(None, 19, 19, 728)	0	block11_sepconv2_bn[0][0]
block11_sepconv3 (SeparableConv2)	(None, 19, 19, 728)	536536	block11_sepconv3_act[0][0]
block11_sepconv3_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block11_sepconv3[0][0]
add_9 (Add)	(None, 19, 19, 728)	0	block11_sepconv3_bn[0][0]
			add_8[0][0]
block12_sepconv1_act (Activation)	(None, 19, 19, 728)	0	add_9[0][0]
block12_sepconv1 (SeparableConv2)	(None, 19, 19, 728)	536536	block12_sepconv1_act[0][0]
block12_sepconv1_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block12_sepconv1[0][0]
block12_sepconv2_act (Activation)	(None, 19, 19, 728)	0	block12_sepconv1_bn[0][0]
block12_sepconv2 (SeparableConv2)	(None, 19, 19, 728)	536536	block12_sepconv2_act[0][0]
block12_sepconv2_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block12_sepconv2[0][0]
block12_sepconv3_act (Activation)	(None, 19, 19, 728)	0	block12_sepconv2_bn[0][0]
block12_sepconv3 (SeparableConv2)	(None, 19, 19, 728)	536536	block12_sepconv3_act[0][0]
block12_sepconv3_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block12_sepconv3[0][0]
add_10 (Add)	(None, 19, 19, 728)	0	block12_sepconv3_bn[0][0]
			add_9[0][0]
block13_sepconv1_act (Activation)	(None, 19, 19, 728)	0	add_10[0][0]
block13_sepconv1 (SeparableConv2)	(None, 19, 19, 728)	536536	block13_sepconv1_act[0][0]
block13_sepconv1_bn (BatchNormalization)	(None, 19, 19, 728)	2912	block13_sepconv1[0][0]
block13_sepconv2_act (Activation)	(None, 19, 19, 728)	0	block13_sepconv1_bn[0][0]
block13_sepconv2 (SeparableConv2)	(None, 19, 19, 1024)	752024	block13_sepconv2_act[0][0]
block13_sepconv2_bn (BatchNormalization)	(None, 19, 19, 1024)	4096	block13_sepconv2[0][0]

conv2d_3 (Conv2D)	(None, 10, 10, 1024)	745472	add_10[0][0]
block13_pool (MaxPooling2D)	(None, 10, 10, 1024)	0	block13_sepconv2_bn[0][0]
batch_normalization_3 (BatchNormalization)	(None, 10, 10, 1024)	4096	conv2d_3[0][0]
add_11 (Add)	(None, 10, 10, 1024)	0	block13_pool[0][0]
			batch_normalization_3[0][0]
block14_sepconv1 (SeparableConv2)	(None, 10, 10, 1536)	1582080	add_11[0][0]
block14_sepconv1_bn (BatchNormalization)	(None, 10, 10, 1536)	6144	block14_sepconv1[0][0]
block14_sepconv1_act (Activation)	(None, 10, 10, 1536)	0	block14_sepconv1_bn[0][0]
block14_sepconv2 (SeparableConv2)	(None, 10, 10, 2048)	3159552	block14_sepconv1_act[0][0]
block14_sepconv2_bn (BatchNormalization)	(None, 10, 10, 2048)	8192	block14_sepconv2[0][0]
block14_sepconv2_act (Activation)	(None, 10, 10, 2048)	0	block14_sepconv2_bn[0][0]
avg_pool (GlobalAveragePooling2D)	(None, 2048)	0	block14_sepconv2_act[0][0]
predictions (Dense)	(None, 1000)	2049000	avg_pool[0][0]

Total params: 22,910,480

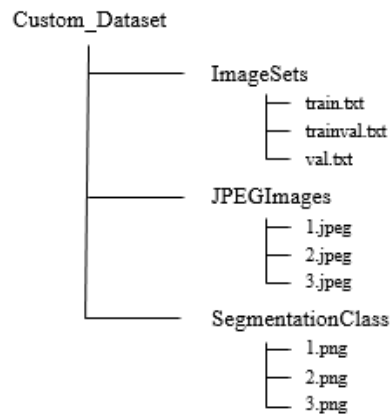
Trainable params: 22,855,952

Non-trainable params: 54,528

2. Training Custom Dataset using Transfer Learning

For training the DeepLab model on a custom dataset, it is required to convert the data to TFRecord format. TFRecord is a simple binary file format for storing data which helps in efficient computing because it takes lesser disk space and can be read faster.

Firstly, we need to store our data as per the below-mentioned directory structure:



The train.txt, val.txt, and trainval.txt contains the names of image patches to be used for training, validation, and all images combined (training+validation) respectively. The JPEGImages folder should contain the false-color composite image patches for training and validation. The SegmentationClass folder should contain the annotated labels.

Once, the directory is set, it is required to run the build_voc2012_data.py with updated flag values as per the directory. This converts the data to TFRecord format and it is saved to the location pointed by '- output_dir'. The segmentation_dataset.py requires updation (addition of _SAR_INFORMATION) as shown below:

```

# These number (i.e., 'train'/'test') seems to have to be hard coded
# You are required to figure it out for your training/testing example.
_ADE20K_INFORMATION = DatasetDescriptor(
    splits_to_sizes={
        'train': 20210, # num of samples in images/training
        'val': 2000, # num of samples in images/validation
    },
    num_classes=151,
    ignore_label=0,
)

_SAR_INFORMATION = DatasetDescriptor(
    splits_to_sizes={
        'train': 18,
        'trainval': 20,
        'val': 2,
    },
    num_classes=4,
    ignore_label=255,
)

_DATASETS_INFORMATION = {
    'cityscapes': _CITYSCAPES_INFORMATION,
    'pascal_voc_seg': _PASCAL_VOC_SEG_INFORMATION,
    'ade20k': _ADE20K_INFORMATION,
    'SAR': _SAR_INFORMATION,
}

```

Additional Code

Once this is done, we can run train.py, and vis.py codes using the commands given on the GitHub page of DeepLabv3+ for training, evaluation, and visualization of results. Apart from this one more modification has been done in the visualization part by adding a custom colormap for better inference of the segmentation results. The python code for the same can be found below:

```

from __future__ import absolute_import
from __future__ import division
from __future__ import print_function
import numpy as np
from six.moves import range

# Dataset name.
_CUSTOM = 'SAR'
# Max number of entries in the colormap for each dataset.
_DATASET_MAX_ENTRIES = {
    _CUSTOM: 4,
}

def create_custom_label_colormap():
    colormap = np.zeros((5, 3), dtype=np.uint8)
    colormap[0] = [120, 255, 120] #GROUND
    colormap[1] = [0, 100, 0] #FOREST
    colormap[2] = [255, 175, 175] #URBAN
    colormap[3] = [0, 0, 255] #WATER
    return colormap

def get_custom_name():
    return _CUSTOM

def create_label_colormap(dataset=_CUSTOM):
    if dataset == _CUSTOM:
        return create_custom_label_colormap()
    else:
        raise ValueError('Unsupported dataset.')

```

```

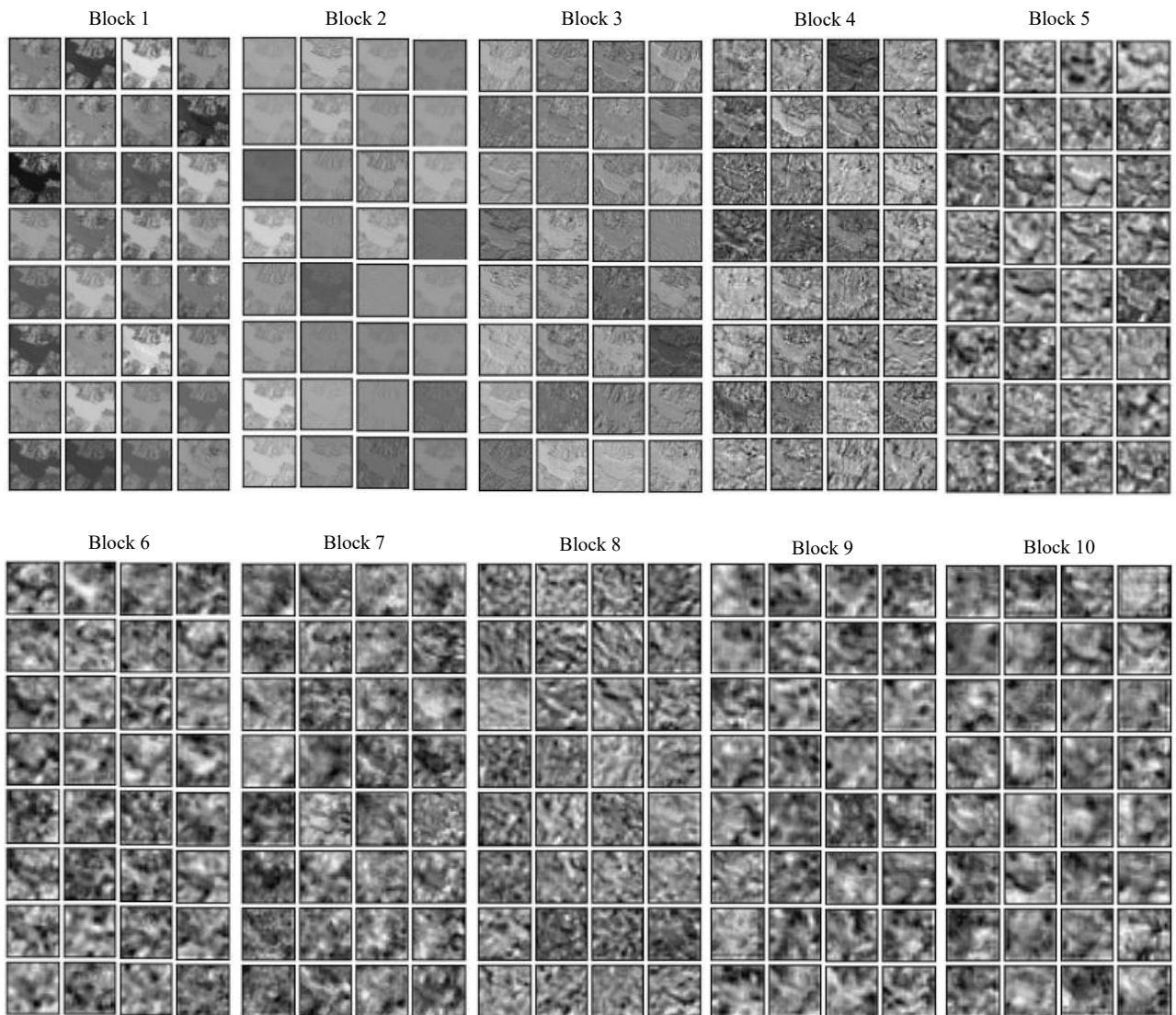
def label_to_color_image(label, dataset=_CUSTOM):
    if label.ndim != 2:
        raise ValueError('Expect 2-D input label. Got {}'.format(label.shape))

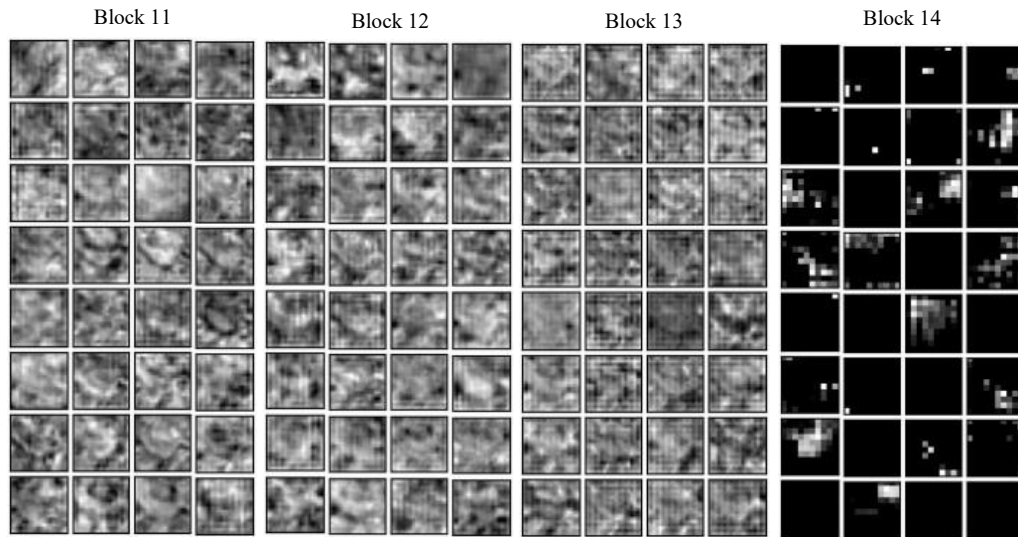
    if np.max(label) >= _DATASET_MAX_ENTRIES[dataset]:
        raise ValueError('label value too large: {} >= {}'.format(np.max(label),
            _DATASET_MAX_ENTRIES[dataset]))
    colormap = create_label_colormap(dataset)
    return colormap[label]

def get_dataset_colormap_max_entries(dataset):
    return _DATASET_MAX_ENTRIES[dataset]

```

3. Feature Visualization





Visualization of the Feature Maps extracted block-wise using the Xception Model corresponding to the 1st hidden layer of every block for test patch 1.

Rajat Garg

Development of machine learning algorithm for improved classification of natural and man-made targets using PolSAR ...

 Quick Submit Quick Submit UPES: University of Tomorrow

Document Details

Submission ID

trn:oid:::1:3309644443

Submission Date

Aug 7, 2025, 4:46 PM GMT+5:30

Download Date

Aug 8, 2025, 4:02 PM GMT+5:30

File Name

Thesis_Final_Rajat_Garg_-for_plag_run.pdf

File Size

4.0 MB

96 Pages

28,619 Words

154,660 Characters

8% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Small Matches (less than 14 words)

Exclusions

- 6 Excluded Sources

Match Groups

-  **59 Not Cited or Quoted 8%**
Matches with neither in-text citation nor quotation marks
-  **3 Missing Quotations 0%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 5%  Internet sources
- 2%  Publications
- 5%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.