

**DEEP LEARNING MODELS FOR CHANGE
DETECTION ANALYSIS USING GEO-SPATIAL
DATA**

Thesis

Submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

NAIK NITESH NAVNATH



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

SURATHKAL, MANGALORE - 575 025

October , 2023

**DEEP LEARNING MODELS FOR CHANGE
DETECTION ANALYSIS USING GEO-SPATIAL
DATA**

Thesis

Submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

NAIK NITESH NAVNATH

(197091CS004)

Under the guidance of

Prof.(Dr.) K Chandrasekaran

Dr. M Venkatesan



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

SURATHKAL, MANGALORE - 575 025

October, 2023

DECLARATION

by the Ph.D. Research Scholar

I hereby declare that the Research Thesis entitled **Deep Learning Models For Change Detection Analysis Using Geo-Spatial Data** which is being submitted to the **National Institute of Technology Karnataka, Surathkal** in partial fulfilment of the requirements for the award of the Degree of **Doctor of Philosophy** in Department of Computer Science and Engineering is a bonafide report of the research work carried out by me. The material contained in this Research Thesis has not been submitted to any University or Institution for the award of any degree.

 17.10.2023

Naik Nitesh Navnath, 197091CS004

Department of Computer Science and Engineering

Place: NITK, Surathkal.

Date: October 17, 2023

CERTIFICATE

This is to certify that the Research Thesis entitled **Deep Learning Models For Change Detection Analysis Using Geo-Spatial Data** submitted by **Naik Nitesh Navnath** (Register Number: 197091CS004) as the record of the research work carried out by him, is accepted as the Research Thesis submission in partial fulfilment of the requirements for the award of degree of **Doctor of Philosophy**.

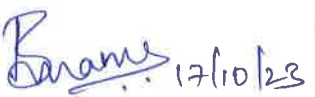

(Dr. K Chandrasekaran)

Dr. K. CHANDRA SEKARAN Ph.D.
Professor, Dept. of Computer Sc. & Res
Chairman - EECS Division
NITK, Surathkal, Sririvasnagar - 575 025
Mangalore, Karnataka State, India.
(Signature with Date and Seal)


(Dr. M Venkatesan)

Research Guide

(Signature with Date and Seal)


Chairman - DRPC

(Signature with Date and Seal)

DUCC - DPGG - DRPC
Dept. of Computer Science & Engineering
NITK - Surathkal
Srivasnagar - 575 025

ACKNOWLEDGEMENTS

With great pleasure, I thank the people who have supported and helped me so much throughout my research. First and foremost, I would like to express my sincere gratitude to my research supervisors, Dr. K. Chandrasekaran and Dr. M. Venkatesan, for constantly supporting me throughout my research journey with their knowledge, experience, and patience. I am deeply grateful to them for their valuable guidance, timely suggestions, and advice. Beyond the subject area of this thesis, I have learned a lot from them that will be helpful in the later stages of my career and life.

I sincerely thank the members of my Research Progress Assessment Committee (RPAC), Dr. Jeny Rajan and Dr. Anand Kumar M, for their insightful assessment and constructive suggestions to improve my research work further. I thank the Chairman of the Doctoral Research Programme Committee (DRPC), the Head of the Department, Dr. Manu Basavaraju, and the Secretary (DRPC), Dr. B. R. Chandavarkar, for their continuous encouragement. I also sincerely thank Dr. Shashidhar G.Koolagudi, Dr Alwyn Roshan Pais Associate Professors and Heads of the Department (during my period of study), teaching and non-teaching staff of the Computer Science and Engineering department, for their enormous help and support throughout my research. Nevertheless, I am also grateful to the technical and administrative staff of the department and institute for their timely service and cooperation in carrying out the research work.

I also extend my sincere gratitude to the All India Council for Technical Education, Government of India, Directorate of Technical Education, Government of Goa, and my parent institution Goa College of Engineering Farmagudi Goa, for providing me the opportunity to do the research work under the Quality Improvement Programme of Government of India. I would also like to thank the faculty and staff at the National Remote Sensing Agency Indian Space Research Organization, Government of India,

for providing us with the datasets for productively carrying out our research work.

Special thanks to my friends Dr.Alkha Mohan, Dr.Swathi P, and G Victor Daniel for being with me throughout my research work as a pillar of incredible support. I warmly thank my laboratory colleagues Sadhana, Aarabhi, Saraswati, Sneha, Kalinatha, Sandeep, Sarita, and Pradeep for their encouragement and our good times in the department. I thank my senior doctoral students and friends for their constructive criticism, knowledge transfer, and valuable cooperation.

This thesis would not have been possible without the exceptional support of my family. I am indebted to my late father, mother, siblings, in-laws, and supportive and caring wife for their unconditional love, inspiration, and encouragement. Finally, I thank the Almighty for granting me the health and strength to conduct this research.

Naik Nitesh Navnath

ABSTRACT

Due to enhancements in space technology, the number of sensors is increasing gradually, contributing towards the high availability of remote sensing images. Multiple satellites have been placed into orbit around the Earth, and more launches are anticipated. There is a need to monitor the surface of the Earth continuously. Every day, many images with high spatial and temporal resolution are captured. These contribute to acquiring a large amount of remote sensing data and a need to analyze this available data on time. These various possibilities contribute to the known revolution of Big Data in remote sensing. Numerous global and local environmental monitoring applications are made possible by the tremendous volume of images. Assessment of climate change, monitoring of disasters, and urban planning are only a few examples of uses.

Determining changes in land use and land cover is crucial for conserving various natural resources and the environment. It can reveal how humans utilize the land in a particular area. Learning how land use patterns have changed worldwide has become necessary to address global climate change and promote sustainable development. This research proposes novel approaches and methods for autonomously utilizing the remote sensing data collected by the increasing number of sensors. There is a high requirement to offer new strategies to extract information from these remote-sensing images in a reliable manner.

The primary focus of this thesis is on Change Detection (CD) methods that identify regions in remote sensing images where the land cover or land use has changed. The CD is the first step in comprehending the Earth's surface's dynamics and its evolution. This thesis investigates methods for better information extraction by utilizing the temporal correlation found in bitemporal and multitemporal image time series. Three significant innovative contributions to the state of the art are presented in the thesis. The spatial-temporal information is modeled with various pre-processing and clustering techniques and integrated using deep learning approaches. These results in a significantly more

accurate change map identifying when, where, and what land cover changes have occurred. The first contribution consists of a novel CD framework for bitemporal image CD. The hybrid CD approach using superpixel segmentation, fuzzy-based clustering, and lightweight deep learning is used to exploit the changes in remote sensing data. The second contribution presents a novel hybrid encoder-decoder model for land use and land cover CD that considers spatial and temporal aspects of binary and multiclass changes in remote sensing images. The final contribution presents an iterative method for enhancing the overall land cover classification performance for every pair of images defined inside a time series.

For the five datasets used in this research work, the overall accuracy of the proposed methods has improved significantly, increasing above 85%, when both space and time are considered with the maximum likelihood. These results indicate that combining space and time domains has immensely improved the accuracy of temporal CD analysis and can produce high-quality land cover prediction maps. A thorough qualitative and quantitative analysis complements the outcomes of the experiment.

Keywords: Environmental monitoring, Change Detection, Land use and Land Cover, Time-series, Deep Learning, Prediction.

CONTENTS

| | |
|---|-------------|
| List of Figures | ix |
| List of Tables | xii |
| List of Abbreviations | xiii |
| 1 Introduction | 1 |
| 1.1 Background and Motivations | 1 |
| 1.2 Change Detection in Remote Sensing Images | 4 |
| 1.3 Approaches to Change Detection | 6 |
| 1.4 Applications of Remote Sensing in Change Detection for Earth Resources Management | 6 |
| 1.5 Highlights of the Present Research Work | 7 |
| 1.6 Brief Overview of Thesis Contributions | 8 |
| 1.6.1 Hybrid Change Detection Framework with Clustering and Lightweight Deep Learning | 8 |
| 1.6.2 Spatio-Temporal Feature-Based Bitemporal Image Change Detection | 8 |
| 1.6.3 Multitemporal Time Based Land Cover Prediction | 8 |
| 1.7 Organization of the Thesis | 9 |
| 2 Literature Review | 11 |
| 2.1 Change Detection Process | 11 |
| 2.2 Categories of Change Detection | 13 |
| 2.2.1 Bitemporal Change Detection Techniques | 14 |
| 2.2.2 Multitemporal Change Detection Techniques | 16 |
| 2.3 Change Detection using Unsupervised Techniques: A Review | 18 |

| | | |
|----------|--|-----------|
| 2.4 | Land Use Land Cover Change Detection with Deep Learning Techniques: A Review | 21 |
| 2.5 | Time Series Analysis using Deep Learning Techniques: A Review | 24 |
| 2.6 | Research Gaps | 33 |
| 2.7 | Problem Statement and Research Objectives | 34 |
| 2.8 | Summary | 34 |
| 3 | Bitemporal Image Change Detection for Remote Sensing Images using Un-supervised Deep Learning Techniques | 35 |
| 3.1 | Introduction | 35 |
| 3.2 | Proposed Hybrid Change Detection Framework | 37 |
| 3.2.1 | Pre-processing and Generation of the Difference Image | 38 |
| 3.2.2 | Supapixel Segmentation using SLIC Algorithm | 39 |
| 3.2.3 | Clustering using Fuzzy C-Means Algorithm | 39 |
| 3.2.4 | Training and Testing | 40 |
| 3.2.5 | Lightweight Deep Convolutional Neural Network for Classification | 40 |
| 3.3 | Study Area and Data Description | 41 |
| 3.3.1 | Alappuzha District Kerala,India | 41 |
| 3.3.2 | Earth Dataset of a Small Region Near Paris,France | 42 |
| 3.4 | Results and Discussions | 43 |
| 3.4.1 | Experimental Setup | 43 |
| 3.4.2 | Performance Evaluation Measures | 43 |
| 3.4.3 | Comparative Methods | 45 |
| 3.4.4 | Experiments on the LISS-III Dataset | 46 |
| 3.4.5 | Experiments on the Earth Dataset | 46 |
| 3.5 | Summary | 48 |
| 4 | A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques | 51 |
| 4.1 | Introduction | 51 |
| 4.2 | Overall Framework | 53 |
| 4.2.1 | Proposed Change Detection Framework | 55 |

| | | |
|----------|---|------------|
| 4.2.2 | Attention Mechanism | 57 |
| 4.2.2.1 | Self Attention | 58 |
| 4.2.2.2 | Dual Attention | 61 |
| 4.2.3 | Model Parameters | 65 |
| 4.3 | Study Area and Data Description | 67 |
| 4.3.1 | Dakshina Kannada | 67 |
| 4.3.2 | Goa | 69 |
| 4.3.3 | Data Collection | 70 |
| 4.4 | Results and Discussions | 71 |
| 4.4.1 | Experimental Setup | 71 |
| 4.4.2 | Performance Evaluation Parameters | 71 |
| 4.4.3 | Comparative Methods | 72 |
| 4.4.4 | Result Analysis | 74 |
| 4.4.4.1 | Experiments on the Dakshina Kannada Dataset | 74 |
| 4.4.4.2 | LULC Change Analysis for Dakshina Kannada Region | 75 |
| 4.4.4.3 | Experiments on the Goa Dataset | 89 |
| 4.4.4.4 | LULC Change Analysis for Goa Region | 90 |
| 4.5 | Summary | 105 |
| 5 | A Multitemporal Technique for Land Cover Classification in Image Time Series | 107 |
| 5.1 | Introduction | 107 |
| 5.2 | Overall Framework | 108 |
| 5.2.1 | Gated Recurrent Unit | 108 |
| 5.2.2 | Temporal CNN | 109 |
| 5.2.3 | GRU+Temporal CNN | 110 |
| 5.2.4 | Attention on Temporal CNN and GRU | 110 |
| 5.3 | Proposed Framework | 111 |
| 5.4 | Study Area and Data Description | 113 |
| 5.5 | Results and Discussion | 115 |
| 5.5.1 | Performance Evaluation Parameters | 115 |
| 5.5.2 | Comparative Methods | 115 |

| | | |
|----------|---|------------|
| 5.5.3 | Result Analysis with the Proposed Framework | 117 |
| 5.5.3.1 | Univariate+ Multivariate + Coordinates | 117 |
| 5.5.3.2 | Univariate+ Multivariate | 118 |
| 5.5.3.3 | Univariate + Multivariate (LSTM) + Coordinates . . . | 118 |
| 5.6 | Summary | 123 |
| 6 | Conclusions and Future Scope | 125 |
| 6.1 | Future Scope | 127 |
| | Bibliography | 129 |
| | Publications | 148 |

LIST OF FIGURES

| | | |
|------|---|----|
| 1.1 | Spatial, Spectral and Temporal Resolutions. | 3 |
| 2.1 | Bitemporal Image Change Detection. | 15 |
| 3.1 | Proposed Hybrid Change Detection Framework. | 38 |
| 3.2 | Light Weight Deep Convolutional Neural Network. | 41 |
| 3.3 | LISIII Dataset for the Alappuzha Region (Bhuvan). | 42 |
| 3.4 | Earth Dataset for a City near Paris. | 43 |
| 3.5 | Change Maps generated for LISS-III (Alappuzha) dataset. | 47 |
| 3.6 | Change Maps generated for Earth (Paris) dataset. | 48 |
| 3.7 | AUC Curve. | 48 |
| 4.1 | Overall Framework of the Proposed Approach. | 54 |
| 4.2 | Spatio Temporal Encoder Decoder Network. | 56 |
| 4.3 | Self Attention Mechanism. | 58 |
| 4.4 | Self-Attention Module incorporated in STEDSAN. | 60 |
| 4.5 | Channel and Spatial Attention Modules. | 66 |
| 4.6 | Location Map of the Study Area (Dakshina Kannada.) | 68 |
| 4.7 | LULC Data and Description of each LULC Class (Dakshina Kannada) | 68 |
| 4.8 | Location Map of the Study Area (Goa). | 69 |
| 4.9 | LULC Data and Description of each LULC Class (Goa). | 70 |
| 4.10 | Change Maps generated for Dakshina Kannada dataset. | 75 |
| 4.11 | Multiclass Change Map of the Dakshina Kannada District. | 77 |
| 4.12 | Area Wise Changes in LULC Classes (Built-up and Kharif Crop). | 78 |
| 4.13 | Area Wise Changes in LULC Classes (Rabi Crop and Zaid Crop). | 79 |

| | | |
|------|--|-----|
| 4.14 | Area Wise Changes in LULC Classes (Double/Triple Crop and Current Fallow). | 80 |
| 4.15 | Area Wise Changes in LULC Classes (Plantation and Evergreen Forest). | 81 |
| 4.16 | Area Wise Changes in LULC Classes (Deciduous Forest and Degraded/Scrub Forest). | 82 |
| 4.17 | Area Wise Changes in LULC Classes (Littoral Swamp and Grassland). . | 83 |
| 4.18 | Area Wise Changes in LULC Classes (Wasteland and Waterbodies max). | 84 |
| 4.19 | Area Wise Changes in LULC Classes (Waterbodies min). | 85 |
| 4.20 | Change Maps generated for Goa dataset. | 89 |
| 4.21 | Multiclass Change Map for the Goa Region. | 96 |
| 4.22 | Area Wise Changes in LULC Classes (Built-up and Kharif Crop). . . . | 97 |
| 4.23 | Area Wise Changes in LULC Classes (Rabi Crop and Zaid Crop). . . . | 98 |
| 4.24 | Area Wise Changes in LULC Classes (Double/Triple Crop and Current Fallow). | 99 |
| 4.25 | Area Wise Changes in LULC Classes (Plantation and Evergreen Forest). | 100 |
| 4.26 | Area Wise Changes in LULC Classes (Deciduous Forest and Degraded/Scrub Forest). | 101 |
| 4.27 | Area Wise Changes in LULC Classes (Littoral Swamp and Grassland). . | 102 |
| 4.28 | Area Wise Changes in LULC Classes (Wasteland and Waterbodies max). | 103 |
| 4.29 | Area Wise Changes in LULC Classes (Waterbodies min). | 104 |
| 5.1 | Full approach for Time Series Analysis of SITS. | 109 |
| 5.2 | Bidirectional Gated Recurrent Unit. | 110 |
| 5.3 | Attention on Temporal CNN. | 111 |
| 5.4 | Proposed Framework of Univariate + Multivariate + Pixel Coordinates for Land Cover Classification. | 112 |
| 5.5 | Reunion Island. | 114 |

| | | |
|-----|--|-----|
| 5.6 | Confusion Matrix of Land Cover Classification for the Time Series Dataset using the DL Models: (a) Bidirectional GRU (b) Temporal CNN (c) GRU + Temporal CNN (d) Attention on Temporal CNN (e) Attention on Temporal CNN + GRU (f) Univariate + Multivariate + Coordinates (g) Univariate + Multivariate (h) Univariate + Multivariate (LSTM) + Coordinates. | 119 |
| 5.7 | F1 Scores of Land Cover Classification for the Time Series Dataset using the DL Models: (a) Bidirectional GRU (b) Temporal CNN (c) GRU + Temporal CNN (d) Attention on Temporal CNN (e) Attention on Temporal CNN + GRU (f) Univariate + Multivariate + Coordinates (g) Univariate + Multivariate (h) Univariate + Multivariate (LSTM) + Coordinates. | 120 |

LIST OF TABLES

| | | |
|-----|---|-----|
| 2.1 | Summary of Research Works on Unsupervised Change Detection. . . . | 20 |
| 2.2 | Summary of Research Works on LULC Change Analysis using Deep Learning Techniques. | 23 |
| 2.3 | Summary of Research Works on Time Series Prediction. | 31 |
| 2.4 | Summary of Research Works on Time Series Classification. | 32 |
| 3.1 | Lightweight CNN Model Parameters. | 41 |
| 3.2 | Comparative Analysis of Various Methods on LISS-III (Alappuzha) Dataset. | 46 |
| 3.3 | Comparative Analysis of Various methods on Earth (Paris) Dataset. . . | 47 |
| 4.1 | Details of the Dataset (Goa and Dakshina Kannada Region). | 71 |
| 4.2 | Quantitative Performances of Different Methods on the Dakshina Kannada Dataset. | 74 |
| 4.3 | LULC Changes in terms of Area for Dakshina Kannada Region from 2005-2010. | 86 |
| 4.4 | LULC Changes in terms of Area for Dakshina Kannada region from 2010-2015. | 87 |
| 4.5 | LULC Changes in terms of Area for Dakshina Kannada region from 2015-2018. | 88 |
| 4.6 | Quantitative Performances of Different Methods on the Goa Dataset. . . | 90 |
| 4.7 | LULC Changes in terms of Area for Goa Region from 2005-2010. . . . | 93 |
| 4.8 | LULC Changes in terms of Area for Goa Region from 2010-2015. . . . | 94 |
| 4.9 | LULC Changes in terms of Area for Goa Region from 2015-2018. . . . | 95 |
| 5.1 | Class distributions in the Training and Testing data for TiSeLaC Dataset. | 114 |

| | | |
|-----|--|-----|
| 5.2 | Precision, Recall and F1 score values for the GRU, Temporal CNN and Attention Models on Temporal CNN. | 121 |
| 5.3 | Precision, Recall and F1 score values for the Attention on Temporal CNN+GRU, Univariate+Multivariate+Coordinates, Univariate+Multivariate and Univariate+Multivariate(LSTM)+Coordinates. | 122 |

LIST OF ABBREVIATIONS

| <u>Abbreviations</u> | <u>Expansion</u> |
|-----------------------------|--|
| ADA | Adaptive Differential Algorithm |
| AwiFs | Advanced Wide Field Sensor |
| AUC | Area Under Curve |
| BU | Builtup |
| CA | Cellular Automata |
| CF | Current Fallow |
| CD | Change Detection |
| CNN | Convolutional Neural Network |
| D/SF | Degraded/Scrub Forest |
| D/TC | Double/Triple Crop |
| DF | Deciduous Forest |
| DF | Decision Tree |
| DBN | Deep Belief Network |
| DL | Deep Learning |
| DNN | Deep Neural Network |
| ESN | Echo State Network |
| EF | Evergreen Forest |
| FN | False Negative |
| FP | False Positive |
| FCN | Fully Convolutional Network |
| GA | Genetic Algorithm |
| GAN | Generative Adversarial Networks |
| GIS | Geographical Information Systems |
| GL | Grassland |
| ISRO | Indian Space Research Organization |
| KC | Kharif Crop |
| LCC | Land Cover Classification |
| LISS | Linear Imaging Self Scanner |
| LS | Littoral Swamp |
| LSTM | Long Short Term Memory Network |
| LULC | Land Use and Land Cover |
| LULCC | Land Use and Land Land Cover Change |
| LULCCD | Land use and Land Cover Change Detection |

| <u>Abbreviations</u> | <u>Expansion</u> |
|-----------------------------|---|
| MC | Markov Chain |
| MIOU | Mean Intersection Over Union |
| ML | Machine Learning |
| MLP | Multilayer Perceptron |
| NRSC | National Remote Sensing Centre |
| OA | Overall Accuracy |
| PA | Producer Accuracy |
| PCA | Principal Component Analysis |
| PL | Plantation |
| RC | Rabi Crop |
| RADAR | Radio Detection and Ranging |
| RF | Random Forest |
| RNN | Recurrent Neural Network |
| ReLU | Rectified Linear Unit |
| RS | Remote Sensing |
| SA | Self Attention |
| SAR | Synthetic Aperture Radar |
| SAX | Symbolic Aggregate Approximation |
| SITS | Satellite Image Time Series |
| SOTA | State of the Art |
| ST | Spatial-Temporal |
| STEDSAN | SpatioTemporal Encoder Decoder Self Attention Network |
| STEDDAN | SpatioTemporal Encoder Decoder Dual Attention Network |
| SVM | Support Vector Machine |
| TN | True Negative |
| TP | True Positive |
| TCN | Temporal Convolutional Networks |
| UA | User Accuracy |
| WL | Wasteland |
| Wmx | Waterbodies max |
| Wmn | Waterbodies min |
| Xgboost | Extreme gradient boosting |
| ZC | Zaid Crop |

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND AND MOTIVATIONS

The most stunning planet we can see from space is Earth, which extraordinarily sustains life. However, the rate at which Earth's environmental factors alter poses a hazard to life. Therefore, it is the responsibility of the global community to protect the planet Earth and limit the severity of numerous changes occurring on its surface. Understanding the physical, chemical, and biological processes that affect the Earth's system is necessary to comprehend the changes occurring there. Understanding the effects of both natural and artificial variables is also essential. It necessitates constant observation of changes taking place on the Earth's surface. By recognizing change, a closer eye can be kept on the resources of the Earth. These might result in more effective planning and sustainable resource use.

However, the stress on finite natural resources is tremendous due to population growth and rising consumption. Food, minerals, fuels, and other necessities are in high demand when there is a large population. Additionally, some natural resources are depleted by regular natural disasters, including droughts, floods, landslides, forest fires, earthquakes, etc., (Abbas Khan et al. 2019). Although the industrial period promised a comfortable existence, it also caused pollution and the devastation of habitats. To meet the ever-increasing demands of the people, it is an urgent need to manage our natural resources while also taking precautions to protect the Earth's ecology. Additionally, care should be taken to ensure that the requirements of the coming generation can also

be satisfied.

The advancement of remote sensing technology and its use with space-based sensors have demonstrated that it can support the monitoring and sustainable management of various natural resources. Remote sensing is learning about an object without touching it, such as the human eye or camera. But these days, “remote sensing” generally refers to identifying different features on the planet’s surface by monitoring electromagnetic radiation interactions with the targets. Since resource management is the focus of the majority of remote sensing applications, the United Nations (via resolutions of the general assembly) has adopted the definition of remote sensing as follows:

“Remote sensing means sensing of the earth’s surface from space by using the properties of the electromagnetic wave emitted, reflected or diffracted by the sensed objects, to improve resource management, land use and the protection of the environment.”

Thus, remote sensing can give us the inputs required in the form of reliable, timely, and precise information for managing the Earth’s resources and making the right decisions. These can help us better understand how the world works and lessen the adverse effects of our local, regional, and international activities.

Remote sensing image analysis has become an emerging area nowadays. Various remote sensing images like RADAR, multispectral, and multitemporal images are used for different applications based on their spatial, spectral, and temporal resolutions, as shown in figure 1.1, with multiple existing data collection techniques. LULCCD (Land Use and Land Cover Change Detection) and analysis through remote sensing imagery are fundamental research topics in the remote sensing community. Due to the limited spatial resolution of optical remote sensing imagery, pixel-centric spectral-based methods are the mainstream of traditional LULCCD and classification results. However, the rapid development of high-resolution remote sensing imagery brings opportunities to dig into more complex spatial patterns. Geographic object-based image analysis (GEO-BIA) has thus become a new paradigm for LULCCD. It initially divides remote sensing imagery into segmented objects and then classifies them.

Today, many imaging sensors can continuously acquire remotely sensed images

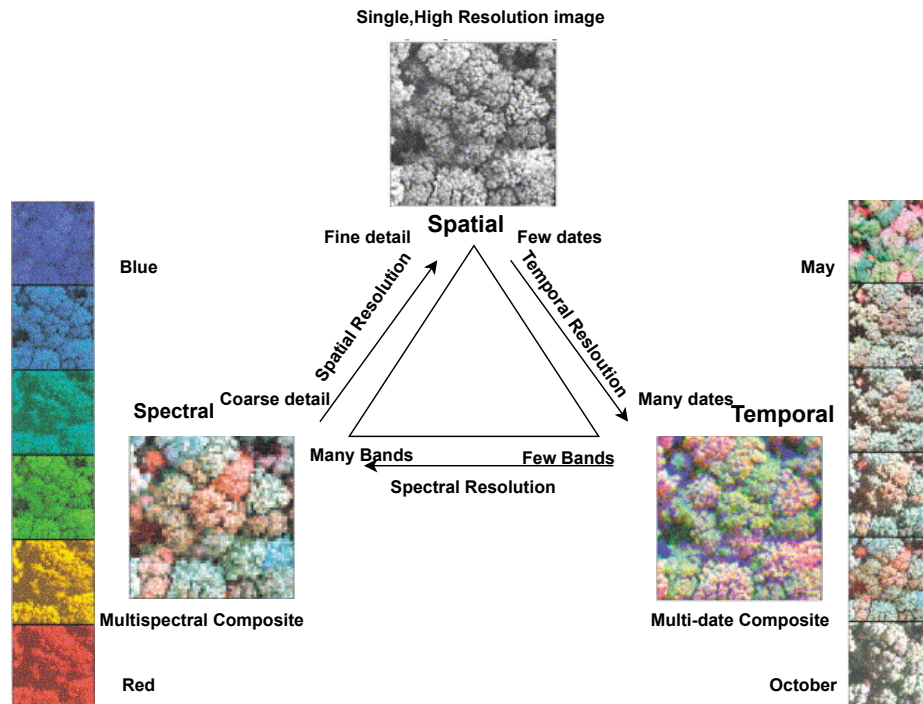


Figure 1.1: Spatial, Spectral and Temporal Resolutions.

source:<https://www.sciencedirect.com/science/article/pii/S0034425700001590>

[//www.sciencedirect.com/science/article/pii/S0034425700001590](https://www.sciencedirect.com/science/article/pii/S0034425700001590)

of a given area on the ground. According to the trend, a more extended image time series will be obtained and made available for usage in the future. Due to these factors, many images of the same location might be routinely gathered. Additionally, many historical discoveries from a particular investigative area are already kept in sizable archives available to the public for free.

The generated image time series are characterized by time-varying quality and higher geometrical and temporal resolution. Precision agriculture, geohazard prevention, innovative governance, and quick emergency response are just a few of the applications they can revolutionize and make possible (Mulla 2013) (Barrett 2013) (Sawaya et al. 2003) (Jensen 2009). However, with the right tools to automatically and reliably extract data from this massive volume of images, there is a huge possibility of being able

to see the data gathered by these sky-gazing remote sensing applications.

This work focuses on CD approaches because of the significance of information extraction from multiple images. Within a pair or set of images taken over the same area, the CD is a method that distinguishes between changed and unchanged pixels. The CD is crucial since it is one of the first processes that can be automatically applied to pairs of images over a long period to comprehend the processes inside the numerous images. CD can make large geographic datasets more usable and maintainable inside this framework. Recently, the advancement of deep learning techniques empowers the ability to learn more high-level semantic features from images. These also significantly benefit the development of LULCC studies.

1.2 CHANGE DETECTION IN REMOTE SENSING IMAGES

CD (Change Detection) in remote sensing refers to identifying changes in an area's characteristics over time using remotely-sensed data. These can be achieved using satellite or aerial imagery collected at different points in time and other types of remote sensing data such as radar or lidar. Various CD techniques have been developed in the past years, and efforts have been made to produce a comprehensive summary and review of these methods. As one of the pioneers' works (Singh 1989) review classified CD methods into multiple categories. CD approaches can be characterized into two broad groups: bitemporal CD and multitemporal CD, also known as time series analysis-the former measures change based on a 'two-epoch' timescale, i.e., comparing two dates. The latter analyses the changes based on a 'continuous' timescale, i.e. the focus of the analysis is not only on what has changed between dates, but also on the progress of the change over the period. Most CD methods belong to the bitemporal CD approach. Almost all classifications for CD algorithms are based on bitemporal CD with little attention to temporal trajectory analysis or multitemporal CD.

For bitemporal CD, algorithms can be attributed to one of the three approaches, namely, directly comparing different data sources (direct comparison method), comparing extracted information (post-analysis comparison method) and integrating all data sources into a uniform model (uniform modelling method). For multispectral remotely-

sensed images, transformation is often a necessary procedure. The detection elements of the direct comparison method include pixels, basic image features, and transformed features. The texture features and edge features are always taken as basic image features. Based on the two most widely-used object extraction methods, namely, image classification and feature extraction, comparison between objects after classification and feature extraction are typical for the post-analysis comparison method.

Compared with bitemporal CD, the temporal trajectory analysis emphasizes discovering the changing trend by constructing multitemporal data's 'curves' or 'profiles.' From the viewpoint of processing methods, temporal trajectory analysis can be decomposed into bitemporal CD, and then relative post-processing is implemented after the bitemporal CD. On the other hand, the so-called long time-series analysis method can be employed for temporal trajectory analysis. Another critical application of temporal trajectory analysis is real-time CD, such as video image sequences analysis.

Analysis of Change is concerned with two basic types of data:

1. Quantitative-differences in degree (continuous data), e.g., DEM(Digital Elevation Model)
2. Qualitative-differences in kind (discrete data) - e.g., LULC

The temporal ordering inside the given images makes them particularly interesting to study. Techniques will differ depending on whether the data is quantitative or qualitative pairwise (change) or time-series comparisons. It is thus necessary to propose CD and time-series prediction algorithms robust to these changes. To take into account the intra-class variability (i.e., the differences between the data/time series of the same class), the classifiers must also generalize information very well.

Deep learning techniques have recently achieved breakthroughs in various computer vision tasks, including image classification, object detection, and semantic segmentation. Meanwhile, this technology has quickly been adopted for remote sensing image applications. For instance, semantic segmentation classification at a pixel level has been proven to have great potential in land cover classification. The current state-of-the-art

(SOTA) algorithms at an object level, such as YOLO (You only look once) and Faster-RCNN, were also used for land cover object detection. However, the existing deep learning-based approach to high-resolution remote sensing data is still in its infancy, and a holistic approach still needs to be developed. Therefore, this work systematically examines how well the deep learning-based techniques perform on high and medium spatial resolution remotely sensed data considering the temporal domain, land use and land cover classification, and CD, compared to traditional approaches.

1.3 APPROACHES TO CHANGE DETECTION

CD can monitor a wide range of phenomena, including LULCC, urbanization, and deforestation, assessing the impacts of natural disasters such as earthquakes and hurricanes, and detecting changes in vegetation. There are several approaches to CD in remote sensing. The most used ones are listed here:

1. Image differencing: This involves subtracting two images of the same area taken at different times and thresholding the resulting difference image to identify change areas.
2. Multitemporal image analysis: This involves comparing images of the same area taken at different times using various statistical techniques, such as principal component analysis or spectral mixture analysis.
3. Object-based image analysis: This involves segmenting an image into distinct objects or features and comparing them across multiple images to detect changes.
4. Machine/Deep learning approach: These approaches involve training a machine/deep learning model on a set of labeled change/no-change examples and using the trained model to detect changes in new images.

1.4 APPLICATIONS OF REMOTE SENSING IN CHANGE DETECTION FOR EARTH RESOURCES MANAGEMENT

Remote sensing has a wide range of applications in earth resources management, including:

1. Land use and land cover mapping: Remote sensing can be used to map and moni-

tor changes in land use and land cover, including urbanization, deforestation, and agriculture.

2. Natural resources management: Remote sensing can map and monitor resources such as forests, water bodies, and minerals and track their distribution and condition changes over time.
3. Environmental monitoring: Remote sensing can be used to monitor the health and condition of natural resources, including forests, wetlands, and coral reefs, and to detect and track changes that human activities or natural disasters may cause.
4. Disaster management: Remote sensing is also used to monitor and track natural disasters such as floods, hurricanes, and earthquakes and to assess the extent of damage and the resources needed for recovery.
5. Agriculture and forestry: Remote sensing can be utilized to monitor crop health and yield and to track the growth and condition of forests.
6. Climate and weather monitoring: Remote sensing can monitor and track changes in atmospheric and climatic conditions, including temperature, precipitation, and atmospheric composition.
7. Mineral exploration: Remote sensing is incorporated to map and identify the presence of minerals and other resources, such as oil, gas, and coal.
8. Agricultural monitoring: Remote sensing can monitor crop health, assess irrigation needs, and optimize fertilization and pest management.
9. Water resources management: Remote sensing can monitor and assess the quality and quantity of water resources, including surface water and groundwater.
10. Coastal and marine resource management: Remote sensing can be utilized to monitor and assess the health of coastal and marine environments, including coral reefs, and to track the movement of marine animals.

1.5 HIGHLIGHTS OF THE PRESENT RESEARCH WORK

- An in-depth literature survey on challenges in CD in remote sensing images and various CD models proposed for remote sensing data.
- Introducing unsupervised lightweight deep learning technique for CD in bitem-

poral images.

- Proposing a multi-scale hybrid spatiotemporal model with attention mechanisms for LULCCD.
- Presenting a time series analysis model for land cover classification using multiple CNN and LSTM models.

1.6 BRIEF OVERVIEW OF THESIS CONTRIBUTIONS

1.6.1 Hybrid Change Detection Framework with Clustering and Lightweight Deep Learning

CD is crucial in geospatial data processing, and proposed linear techniques fail to preserve the complexities arising in medium and high-resolution remote sensing data. The proposed nonlinear methods are time-consuming for massive datasets. The hybrid CD approach using superpixel segmentation, fuzzy-based clustering, and lightweight deep learning is used to exploit the changes in remote sensing data. Experiments were performed on two datasets using the proposed hybrid CD framework. The results demonstrate that the proposed framework reduces computational complexity with better classification accuracy.

1.6.2 Spatio-Temporal Feature-Based Bitemporal Image Change Detection

Integrating spatial and temporal features increases the CD performance for multitemporal images. The novel hybrid encoder-decoder model in the presented research considers spatial and temporal aspects for bitemporal image CD. The attention mechanism is incorporated to exploit essential features relevant to CD. The proposed model is compared with various SOTA techniques on two LULC datasets. It is encountered that the proposed model is computationally less complicated and achieves good accuracy for the process of multiclass CD.

1.6.3 Multitemporal Time Based Land Cover Prediction

Multitemporal data analysis is carried out using SITS (Satellite Image Time Series) dataset. The proposed architecture uses the characteristics of univariate, multivariate, and pixel coordinate data for the land cover classification of SITS. Multiple deep learning methods are compared with the proposed framework, which outperforms the ex-

isting models and provides higher classification accuracy and F1 scores for each land cover class.

1.7 ORGANIZATION OF THE THESIS

The thesis advances in 6 chapters. An outline of each chapter is given below.

- **Chapter 1: The Introduction** section covers the need for CD and the difficulties encountered during the change analysis process from remote sensing data. The difficulties in processing remote sensing images, such as bitemporal and multi-temporal, are discussed. Motivation, applications, and challenges in remote sensing data based CD are also addressed. The chapter ends with a brief overview of research contributions and a thesis outline.
- **Chapter 2: Literature Review** section mainly consists of a detailed review of the CD techniques proposed using supervised and unsupervised methods like classification, prediction using LULC data with inclusion of spatiotemporal analysis of time series images. The section discusses the identified research gaps and the scope of the research work.
- **Chapter 3: CD using unsupervised techniques** includes the proposed hybrid CD models and their design details and result analysis. The section contains the details of the experiments and comparisons carried out to evaluate the model. The analysis of the result is discussed in this section with appropriate conclusions.
- **Chapter 4: Land Use Land Cover CD using Supervised Techniques** covers a novel multiscale spatiotemporal deep learning model for CD. The presented model's performance is compared against the available SOTA techniques and analyzed the results to reach conclusions.
- **Chapter 5: Spatio-temporal analysis of satellite image time series** discuss the model and design of a univariate, multivariate, and pixel coordinates time-based prediction model for land cover classification of time series data. The performance of the model is presented with appropriate analysis and discussions.

- **Chapter 6: Conclusions and Future Scope** chapter summarizes the contributions and findings of this research work. This chapter also provides insights into the future scope and directions for CD from remote sensing data.

CHAPTER 2

LITERATURE REVIEW

This section offers a critical overview of the studies that have already been done on CD in remote sensing. A concise description of this topic is provided in particular. Then, the emphasis is shifted to review the CD literature using two types of remote sensing images (bitemporal CD) or more (multitemporal CD). CD techniques focusing on extracting rich semantic meaning from the remote sensing data to make substantial spatial databases more maintainable and usable are considered in this review.

2.1 CHANGE DETECTION PROCESS

CD identifies a group of pixels that differs significantly between two or more multitemporal images (Radke et al. 2005). According to the definition in (Janssen and Vanderwel 1994), spatial-based technologies for remote sensing detect changes when the spatial objects (pixel groups or polygons) evolve as totally different, shrink or expand, shift positions or fragments, or gets merge. Finding differences in an object or phenomenon's condition by monitoring it at various times has also been described as "change detection" (Coppin et al. 2002).

Many communities are interested in CD for multitemporal data because of its applicability in many areas, including remote sensing, medical image analysis, surveillance, infrastructure monitoring, etc. The multitemporal images of the same geographic area are processed collectively in remote sensing to evaluate and quantify the degree of change that has taken place on the Earth's surface between the two dates. It indi-

cates that multitemporal data may be utilized to quantify the temporal impacts. The fundamental justification for using remote sensing data to identify surface changes is that these changes may result in changes in radiance values. These changes are to be distinguished from radiance changes brought on by other variables, such as varying atmospheric conditions, etc. (Ingram et al. 1981).

According to (Lu et al. 2004), successful CD research should identify the changing area, change rate, the spatial distribution of the altered types, and the accuracy of the CD procedure. The different CD techniques can be categorized according to the data transformations they use or the analysis they perform to identify the areas of substantial change (Singh 1989). The fundamental method of change identification may be based on comparing the outcomes of the independent classification of the images from the two dates or the concurrent analysis of multitemporal data (Singh 1989).

When using remote sensing, the information is gathered by sensors and then recorded in images typically examined for changes brought on by a change in the land cover or land use usage. CD aims to reject “unimportant” changes while detecting “major” changes (Radke et al. 2005). The latter could result from using various acquisition technologies, meteorological and climatic conditions, or sensor noise. Due to these factors, multiple processes are typically carried out to prevent identifying changes that are unimportant to the application. Usually, before using CD on Earth observation data in an operational scenario, images are intended to be processed to limit the influence of these factors on the images concerning the desired changes for the current application. The following set of processes is to be carried out.

- The images should be geometrically corrected and co-registered at the sub-pixel level. (Dai and Khorram 1999),(Coppin and Bauer 1994) (Townshend et al. 1992),(Bruzzone and Cossu 2003),(Dai and Khorram 1998),(Bovolo et al. 2011).
- The images are atmospherically corrected to top-of-atmosphere reflectance or, in the case of optical images, preferably to surface reflectance (Robinove 1982) (Coppin and Bauer 1996) (Hall et al. 1991).

A pre-processing of this kind makes it possible to analyze changes quantitatively and

consistently. CD is usually concerned with detecting three categories of changes (Verbesselt et al. 2010).

1. Rapid categorical changes, or a sharp change from one land cover class to another. Wildfires, deforestation, floods, and urbanization are a few examples of these inter-class types of changes.
2. Seasonal variations linked to the annual life cycles of plants or other periodic events.
3. Slow, gradual transitions that reflect intraclass variability. A specific instance is when the LULC remains constant over time, but the spectral or backscattering signature changes. These alterations could be brought about by inter-annual climatic changes or gradual changes in land management. According to how they behave over time, changes can also be categorized. They may be permanent or recurring. They can also be classified according to whether a phenomenon is natural or caused by humans (Olson et al. 2004).

2.2 CATEGORIES OF CHANGE DETECTION

The literature on remote sensing frequently discusses CD. Numerous authors have written articles on this subject, and multiple surveys and reviews have been published over the years (Pandey et al. 2021),(Lv et al. 2021),(Shi et al. 2020) (Hussain et al. 2013). This enormous body of literature may undoubtedly be seen from various angles. The multiple categories of using CD approaches in remote sensing are as follows:

1. CD approaches are based on pixels or objects, depending on the constituents on which they depend. Object-based CD organizes pixels into regions with uniform spectral or texture information.
2. CD approaches for active or passive sensors depend on the input data type.
3. CD approaches can be single-scale or multi-scale depending on the scale on which changes are examined.
4. CD techniques based on a supervised or unsupervised approach: based on the type of changes made, the information about when they happened, and whether or not they employ any ground truth data about changes made on the ground.

The basis for supervised CD is knowledge of either changes or the land use for the individual images. With the help of this ground truth data, a trained classifier can detect changes and identify class transitions.

These CD approaches are independent of sensor discrepancies or subpar atmospheric corrections and more resistant to pre-processing data. The disadvantage is the requirement for reference data which should be gathered through an expensive in-situ campaign or photo/image interpretation. Long image time series cannot be effectively used to implement the collection of reference samples. They cannot be timely updated worldwide, which is more significant. The post-classification comparison of the classification outcomes is a common step in supervised CD approaches. Other supervised technique types rely on direct multirate classification, which directly categorizes the multitemporal stack of images into land covers of static and dynamic types (Im and Jensen 2005),(Deilami et al. 2015), (Tewkesbury et al. 2015).

Contrarily, unsupervised CD methods only use the data in multitemporal images. The information on the type of change is typically implicitly available and must be reconstructed using earlier knowledge of the kind of sensor and the study area. The CD techniques are classified into two categories, as discussed below:

2.2.1 Bitemporal Change Detection Techniques

Based on comparing two images taken at two different time intervals. These CD approaches are better suited for the identification of abrupt category changes. Given a set of two images with spatial dimensions represented as $I_1(i,j)$ and $I_2(i,j)$ captured over a similar geographical location at different instants of time t_1 and t_2 , including the ground truth and reference image. Our aim is to determine the changed and unchanged areas between these two sets of images. The change map (CM) generated from I_1 and I_2 and $CM_{P(i,j)}$ is the set of values that are being changed at location (i,j) are represented in equation 2.1 as

$$CM_{P(i,j)} = \begin{cases} 1, & \text{changed} \\ 0, & \text{unchanged} \end{cases} \quad (2.1)$$

The estimation of the probability of the changes in the pixel $P(i,j)$ can be formulated as in equation 2.2 as

$$CM_{P(i,j)} = P_d(c_{i,j} = 1 | I_1(i, j), I_2(i, j), \Theta) \quad (2.2)$$

where P_d is the dense probability prediction problem and $c(i,j) = 1$ denotes the pixels that are being changed and the model's parameter set is represented as Θ . The general technique of bitemporal image CD is as shown in figure 2.1 with a brief discussion on each module.

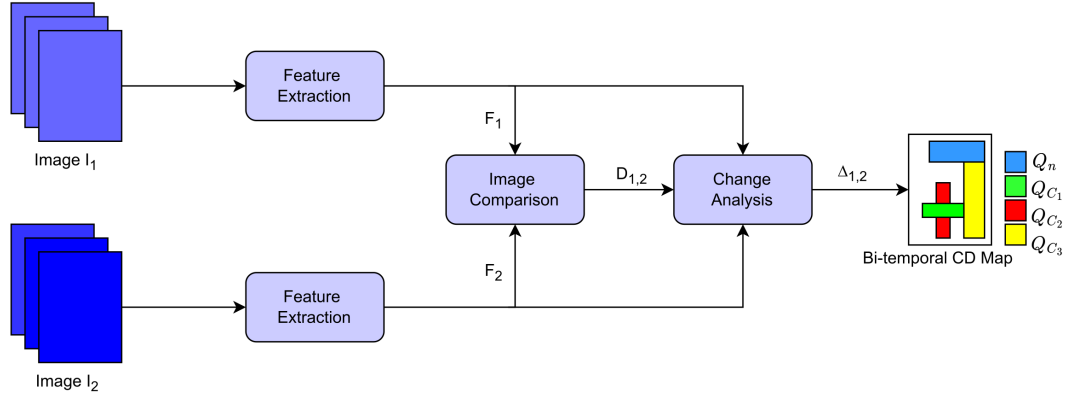


Figure 2.1: Bitemporal Image Change Detection.

Feature Extraction Features are extracted between images I_1 and I_2 to emphasize multitemporal information and stored in F_1 and F_2 , respectively.

Image Comparison For the assessment of the changes, features are compared. A change index, $D_{1,2} = d(F_1, F_2)$ is created by comparing features using mathematical operators. It evaluates changes in the feature space apparent in the spectral signature or the backscattering coefficient in optical images.

Change Analysis A change map $\Delta_{1,2} = f(D_{1,2})$ is created by analyzing the information extracted from $D_{1,2}$. The change index detects changes at the pixel/object, spatial context, and single or multiple scale levels. In the final change map $\Delta_{1,2}$, labels are allocated to classes in the following way: $\Omega = \{Q_n, Q_{c_1}, Q_{c_2}, \dots, Q_{c_K}\}$, where Q_n is the

no-change class and c_k is the k-th class in the general multiclass scenario ($K > 1$). In binary approaches ($K = 1$), the pixels in 1, 2 can belong to either the change class or the no-change class.

2.2.2 Multitemporal Change Detection Techniques

These methods can analyze many co-registered images ($N \gg 2$) acquired at various times. They are also known as SITS due to the increasing availability of image time series. Given, the historical daily time series dataset over n variables in $V = \{V_1, V_2 \dots V_n\}$ corresponding to a set of k locations as in equation 2.3

$$Loc^k = \{Loc_k^1, Loc_k^2, \dots, Loc_k^k\} \quad (2.3)$$

in a spatial region R for previous t years: $\{y_1, y_2, \dots, y_t\}$.

Also given, the spatial attribute information as in equation 2.4

$$SA = \{SA_1^{loc}, SA_2^{loc}, \dots, SA_r^{loc}\} \quad (2.4)$$

regarding each location $loc \in Loc^k$.

The problem is to determine the daily state/conditions of the variables in V for any location $x \notin (Loc^k \cup Loc^u)$ for future i years as in equation 2.5

$$\{y_{(t+1)}, y_{(t+2)}, \dots, y_{(t+i)}\} \quad (2.5)$$

when the spatial attributes of x are observed as $SA_1^x, SA_2^x \dots, SA_r^x$.

Here, Loc^u is a set of z new locations $\{Loc_u^1, Loc_u^2, \dots, Loc_u^z\}$ such that $Loc_j^u \in Loc^k$ for $j = 1$ to z, and i is a positive integer, i.e. $i \in \{1, 2, 3, \dots\}$.

The most effective bitemporal CD approaches are reviewed and suggested in section 2.2.1 using a general block diagram with three stages (see figure 2.1). Due to the increase in the availability of Earth observation images and data regulations, providing unrestricted access to satellite images with no cost for the final user enables new classes of CD methods utilizing SITS with the number of images as $N \gg 2$.

In this framework, CD approaches based on deep learning work with SITS, a collection of images of the same scene taken by one or more sensors at various periods. Different multitemporal CD approaches may be classified depending on whether the time-series images are arranged according to the sequence of their acquisition (Petit-jean et al. 2012). Some methods do not employ SITS's temporal data to distinguish

between several series of images. These techniques make use of all the available time-series images. Still, they need to consider their temporal order or verify the consistency of the temporal information within the image time series, limiting multitemporal information utilization. Most SITS processing approaches perform a multitemporal analysis of the behavior of the reflectance or backscattering values by extracting a temporal profile time series (TS) for each pixel of the study area.

Due to the greater availability of data, seasonal or inter-annual time series analysis is primarily discussed in the literature at medium resolution (like 300m - 400m). These methods can only accurately assess minute changes obtained by high-resolution acquisitions if they are made for TS with coarse spectral and spatial resolution. New strategies have been suggested to cope with the recently accessible SITS with an improved trade-off between spatial and temporal resolutions due to the launch of remote sensing missions capturing images at higher revisit times (Solano-Correa et al. 2017),(Solano-Correa et al. 2018),(Bruzzone et al. 2017).

Other unsupervised CD methods are based on the sequential analysis and forecasting of the temporal profiles for TS collected by optical sensors. They are made to operate in real-time or close to real-time CD with a short detection latency and a fixed false alarm rate. Since they are highly non-stationary with seasonal oscillations and inter-annual changes, remotely sensed 1D-TS are challenging to anticipate.

However, several methods have been put forth in the remote sensing literature to address this problematic issue, particularly in the context of land cover time series analysis. Whenever the observations deviate from the forecast made from historical time series data, change is recognized using this method. The TS's stationarity is a common presumption (Aminikhanghahi and Cook 2017). However, in the general scenario, stationarity is lost due to land-cover dynamics (Kleynhans et al. 2012). To overcome this issue, the time series can be transformed into roughly stable parameters in case of no change. Finally, modeling the backdrop image has been used to build unsupervised CD approaches for SITS (Coppin et al. 2002).

For decades, processing multitemporal images and CD has been an active research

field in remote sensing. Although many successful application cases have been reported on monitoring and detecting environmental change, there are enormous challenges in applying multitemporal imagery to derive timely information on the Earth's environment and human activities. In recent years, significant progress has been observed in overcoming technological obstacles by developing new platforms and sensors. The broader availability of extensive archives of historical images also makes long-term CD and modeling possible. Such a development stimulates further investigation into developing more advanced image processing methods and new approaches to handling image data in the time dimension. Over the past years, researchers have put forward large numbers of CD techniques for remote sensing images and summarized or classified them from different viewpoints(Singh 1989)(Liu and Zhou 2004). It has been generally agreed that CD is a complicated and integrated process. No existing approach is optimal and applicable to all cases. The categorization of the related work was done as given below:

- CD using unsupervised deep learning techniques when labeled data is unavailable.
- A deep learning approach for LULC multiclass CD in satellite images.
- A multitemporal technique for land cover classification in SITS using deep learning.

2.3 CHANGE DETECTION USING UNSUPERVISED TECHNIQUES: A REVIEW

The supervised learning techniques need to gather a large amount of training data as the labeled sample for the CD task. With supervised learning, a bilateral semantic fusion Siamese network (BSFNET) is proposed in (Du et al. 2021) that is trained by a scale-invariant sample balance method for pixel-wise CD from multitemporal images. In a semi-supervised approach, a small amount of labeled data and a large amount of unlabeled data are used, contributing to supervised and unsupervised learning for the model to learn and make predictions from the multitemporal data. A semi-supervised graph convolutional neural network (GCN) is used in (Saha et al. 2020b) for the enhancement

of spectral and spatial features in a multitemporal set of images.

A variety of unsupervised techniques are applied on remote sensing image analysis based on principal component analysis, and low-rank (Zhang et al. 2021b) and weighted change vector analysis with improved Markov random field (Fang et al. 2021). In (Guo et al. 2021) and (Shi et al. 2021), multiple CD approaches with spatio-spectral channel augmentation, and generative representation learning are discussed. Concerning time-series images numerous unsupervised methods like LSTM (Saha et al. 2020a) LSTM with Generative adversarial network-based one class classification (Jian et al. 2021), transfer learning-based bilinear convolution network (Zhan et al. 2021), and adversarial learning (Zhao et al. 2020) for CD are being proposed.

Many CD frameworks are being proposed in the literature for various types of images like SAR, Optical and Multispectral images. SAR images are mainly analyzed utilizing unsupervised CD techniques with superpixel segmentation and Fuzzy C-means clustering (FCM). In (Gao et al. 2016) they employed Gabor wavelets and FCM based on PCANet to perform CD in SAR images.

An adaptive discrete wavelet transform was adapted in (Jakka et al. 2019) to perform the CD. Superpixel feature extraction with the inclusion of contractive autoencoder with SLIC was employed in (Lv et al. 2018) for SAR images. A DCNN was designed for CD without any preprocessing operations with spatial fuzzy clustering in (Li et al. 2019b). Saliency guidance was provided in (Geng et al. 2019) with FCM clustering using deep learning for CD in SAR images. CD for SAR images with radial basis function and DCNN was proposed in (Pandeewari et al. 2021). Unsupervised CD was applied in (Attoui and Najah 2021) for SAR images using FCM clustering and Deep belief network. Patch-based CD method for SAR images with a pre-trained multilayer fusion network was proposed in (Shu et al. 2021). Table 2.1 displays the summary research work done in Unsupervised change detection based frameworks.

Table 2.1: Summary of Research Works on Unsupervised Change Detection.

| Research work | Approach Used | Limitations |
|--------------------------|--|---|
| (Gao et al. 2016) | PCA+Fuzzy Cmeans | Change point targets not detected |
| (Gong et al. 2017) | Superpixel segmentation+DNN | Registration of image pairs from different sensors is a challenge |
| (Lv et al. 2018) | Stacked contractive autoencoder | Fine tuning not done to extract deeper features |
| (Zhang et al. 2018) | Kmeans+Bottomup aggregative clustering | Non reliability in gaussian assumption |
| (Jakka et al. 2019) | DWT+Grey wolf optimization+F cmeans | Need of high processing power due to wavelet transform |
| (Li et al. 2019b) | CNN+Fuzzy spatial clustering | Heterogenous images not considered |
| (Geng et al. 2019) | Saliency+HFCM+Fisher Autoencoder | Deficiency in end to end feature extraction |
| (Lei et al. 2019) | Multiscale Superpixel segmentation+SDAE | Noise generated may affect the change detection process |
| (Dong et al. 2020) | Discriminative adversarial DNN | High possibility of generation of fake images while training |
| (Meng et al. 2020) | Fuzzy clustering+CWNN | Performance depends on the accuracy of preclassification |
| (Zhao et al. 2020) | Attention gates+GAN adaptation network | Robustness recognition of target changes not achieved |
| (Saha et al. 2020a) | Encoder Decoder LSTM | Multiclass changes in time series not addressed |
| (Saha et al. 2020b) | Graph convolutional neural network | Segmentation mask using graphs may lead to inaccuracy |
| (Du et al. 2021) | Bilateral semantic fusion siamese network | Scale-invariant sample balance loss function may demand better classification |
| (Zhang et al. 2021b) | PCA+Low rank prior | Features not extracted using CNN to differentiate between changed and unchanged superpixels |
| (Fang et al. 2021) | CVA+improved markov RF | Handcrafted features only including the spectral bands and morphological profiles were used |
| (Guo et al. 2021) | Matrix factorization+ Channel Augmentation | Matrix factorization may suffer from cold start problem η and γ |
| (Shi et al. 2021) | Generative representation learning network | Optimization through a self-adaptive way and may not converge to its stable state |
| (Jian et al. 2021) | (GANs)-based one-class classification | Training only with the unchanged data may lead to improper classification |
| (Zhan et al. 2021) | Transfer Learning-Based Bilinear CNN | Effective on images only from one sensor |
| (Pandeewari et al. 2021) | Deep CNN+ radial basis function | Robust features not utilized during the extraction process |
| (Attoui and Najah 2021) | DBN+improved Fuzzy cmeans | Preclassification using existing morphological reconstruction |
| (Shu et al. 2021) | Two stage patch based deep CNN | Threshold value α constant during training process |
| (He et al. 2021) | Multiscale visual saliency (MVSF) | Observation scale into the framework not incorporated |
| (Zhang et al. 2021a) | Superpixel sampling networks+UNet | Fusion and superpixel convolutions not a main perspective |

2.4 LAND USE LAND COVER CHANGE DETECTION WITH DEEP LEARNING TECHNIQUES: A REVIEW

Deep learning techniques have emerged as a growing field with various architectures like CNN, RNN, and a hybrid model for LULCCD. The current research involves analyzing the bitemporal images of a similar location to perform pre and post-change analysis for better future planning and development of any study area. (Panuju et al. 2020) investigated land cover dynamics employing multitemporal and bitemporal CD as a two-dimensional technique, focusing on multispectral images as a precondition for the study. In (Wang et al. 2021), an attention-based, highly supervised network (ADS-Net) detects remote sensing image variations. A dual-stream fully CNN identifies features at multiple layers from bitemporal images. Then, a channel and spatial feature fusion attention module are added in the network's decoding phase. The features of each layer are integrated to provide a variety of prediction maps for multiple supervision modules. A basic CNN is used for LULC classification in (Verma and Jana 2019) to map complex urban structures at a finer scale in the Mumbai region. Similarly, Deep CNN dubbed AlexNet and VGGNet were incorporated for the similar task in (Gharbia et al. 2020).

The authors in (Albert et al. 2017) employed large-scale satellite imaging data and cutting-edge computer vision techniques based on deep CNN to investigate trends in land usage in urban districts. In (Ekim and Sertel 2021), the Stochastic Ensemble, Stochastic Weight Averaging (SWA), and Fast Geometric Ensemble (FGE), Deep Neural Network Ensemble (DNNE) approach for constructing a comparison analysis for the LULC classification task is discussed. The Multi-Layer Perceptron–Markov Chain Model for analysing and predicting changes in New Jersey was performed in (Ngoy et al. 2021) for the projection of the 2015 LULC and was confirmed by actual data to produce a 2100 LULC. Multispectral remote sensing classification using deep learning algorithms was presented in (He and Wang 2021). The association between characteristics and spectra were investigated using correlation analysis. The contourlet was used to estimate texture features and create spectral–texture characteristics of categorized images after the specification of the parameters of the deep belief network.

Constrained extreme learning classifiers based on cascaded deep CNN of VGG-S and CaffeNet for land-use classification are proposed in (Liang et al. 2020). (Rajesh et al. 2020) described object-based CNN with deep Features for LULC Mapping of LISS-IV Imagery. A compressed SITS representation was created in (Kalinicheva et al. 2020) with a multi-view 3D convolutional autoencoder for the above task. First, a single segmentation map for the whole SITS is produced. The retrieved spatiotemporal objects are then grouped using their descriptions for the unsupervised SITS clustering. With multitemporal and multispectral Sentinel-2 satellite data, the study in (Sefrin et al. 2021) focused on land cover classification and CD. They used two alternative deep learning architectures and a few pre-processing procedures to tackle the difficult task of detecting land cover change. The pre-processing phase defined an excluded class and dealt with temporal water shoreline changes. An FCN was used and combined with LSTM networks. The FCN could only accept monotemporal input data, and after pairing with an LSTM, it also handled sequential multitemporal data. In (Rousset et al. 2021), authors employed neo channels with encoder-decoder networks based on deep lab architecture to assess deep learning techniques for LULC classification.

(Campos-Taberner et al. 2020) utilized a RNN , known as a 2-BiLSTM (Bidirectional LSTM) network for land use classification for Sentinel-2 time-series data. Based on a generative adversarial network, a semisupervised CNN for CD (SemiCDNet) is proposed in (Peng et al. 2020). The labeled and unlabeled data are provided as input to the segmentation network to construct initial predictions and entropy maps. For the utilization of most of the features of the unlabeled data, two discriminators were employed to ensure that the segmentation maps and entropy maps between these data have the same distribution of features. For multitemporal remote sensing images, (Du et al. 2019) presented a unique CD technique named Deep Slow Feature Analysis (DSFA) in their study which incorporated two DNN's to project original bitemporal input images onto a new feature space. SFA is enhanced to extract the most invariant components of unaltered pixels, which are then suppressed in changing regions to emphasize modified elements. Table 2.2 displays the summary of research work done in LULC CD based frameworks.

Table 2.2: Summary of Research Works on LULC Change Analysis using Deep Learning Techniques.

| Research work | Methodology | Limitations |
|-------------------------------|--|---|
| (Albert et al. 2017) | Deep Convolutional neural network | Physical elements not analyzed deeply in urban cities |
| (Du et al. 2019) | Deep slow feature analysis | Inaccurate to multiclass changes for continuous signals |
| (Verma and Jana 2019) | Convolutional neural networks | High computational cost |
| (Gharbia et al. 2020) | Deep Convolutional neural network | Patch wise classification may reduce the accuracy |
| (Liang et al. 2020) | Cascaded deep convolutional neural networks | Discriminative deep features with fusion mode |
| (Rajesh et al. 2020) | Convolutional Neural Network+Object based classification | Fails to classify categories with small number of samples for urban areas |
| (Campos-Taberner et al. 2020) | Bi-directional long short term memory | Much slower model and requires more time for training. |
| (Peng et al. 2020) | semisupervised convolutional network+GAN | Inclusion of self-supervised learning |
| (Kalimicheva et al. 2020) | Multiview 3D convolutional autoencoder | Clustering of linear objects may lead to higher accuracy |
| (Panuju et al. 2020) | Convolutional neural networks | Non identification of from-to changes |
| (Wang et al. 2021) | Attention based deeply supervised network | Multiclass changes not detected since use of existing pretrained network |
| (Ekim and Sertel 2021) | Deep Neural Network Ensemble (DNNE) | Multitraining scheme may affect the classification performance |
| (Ngoy et al. 2021) | Multilayer Perceptron and Markov chain | Terret software used for IULc analysis |
| (He and Wang 2021) | Deep belief network+Counterlet transform | Analysis of spatial+spectral features |
| (Sefrin et al. 2021) | FCN+LSTM | Vanishing gradient problem |
| (Rousset et al. 2021) | AlexNet+ResNet+SegNet+DeepLabv3 | Pre-trained networks may lead to non robust changes |

2.5 TIME SERIES ANALYSIS USING DEEP LEARNING TECHNIQUES: A REVIEW

The following papers represented how to model DNN for time series prediction/forecasting. In (Grover et al. 2015) they discussed the combination of trained predictive models in a discriminative approach with a DNN that performs modelling of the joint statistics on a set of variables related to weather. The base model is incorporated with three features like temporal mining, inter variable interactions and spatial interpolation. In (Ghaderi et al. 2017) they used RNN with DL called as Deepforecast for the prediction of wind speed which is spatio temporal in nature. They basically made use of graph, whose nodes consist of entities which generate the data and the edges carries out the task of modelling those nodes which are in interaction with each other. In (Borovykh et al. 2017) authors have analyzed and presented a conditional time series forecasting method which is purely based on CNN by making use of wavenet architecture which allows to extract temporal relationships between time series. It uses dilated convolutions for the multivariate time series data. A general model for time series regression with probability is presented in (Wen et al. 2017). They accommodated both the covariates of static and temporal data which are to be learned across multivariate series, shifting of the seasonality and other event spikes which are being planned for future.

A Significance-Offset CNN is developed in (Binkowski et al. 2018), which is a deep CNN for performing regression of asynchronous time series which is multivariate in nature. The model uses autoregressive system of weighting where in the final prediction is represented as a weighted sum of regressor which are adjusted and the weights are the data which are dependent on the functions learnt through a CNN. In (Rangapuram et al. 2018) they combined state space models with DNN networks for forecasting the time series. It allows to handle small data and also to learn complex relationship and patterns from raw time series for a given larger data. The work in (Li et al. 2017) modelled the traffic flow as a process of diffusion on a directed graph and they introduced diffusion convolutional recurrent neural network (DCRNN) for prediction of traffic data that involves both spatial and temporal dependency in the flow of the traffic. The architecture made use of random walks in a bidirectional manner for

extracting the spatial dependency and to capture the temporal dependency, it made use of encoder decoder architecture with the features of scheduled sampling.

An empirical assessment of generic convolutional and recurrent neural architectures was presented in (Bai et al. 2018) for task related to sequence modelling. They made use of TCN which combines residual connections and dilations with the incorporation of causal convolutions which are needed for prediction in an autoregressive manner. In (Lim 2018) they introduced recurrent marginal structural network in the field of epidemiology which uses the property of sequence-to-sequence architecture for predicting the expected outcome to a series of treatments which were planned in advance. A hybrid model is being developed in (Lim et al. 2019) called as deep momentum networks that combines trading rules based on DL into the volatility scaling architecture of time series momentum. The model also tries to learn the estimation of the trends and sizing of the position in a data driven approach by optimizing the sharpe ratio of the signal. A transformer architecture is being applied in (Li et al. 2019a) for the prediction of time series data by incorporating a convolutional SA network with the production of queries and keys with the involvement of causal convolution for the better understanding of the local context.

In (Lim et al. 2021) they proposed a Temporal Fusion Transformer (TFT) which incorporates architecture based on attention levels and a high-performance forecasting which is multi-horizon in nature with explainable insights into temporal dynamics. The TFT made use of recurrent layers for processing at local level and the long-term dependencies processing was done by the interpretable SA layers. The authors designed a time series model in (Salinas et al. 2019) based on RNN in combination with Gaussian copula process which has a covariance structure with low rank which allows to lower the computational complexity and also handle marginal distributions related to non-gaussian processes. A conditional generative model for multivariate data is developed in (Wen and Torkkola 2019). They made use of DNN to parameterize the data. The multivariate time series data is solved by learning the properties of quantile functions and a conditional copula to be incorporated with the latent uniform random variables. A DL framework for multi horizon prediction of time series data with the advent of

temporal attention mechanism that also captures the latent patterns in the historical data is being adapted in (Fan et al. 2019). It also used a multimodal fusion approach to combine features from various parts of the history to effectively represent the future forecast. In (Siddiqui et al. 2019) they proposed the concept of demystifying convolutional DL models for analysis of time series data. A framework named (TSViz) for testing the interpretability of time series analysis models based on DL is being designed. The architecture identifies the parts of the input that are accountable for the prediction, performing clustering of the filters, interpreting the variation learnt through inverse optimization by the network and analyses of the network was done to check for its robustness if any against the adversarial noise.

A global local framework for prediction of the time series using DNN is being adapted in (Wang et al. 2019b) which allows to uniquely characterize the time series in an exchangeable manner. In (Sen et al. 2019) they proposed a model for deep forecasting of high dimensional data called as DeepGLO which has the feature of thinking at global level and acting at local level. DeepGLO is a hybrid model design which consist of a TCN with regularization along with a global matrix factorization model inclusive of a temporal network that has the ability to extract local properties of each of the time series and its associated covariates. The authors in (Smyl 2020) incorporated a dynamic computational graph neural network model that allows a standard exponential smoothing model to be mingled with LSTM into a common framework consisting of hybrid and hierarchical method of forecasting. A novel framework called as GNET is being presented in (Li et al. 2020) which is sequential DL framework designed for G computation that can process complex time series while the imposition of assumptions of minimal modelling and also gives an estimate of the individual or population level treatment effects which varies with time. The authors in (Lim et al. 2020) introduced a Recurrent Neural Filter (RNF), which uses the architecture of recurrent autoencoder that predicts dissimilar representations for each step of the Bayesian filtering which is being captured or extracted by a series of encoders and decoders.

In (Lv et al. 2014) they Nade use of stacked autoencoder for feature learning from the time series data of traffic flow for the prediction of road segment level. In (Soua

et al. 2016) they made use of DBN to forecast the traffic flow of future based on the observations of the previous traffic flow. They considered traffic flow data as a freeway in terms of timeseries. In (Rodrigues et al. 2019) consideration of the issue of taxi demand prediction was done and they exhibited a specific area as a time series of the taxi demand. A DL model was designed to learn the features from the historical data of time series related to taxi demand and they combined the features with other features based on context like weather and social media data which can predict what may be the future demand.

In (Liao et al. 2018) an integration was performed of sequence-to-sequence model and LSTM to predict the speed of the traffic of a road segment. The model also took into consideration other features which are external like roads geographical area, social events of public and travel information related activities like crowd queries which are held online. The model in (Cheng et al. 2018b) amalgamated traditional methods of prediction of wind speed in combination with threshold denoising (WTD) and adaptive neuro fuzzy inference system (ANFIS) with a RNN. In (Dvornek et al. 2017) they developed a model using LSTM for classifying individual with autism spectrum disorder (ASD) and it controls directly from the fMRI time-series of the resting state. In (Huang et al. 2017) they developed an unsupervised model known as DCAE deep convolutional auto encoder for feature learning of mid and high-level data from large scale fMRI time series.

The following papers represented how to model DNN for time series classification. The authors in (Simonyan et al. 2013) adapted the use of classification of image models using deep convnets which tries to generate an artificial image which is an illustration of the class of interest. Another operation which they performed was saliency map with respect to a specific class which highlights the specific areas of the image given which is discriminative with reference to a given class. This allows the graph cut-based object segmentation and also performed deconvolution procedure to generalize the gradient based visualization techniques. In (Krizhevsky et al. 2017) they performed classification of images by making use of deep CNN by activation function like ReLu also reduced overfitting with the features such as data augmentation and dropout on a

highly challenging dataset. The temporal aspect was not addressed and they proposed to use very large deep CNN for understanding the temporal structure like in videos as compared to static images.

In (Santos and Kern 2016) a review of early TSC (Time Series Classification) approaches is provided using DL models this work concentrates not only on theoretical results and frameworks but also how different datasets of time series are used for early time series classification. The proposed framework in (Wang and Oates 2015) in an offline manner to exclusively encode temporal patterns as spatial in the form different types of images which are known as Gramian Angular Fields (GAF) and Markov Transition Fields (MTF). They made use of tiled CNN to extract high level features from the individual GAF, MTF, and GAF-MTF images. In (Zheng et al. 2016) the learnt features are applied into a multilayer perceptron (MLP) for classification. They proposed a deep CNN (MC-DCNN), for classification of multivariate time series data. This model first extract features from single univariate time series in each channel, and mixes information from all the channels as feature illustration or representation at the final layer. In (Wang et al. 2017) the paper proposed a cycle DBN model for the classification of multivariate time series in comparison to the performance of the DBN and KNN (K-Nearest Neighbor). The presentation learning capability of DBN and the time series data correlation is being utilized in an efficient manner. They used a cyclic approach of feedback to classify and extract features from the data. The work in (Mittelman 2015) considers design of a FCN that makes use of an operation on causal filtering and it permits for the rate of the output signal to be similar to that of the input signal. They further processed an undecimated FCN (UFCN), which is inspired by the undecimated wavelet transform.

Stacked LSTM autoencoder networks were made used in (Mehdiyev et al. 2017) in an unsupervised or self-supervised manner. The compressed depiction of the time-series data attained from LSTM autoencoders are then offered to Deep Feedforward Neural Networks for classification. They incorporated cost sensitive learning due to the imbalanced nature of the time series data. The work in (Hatami et al. 2018) they made use of recurrence plots (RP) which performs the conversion of time series to 2D texture

images and then a deep CNN classifier is applied on the images for classification of the time series. The combination of RP and CNN is termed as texture image recognition task. In (Cui et al. 2016) they proposed a novel end to end model called as multiscale CNN (MCNN) that carries out the task of feature extraction and classification in one single framework by performing transformations on the time series. It performs extraction of features at various scales and frequencies. It exploits the CNN to learn features in both time and frequency domains. An adaptive cost-sensitive learning strategy was developed in (Geng and Luo 2018) to modify temporal information using DL models to tackle imbalanced time series classification problems. They made use of deep CNN by using five neural networks to carry out the classification task.

A data augmentation technique being proposed in (Fawaz et al. 2018a) which uses dynamic time warping distance to overcome overfitting small time series datasets with the inclusion of the weighted version of the DTW (Dynamic Time Warping) Barycentric Averaging technique. The space induced by the DTW and the learned features of the CNN were used to time-invariant features which are being used for classification. In (Liu et al. 2018) a tensor scheme along with a DL model known as multivariate CNN (MVCNN) is being developed for classification of multivariate time series data and it also handles data which has lagged features. Convolution operation is applied to exploit the local interactions amongst the variables.

In (Fawaz et al. 2018b) they investigated how to apply transfer learning on deep CNN on the target dataset by fine turning the network that is being pretrained on similar kind of source dataset for the TSC task. They combined classic DTW and CNN by exploiting it with the features of transfer learning. The work in (Gong et al. 2018) devised an algorithm known as MOMM (Multiobjective Model-Metric) learning for approximation and classification of time series data. In MOMM a recurrent network is used as a filter for the temporal data based on that a generative model is being learnt for every time series which is a representation of that series. The model works on both univariate and multivariate time series datasets. It also incorporates the features such as weight optimization, network size and the distance metric while learning the representations. A meticulous approach for learning features for human activity recognition

problem is being discussed and developed in (Yang et al. 2015). The time series raw signals are considered as higher-level abstract representations by the learned features from the deep CNN architecture. The framework tackles multichannel time series data by design of a unified model consisting of feature extraction and classification. In (Song et al. 2020) a new model is designed based on the deconvolutional networks and SAX discretization to learn the representation for multivariate time series. A new model is being discussed to learn the multivariate time series with the incorporation of deep CNN with unsupervised learning and SAX discretization. They designed a network framework exclusively to extract the cross-channel correlation with the operation of deconvolution, which forces the pooling operation to perform the task of dimensionality reduction along each setting in the individual channel.

In (Martinez et al. 2018) they introduced an early classifier approach which incorporates reinforcement learning agent called as deep Q-network at an end-to-end level. They defined a suitable series of states and actions with the inclusion of a special reward function which targets at determining the negotiation between earliness and the classification accuracy. Timenet a multilayered RNN known as timenet is designed in (Malhotra et al. 2017) which is used in a unsupervised way to perform extraction of features from time series. It is the encoder network of the designed autoencoder as deep RNN which is based on sequence-to-sequence models that converts time series of varying length into vector representations of fixed dimensions. In (Wang et al. 2016b) they presented an Earliness-Aware Deep Convolutional Networks (EA-ConvNets), which is an end-to-end DNN for early classification of time series data. The architecture learns the features by using deep hierarchy of shapelets which captures the salient properties in each time series in combination with a dynamic truncation model which helps in focusing on the early parts of each of the time series. The algorithm proposed in (Wang et al. 2016a) made use of classification of MTS data by making use of a RNN and adaptive differential algorithm. Firstly, they used RNN for training the MTS sample into various state clouds. Table 2.3 and 2.4 provides the summary of the prediction and classification of time series data using DL techniques.

Table 2.3: Summary of Research Works on Time Series Prediction.

| Research Work | Approach Used | Limitations |
|--------------------------|---|---|
| (Grover et al. 2015) | Deep Belief Network | Time component not considered with large distance, No guide sensing for weather prediction |
| (Ghaderi et al. 2017) | Recurrent Neural Network with LSTM | Cannot predict the forecast of individual graph node at time t. |
| (Borovykh et al. 2017) | Wavenet (CNN +Dilated Convolution) | Inability to learn long term dependencies on intraday data. |
| (Wen et al. 2017) | LSTM with forking sequences | Explicit multivariate prediction and joint distributions modelling not addressed |
| (Binkowski et al. 2018) | Significance offset CNN+ Autoregressive model | Assumption of independent offset values for each past observation |
| (Rangapuram et al. 2018) | Deep State Space Models | Not working for all instances of state space models, missing non-gaussian likelihood |
| (Li et al. 2017) | Diffusion Convolutional RNN(DCRNN) | Underlying graph structure prediction not based on moving objects (evolving data) |
| (Bai et al. 2018) | Temporal Convolution Network (TCN) | More memory during evaluation,Potential parameter change for a transfer of domain. |
| (Lim 2018) | Recurrent Marginal Structural Network | Cannot handle Biased treatment responses over time |
| (Lim et al. 2019) | Deep Momentum Networks | Non stationarity data cannot be handled, missing of time series momentum |
| (Li et al. 2019a) | Transformer + Convolutional self-attention | Low sparsity strategy in self-attention |
| (Lim et al. 2021) | Temporal Fusion Transformer (TFT) | Inability to analyze crucial variables for prediction, non-identification of significant regime changes |
| (Salinas et al. 2019) | RNN+ Gaussian Copula | Harder to overfit |
| (Wen and Torrkola 2019) | Deep Generative Quantile Copula Model | Quantile and copula part not parameterized by flow-based models. |
| (Fan et al. 2019) | LSTM+ Deep Attention model | Unable to predict from the Entire time period history |
| (Siddiqui et al. 2019) | TSviz (Visualization of time series) | Unstable predictions due to noise factor may occur |
| (Wang et al. 2019b) | Deep Factor Models with random effects | No comparison with variational dropout or deep ensemble non probabilistic models for uncertainty |
| (Sen et al. 2019) | DeepGlo(Global Matrix Factorization+ TCN) | Global and local features may be wrongly predicted depending on size of dataset |
| (Smyl 2020) | Dynamic Computation Graph with Deep Ensemble | Assumption of continuity for in sample and out sample prediction of the time series |
| (Li et al. 2020) | G-Net (Gcomputation + Sequential DNN) | Incorporation of prior causal knowledge missing |
| (Lim et al. 2020) | RNN + Bayesian Filtering | RNN parameters are fixed once training phase finishes |

Table 2.4: Summary of Research Works on Time Series Classification.

| Research work | Approach used | Limitations |
|-------------------------|--|---|
| (Lv et al. 2014) | Stacked Auto Encoder | Prediction layer uses only logistic regression |
| (Soua et al. 2016) | DBN + Dempster-Shafter Theory | Works efficiently only for stream data |
| (Rodrigues et al. 2019) | DL-LSTM | Inability to perform data fusion |
| (Liao et al. 2018) | Seq2Seq Deep Neural Network | Non accurate forecast on large datasets |
| (Cheng et al. 2018b) | Wavelet Thresholding+RNN+Neuro Fuzzification | Less utilization of RNN as submodel |
| (Dvornek et al. 2017) | LSTM | Need of huge resources for training |
| (Huang et al. 2017) | Deep Convolutional Autoencoder | Learns only hierarchical dynamical connectivity |
| (Santos and Kern 2016) | Tile Convolution Neural Network | Modelling of time series through GAF and MTF images not done |
| (Zheng et al. 2016) | Multichannel Deep CNN (MC-DCNN+ MLP) | Initialization and momentum parameters not considered |
| (Wang and Oates 2015) | Cycle Deep Belief Network | Each RBM (Restricted Boltzmann Machine) has to be trained independently |
| (Mittelman 2015) | Undecimated FCN(UFCN) | Only capture dependencies that occur within the overall extent of the causal filters. |
| (Mehdiyev et al. 2017) | Stacked LSTM Encoder Network+ MLP | Longer time for training the pretraining and the fine-tuning phase |
| (Hatami et al. 2018) | CNN + Recurrence Plots | Inability to classify small sample size of data over time |
| (Cui et al. 2016) | Multiscale CNN (MCNN) | Designed only for univariate time series, |
| (Geng and Luo 2018) | LSTM+ Temporal Fully CNN | Cost sensitive strategy not extended to multiclassification task |
| (Fawaz et al. 2018a) | Deep Residual Networks +DTW | Variant weighting schemes for the DTW based data augmentation technique not done |
| (Liu et al. 2018) | Multivariate CNN+ Tensor | Imbalance problem in the preprocessing stage not done |
| (Fawaz et al. 2018b) | Fully Convolutional Neural Network | Optimization algorithm may get stuck in local optimum for a bad source dataset. |
| (Gong et al. 2018) | MOMM (Multiobjective Model Metric RNN) | Less improvement in computational efficiency |
| (Yang et al. 2015) | Deep CNN | SoftMax function used does not support null rejection |
| (Song et al. 2020) | Deconvolutional Network +SAX | Integration of grammar induction approach not done |
| (Martinez et al. 2018) | Deep Q Network | No dynamic adjustment of the reward function parameters over training data |
| (Malhotra et al. 2017) | TimeNet (Multilayered Deep RNN) | Prone to overfitting |
| (Wang et al. 2016b) | Earliness-Aware Deep CNN | Works only on univariate data ,Prediction layer uses only logistic regression |
| (Wang et al. 2016a) | ESN+ ADEA | Fruit fly optimization and genetic simulated annealing could have been done |

2.6 RESEARCH GAPS

The satellite remote sensing imagery is significantly profuse in space, providing detailed view of large areas. However, these are relatively scarce with respect to time. Data from fixed sensors are plentifully available over time, though these provide relatively little detail in space due to limitation in the number of spatially distributed sensors. The recent advancement in satellite and remote sensing technology has led to explosive growth in spatial and spatiotemporal data. Extracting useful and interesting information or patterns from these huge amount of data is also an added challenge in this regard. Considering also the spatial evolution in order to differentiate among the different classes is crucial since the phenomena to be dealt with are fully spatiotemporal. The following gaps needs to be addressed.

- Urgent need to have rapid access to extended, temporal satellite bi-temporal images and multitemporal data for identifying patterns of change.
- The DNN (Deep Neural Network) models applied in the domain of ST prediction generally appear with additional functionalities like feature-level data fusion, convolution, etc., which are mainly utilized to learn the spatial and temporal dependencies. However, because of the utilization of many parameters, they often suffer from significant memory consumption and overfitting problems.
- DNNs typically require multitemporal data to be discretized at regular intervals, making it challenging to forecast datasets where observations can be missing or arrive at random intervals. The multitemporal data usually do not contain spatial information. Thus, the spatial correlations among the data are not explicitly considered in deep learning-based prediction/classification models.
- Continuous monitoring of land is needed for accurate prediction, and most of the existing prediction models tend to fail to provide continuous monitoring of the environment due to scarcity of data.
- Need for the development of multidimensional space time architecture for efficient change detection and time series analysis and data assimilation for environ-

mental characterization and scientific analysis.

2.7 PROBLEM STATEMENT AND RESEARCH OBJECTIVES

Geo-spatial data usually consists of spatio-temporal phenomena that are complex and difficult to understand. It demands new methods to efficiently handle this data that enable temporal analysis while accounting for the spatial context. Therefore, the research aims to design novel approaches for CD analysis for geo-spatial data using deep learning methods for land use and land cover mapping, analysis, and monitoring.

The objectives of the work are:

1. To design a unsupervised learning technique to filter, analyze and extract new and relevant patterns for change detection from multi-sensor and multitemporal satellite remote sensing images.
2. Develop supervised methods for spatial and temporal monitoring of land use and land cover data using images of high temporal and medium spatial resolution by a comparative study with various performance measures.
3. To develop a prediction and automatic/semiautomatic classification system for satellite image time series for land cover classification data using machine/deep learning techniques.

2.8 SUMMARY

This chapter elaborates on a detailed discussion about the CD process and the types of CD. This chapter critically reviewed the latest work on multitemporal image CD, classification, and prediction techniques. A detailed review of all the ML and DL models proposed for CD is also discussed in this chapter. The existing research gaps are listed, and the current research work's problem statement with research objectives is included in one section of the chapter. The proposed methodology for each goal/objective is explained in forthcoming chapters.

CHAPTER 3

BITEMPORAL IMAGE CHANGE DETECTION FOR REMOTE SENSING IMAGES USING UNSUPERVISED DEEP LEARNING TECHNIQUES

3.1 INTRODUCTION

Due to outstanding performance in representing information and mining relevant knowledge, DL methods have seen tremendous growth in pursuance of pattern recognition and computer vision for assessment of remote sensing images. DL-based CD methods can be generally classified into two categories for generating the change maps, i.e: i) Comparison based on in-depth features and ii) Classification based on CNN. In the former, the difference image (DI) is analyzed to determine the change regions generated by calculating the distance pixel-wise between the pairs of the deep features. In the latter, the classification technique aims to learn the elements of change, unchanged and intermediate data between the two bitemporal images used for training and testing the model. Dealing with changes in smaller regions or areas from multitemporal images is challenging in DL using unsupervised learning algorithms. A challenge of this nature manifests itself in several stages, as follows:

1. The production of DI : A maximum number of DI generation methods use a rectangular window to characterize spatial information locally. Smoothing out may occur during changes in small regions or amongst the image's more refined details between the unchanged and changed areas in which the pixels are considered to be changed are complex and challenging to recognize.

3. Bitemporal Image Change Detection for Remote Sensing Images using Unsupervised Deep Learning Techniques

2. Analysis of the DI based on clustering : Most of the clustering algorithms like K-means or Hierarchical clustering approaches didn't succeed in generating accurate outcomes for the DI analysis, owing to the use of a cumulative optimization objective function, which may force the prototype cluster of the minority class i.e., changed pixels to move to the majority class i.e., unchanged pixels, specifically for distributions of the data that are imbalanced.

3. Classification and training using a classifier based on deep learning : When dealing with small regions, the number of changed or modified pixels is less than that of those that stay unchanged. The pseudo label samples used for training the change class may be insufficient to improve DL classifier training, resulting in poor classification accuracy.

The analysis of the DI is a crucial step that assists in converting the CD process into a task of classifying binary values, such as clustering by k-means or the operation of thresholding. Fuzzy C- Means (FCM) is a popular clustering technique that can be applied optimally to detect changes in optical images. According to recent research, the DI has to be classified into three classes: higher probability unchanged, intermediate, and higher probability changed. The pixels in the intermediate category are difficult to separate using a particular clustering approach. Still, they may be discriminated against using a classifier based on DL like a deep CNN.

CNN or ConvNet have contributed to remote sensing image analysis, which can automatically extract useful features from the raw dataset in a hierarchical learning fashion without the manual development of complex features. For the task of CD, the upper layers of information representation that are being learned and tested by ConvNet can intensify the parameters of the given input that are important for discrimination and limit the irrelevant variations, thus contributing to emphasizing the changes that have occurred concerning the ground observations.

There is a need to design a CD framework for analyzing optical remote sensing images of medium and very high resolutions (VHR). Hence, this work proposes a robust and unified Superpixel based Parallel Fuzzy Clustering approach using unsupervised

learning in optical remote sensing images. The significant contributions of our research work are outlined as follows.

- The full benefit of superpixels containing neighborhood information using the Simple Linear Iterative clustering (SLIC) method of segmentation is utilized in the proposed approach.
- Incorporation of a parallel FCM clustering with lightweight deep convolutional neural network (LWDCNN) model to extract the deep spatial-spectral-temporal features from the remote sensing images, further improving CD performance for small areas.
- The proposed model is unsupervised and can cater to CD tasks for images of medium and very high spatial resolutions.

For the analysis of more discriminative features, the extracted feature pair of the fused and reconstructed image is divided into three clusters based on FCM Clustering. The changed and unchanged samples from the bitemporal images are incorporated as labeled data to train the network. The intermediate samples are utilized as labeled data to test the network by feeding the features to the softmax classifier to categorize the final change map.

The chapter is organized as follows: Section 3.1 introduces the motivation behind using unsupervised technique. Section 3.2 briefs preliminary concepts and the details of each module of the proposed CD framework . Section 3.3 describes the data and consideration of the study area . Section 3.4 details the experimental setup and analysis of results, and Section 3.5 summarizes the proposed method and its significance.

3.2 PROPOSED HYBRID CHANGE DETECTION FRAMEWORK

The first framework is being divided into multiple phases and the proposed architecture of unsupervised change detection is as shown in figure 3.1. The following sequence of operations are being followed for the change detection in the bi-temporal images.

- Pre-processing and generation of the difference image

3. Bitemporal Image Change Detection for Remote Sensing Images using Unsupervised Deep Learning Techniques

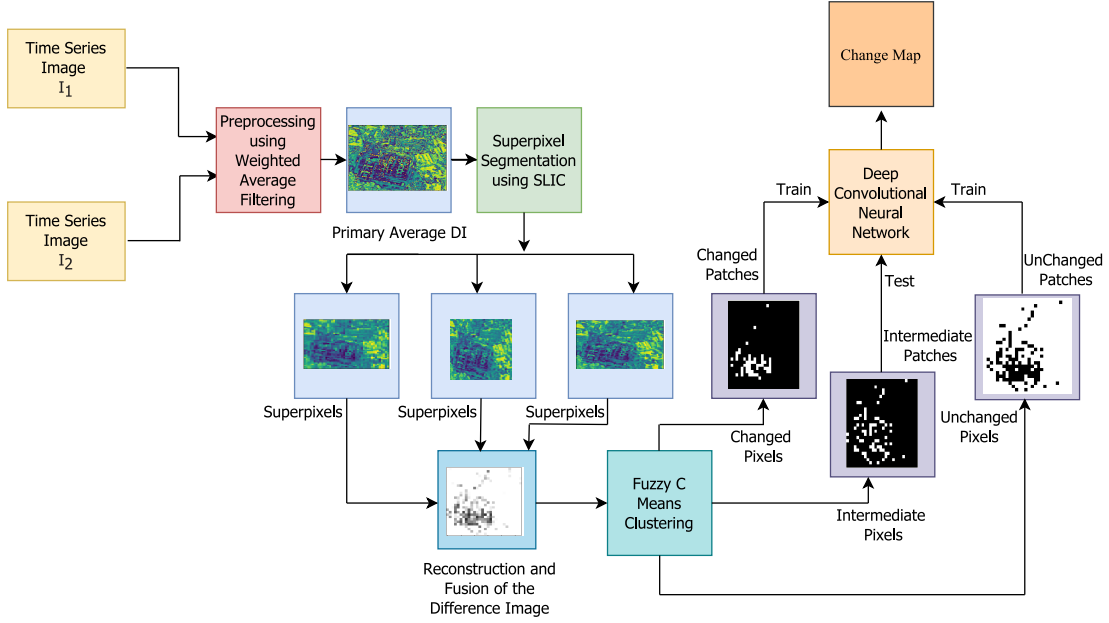


Figure 3.1: Proposed Hybrid Change Detection Framework.

- Superpixel segmentation using SLIC algorithm
- Clustering using Fuzzy C-means
- Training and testing
- Deep convolutional neural networks for classification.

3.2.1 Pre-processing and Generation of the Difference Image

A primary DI (Difference Image) as in equation 3.1 is produced to analyze the changes to demonstrate the degree of differences between a given set of bi-temporal images.

$$X_D = \sqrt{\sum_{b=1}^B (X_1^b - X_2^b)^2} \quad (3.1)$$

where X_m^b denotes the values of the pixels in the b^{th} ($b = 1, 2, 3..B$) band of image X_m ($m=1,2..$).

Structural Similarity Index (SSI) : The equation 3.2 is used to calculate the similarity between the two images for the generation of the primary DI where μ_x , μ_y are the averages of x_i and y_i , σ_x^2 , σ_y^2 is the variance of x_i and y_i , σ_{x_y} being the covariance of x

and y and c_1, c_2 being the two variables for the stabilization with the weak denominator.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_x\sigma_y + c_2)}{(2\mu_x^2\mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3.2)$$

3.2.2 Superpixel Segmentation using SLIC Algorithm

Simple linear iterative clustering (SLIC) algorithm creates superpixels by clustering pixels in the image plane based on their similarity, color, and proximity. D_s is the sum of the lab and XY plane distances, normalized by the grid interval S . In D_s , the variable m , allows us to modify the compactness of a superpixel. The segmentation scale parameter m is set to 10, which means 10 x 10 pixels for each of the superpixels, allowing for a good fit between the segmentation acceleration and the final accuracy of the generated superpixels. Algorithm 3.1 shows the pseudocode representation of proposed hybrid dimensionality reduction technique. The algorithm as developed in (Achanta et al. 2012) for superpixel segmentation with SLIC was being used for the segmentation purpose in our proposed model.

$$d_{lab} = \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} d_{xy} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} D_s = d_{lab} + \frac{m}{S} d_{xy}, \quad (3.3)$$

Algorithm 3.1: Superpixel Segmentation with SLIC.

- 1: Initialize cluster centers in $C_k = [l_k, a_k, b_k, x_k, y_k]^T$ by sampling pixels at regular grid steps S .
- 2: Perturb cluster centers in an $n \times n$ neighborhood, to the lowest gradient position.
- 3: **repeat**
- 4: **for** each cluster center C_k **do**
- 5: Assign the best matching pixels from a $2S \times 2S$ square neighborhood around the cluster center according to the distance measure as in Eqn.3.3 .
- 6: **end for**
- 7: Compute new cluster centers and residual error E (L1 distance between previous centers and recomputed centers)
- 8: **until** $E \leq$ threshold
- 9: Enforce Connectivity.

3.2.3 Clustering using Fuzzy C-Means Algorithm

The FCM method divides a bounded collection of n elements $X = \{x_1, x_2, \dots, x_n\}$ into sets of C fuzzy clusters as per the stated criterion. The algorithm returns a list of cluster centers $C = \{c_1, c_2, \dots, c_c\}$ and a partition matrix for a finite data collection. Each element

3. Bitemporal Image Change Detection for Remote Sensing Images using Unsupervised Deep Learning Techniques

specifies the degree to which an element belongs to a cluster. Where $W = w_{i,j} \in [0, 1]$, $i=(1,2\dots n)$, $j=(1,2\dots c)$ and $w_{i,j}$ denotes the degree to which element x_i belongs to cluster c_j . The objective function is calculated as in equation 3.4 as:

$$\operatorname{argmax}_C \sum_{i=1}^n \sum_{j=1}^n w_{i,j}^m \|x_i - c_j\|^2 \quad (3.4)$$

The $w_{i,j}$ is computed using equation 3.5 :

$$w_{i,j} = \frac{1}{\sum_{k=1}^c} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}} \quad (3.5)$$

The reconstructed and fused difference image with parallel FCM technique is used to categorize the DI into sets of three clusters.

3.2.4 Training and Testing

The clusters are denoted as change class θ_c that represents pixels of high intensity, unchanged class θ_u , pixels of low intensity, and the intermediate class θ_i that are the hard pixels being classified to determine final change or unchanged class. The pixels that exhibit a high probability of being changed are denoted as θ_c and unchanged as θ_u these two clusters, i.e., θ_c and θ_u are being nominated as training samples to train the model. The pixels θ_i will be tested and further classified by Deep Convolutional Neural Network (DCNN). Let $T_p^{I_1}$ denotes a group of pixels being focused at pixel p in image I_1 and $T_p^{I_2}$ represents a group of pixels being focused at pixel p in image I_2 . Each group of pixels is of size $p \times p$. The concatenation of both the images is performed to create a new image patch T_p with a size of $2z \times z$. Let N denotes the group of pixels that belongs to θ_c and θ_u , which contributes to obtaining the sample images T_p , $p=1,2,3\dots N$.

3.2.5 Lightweight Deep Convolutional Neural Network for Classification

The CNN aims to learn distinct features from multitemporal images, and this set of features can distinguish between change and unchanged information. The lightweight DCNN, is a network derived from CNN, which contains multiple convolutional, max-pooling, and fully connected layers. The architecture of the DCNN consist of three convolutional layers namely Conv1, Conv2, and Conv3, and two max-pooling layers as shown in figure 3.2. The model parameters are shown in Table 3.1

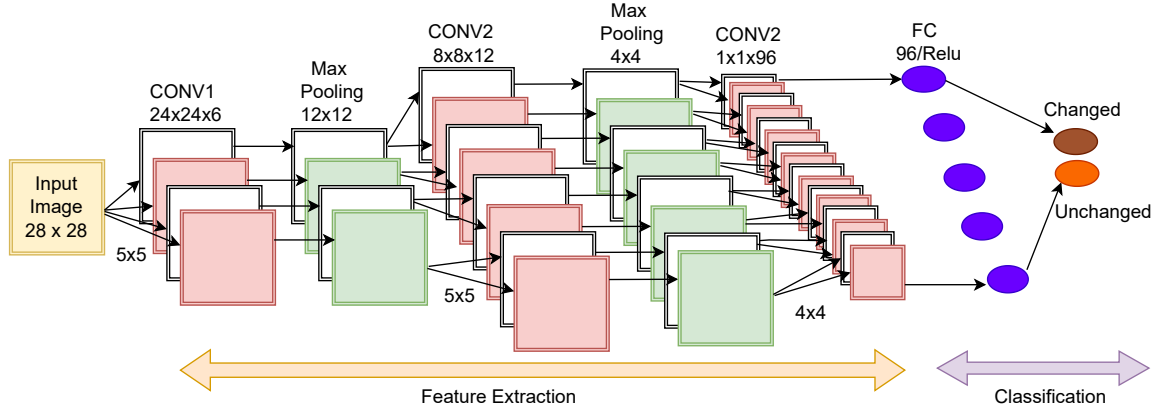


Figure 3.2: Light Weight Deep Convolutional Neural Network.

Table 3.1: Lightweight CNN Model Parameters.

| Layer Type | Output Shape | Parameters |
|--------------------------|--------------|---------------|
| Input Layer | 28,28,3 | — |
| Conv1 | 28,28,6 | 10374 |
| Pool1 | 3,3,6 | 0 |
| Dropout | 3,3,6 | 0 |
| Conv2 | 3,3,12 | 4620 |
| Pool2 | 1,1,12 | 0 |
| Dropout | 1,1,12 | 0 |
| Conv2 | 1,1,96 | 1248 |
| Dropout | 1,1,96 | 0 |
| Flatten | 96 | 0 |
| Dense | 96 | 9312 |
| Dropout | 96 | 0 |
| Dense | 2 | 194 |
| Total Trained Parameters | | 25,748 |

3.3 STUDY AREA AND DATA DESCRIPTION

3.3.1 Alappuzha District Kerala, India

The present study was conducted for Alappuzha, Kerala's smallest district and a tourist attraction in India. It is well-connected to other parts of Kerala via waterways. Alappuzha (Alleppey) is located on a peninsula between the Arabian Sea and a network of rivers that pour into it. The experiments were performed on these datasets with medium spatial resolutions. The data set used for analysis was bitemporal images from the

3. Bitemporal Image Change Detection for Remote Sensing Images using Unsupervised Deep Learning Techniques

Geoportal website, Bhuvan, National Remote Sensing Agency, Indian Space Research Organization, India. We acquired the remote sensing images from the Resourcesat2 Linear Imaging Self Scanning (LISS-III) sensor over the small area of the Alappuzha district in Kerala, India. The LISS-III data sets are categorized into four spectral bands: 0.52-0.59 microns (B2), 0.62-0.68 microns (B3), 0.77-0.86 microns (B4), and 1.55-1.70 microns (B5). There are three visible bands and one near-infrared band. LISS-III data has a spatial resolution of 23.5 m. The images are as shown in figure 3.3 (a) acquired on 05-02-2009 and figure 3.3 (b) acquired on 10-12-2019 with the reference image in 3.3(c). Both the images are of similar spatial resolution of 23.5m/pixel (medium resolution) and consist of 257×257 pixels.

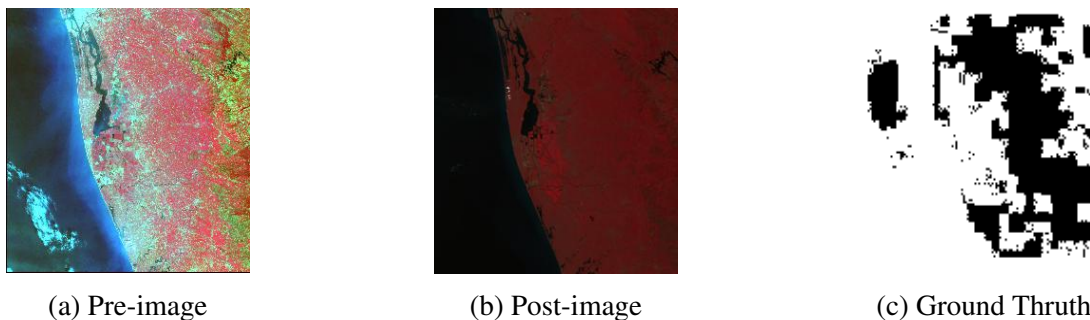


Figure 3.3: LISIII Dataset for the Alappuzha Region (Bhuvan).

3.3.2 Earth Dataset of a Small Region Near Paris,France

The second dataset used was the bitemporal optical time-series images for a city near the Paris location called as Earth dataset from Baudhuin et al. , the images are as shown in figure 3.4(a) acquired on 24-03-2010 and 3.4(b) acquired on 17-12-2019 with reference image in 3.4(c). The experiments were performed on these datasets consisting of very high-resolution images. They are of 50cm/pixel spatial resolution with a size of 650x 650 pixels. The dataset consists of three bands that describe the seasonal variations. As seen in Figures 3.3 and 3.4, significant variations occurred in the two sets of bitemporal images. Both the datasets used were unique. The first study area is an ecologically small, fragile area in India. The second study region consisted of the green land converted mostly to buildings for a small area near Paris. The final change maps for the bitemporal images were obtained, with white pixels signifying a change and

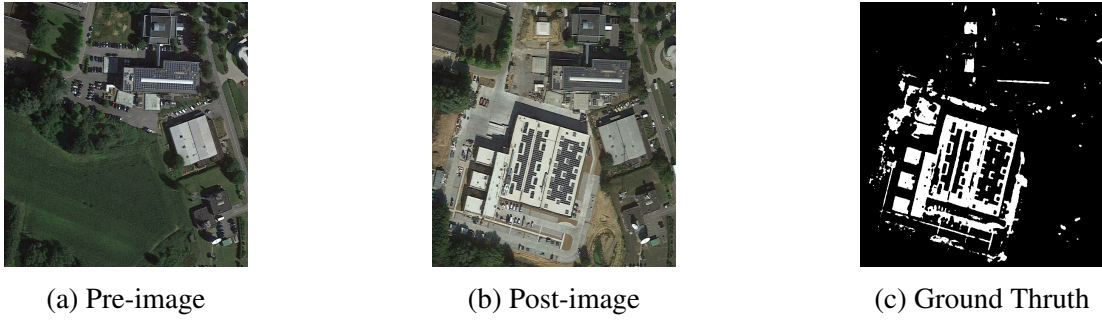


Figure 3.4: Earth Dataset for a City near Paris.

black pixels representing unchanged.

3.4 RESULTS AND DISCUSSIONS

3.4.1 Experimental Setup

The proposed method is evaluated by experiments in universally available optical image datasets. The entire set of experiments are executed on a computer with Core Intel i7-8700k CPU, 16-GB RAM, and NVIDIA GeForce GTX 1090Ti GPU. using python3 programming implementation. Two remote sensing datasets as described above are used to evaluate the performance of the proposed method and SOTA techniques.

3.4.2 Performance Evaluation Measures

For the evaluation of our proposed method, numerous experiments were performed on the two optical image datasets. A detailed explanation of evaluation measures are as follows:

Various performance measures are used like Percentage of correct classification (PCC), as in Eqn .3.6, Kappa Coefficient (KC), and Area under the receiver operating characteristics (AUC).

$$PCC = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.6)$$

The kappa coefficient(KC) as in Eqn. 3.7 with PRE (proportion of expected agreement) in Eqn.3.8 is a parameter that is used to measure the classification outcome accurately. Higher values of kappa signify higher accuracy. The kappa coefficient is given by

$$Kappa(KC) = \frac{PCC - PRE}{1 - PRE} \quad (3.7)$$

$$PRE = \frac{(TP + FP).(TP + FN)}{(TP + TN + FP + FN)^2} + \frac{(FN + TN).(FP + TN)}{(TP + TN + FP + FN)^2} \quad (3.8)$$

Area under ROC curve: The area under the ROC curve or precisely area under the curve in classification performance evaluation is an aggregate measure over all possible classification thresholds. It was used when the model ranks random positives over random negatives. It measures how well the prediction ranked than its absolute value; therefore, AUC is scale-invariant. AUC value ranges from 0 to 1.

The Area Under the Curve (AUC) as indicated in Eqn.3.9 and Eqn.3.10 with Sensitivity and Specificity is a summary of the ROC curve that measures a classifier's ability to distinguish between classes. The AUC indicates how well the model distinguishes between positive and negative classes. The greater the AUC, the better the performance of the model.

$$Sensitivity = Recall = \frac{TP}{TP + FN} \quad (3.9)$$

$$Specificity = \frac{TN}{FP + TN} \quad (3.10)$$

AUC is calculated as the Area Under the $Sensitivity(TPR) - (1 - Specificity)(FPR)$ Curve. The F1 score is calculated using Eqn.3.11.

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \quad (3.11)$$

Overall accuracy: Overall accuracy is a measure that states how many samples are correctly mapped to their corresponding class. It is the most comfortable measure to find the performance of a classifier. Consider a dataset that has N number of samples and C class labels, then the confusion matrix M of classification is a square matrix of size $C \times C$. Overall accuracy is measured using the equation 3.12

$$OA = \frac{\sum_{i=1}^C M_{ii}}{N} \quad (3.12)$$

Diagonal elements M_{ii} in the confusion matrix is the number of samples correctly classified, and the overall accuracy of a classifier is always measured on a percentage scale.

3.4.3 Comparative Methods

The comparison and analysis are performed with the current SOTA methods proposed in the CD domain. A unique CD framework for high-resolution remote sensing images was proposed in Gong et al. (2017). It incorporated superpixel-based change feature extraction using superpixel segmentation and neural network-based hierarchical difference representation learning (SBDN) to classify the changed and the unchanged pixels to generate the final change map.

Multi-scale superpixel segmentation (MSDNN) CD technique based on stacked denoising autoencoders (SDAE) is proposed in Lei et al. (2019). SDAE was utilized to analyze the difference representation between superpixels of bitemporal images in the proposed method.

In He et al. (2021), the authors developed a unique multi-scale analysis framework for unsupervised CD based on multi-scale visual saliency coarse-to-fine fusion (MVSF). Superpixels were used as primitives in MVSF to analyze the DI obtained using the change vector analysis approach. The global contrast of each superpixel was then used to construct multi-scale saliency maps at the superpixel level.

A weighted fusion technique was devised to add multi-scale saliency at the pixel level. The authors in Zhang et al. (2018) merged DRL (difference representation learning) and unsupervised clustering into a unified model with superpixel segmentation (HIDRL-Net). Superpixel segmentation was employed to organize individual pixels into homogeneous zones, which simplifies the multi types CD work, and patches centered at the superpixels are being used to represent them for object-based change analysis.

The research in Zhang et al. (2021a) proposed an end-to-end superpixel-enhanced CD network (ESNet) for VHR images, which combined differentiable superpixel segmentation and a deep CNN for VHR images. For feature extraction and superpixel segmentation of bitemporal image pairings, two weight-sharing superpixel sampling networks (SSNs) have been developed. The varied information is then mined using a UNet-based Siamese neural network.

3.4.4 Experiments on the LISS-III Dataset

The application of various methods for CD on the LISS III dataset is demonstrated in figure 3.5 (a)-(f). SBDN 3.5(a) and MSDNN 3.5(b) results have not produced a good map, and multiple regions are not detected and classified. The change map shows that the MVSF3.5(c) and HIDRLNet3.5(d) method has exhibited good but not complete changes compared with the ground truth image. ESCNet 3.5 (e) method has also been performed very effectively, but detecting small changes is not done accurately. The map of the proposed approach reveals the change areas without any noise, as depicted with 3.5(f) in comparison to all of the above methods. The statistics in Table 3.2 show that our proposed method has performed well compared to other CD methods.

Table 3.2: Comparative Analysis of Various Methods on LISS-III (Alappuzha) Dataset.

| Methods | LISSIII Dataset | | | |
|----------|-----------------|-------|--------|-------|
| | PCC(%) | KC(%) | AUC(%) | F1(%) |
| SBDN | 91.25 | 61.35 | 78.2 | 43.71 |
| MSDNN | 94.56 | 72.56 | 68.5 | 48.34 |
| MVSF | 89.66 | 77.45 | 69.8 | 51.02 |
| HIDRLNet | 95.33 | 63.58 | 79.1 | 59.34 |
| ESCNET | 96.21 | 75.69 | 80.1 | 67.45 |
| Proposed | 98.28 | 81.64 | 84.6 | 74.57 |

3.4.5 Experiments on the Earth Dataset

Figure 3.6(a)-(f) demonstrates the CD results generated by various approaches on the Earth dataset. SBDN3.6(a) has detected very few changes, and partial analysis of change areas by the MSDNN is seen in 3.6(b). MVSF3.6(c) and HIDRLNet3.6(d) has shown the change regions but is not adequate to the ground truth, including the ESCNET in 3.6(e). The proposed 3.6(f) method has meticulously identified all the changed regions in small areas without the inclusion of any noise. The statistics in Table 3.3 show that our proposed method is better than that of other CD methods. The PCC, KC, AUC, and F1 score results for both datasets have increased, and the use of self-supervision as a data-centric the approach has demonstrated the efficiency of the proposed framework. Using superpixel segmentation, clustering, and bitemporal

patches of images collected to train the model with self supervision have shown significant performance and improvement in the results. The AUC for both the datasets is as shown in figure 3.7(a) and 3.7(b).

Table 3.3: Comparative Analysis of Various methods on Earth (Paris) Dataset.

| Paris Dataset | | | | |
|---------------|--------|-------|--------|-------|
| Methods | PCC(%) | KC(%) | AUC(%) | F1(%) |
| SBDN | 93.54 | 69.47 | 66.1 | 38.28 |
| MSDNN | 94.44 | 71.43 | 58.9 | 45.09 |
| MVSF | 88.37 | 80.41 | 70.1 | 41.25 |
| HIDRLNet | 92.10 | 75.25 | 67.3 | 52.11 |
| ESCNET | 94.22 | 78.50 | 81.2 | 58.78 |
| Proposed | 96.43 | 82.30 | 85.7 | 64.06 |

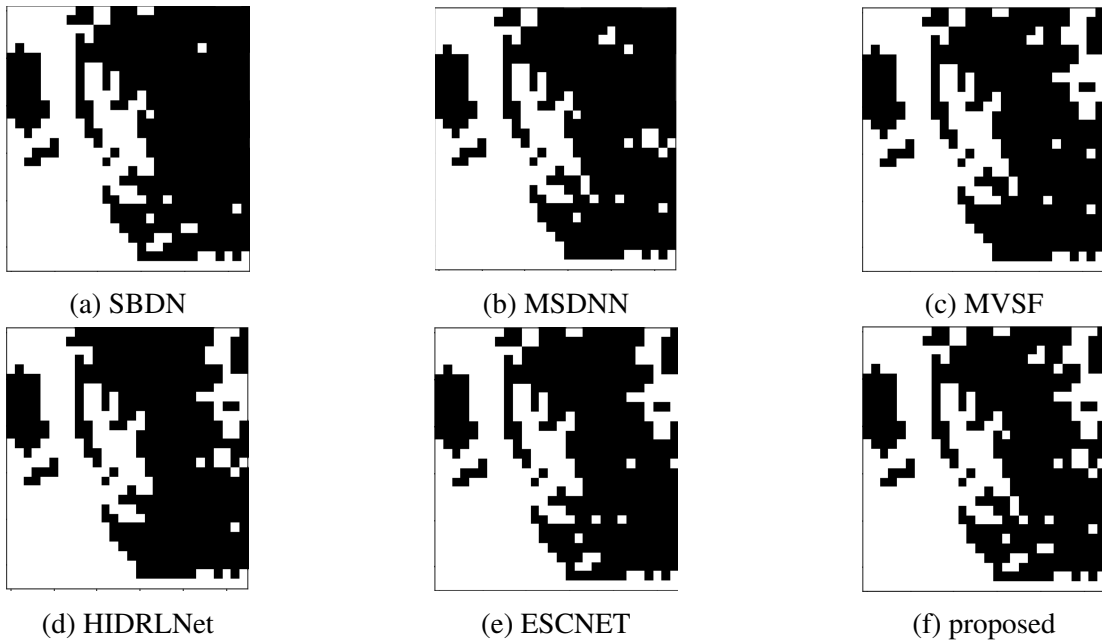


Figure 3.5: Change Maps generated for LISS-III (Alappuzha) dataset.

3. Bitemporal Image Change Detection for Remote Sensing Images using Unsupervised Deep Learning Techniques

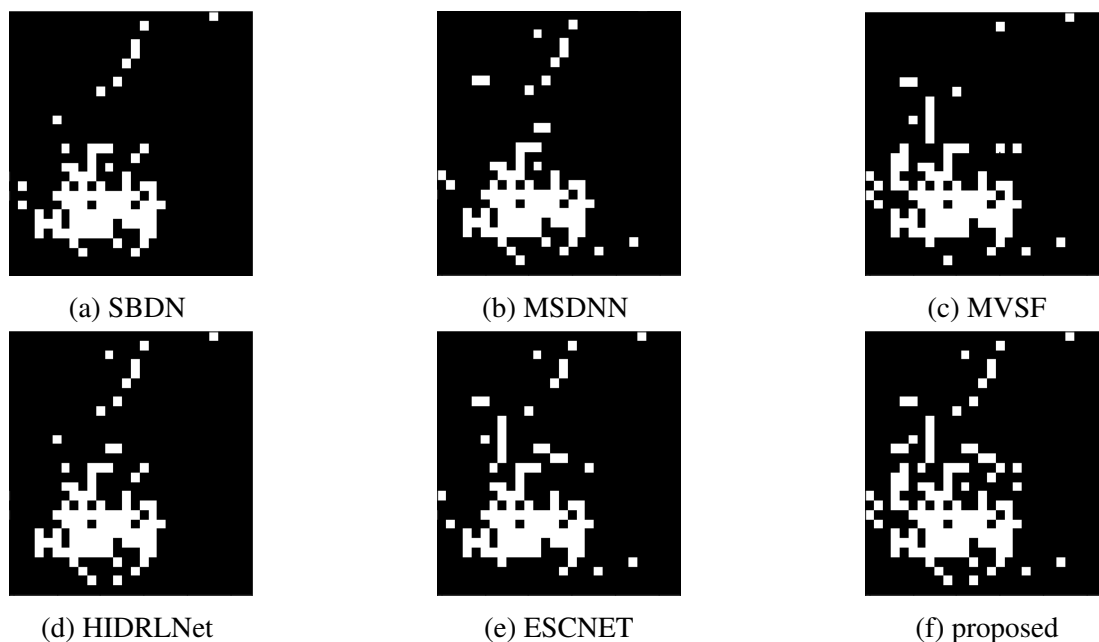


Figure 3.6: Change Maps generated for Earth (Paris) dataset.

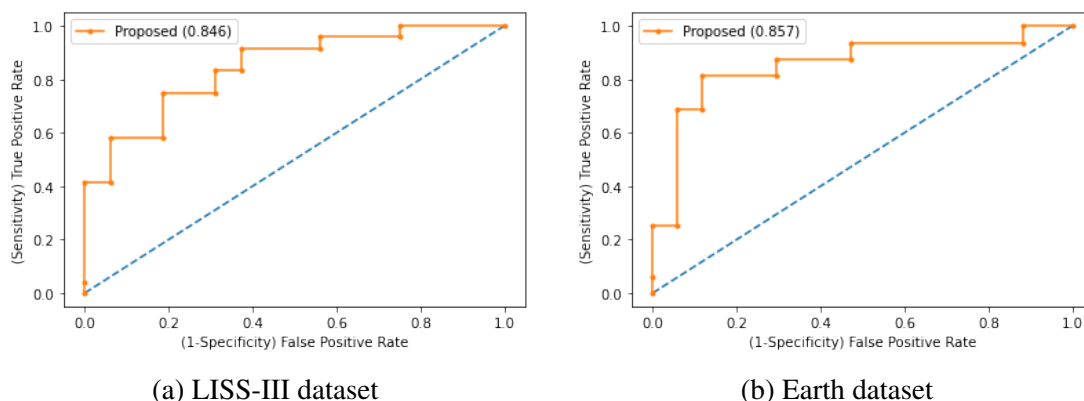


Figure 3.7: AUC Curve.

3.5 SUMMARY

In addressing issues like performing CD on optical remote sensing datasets of a small area, there is a high possibility of generating noise and distortion in the images, which may decline CD accuracy. For locations designated as small areas, the number of changed pixels may be false compared to the actual number of real changed pixels due to the closeness of the pixel values in the images, which is a big challenge addressed in this chapter. Existing methods are mainly proposed for a large set of areas. This chapter proposed an accurate CD framework due to the richness of DL. In the phase of DI

generation using superpixel segmentation, the developed fused image due to the property of superpixels focused on enhancing the gap between the changed and unchanged areas. It also helped decrease the interference caused by the pixels, which are changed and labeled as false.

While performing DI analysis in the clustering phase using preliminary centers as constraint terms in the objective function, the novel parallel FCM algorithm was utilized to preserve the clustering algorithm in the right direction during the optimization phase. To meticulously control the imbalance between the changed and unchanged classes from the bitemporal images, the training and testing samples were generated from the image patches of the changed and unchanged pixels. This strategy enriched the CD process by training and testing the proposed light weight DCNN with the intermediate patches and final changed class. With the incorporation of patches of the same area, the framework could effectively capture the changed class's distinguishable features. The evaluation of each module in the framework was done in a precise manner by comparing it with several SOTA approaches. The experimental analysis demonstrates the effectiveness and usefulness of each component in the proposed framework. In the future, an effort will be made to enhance the framework with the facility of limited supervision using deep learning to overcome the drawbacks of supervised learning.

CHAPTER 4

A DEEP LEARNING APPROACH FOR MULTI-CLASS CHANGE DETECTION IN LAND USE LAND COVER DATA USING SUPERVISED TECHNIQUES

4.1 INTRODUCTION

Land is the most crucial component for analyzing Earth's change. Handling the increasing load on land use demands a productive policy for enduring land use and economic growth. Investigation of local and global indicators established on land use maps could disclose information on sustainable development progress (Camilleri et al. 2017). LULC changes, on the other hand, contribute to severe environmental concerns in various regions of the world. It is necessary to detect and mitigate land deterioration and improve the land rehabilitation process to detect and monitor changes on the surface of the Earth (Nedd et al. 2021). With industrialization and globalization, human exploitation of land has increased globally, posing a severe environmental hazard (Wijaya et al. 2018) (Antrop 2009). As a result, determining the change in LULC has become a critical issue that needs attention globally (Ojima et al. 1994).

The physical component of the Earth's surface (land cover) and how we use the land (land use) are difficult to distinguish in environmental monitoring and many other subdomains, according to (Hansen et al. 2022) worldwide. Field surveys or satellite image analysis can achieve this via remote sensing techniques. While surveys in the field are more comprehensive and definitive, they are expensive projects that take a

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

long time to update (MohanRajan et al. 2020).

Deep learning (DL) methods are currently the talk of the town due to their extensive applications in image processing, natural language processing, signal processing(Xie et al. 2017),etc. DL has also emerged as a strong baseline in earth observation and remote sensing domain to analyze SAR, optical, VHR(very high resolution), and hyperspectral data (Zhu et al. 2017)(Bai et al. 2022). CD methods involved with remote sensing data consist of the feature extraction process and identification of the change areas. The essential features like texture, color, and context are analyzed in the extraction stage. Identifying the extracted features for determining change points is performed by designing hybrid or using pre-trained deep learning models in the set of multitemporal images.

To solve these challenges, a DL-based approach called a ST encoder-decoder network is developed with attention mechanisms to examine historical LULCC for a given region using satellite images. This analysis will be helpful for city planning and environmental management in fast-increasing cities, saving time and energy. The major contributions of our research work are as follows:

- A dual branch end-to-end encoder-decoder network is designed with self and dual attention modules that accept two LULC maps as input and extract features independently to identify and analyze the changes.
- The class imbalance problem is overcome by the attention module based network trained directly without the pre-training models.
- The architecture learns the multiscale features from both the images and the low-level features are preserved through the long-range connections between the encoder and the decoder module.
- The proposed framework is evaluated on a LULC dataset for the Dakshina Karnataka and Goa region, and the model achieved an overall accuracy of 94.11% and 94.93%.

The chapter is organized as follows: Section 4.1 is for introduction, Section 4.2

briefs the overall framework and provides description of each component. Section 4.3 describes the data and its collection with the study area considered, and Section 4.4 presents the results and its detailed analysis using various evaluation parameters, and Section 4.5 summarizes the proposed method and its significance.

4.2 OVERALL FRAMEWORK

In general, the overall framework is depicted in figure 4.1. The proposed CD approach for LULC remote sensing datasets consists of three basic steps: (1) pre-processing, (2) training sample generation, and (3) end-to-end CD learning. The presented method uses a supervised binary CD framework to create a change map in binary format (i.e., a change/no-change map) from two images obtained from the same location simultaneously. In the following three subsections, the further steps are discussed.

Pre-processing The pre-processing phase plays a crucial role in CD for remote sensing data. Image processing methods like co-registration and denoising are used to make the images as comparative as possible. This stage also deals with eliminating noise and scale variations, if any.

Training sample generation The initial step in the model design process is creating training data to build the DL model. A training data set is a group of samples used to train and develop a model. Several training samples were extracted by the individual selection of pixel locations belonging to each class from the LULC data. The distinct images were used to utilize each LULC class location for training the model. The proposed DL model for CD “learns” from training samples, each of which relates an example input to the related output. Each training dataset for the final models included multiple samples corresponding to several images of every location from LULC data. For our study, LULC maps of thirteen years, from 2005 to 2018, were used for the CD process, and the proposed model was applied to identify the LULC changes. Ten images from 2005-06 to 2014-15 were used for training, and three images from 2015-16 to 2017-18 were used for testing and validation.

The different LULC maps of different years were utilized to derive the signature

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

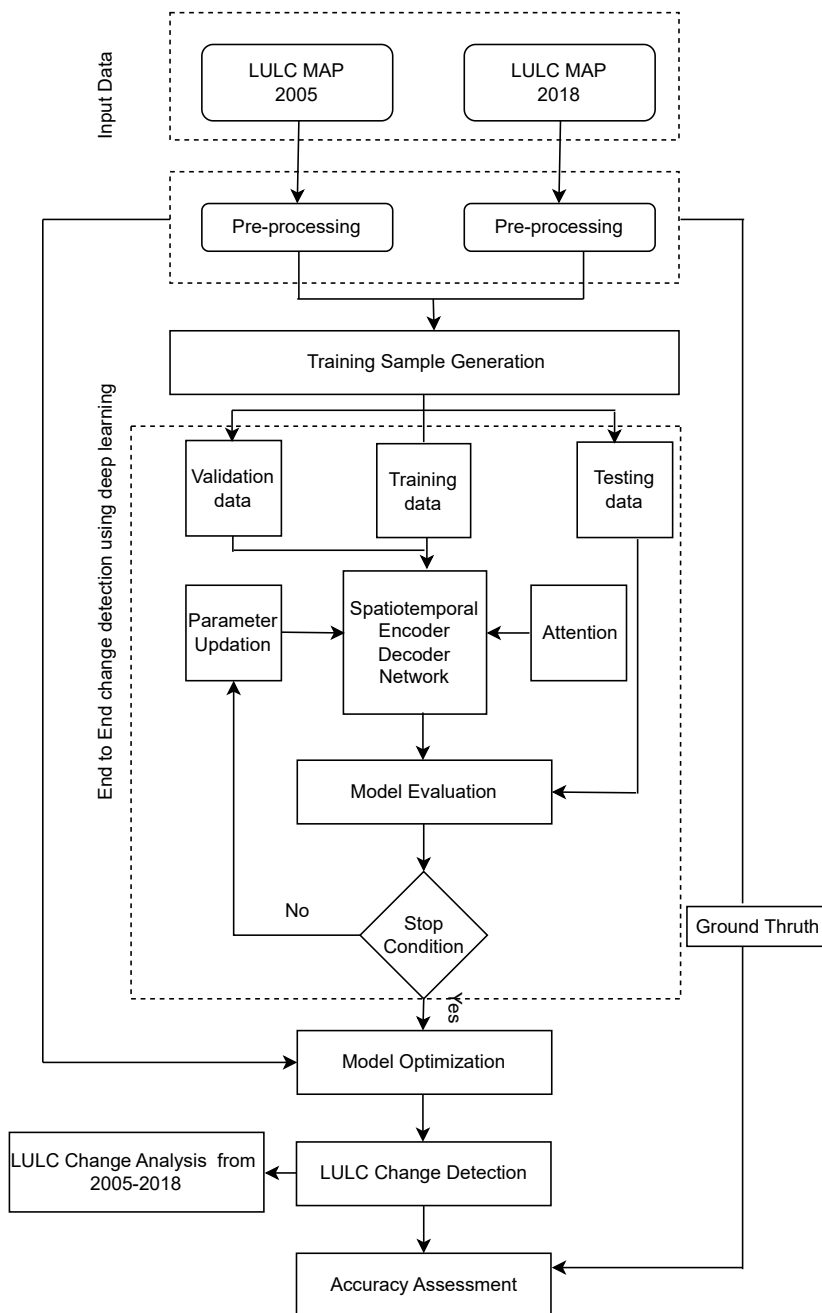


Figure 4.1: Overall Framework of the Proposed Approach.

classes from the same LULC class during the training, testing and validation process. The derived signatures were also utilized with the four models for comparison during this training process. Every image's samples were divided in a similar way, and this division is maintained for every training. Thus no training sample was ever employed

for validation, and vice versa. A unique, weighted sparse categorical cross-entropy loss function was utilized for all model training. However, given the enormous training data, the configuration of DL routines was done to last a fixed number of epochs. The proposed model is being trained and evaluated using the testing data by updating and fine-tuning the parameters needed to generate the final binary change map.

4.2.1 Proposed Change Detection Framework

In this section, the details of the proposed architecture are elaborated as shown in figure 4.2. Four encoders and four decoders are developed for the task of CD. The criteria for fixing the number of encoders and decoders was done based on the proposed architectures in (Peng et al. 2021) (Bandara and Patel 2022) (Chen et al. 2022). An explanation of how the self and dual attention mechanisms were applied at the decoder level to identify ST relationships for detecting changes in the bitemporal images based on the proposed architecture led to two novel architectures, STEDSAN and STEDDAN.

Encoder Network At the encoder the dual branch feature extraction module (FEM) contains four blocks of the 2D convolution layer of size 3x3 and performs strided convolutions with a stride of 1 and 2 with downsampling, including the batch normalization function. This dual branch encoder network takes both the bitemporal images as separate input for the processing without losing valuable information, leading to faster convergence and making the network more stable. For the encoder, the feature maps extracted in the n^{th} layer of FEM at time instant t_1 and t_2 are used as $f_n^{t_1}, f_n^{t_2}$ $\mathbf{R}^{w_n \times h_n \times d_n}$ for $n=1 \dots N$. The extracted feature maps are represented as $f_N^{t_1}$ for X^{I_1} and $f_N^{t_2}$ for X^{I_2} . The information which is being changed or unchanged is inferred from $f_N^{t_1}$ and $f_N^{t_2}$. To achieve the difference image from the change information being learnt and the combined feature illustration of the extracted two images from the outputs of the last layer of each block of convolution, the task of concatenation of the two features is incorporated. The task of concatenation is performed to concatenate the two features $(f_1^{t_1}, f_1^{t_2}), (f_2^{t_1}, f_2^{t_2}), (f_3^{t_1}, f_3^{t_2}), (f_4^{t_1}, f_4^{t_2}), \dots, (f_N^{t_1}, f_N^{t_2})$ is carried out in an effective way. The encoder generates multilevel features consisting of high-resolution coarse features and

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

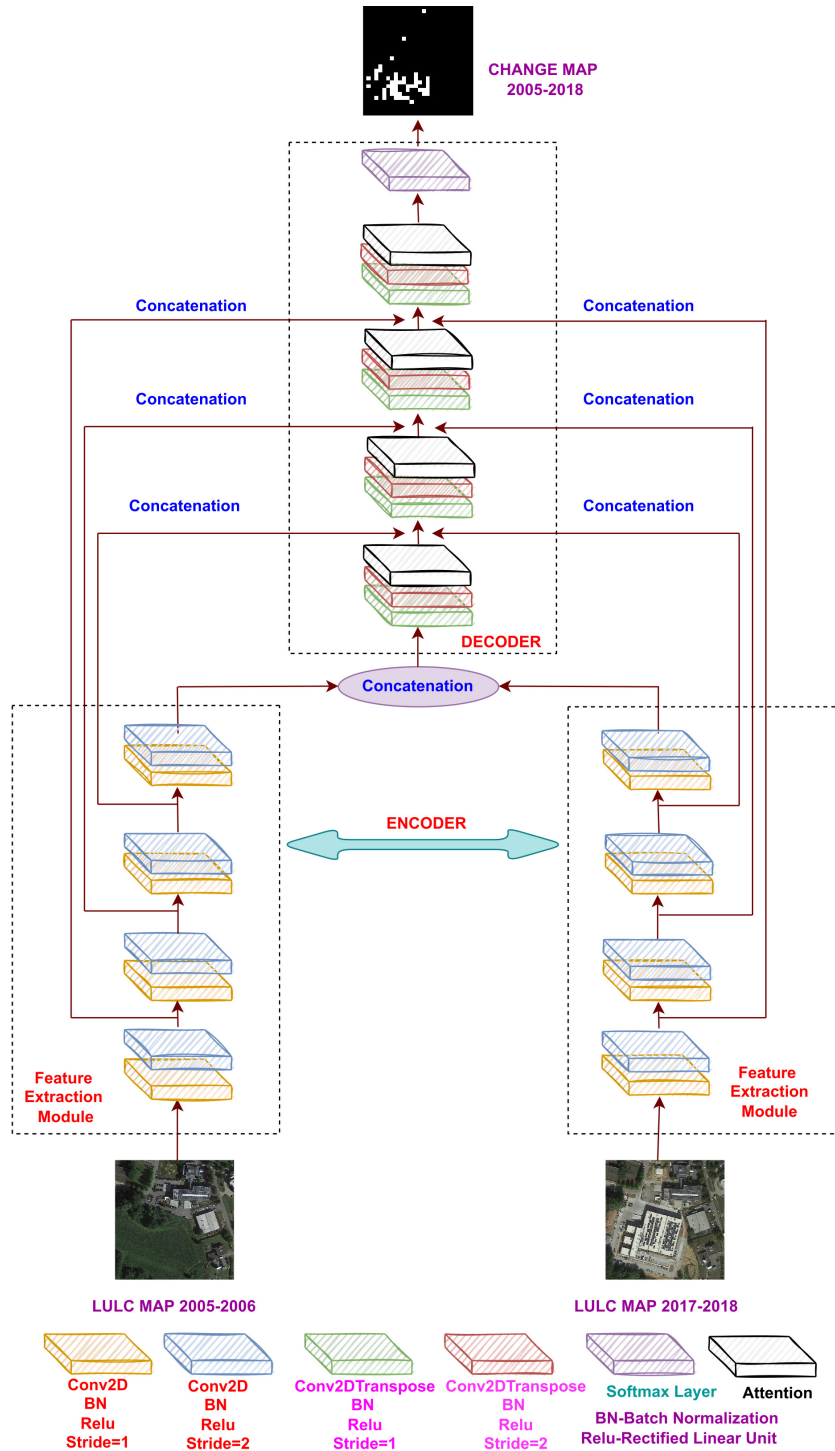


Figure 4.2: Spatio Temporal Encoder Decoder Network.

low, fine-grained resolution features needed for the CD process. The difference between multilevel features of the pre-change and post-change images is computed. The encoder block is given in Equation 4.1.

$$F_N^T = BN((CONV2D_{3 \times 3}(cat(f_n^{t_1}, f_n^{t_2}))) \quad (4.1)$$

The features are concatenated and provided as convolved feature maps to the decoder. Since it is a probability prediction problem, for determining the changes in every pixel a decoder like architecture is successively inducted for carrying out this task as explained in the next section.

Decoder network The decoder consists of four sets of deconvolutional (transposed convolution) blocks, each consisting of a deconvolutional layer of size 3x3 with a stride of 1 and 2 with the upsampling operation. Each module's various intermediate change maps are represented as F_N^T where (N=1..4) are being weighted, upsampled and fused as the final ultimate change map. The activation function is included after every deconvolutional layer except for the last deconvolutional layer. To achieve this, a softmax layer is added to develop the final binary change map that will correspond to determine the changes in each class label for the prediction of the final map. With such a dual-stream architecture of encoder-decoder, the output will be the same size as the input, causing no harmful effects on the pixels and finer details of the image will also be preserved. The concatenated extracted feature pairs at each stage by strengthening the change information through the SA modules, are explained in the next section.

4.2.2 Attention Mechanism

The attention mechanism is incorporated into the DL architectures. It refers to a process through which a network can weight features according to their importance to a task and use that weighting to aid in completing the task (Mnih et al. 2014). For image CD the main goal of the attention technique was to determine the weight distribution in the image and assign various weights to various significant locations. By decreasing the weight of features irrelevant to the target and increasing the weight of relevant features to promote learning, the attention mechanism inhibits the learning of such features.

Multiple attention models are being proposed in the area of DL for computer vision like soft, spatial, channel, mixed, self, and hard attention, as summarized in (Yang 2020)(Brauwera and Frasincar 2021).

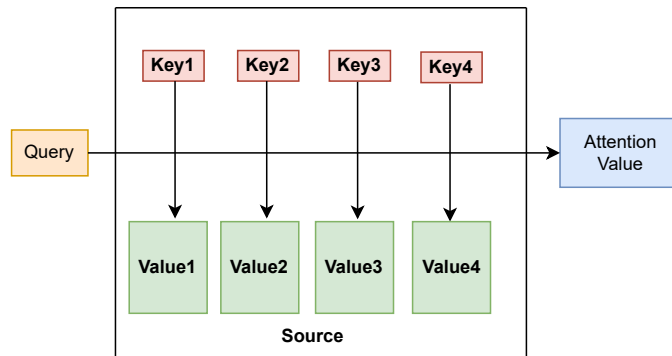


Figure 4.3: Self Attention Mechanism.

4.2.2.1 Self Attention

SA (Self Attention) for sequential data(Vaswani et al. 2017) is a mechanism for focusing attention that links various positions of a single sequence to determine how each position of the sequence should be represented. SA is one kind of attention technique built on encoder-decoder architecture and incorporated in our proposed approach as employed in (Ramachandran et al. 2019). There is no recurrence operation being used in SA. Weighted correlations exist between the components of the input sequence. As per figure 4.3, three terms are considered query, key, and value to clarify the fundamental concept of the SA mechanism. Consider a database with several key-value pairs. The database element must be located for a new query that most closely matches it. These can be done by determining how closely the query resembles each key in the database. This concept is the foundation for the SA process, which determines the correlations between various components. In particular, the query and key are derived from the same source in the SA method. SA calculates queries(Q), keys (K), and values (V) for the images in a given hidden sequence H using a linear transformation of the input sequence in equation 4.2, as follows:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4.2)$$

where d_k is the scaling factor, K denotes dimensionality, $Attention(Q, K, V)$ denotes the attention matrix ($N \times d_k$), and N denotes the number of items in the input sequence. In our case, while taking the inputs for the study for the SA mechanism, the sensitivity analysis was done on how to compare and contrast the effect of each input on the output. Typically, the most common approach is to hold all the attributes at their mean value while varying just one of the inputs to assess the effect of changing just one variable. In this example, one attribute is varied at a time at multiple steps to assess the overall sensitivity of each variable. Adding half a step means the midpoint value is used between the mean and maximum for a single variable, and adding a full step means the maximum value for that variable will be used. This process was carried out to supply the required values effectively in our proposed approach. The weighted sum of the value vectors yields the output vector, with each value vector's weight determined by an affinity function based on the query and its corresponding key.

The attention value is derived by weighting and summing all values. The weight coefficient computation reflects the process of concentrating on crucial information. The mechanism focuses more on the value corresponding to the weight, representing the information's significance. In contrast, the value is the information that corresponds to the weight.

In our proposed architecture, the source comes from the encoder layer, and the target is the result of the decoder layer. The source's elements comprise several "key, value" data pairs. The design of the SA mechanism included in our proposed STED-SAN network is depicted in the figure 4.4. When the query, key, and value tensors are individually extracted from the input feature tensor by three different 1×1 convolutional layers, a query as a vector is described at a given point in the query (or key, value) tensor. The bitemporal image feature mappings in the temporal dimension are concatenated to create the feature tensor. Learning an attention function that translates a query vector, a set of key-value vector pairs, and a set of output vectors is the main goal of the SA module for ST analysis.

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

To calculate E, a set of key vectors, value vectors, and query vectors are created from an input tensor. Each output vector is then created by learning the weighted sum of the values. The weight given to each value depends on how similar the query is to the input vector and the associated key. The features are calculated from the input images given as $A \in \mathbb{R}^{C \times H \times W}$, where C, H, and W are the input image's channel, height, and width, as seen in the figure 4.4. The computation of keys, values, and queries from A occurs initially. Feeding them to 1×1 convolution layers, the input feature tensor A is first transformed into two feature tensors, Q (query) and K (key), where $Q, K \in C \times H \times W$.

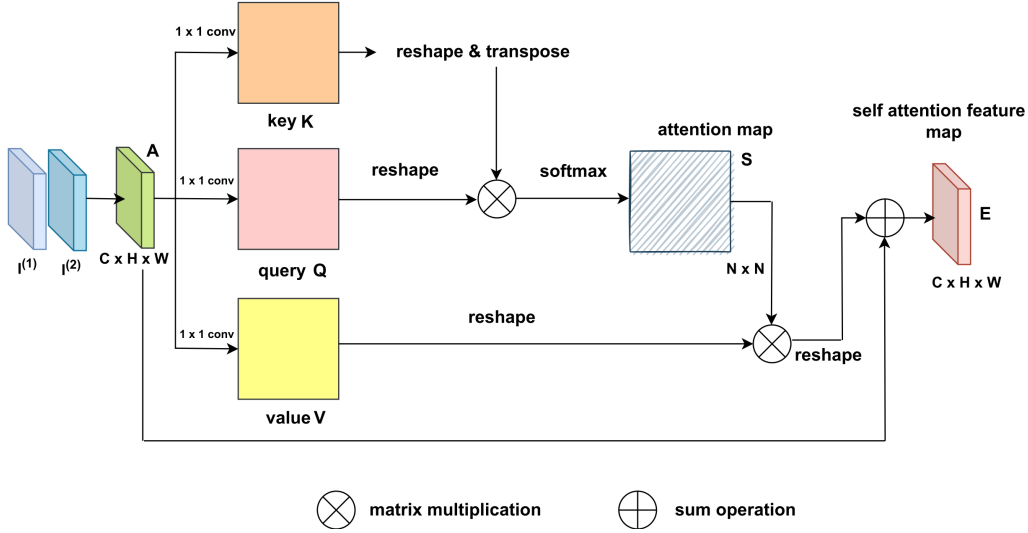


Figure 4.4: Self-Attention Module incorporated in STEDSAN.

This tensor is subjected to the reshape operation to produce $\mathbb{R}^{C \times N}$, where $N=H \times W$ is the number of pixels. The attention is later calculated using the key matrix and the query matrix. The key matrix K^T and the query matrix Q are multiplied, and each element is divided by the square root of d_k . A softmax layer is then applied following the given equation to produce the initial attention map $S \in \mathbb{R}^{N \times N}$. The similarity between the i^{th} key and the j^{th} query is represented by the element S_{ji} in the similarity matrix as given in equation 4.3. Here, S_{ji} stands for the i^{th} position impact on the j^{th} position or the influence of position j on the feature tensors at position i. As a result, the

dependency linkage between any two components of the wider context is expressed.

$$s_{ji} = \frac{\exp(K_i \cdot Q_j)}{\sum_{i=1}^N \exp(K_i \cdot Q_j)} \quad (4.3)$$

Similarly, A is provided into another 1×1 convolutional layer to generate a new feature tensor $V \in \mathbb{R}^{C \times N}$. The matrix multiplication operation is performed between V and the transpose of S and reshape the result to $\mathbb{R}^{C \times H \times W}$. In the last stage, the scale parameter α is multiplied, and an element-wise sum operation with feature vector V is performed to obtain the final output $E \in \mathbb{R}^{C \times H \times W}$ as shown in equation 4.4 that is the SA feature map as follows:

$$E_j = \alpha \sum_{i=1}^N (s_{ji} \cdot V_i) + A_j \quad (4.4)$$

α is initially set to 0 in this case and subsequently evolves to assign increasing weight. As observed, the final SA feature map E is a weighted average of features from all positions and the original features.

4.2.2.2 Dual Attention

An attention layer is added in the framework to improve its ability in our proposed encoder-decoder architecture to focus on valuable areas of the inputs. The network is designed with both channel and spatial attention modules to achieve robustness in the CD process (Woo et al. 2018). The network can focus on the feature maps with the help of channel attention that are more significant for developing the final change map. The soft selection of feature maps can be regarded as channel-wise attention. On the other hand, spatial attention allows the networks to highlight the essential spatial parts of the feature maps. The upsampling of the attention maps is performed to have the identical field of view as the inputs. The attention maps assist the network in decomposing the distorted portions of the image and focusing on the image's necessary components independently.

The multiscale information generated by the high-level and low-level stages is combined using a spatial attention block (SAB) to learn more spatially representative fea-

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

tures. The channel attention block (CAB) redistributes channel feature responses to strengthen the critical channel information while suppressing irrelevant channels. Channel attention is globally applied from a spatial viewpoint, while spatial attention works locally. The intermediate feature maps generated from the encoder represented as $F \in R^{C \times H \times W}$ provided as input, and the channel and spatial attention modules produce a 1D channel attention map as $M_c \in R^{C \times 1 \times 1}$ and 2D spatial attention map as $M_s \in R^{1 \times H \times W}$ as explained in the following sections. The entire process of the channel and spatial attention is summarized as follows as given in equation 4.5:

$$\begin{aligned} F' &= M_c(F) \otimes F, \\ F'' &= M_s(F') \otimes F' \end{aligned} \tag{4.5}$$

where \otimes represents element-wise multiplication. During the multiplication process, the element-wise broadcasting (copying) of the attention values in a manner by broadcasting channel attention values along the spatial dimension and spatial attention values along the channel dimension. F'' is denoted as the final refined spatial attention feature map.

Channel attention The channel attention method concentrates on “what” is significant given an input image since each feature map channel is considered a feature detector. The input feature map’s spatial dimension is reduced to compute the channel attention effectively. The channel attention module fine-tunes the integrated features through channel attention maps (CAM). Each channel’s significance is encoded via CAM, whose weights are dynamically recalibrated across the network. Average pooling is adopted for gathering spatial information. The max-pooling operation collects crucial information on distinguishing object properties to infer precise channel-wise attention. As a result, simultaneously average-pooled and max-pooled features are used. Channels appropriate to the CD process are highlighted, while irrelevant channels are concealed by multiplying features f_k ($k=1 \dots K$) with appropriate attention map weights w_k ($k=1 \dots K$). As a result, the channel attention module determines which channel to learn from the combined heterogeneous features. Estimation of the CAM (M_c^F) is described as in equation 4.6:

$$\begin{aligned}
M_c^F &= \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\
&= \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c)))
\end{aligned} \tag{4.6}$$

where F represents input feature maps. Finer attention inference is achievable when both average-pooled and max-pooled features are used simultaneously. The spatial details from the input feature maps are first compressed along the spatial axis using an average pooling (AvgPool) and a max-pooling (MaxPool) operation. After the pooling operations, feature maps will be compressed into two vectors of size $C \times 1 \times 1$, since there are C feature maps of $H \times W$. Additionally, as both of the aggregated channel features belong to the same semantic embedding space while using both pooling methods, a shared MLP is utilized for attention inference to save the parameters.

The vectors are then transmitted to a shared multi-layer perception (MLP) layer. It is argued that the maximum-pooled feature, which encodes the fundamental component’s magnitude, can compensate for average-pooled ones, which more subtly encode global statistics. Therefore, both functionalities were combined simultaneously and used a shared network like MLP for those tasks. The shared MLP’s outputs are blended using element-wise summing. Finally, each channel’s attention weights are assigned using a sigmoid function represented as σ . $W_0 \in \mathbb{R}^{C/r \times C}$ and $W_1 \in \mathbb{R}^{C \times C/r}$ are the MLP weights which are being shared for both inputs and the activation function ReLU is followed by W_0 .

Spatial attention The spatial attention is computed using the channel-wise refined features. A 2D descriptor is initially calculated that encodes channel information at each pixel over all spatial locations to produce a 2D spatial attention map. The raw attention map is then obtained by applying one convolution layer to the 2D descriptor. The inter-spatial relationship among the features generates the spatial attention maps. This module focuses on where is an informative part that is complementary to the channel attention. The spatial attention module refines the channel-by-channel sophisticated features across the spatial context after the channel attention process. In spatial attention maps, each pixel location’s value is recorded the same way as in the feature maps. The spatial attention module is trained to improve its ability to recalibrate the weights

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

of each pixel position dynamically. It receives feedback from ground truth maps iteratively during the network training phase. Eventually, it generates spatial attention maps (SAM), which attribute higher and lower relevance to locations of changing and unchanged pixels accordingly. Notably, a SAM $w_c(c = 1 \dots N)$ is initially calculated with the spatial attention component. Here N signifies the absolute number of pixels of each feature map, and c is the c^{th} pixel location in the generated map. The spatial-wise refinement is then accomplished by element-wise multiplication of each feature map of high level named f in the collection of feature map $f_k(k = 1 \dots K)$ with w_c .

The processed and developed feature maps can be designated as $f_k \otimes w_c$. Pixel changes are highlighted by multiplying with a set of larger weights after the spatial-wise refinement, whereas unaffected pixels(unchanged) are concealed by multiplying with lower weights. As a result, the model can quickly approach the regions being changed and enhance the combined features in the spatial dimension. Employing average and maximum pooling to implement channel pooling over the channel axis is an effective method for generating the 2D descriptor. A high kernel size selection suggests that selecting spatially significant regions requires a large receptive field. Given this, spatial attention was computed using the channel pooling method and a convolution layer with a large kernel size. The spatial attention module uses the average- and maximum-pooled features over the channel axis with a convolution kernel size of 7. The architecture of the channel and spatial attention module incorporated with the proposed approach is provided in figure 4.5. The Spatial attention map M_s^F is calculated as follows using equation 4.7::

$$\begin{aligned} M_s^F &= \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \\ &= \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])), \end{aligned} \quad (4.7)$$

Where $f^{7 \times 7}$ means a convolution operation with the filter size of 7×7 and signifies the concatenate operation. The AvgPool and MaxPool layers on the channel axis initially perform the average and max pooling of the input feature maps F. Then, features of the AvgPool and MaxPool are convolved and concatenated by a convolutional layer $f^{7 \times 7}$. σ is utilized to construct the final M_s^F . The sigmoid function σ normalizes the

final attention map. This attention-based fusion of features can automatically investigate the implication of individual raw image features and image difference features to solve the heterogeneous problem across the channel dimension efficiently. It can also further recalibrate the significance of each pixel location across the spatial dimension to approach interested areas rapidly.

4.2.3 Model Parameters

In any DL approach, the model is trained with several parameters classified as trainable and non-trainable. The term “trainable parameters” refers to variables whose values can be changed based on their gradient (the derivative of the lost/error/cost in relation to the parameter). Those parameters whose values are not optimized as per their gradient are “non-trainable parameters”. The total number of parameters that are being utilized for training the model is 87,689,930 which consist of 87,678,918 trainable and 11,012 non-trainable parameters for the effective generation of the final change map.

Updating of parameters Based on the training dataset, the STEDSAN and STEDDAN architecture is trained. There are numerous parameters that need to be tuned, including size of input patch, dropout rate, quantity of neurons, layers, rate of dilation, size of mini-batch, primary learning rate, and the amount of epochs. In this study, the configuration of the network’s parameters begins with an initial value. The network is then evaluated using accuracy indicators, such as computing the network’s overall accuracy with test data. Finally, new values are assigned, and the parameters are updated.

Stopping condition, model evaluation and optimization The parameters of the model are adjusted iteratively using optimization. In this study, CNN parameters are modified using an Adam Optimizer (Kingma and Ba 2014). The tests performed on the dataset are used to determine the best values for these parameters. The Adam optimizer uses little memory and is computationally effective. To do this, three sorts of sample datasets were created: (1) training data, (2) testing data, and (3) validation data. The training and validation datasets are used to train and assess the network.

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

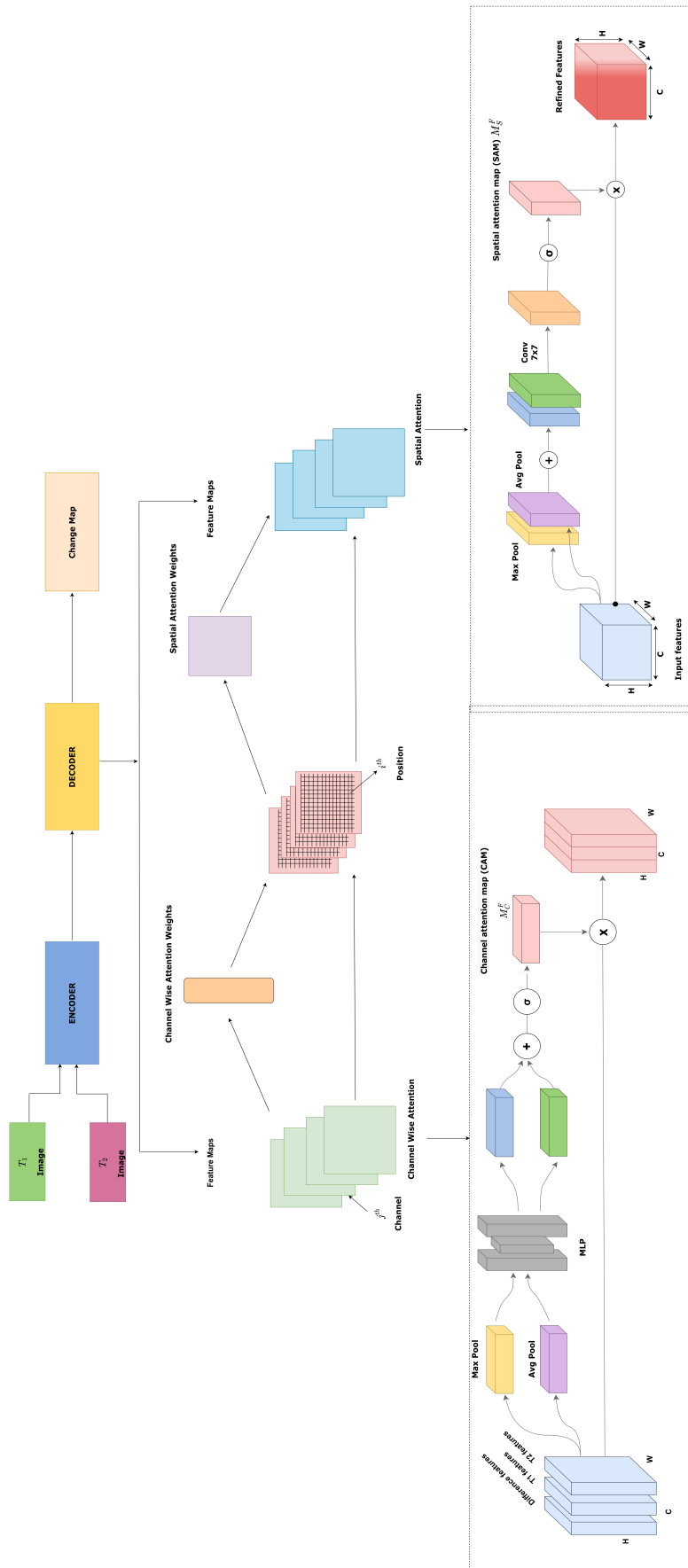


Figure 4.5: Channel and Spatial Attention Modules.

The network's performance on the validation data determines the appropriate weight to use when choosing the parameters. Up to the completion of all epochs, the learning phase remains active. Additionally, the loss function uses the validation dataset to generate the loss value. The network's compliance with the expected output and ground-truth label is used to calculate the loss function. To achieve this, a network is trained using training data. Its error is determined by evaluating the network using a validation dataset. Finally, test data is used to evaluate the network's performance. Until the stopping condition is met, this process is continued. The stopping conditions are established based on obtaining a satisfactory level of accuracy.

4.3 STUDY AREA AND DATA DESCRIPTION

The main aim of the analysis was to select a suitable area for the study's purpose and to obtain the LULC maps of a particular site for the assessment of the data. LULC maps help users interpret the current landscape. Annual LULC data stored in national spatial databases will allow for monitoring agricultural ecosystems, forest conversions, surface water bodies, and other temporal dynamics. The requirement of high-quality spatiotemporal data of the chosen area, classified according to the type of land, for at least ten years was needed to analyze the land cover patterns and changes efficiently and generate a reliable change map for further improving the investigation. The dataset consist of coastal region called Goa and a district from Karnataka, India, known as Dakshina Kannada (formerly South Canara), used to perform the analysis.

4.3.1 Dakshina Kannada

Dakshina Kannada is a district in karnataka India. It consists of 4,559 square kilometre. Its geographical location coordinates are 12.8438 N Latitude and 12.8438 N longitude, with a maximum elevation of approximately 1,115 m above mean sea level. The Arabian Sea shelters it on the west and the western ghats on the east. The climate in this study area is moderately mild and humid in the winter season and dry and hot in the summer season, and receives abundant rainfall during the monsoon. Mangaluru city is the district headquarters, as evident through the vast areas of urbanization. The major rivers of this region are Gurupura, Kumaradhara, Netravathi, Nandini or Pavanje,

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

Phalguni, Shambhavi, and Payaswini join the Arabian sea. The average annual rainfall in this district is 4,030 millimeters. Dakshina Kannada district had a population of 2,083,625, according to the 2011 census of India. The study area location is shown in figure . The location of the study area is as shown in figure 4.6 with the LULC maps from 2005-06 and 2017-18 in figure 4.7(a) and 4.7(b). Figure 4.7 (c) shows the descriptions of the classes and color codes for each of the class labels.



Figure 4.6: Location Map of the Study Area (Dakshina Kannada.)
 source:https://stategisportal.nic.in/stategisportal/Karnataka_BharatMaps/map.aspx

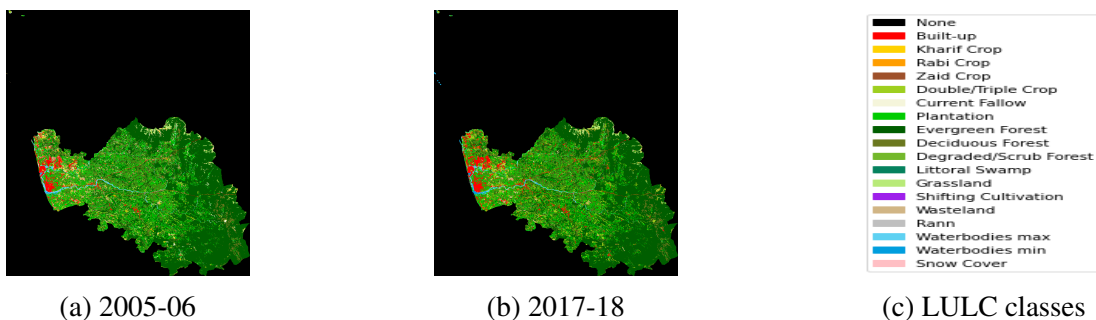


Figure 4.7: LULC Data and Description of each LULC Class (Dakshina Kannada)

4.3.2 Goa

The second study area chosen for our research work is the state of Goa, one of India's smallest states on a coastal belt, covering an area of 3702 square km. Its geographical location coordinates are 15.2993° N Latitude and 74.1240° E Longitude with a maximum elevation of approximately 56 m, above mean sea level (AMSL). The western Ghats geographically separates Goa from the Deccan highlands within the Konkan region on the southwestern coast of India. Goa borders the state of Maharashtra in the northeast and the southeast, Karnataka. It consists of sandy beaches, bays, and rocks covering an area of 160 km falling on the west coast that borders the Arabian sea.



Figure 4.8: Location Map of the Study Area (Goa).

source:https://stategisportal.nic.in/stategisportal/Goa_BharatMaps/map.aspx

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

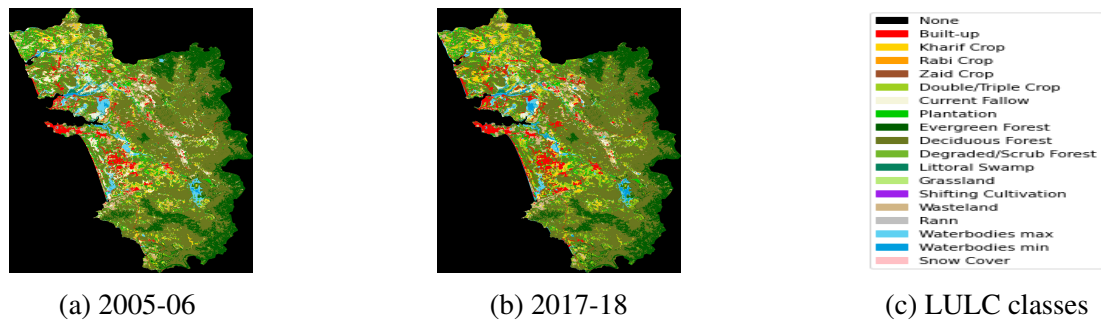


Figure 4.9: LULC Data and Description of each LULC Class (Goa).

The lowlands consist of mango plantations, rice fields, and deciduous forests. The Sahyadri mountains form the natural border between India and Goa, part of the Western Ghats, and a sequence of ridges separates the malabar coast from the deccan plateau. Goa being the tourist destination is evident through the vast areas of urbanization. The major rivers of this region are Terekhol, Mandovi, Baga, Zuari, Colval, Saleri, Mandre, Harmal, Sal, Talpona, and Galjibag that all join the Arabian sea. The climate in this study area is relatively mild and humid in the winter month and dry and hot in the summer, and receives abundant rainfall during the monsoon. The average annual rainfall in this district is 330 centimeters annually. The location of the study area is as shown in figure 4.8 with the LULC maps from 2005-06 and 2017-18 in figure 4.9(a) and 4.9(b). Figure 4.9 (c) shows the descriptions of the classes and color codes for each of the class labels.

4.3.3 Data Collection

The NRSC, Hyderabad, part of the ISRO, provided LULC data for the study. This web-based utility, known as Bhuvan, allows users to explore a set of map-based content and obtain spatial data. Bhuvan generates maps for land resource analysis at the end of each year using multitemporal Indian Remote Sensing Satellite(IRS) Resourcesat-1/2 (P6) AWiFS for time series data. Following a hybrid method, the data is classified into various LULC categories with decision tree (DT), Supervised Maximum Likelihood Classifier (MXL), or both. The number of temporal datasets available during the season, the absence of cloud/haze, the terrain's complexity, and the precision of temporal registration are all factors in choosing a classifier for the generation and classification of

Table 4.1: Details of the Dataset (Goa and Dakshina Kannada Region).

| | |
|------------------------------|-------------------------------------|
| Image Type | LULC Maps |
| Date and number of data used | 2005-06 to 2017-18 (Total 13 maps) |
| LULC classes | 18 (15 active) |
| Format | TIF |
| Resolution | 56 m |
| Sensor | Advance Wide Field Sensor (AWiFs) |
| Scale | 1:250k |
| Source | NRSC, ISRO Hyderabad India |

LULC maps. Using AWiFS data, the fifth cycle built on the categorization and mapping processes employed in earlier cycles are detailed in (NRSC 2010). The details of the dataset are provided in Table 4.1.

4.4 RESULTS AND DISCUSSIONS

4.4.1 Experimental Setup

The proposed method is evaluated by experiments in universally available LULC image datasets. The entire set of experiments are executed on a computer with Core Intel i7-8700k CPU, 16-GB RAM, and NVIDIA GeForce GTX 1090Ti GPU. using python3 programming implementation. Two remote sensing datasets in form of LULC data as described above are used to evaluate the performance of the proposed method and SOTA techniques.

4.4.2 Performance Evaluation Parameters

Overall accuracy (OA), Precision (P), recall(R), F1-score, and mean IOU are five performance metrics used to quantify the proposed method's performance .

1. Overall Accuracy-The overall accuracy suggests how accurately the model detects both the changed as well as unchanged pixels as in equation 4.8

$$OA = \frac{TP + TN}{TP + FP + TN + FN} \quad (4.8)$$

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

2. Precision-The ratio of true positives to the sum of true positives and false positives is precision (P) as in equation 4.9.

$$Precision = \frac{TP}{TP + FP} \quad (4.9)$$

3. Recall-True positives over the sum of true positives and false negatives are used to compute recall (R) values as in equation 4.10. Precision and recall must be higher in an accurate CD architecture.

$$Recall = \frac{TP}{TP + FN} \quad (4.10)$$

4. F1-Score- In binary classification, the F1 score is defined as the harmonic mean of precision and recall as in equation 4.11. It's a statistical metric for evaluating performance.

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (4.11)$$

5. Mean IOU- The percentage of overlap between the predicted image and the corresponding ground truth image is provided by IOU. IOU values are used to measure the average intersection of the predicted image of each class with the labeled picture, and these values are produced using the equation 4.12.

$$MeanIOU = \frac{1}{N_c} \sum_{i=1}^{N_c} \frac{C_{ii}}{T_i + P_i + C_{ii}} \quad (4.12)$$

where N_c is the total number of classes present, C_{ii} is the class-wise predicted pixels, T_i represents the total number of corresponding class-wise pixels in the ground truth, and P_i , is the total number of pixels whose predictions are i.

4.4.3 Comparative Methods

Based on some of the parameters discussed in the evaluation matrix section, the comparison was performed. Several recent CD methods being proposed were chosen for comparison. Existing proposed methods use existing pre-trained convolutional models

like VGG16, Alexnet, VGGNet to train and test the network. One of the most straightforward methods to address class imbalance is to assign a weight to each class. The weight class is automatically defined based on inversely adjusting weights according to class frequencies. The Glorot normal initializer (Glorot and Bengio 2010) is used to initialize all weights. The back-propagation method with the Adam optimizer is used to train the weights. The training, validation, and test data in this study are chosen from 65%, 15%, and 20% of the sample datasets, respectively.

1. IFN (Image Fusion Network) -(Zhang et al. 2020) uses the FC-Siam-Conc design for its image fusion network. The siamese encoder is implemented using VGG-16, and the decoder employs a UNet-based decoder to preserve details. In the decoder, IFN extra introduces channel attention and spatial attention.
2. H-SALENet -It is an encoder and decoder network suggested in (Cheng et al. 2021). H-SALENet uses a deep convolutional module in conjunction with a hierarchical and long-range context augmentation module (HLAM) in the encoder to extract the deep features of bitemporal images. Deep convolution and 2D transformer structured multi-head SA learning improve the representation capability of multi-level and long-range dependent change features. The decoder includes a Laplacian pyramid expansion module (LPEM) that captures change information at many scales while reducing high-frequency information loss, reducing the influence of deep feature resampling on the change map.
3. ADS NET-A dual-stream structure with attention is used in (Wang et al. 2021) to create an encoding–decoding scheme using only convolutions and max-pooling. The encoding stage extracts various level features from bitemporal single date images. The decoding stage adds feature maps from various layers into a deep supervision network with discrete branches to recreate the change map.
4. SIAMGL-This method has effectively extracted representative features from bitemporal HSR RS images using the Siamese architecture with shared parameters in (Zhu et al. 2022). The global hierarchical (G-H) sampling technique was developed to address the issue of an uneven training sample caused by a shortage of

samples. Furthermore, a binary change mask is inserted between the encoder and decoder to limit the influence of the no-change regional background on the regional change foreground, further improving the accuracy of the proposed framework.

4.4.4 Result Analysis

4.4.4.1 Experiments on the Dakshina Kannada Dataset

The proposed model generated the binary change maps for the given two bi-temporal LULC maps for 2005-06 and 2017-18 to determine the LULC class-wise binary changes with black pixels representing no change and white pixels as change. The change maps generated by various SOTA methods are being compared with the proposed approach, as shown in figure 4.10 (a) IFN , 4.10(b) H-SALENet, 4.10(c) ADS-NET, 4.10(d)SIAM-GL ,4.10(e) STEDDAN(Proposed) and 4.10(f) Ground Truth. Visual interpretation and detailed analysis are used to make the comparison. STEDDAN effectively returns large area change regions with complete boundaries and strong internal compactness when large area change areas are found amongst the LULC class. STEDDAN detects the image noises compared to other methods, and these pixels are classified as unaltered areas in the change maps. Visual analysis of STEDDAN change maps reveals outstanding performance, and results match ground truth maps for small and large area LULCCD tasks. The quantitative evaluation is shown in Table 4.2. STEDDAN has achieved the best outcomes, with the highest F1 (82.90%) ,MIOU (74.1%) and OA (94.11 %). SiamGL has gained a very close margin to the proposed method.

Table 4.2: Quantitative Performances of Different Methods on the Dakshina Kannada Dataset.

| Method | F1(%) | MIOU(%) | OA(%) |
|-----------|-------|---------|-------|
| IFN | 69.64 | 51.6 | 88.42 |
| H-SALENet | 72.58 | 64.3 | 92.38 |
| ADS-NET | 71.89 | 66.8 | 92.76 |
| Siam-GL | 78.46 | 71.4 | 93.54 |
| Proposed | 82.90 | 74.1 | 94.11 |

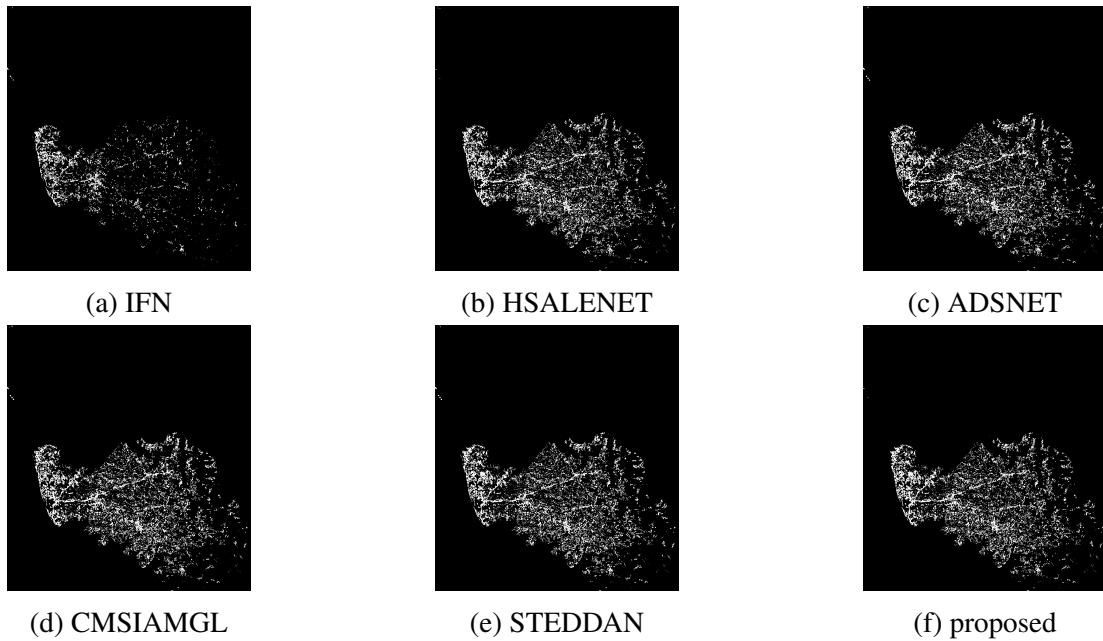


Figure 4.10: Change Maps generated for Dakshina Kannada dataset.

4.4.4.2 LULC Change Analysis for Dakshina Kannada Region

The following analysis shows the different categories of changes in each LULC class in terms of area and percentage of changes for the dakshina Kannada region from 2005-06 to 2017-18. All the changes are shown in Table 4.3, Table 4.4 and Table 4.5. The final multiclass change map showcasing the conversion from one LULC class to another LULC class changes from the year 2005-06 to 2017-18 for the Dakshina Kannada region is shown in figure 4.11, with overall changes shown in each LULC class for the assessment period in each figure separately. The model determined changes in only 15 LULC classes showcasing there was no impact of LULC types like shifting cultivation, rann, and snow cover for the considered area.

- The BU is increased from 135.42 sq km to 205.82 sq km showing an increase in the area with 1.544%. Figure 4.12a displays the area-wise changes showcasing that the ZC, LS, and GL are not converted to the BU class.
- The KC is increased by 0.851% with an area changes from 33.55 sq km to 72.38 sq km. Figure 4.12b displays the area-wise changes showcasing that ZC is not converted to the KC class.

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

- The RC is decreased by 0.158% with an area changes from 14.848 sq km to 7.624 sq km. Figure 4.13a displays the area-wise changes showcasing that ZC,LS, Wmin is not converted to the RC class.
- The ZC is being increased by 0.005% with an area changes from 0 sq km to 0.255 sq km. Figure 4.13b displays the area-wise changes showcasing that D/S F,LS ,Wmax and Wmin are not converted to the ZC class.
- The D/T C is being increased by 0.873% with an area changes from 153.24 sq km to 193.07 sq km. Figure 4.14a displays the area-wise changes showcasing that only ZC is not converted to the D/T C class.
- The CF is being decreased by 1.359% with an area changes from 67.004 sq km to 5.004 sq km. Figure 4.14b displays the area-wise changes showcasing that ZC and LS are not converted to the CF class.
- PL is decreased by 0.008% with an area changes from 928.78sq km to 928.41 sq km. Figure 4.15a displays the area-wise changes showcasing that only ZC is not converted to the PL class.
- The EF is decreased by 0.066% with an area changes from 2441.63 sq km to 2438.58 sq km. Figure 4.15b displays the area-wise changes showcasing that ZC is not converted to the EF class.
- The DF is decreased by 0.062% with an area changes from 555.30 sq km to 552.45 sq km.Only ZC is not converted to DF. Figure 4.16a displays the area-wise changes showcasing that only ZC is not converted to the DF class.
- The D/S F is increased by 0.002% with an area changes from 4.674 sq km to 4.772 sq km. Figure 4.16b displays the area-wise changes showcasing that ZC,LS and GL are not converted to the D/S F class.
- The LS is being increased by 0.0004% with an area changes from 0.0464 sq km to 0.0685 sq km .LULC classes like KC,RC,ZC,CF,GL,WL,Wmin are not converted to LS. Figure 4.17a displays the area-wise changes showcasing that KC,RC,ZC,CF,GL,WL and Wmn are not converted to the LS class.

- The GL is decreased by 0.0083% with an area changes from 30.44 sq km to 30.04 sq km. Figure 4.17b displays the area-wise changes showcasing that ZC,D/S F,LS,Wmx and Wmn are not converted to the GL class.
- The WL is decreased by 1.705% with an area changes from 132.13 sq km to 54.37 sq km. Figure 4.18a displays the area-wise changes showcasing that ZC,LS are not converted to the WL class.
- The Wmx is decreased by 0.088% with an area changes from 58.76 sq km to 54.72 sq km. are not changed to Wmax. Figure 4.18b displays the area-wise changes showcasing that ZC and LS are not converted to the Wmx class.
- The Wmn is increased by 0.181% with an area changes from 3.12 sq km to 11.40 sq km. Figure 4.19a displays the area-wise changes showcasing that ZC,LS and GL are not converted to the Wmn class.

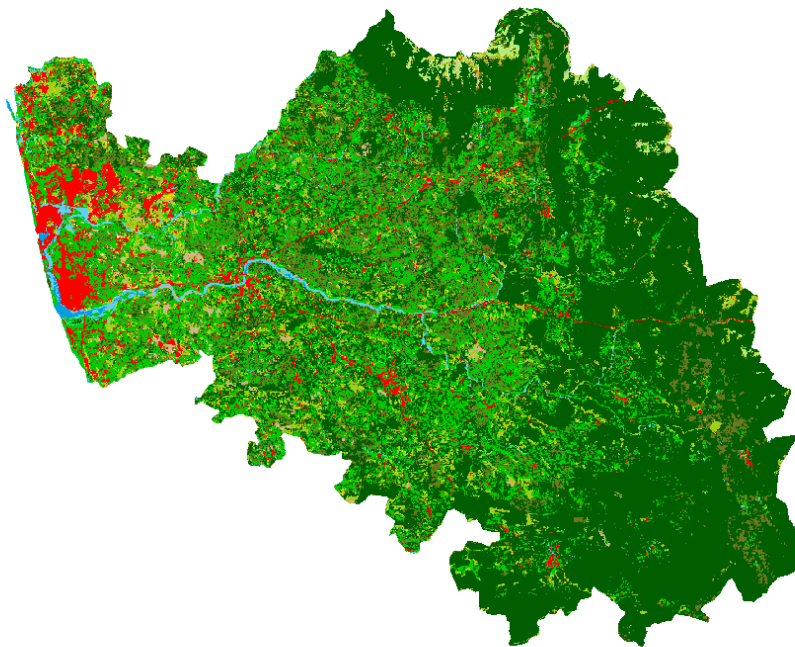
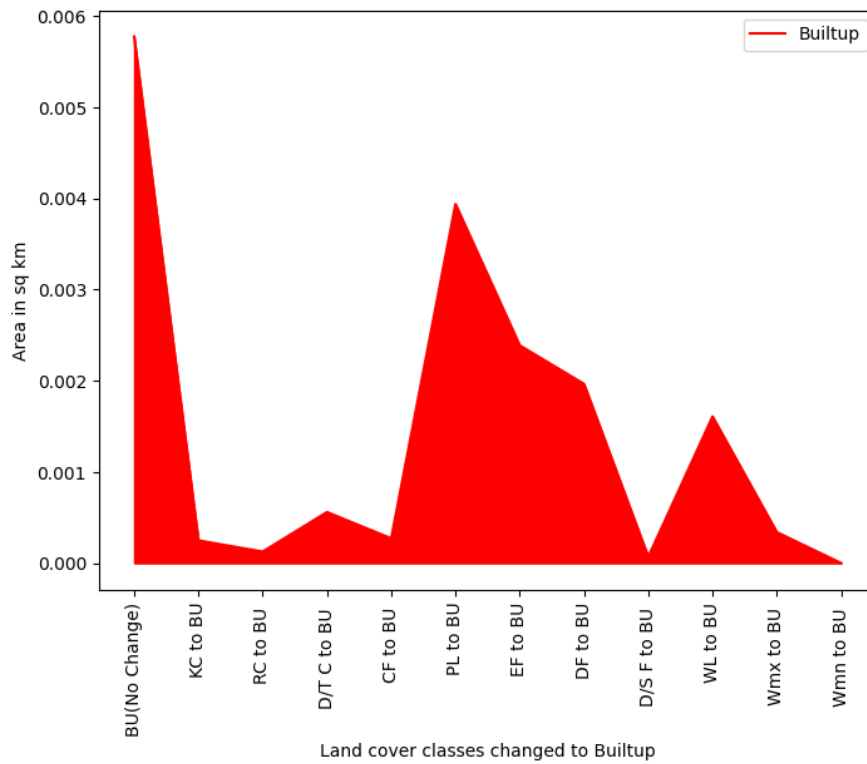
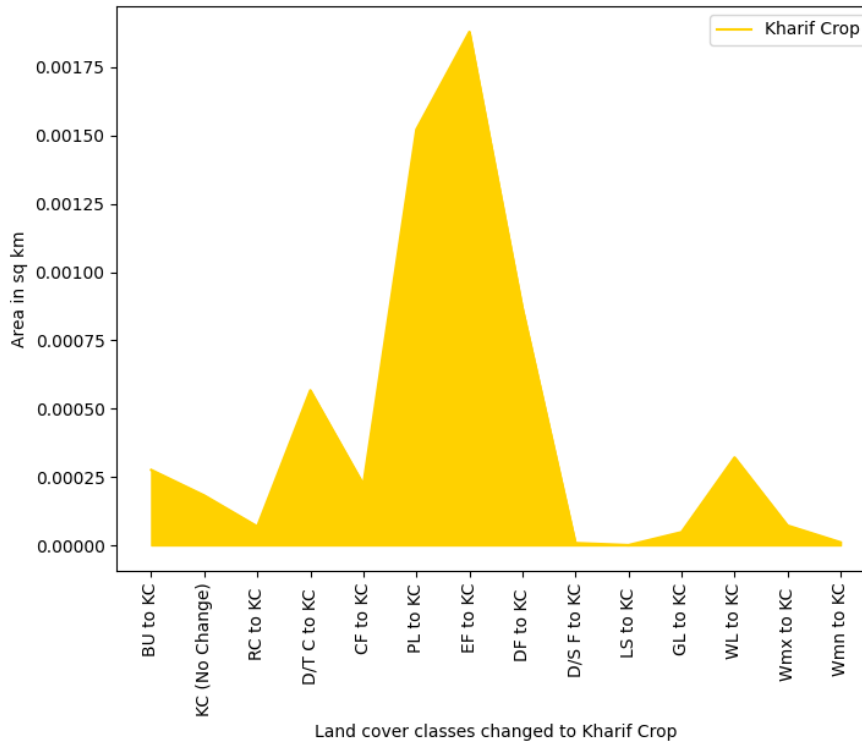


Figure 4.11: Multiclass Change Map of the Dakshina Kannada District.

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

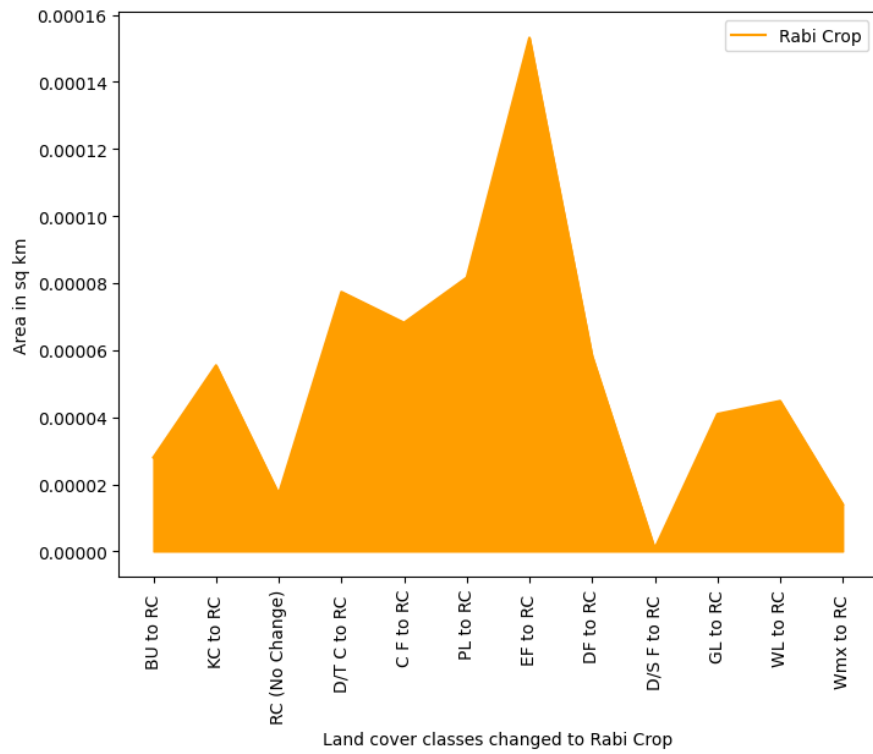


(a) Built-up

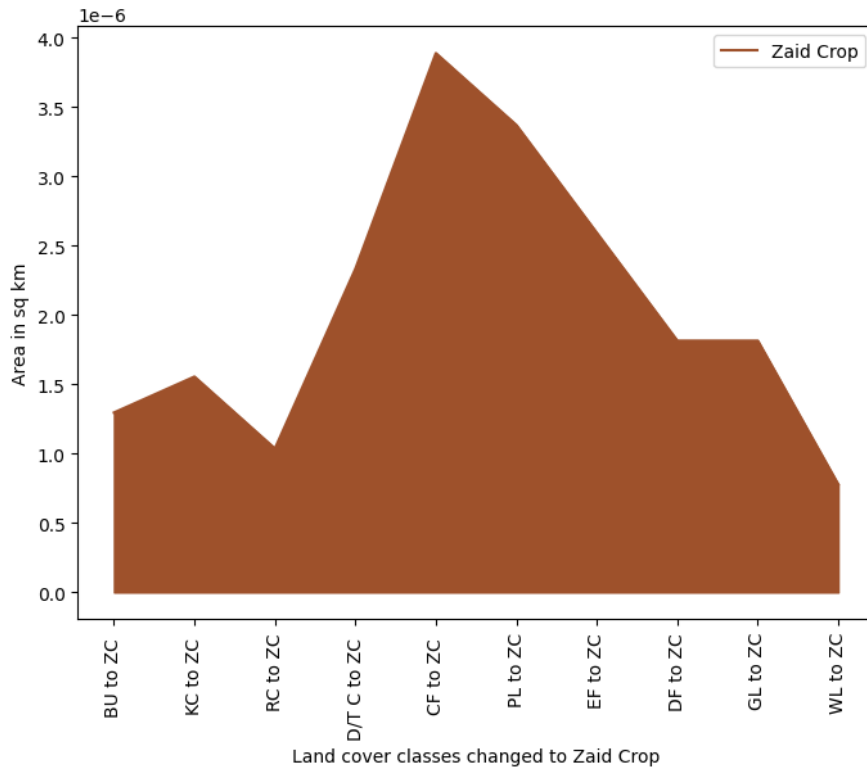


(b) Kharif Crop

Figure 4.12: Area Wise Changes in LULC Classes (Built-up and Kharif Crop).



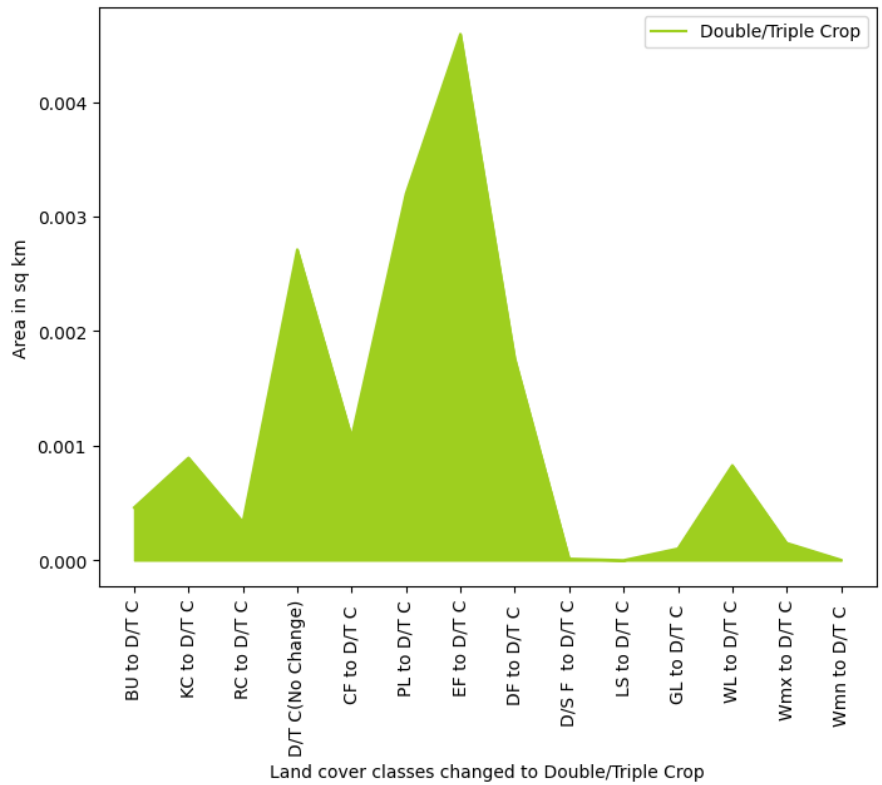
(a) Rabi Crop



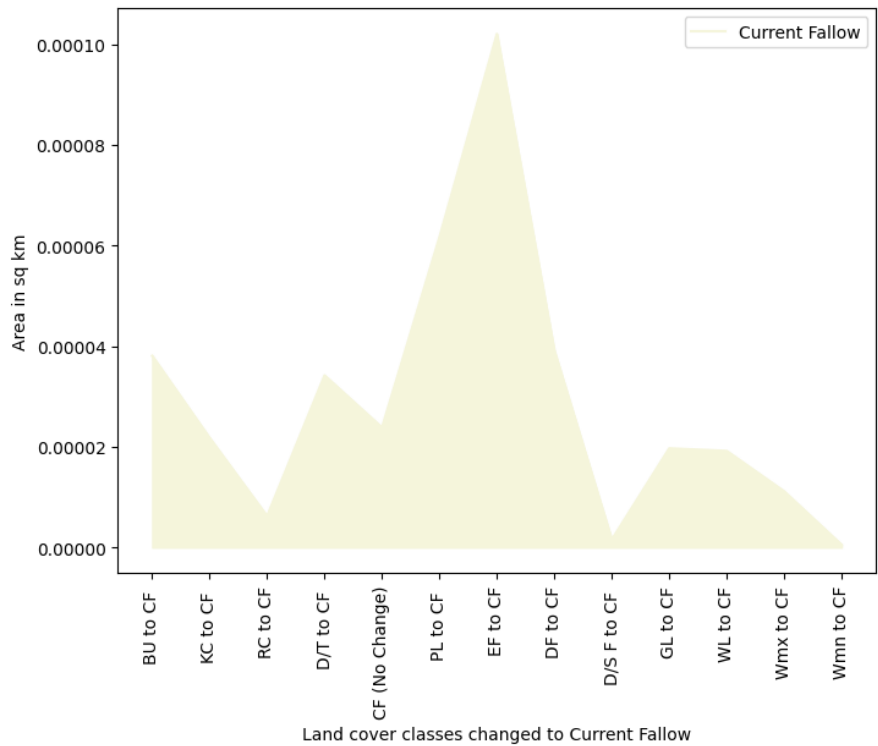
(b) Zaid Crop

Figure 4.13: Area Wise Changes in LULC Classes (Rabi Crop and Zaid Crop).

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

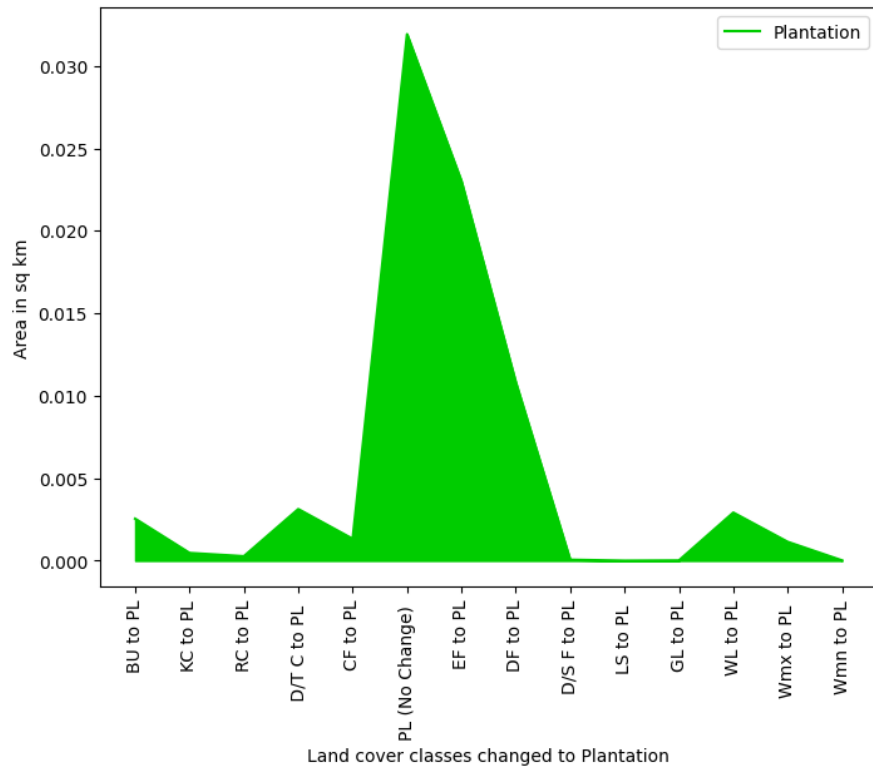


(a) Double/Triple Crop

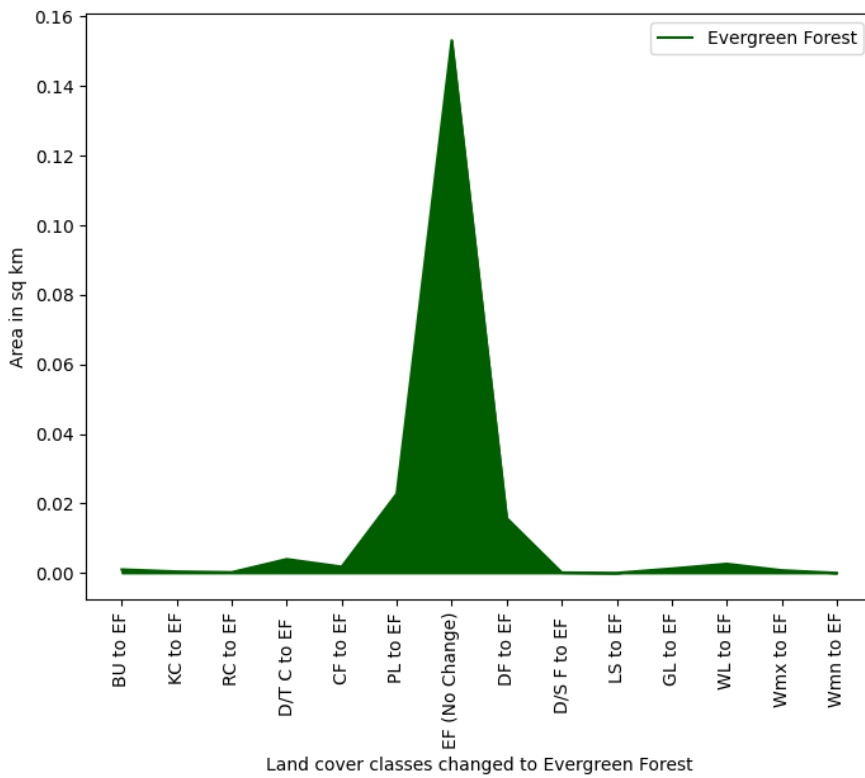


(b) Current Fallow

Figure 4.14: Area Wise Changes in LULC Classes (Double/Triple Crop and Current Fallow).



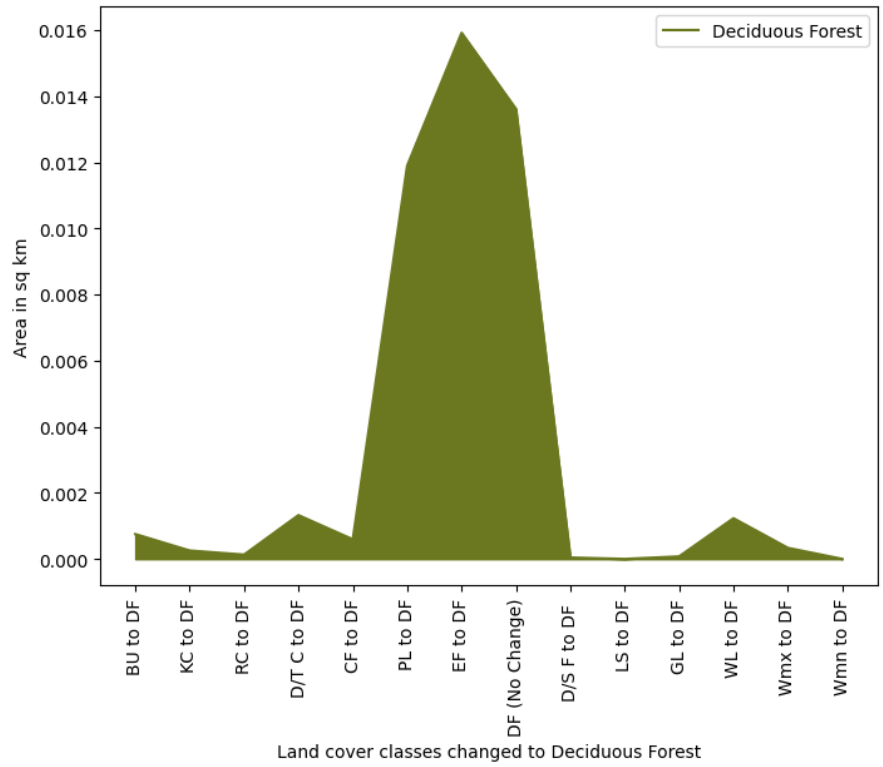
(a) Plantation



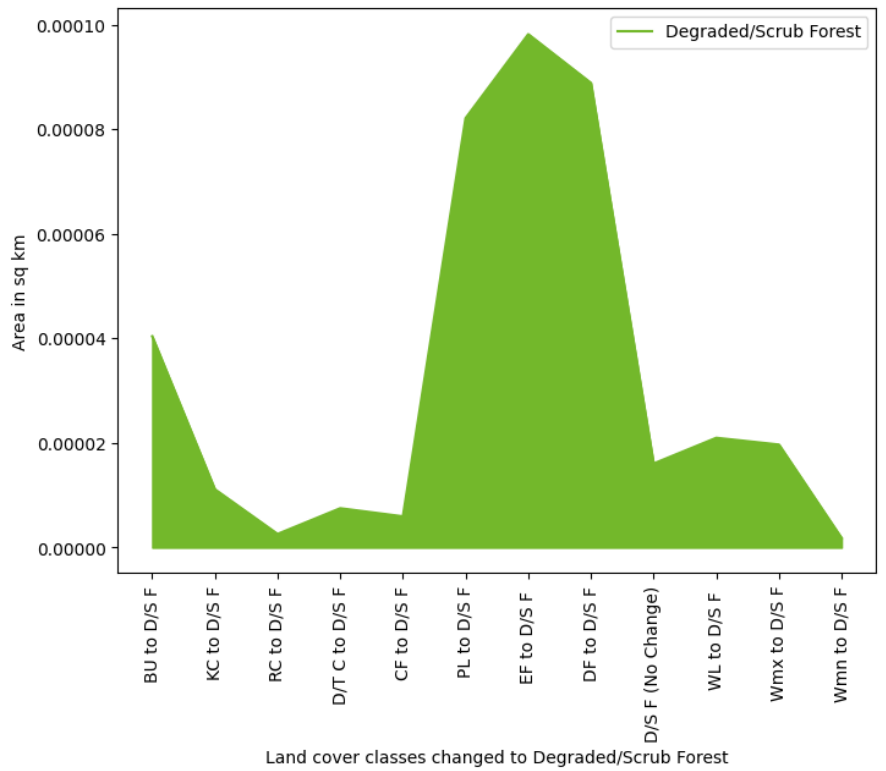
(b) Evergreen Forest

Figure 4.15: Area Wise Changes in LULC Classes (Plantation and Evergreen Forest).

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

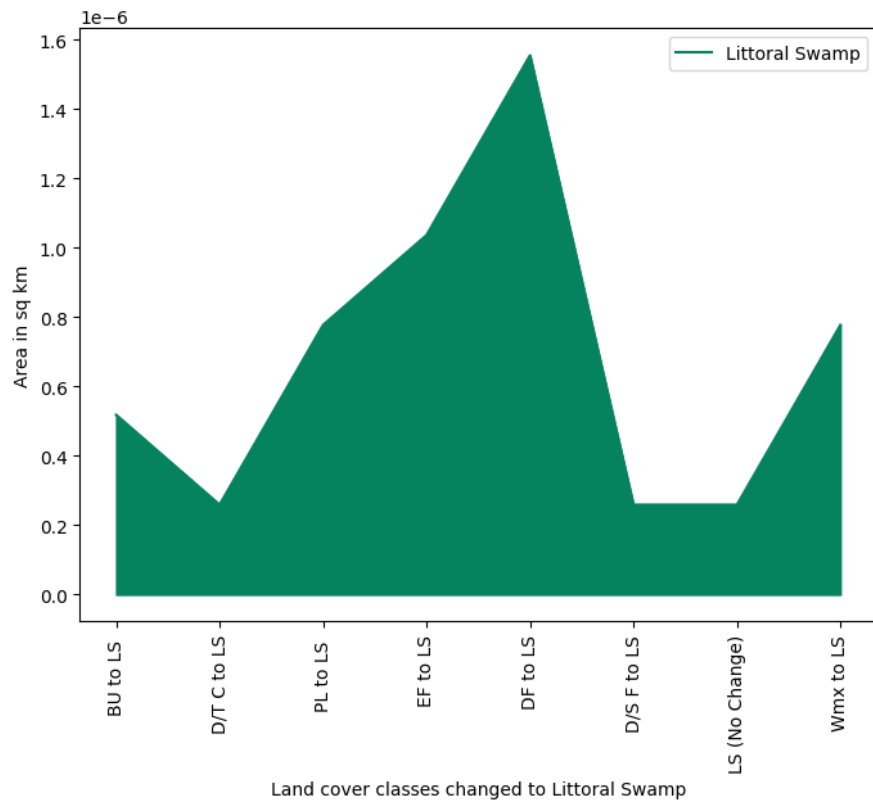


(a) Deciduous Forest

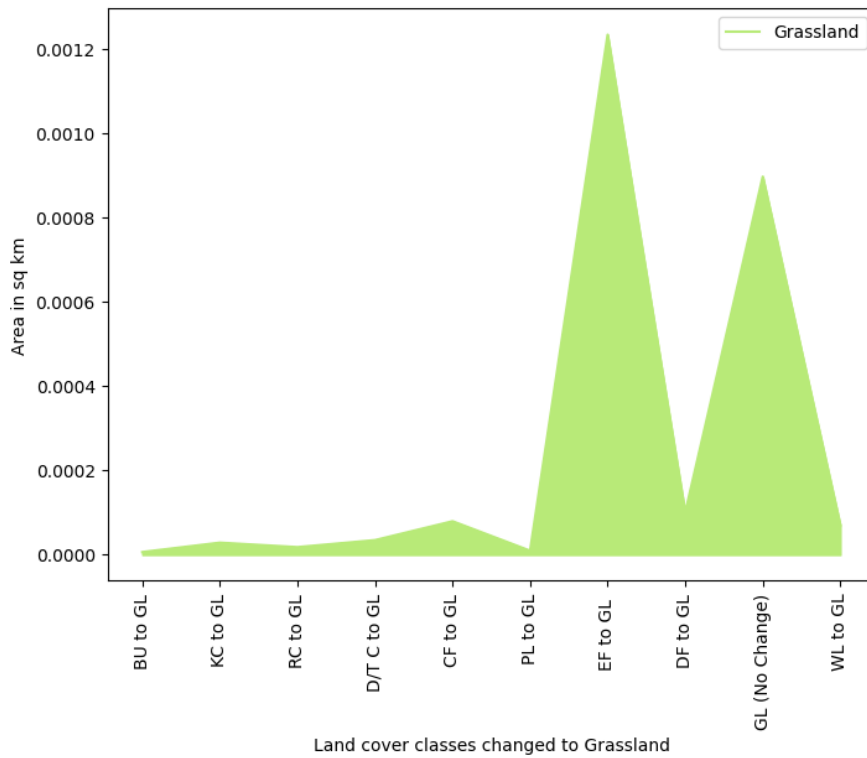


(b) Degraded/Scrub Forest

Figure 4.16: Area Wise Changes in LULC Classes (Deciduous Forest and Degraded/Scrub Forest).



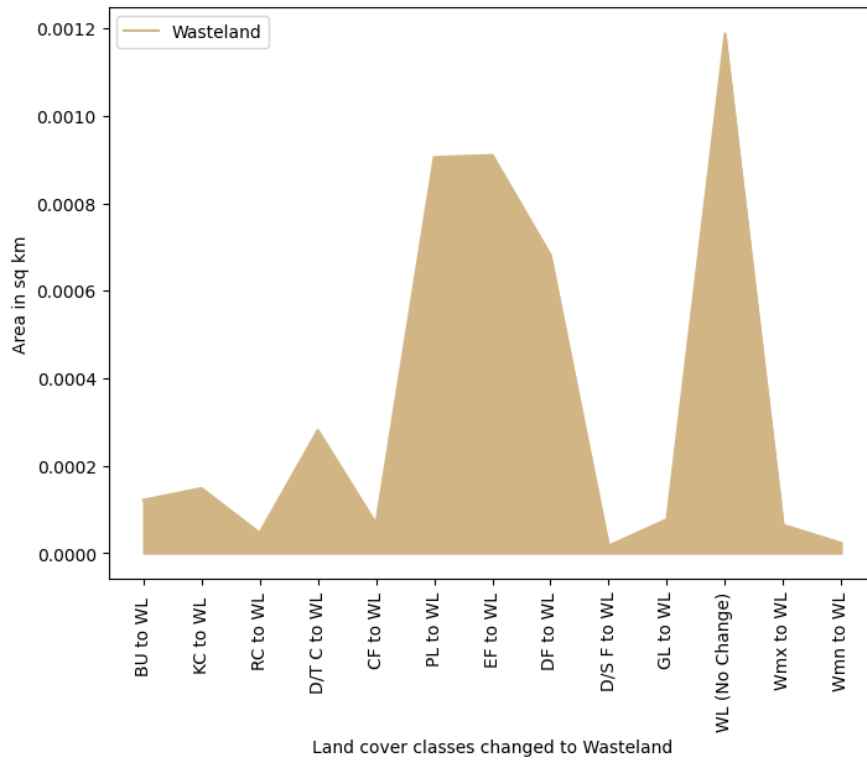
(a) Littoral Swamp



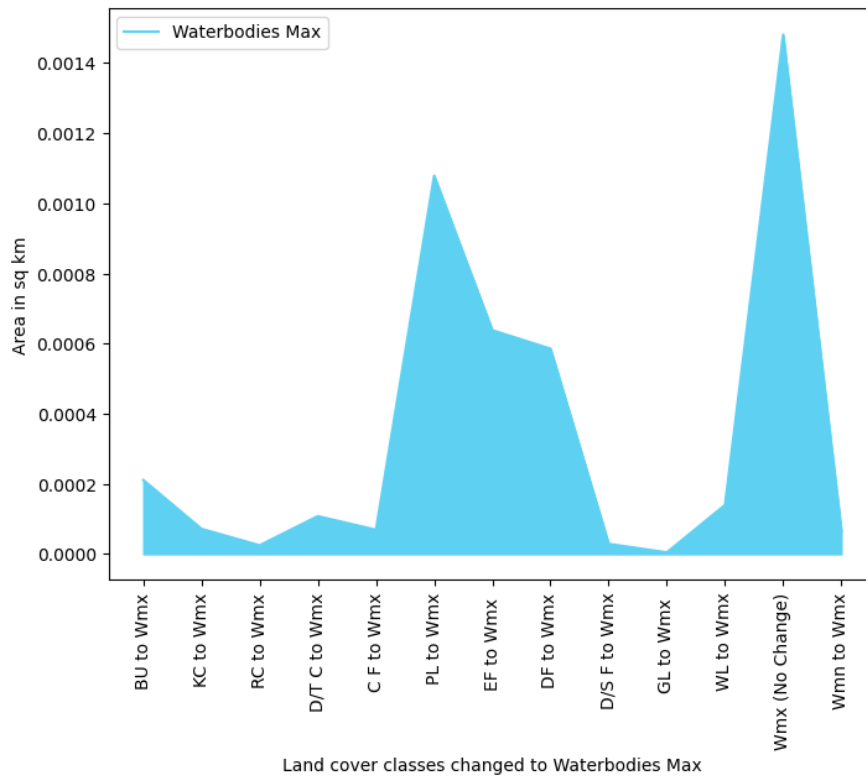
(b) Grassland

Figure 4.17: Area Wise Changes in LULC Classes (Littoral Swamp and Grassland).

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

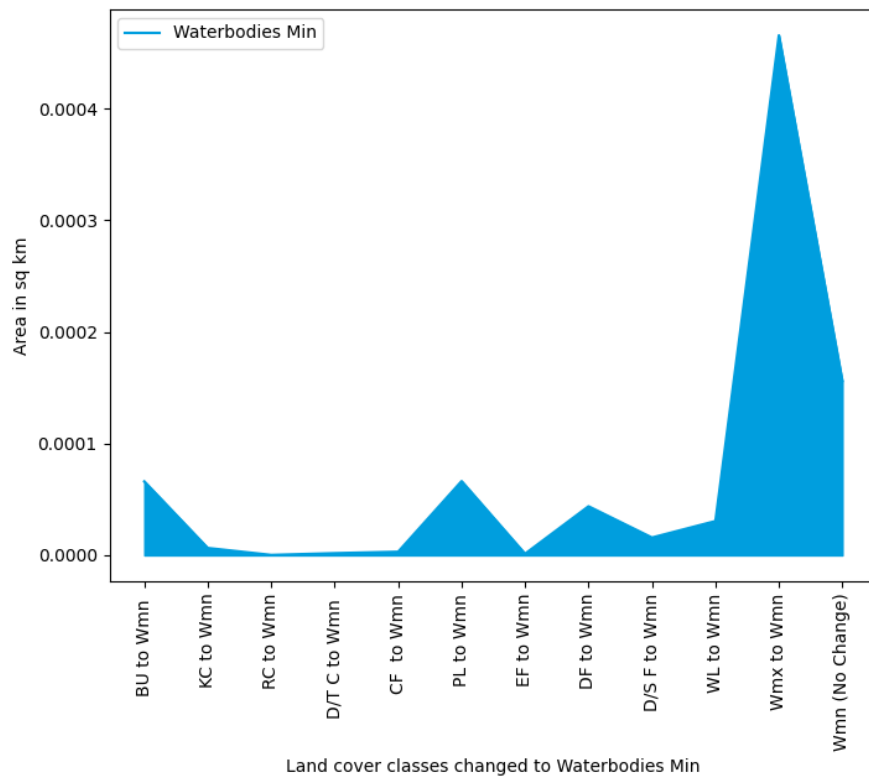


(a) Wasteland



(b) Waterbodies max

Figure 4.18: Area Wise Changes in LULC Classes (Wasteland and Waterbodies max).



(a) Waterbodies min

Figure 4.19: Area Wise Changes in LULC Classes (Waterbodies min).

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

Table 4.3: LULC Changes in terms of Area for Dakshina Kannada Region from 2005-2010.

| Year | 2005 | 2006 | 2006-2007 | 2007-2008 | 2008-2009 | 2009-2010 | | | |
|-------|---------|------------------------|-----------|------------------------|-----------|------------------------|---------|------------------------|-----------|
| Class | Area(%) | Area(km ²) | Area(%) | Area(km ²) | Area(%) | Area(km ²) | Area(%) | Area(km ²) | |
| BU | 2.9705 | 135.4259 | 2.7059 | 123.3629 | 3.4603 | 157.7564 | 3.6752 | 167.5535 | 198.5823 |
| KC | 0.7360 | 33.5577 | 0.7364 | 33.5732 | 0.8154 | 37.176351 | 0.2092 | 9.5401 | 8.9798 |
| RC | 0.3257 | 14.8488 | 0.3790 | 17.2818 | 0.2330 | 10.6235 | 0.3430 | 15.6382 | 1.9718 |
| ZC | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| D/T C | 3.3614 | 153.2464 | 3.3638 | 153.3578 | 3.2709 | 149.1232 | 3.8123 | 173.8033 | 189.5096 |
| CF | 1.4697 | 67.0041 | 1.4513 | 66.1683 | 1.5541 | 70.8548 | 1.5117 | 68.9201 | 68.2980 |
| P | 20.3725 | 928.7834 | 20.3725 | 928.7834 | 20.3725 | 928.7834 | 20.3725 | 928.7834 | 928.7834 |
| EF | 53.5564 | 2441.6381 | 53.5564 | 2441.6381 | 53.5564 | 2441.6381 | 53.5564 | 2441.6381 | 2441.6381 |
| DF | 12.1803 | 555.3020 | 12.1803 | 555.3020 | 12.1803 | 555.3020 | 12.1803 | 555.3020 | 555.3020 |
| D/S F | 0.1025 | 4.6741 | 0.1025 | 4.6741 | 0.1025 | 4.6741 | 0.1025 | 4.6741 | 4.6741 |
| LS | 0.0010 | 0.04643 | 0.0010 | 0.0464 | 0.00101 | 0.04643 | 0.001 | 0.04643 | 0.04643 |
| GL | 0.6677 | 30.4437 | 0.6677 | 30.4437 | 0.6677 | 30.4437 | 0.6677 | 30.4437 | 30.4437 |
| WL | 2.8983 | 132.1354 | 3.1232 | 142.3875 | 2.4259 | 110.5973 | 2.2083 | 100.6764 | 68.7901 |
| Wmx | 1.2889 | 58.764 | 1.06572 | 48.5861 | 1.1721 | 53.4367 | 1.146 | 52.248 | 61.9801 |
| Wmn) | 0.06864 | 3.1294 | 0.2937 | 13.3940 | 0.18739 | 8.5434 | 0.2134 | 9.7320 | 0 |
| | 100 | 4559 | 100 | 4559 | 100 | 4559 | 100 | 4559 | 4559 |

Table 4.4: LULC Changes in terms of Area for Dakshina Kannada region from 2010-2015.

| Year | 2010-2011 | | 2011-2012 | | 2012-2013 | | 2013-2014 | | 2014-2015 | |
|-------|-----------|----------------|-----------|----------------|-----------|----------------|-----------|----------------|-----------|----------------|
| | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Area(%) | Area(km^2) |
| BU | 4.3611 | 198.8238 | 4.4284 | 201.8914 | 4.5039 | 205.3366 | 4.5039 | 205.3366 | 4.5039 | 205.3366 |
| KC | 0.2584 | 11.7812 | 0.2420 | 11.0352 | 0.1709 | 7.7943 | 0.204 | 9.3018 | 1.8888 | 86.1123 |
| RC | 0.2421 | 11.03837 | 0.0155 | 0.7088 | 0.1424 | 6.4942 | 0.3671 | 16.7370 | 1.0228 | 46.6298 |
| ZC | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1646 | 7.5064 |
| D/T C | 3.785 | 172.5589 | 3.6879 | 168.1355 | 3.5155 | 160.273 | 3.3134 | 151.057 | 1.564 | 71.3377 |
| CF | 1.5659 | 71.3903 | 1.8345 | 83.6390 | 1.8630 | 84.9360 | 1.8091 | 82.4813 | 1.3467 | 61.3982 |
| P | 20.372 | 928.7834 | 20.3758 | 928.9351 | 20.3740 | 928.8546 | 20.3740 | 928.8546 | 20.3740 | 928.8546 |
| EF | 53.5564 | 2441.6381 | 53.5564 | 2441.6381 | 53.5564 | 2441.6381 | 53.5564 | 2441.6381 | 53.5564 | 2441.6381 |
| DF | 12.18035 | 555.3020 | 12.1803 | 555.3020 | 12.18035 | 555.3020 | 12.1803 | 555.3020 | 12.18035 | 555.3020 |
| D/S F | 0.1025 | 4.6741 | 0.1025 | 4.6741 | 0.1025 | 4.6741 | 0.1025 | 4.6741 | 0.1025 | 4.6741 |
| LS | 0.0010 | 0.0464 | 0.00101 | 0.04643 | 0.00101 | 0.04643 | 0.00101 | 0.04643 | 0.00101 | 0.04643 |
| GL | 0.6677 | 30.4437 | 0.6677 | 30.4437 | 0.6677 | 30.4437 | 0.6677 | 30.4437 | 0.6677 | 30.4437 |
| WL | 1.5472 | 70.5391 | 1.5386 | 70.1490 | 1.5532 | 70.8146 | 1.5515 | 70.7341 | 1.2574 | 57.3277 |
| Wmx | 1.0430 | 47.5522 | 1.2834 | 58.5101 | 0.8892 | 40.54110 | 1.2831 | 58.5009 | 1.0230 | 46.6422 |
| Wmn | 0.3164 | 14.4278 | 0.08534 | 3.8909 | 0.4792 | 21.8507 | 0.08534 | 3.8909 | 0.34546 | 15.7496 |
| | 100 | 4559 | 100 | 4559 | 100 | 4559 | 100 | 4559 | 100 | 4559 |

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

Table 4.5: LULC Changes in terms of Area for Dakshina Kannada region from 2015-2018.

| Year | 2015-2016 | | 2016-2017 | | 2017-2018 | | 2015-2018 | |
|-------|-----------|----------------|-----------|----------------|-----------|----------------|----------------------|-----------------------------|
| | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Total LULC Change(%) | Total LULC Change(km^2) |
| BU | 4.5222 | 206.1677 | 4.4938 | 204.8758 | 4.5146 | 205.8224 | 1.544 | 1.544 |
| KC | 0.03391 | 1.5460 | 0.4827 | 22.0092 | 1.5877 | 72.3842 | 0.851 | 0.851 |
| RC | 0.07083 | 3.2292 | 0.14986 | 6.8322 | 0.1672 | 7.6246 | -0.158 | -0.158 |
| ZC | 0.0167 | 0.7616 | 0.0228 | 1.0422 | 0.0055 | 0.2551 | 0.005 | 0.005 |
| D/T C | 5.7363 | 261.5187 | 5.3173 | 242.4171 | 4.2349 | 193.07161 | 0.873 | 0.873 |
| CF | 0.2492 | 11.3634 | 0.0915 | 4.1723 | 0.1097 | 5.0043 | -1.359 | -1.359 |
| PL | 20.3598 | 928.2061 | 20.36257 | 928.3296 | 20.3644 | 928.4156 | -0.008 | -0.008 |
| EF | 53.5224 | 2440.0902 | 53.5680 | 2442.1690 | 53.4895 | 2438.5865 | -0.066 | -0.066 |
| DF | 12.1171 | 552.4209 | 12.1966 | 556.04361 | 12.1178 | 552.4525 | -0.062 | -0.062 |
| D/S F | 0.10382 | 4.7334 | 0.10313 | 4.7017 | 0.1046 | 4.772043363 | 0.002 | 0.002 |
| LS | 0.00133 | 0.0609 | 0.00108 | 0.04931 | 0.00150 | 0.06855 | 0.0004 | 0.0004 |
| GL | 0.66046 | 30.1107 | 0.6595 | 30.06787 | 0.6589 | 30.04140307 | -0.0083 | -0.0083 |
| WL | 1.19698 | 54.5703 | 1.1900 | 54.2536 | 1.1926 | 54.3738 | -1.705 | -1.705 |
| WmX | 0.7732 | 35.2517 | 0.7825 | 35.6770 | 1.2002 | 54.7204 | -0.088 | -0.088 |
| Wmn | 0.6354 | 28.9683 | 0.5781 | 26.359 | 0.2501 | 11.4064 | 0.181 | 0.181 |
| | 100 | 4559 | 100 | 4559 | 100 | 4559 | | 4559 |

+ denotes increase in the LULC class, - denotes decrease in the LULC class

4.4.4.3 Experiments on the Goa Dataset

The proposed model generated the binary change maps for the given two bi-temporal LULC maps for 2005-06 and 2017-18 to determine the LULC class-wise binary changes with black pixels representing no change and white pixels as change. The change maps generated by various SOTA methods are being compared with the proposed approach, as shown in figure 4.20 (a) IFN , 4.20(b) H-SALENet, 4.20(c) ADS-NET, 4.20(d)SIAM-GL ,4.20(e) STEDSAN(Proposed) and 4.20(f) Ground Truth. Visual interpretation and detailed analysis are used to make the comparison. STEDSAN effectively returns large area change regions with complete boundaries and strong internal compactness when large area change areas are found amongst the LULC class. STEDSAN detects the image noises compared to other methods, and these pixels are classified as unaltered areas in the change maps. Visual analysis of STEDSAN change maps reveals outstanding performance, and results match ground truth maps for small and large area LULCCD tasks. The quantitative evaluation is shown in Table 4.6. STEDSAN has also achieved the best outcomes, with the highest F1 (82.87%) ,MIOU (74.8%) and OA (94.93 %). Some of the SOTA methods has gained a very close margin to the proposed method.

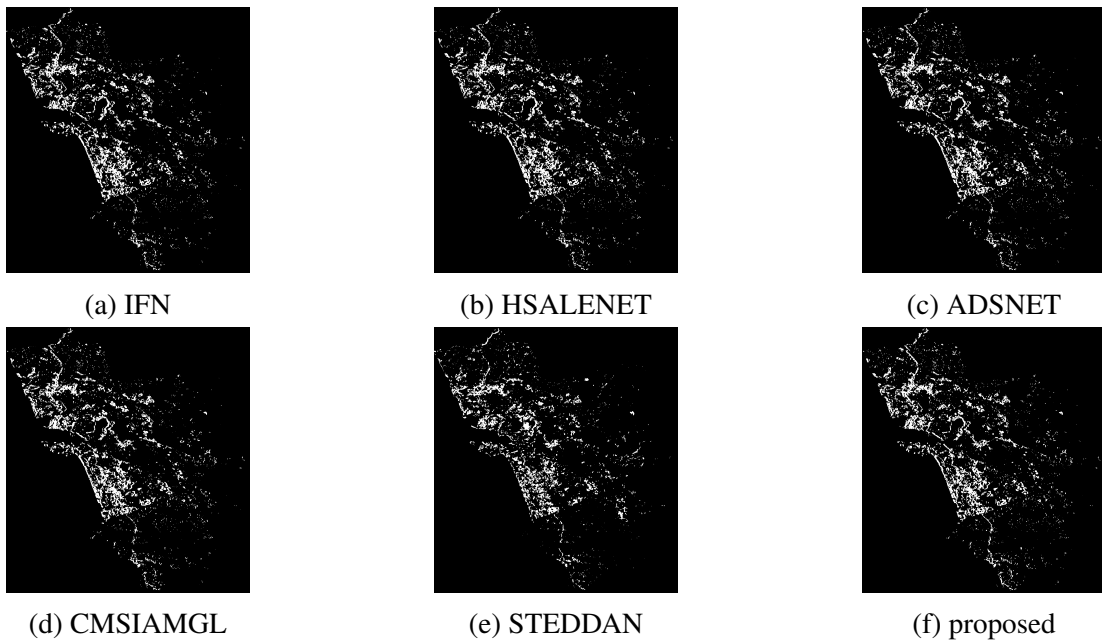


Figure 4.20: Change Maps generated for Goa dataset.

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

Table 4.6: Quantitative Performances of Different Methods on the Goa Dataset.

| Method | F1(%) | MIOU(%) | OA(%) |
|-----------|-------|---------|-------|
| IFN | 70.60 | 52.5 | 89.10 |
| H-SALENet | 72.07 | 63.2 | 91.27 |
| ADS-NET | 72.03 | 67.9 | 92.46 |
| Siam-GL | 79.94 | 70.4 | 93.45 |
| Proposed | 82.87 | 74.8 | 94.93 |

4.4.4.4 LULC Change Analysis for Goa Region

The following analysis shows the different categories of changes in each LULC class in terms of area and percentage of changes for the Goa region from 2005-06 to 2017-18. All the changes are shown in Table 4.7, Table 4.8 and Table 4.9. The model determined changes in only 15 LULC classes as per the LULC maps as shown in figure 4.21, showcasing there was no impact of LULC types like shifting cultivation, rann, and snow cover for the considered area.

- The BU area is increased by 1.259% with an area changes from 156.72 sq km to 203.33 sq km. Figure 4.22a displays the area-wise changes showcasing that the GL and RC are not converted to the BU class.
- The KC is increased by 0.1646% with an area 598 changes from 171.12 sq km to 177.22 sq km. Figure 4.22b displays the area-wise changes showcasing all classes being converted to the KC class.
- The RC is increased by 1.1777% with an area changes from 9.757 sq km to 53.356 sq km. are not converted to RC. Figure 4.23a displays the area-wise changes showcasing that ZC and GL are not converted to the RC class.
- The ZC is increased by 0.103% with an area changes from 0.0314 sq km to 3.847 sq km. Figure 4.23b displays the area-wise changes showcasing that Only GL is not converted to the ZC class.
- The D/T C is increased by 2.42% with an area changes from 118.47 sq km to 208.08 sq km. Figure 4.24a displays the area-wise changes showcasing that only GL is not converted to the D/T C class.

- The CF is decreased by 3.299% with an area changes from 209.54 sq km to 87.41 sq km. are not converted to CF. Figure 4.24b displays the area-wise changes showcasing that ZC and GL are not converted to the CF class.
- The PL is increased by 0.0258% with an area changes from 292.20 sq km to 293.15 sq km. Figure 4.25a displays the area-wise changes showcasing that only GL is not converted to the PL class.
- The EF is decreased by 0.568% with an area changes from 542.02 sq km to 520.96 sq km. Figure 4.25b displays the area-wise changes showcasing that ZC is not converted to the EF class.
- The DF is increased by 0.0029% with an area changes from 1803.82 sq km to 1803.93 sq km. All classes are converted to DF. Figure 4.26a displays the area-wise changes showcasing that all classes are converted to the DF class.
- The D/S F is increased by 0.0094% with an area changes from 31.44 sq km to 31.79 sq km. Figure 4.26b displays the area-wise changes showcasing that ZC and GL are not converted to the D/S F class.
- The LS is increased by 0.0001% with an area changes from 22.601sq km to 22.608 sq km. Figure 4.27a displays the area-wise changes showcasing that ZC and GL are not converted to the D/S F class.
- The GL is decreased by 0.0003% with an area changes from 0.0502 sq km to 0.0386 sq km. Figure 4.27b displays the area-wise changes showcasing that Only three classes are converted into the GL class, i.e., D/T C, EF and DF.
- The WL is decreased by 1.455% with an area changes from 219.37 sq km to 165.48 sq km. Figure 4.28a displays the area-wise changes showcasing that ZC and GL are not converted to the WL class.
- The Wmx is decreased by 0.311 with an area changes from 104.98 sq km to 93.46 sq km. Figure 4.28b displays the area-wise changes showcasing that ZC and GL are not converted to the Wmx class.

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

- The Wmn is increased by 0.471% with an area changes from 19.82 sq km to 37.28 sq km. Figure 4.29a displays the area-wise changes showcasing that ZC and GL are not converted to the Wmn class.

Table 4.7: LULC Changes in terms of Area for Goa Region from 2005-2010.

| Year | 2005-06 | | 2006-2007 | | 2007-2008 | | 2008-2009 | | 2009-2010 | |
|-------|----------|----------------|-----------|----------------|-----------|----------------|-----------|----------------|-----------|----------------|
| | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Area(%) | Area(km^2) |
| BU | 4.2335 | 156.7257 | 5.0074 | 185.3756 | 5.0075 | 185.3788 | 5.2458 | 194.1995 | 5.2559 | 194.5767 |
| KC | 4.6225 | 171.1261 | 4.4448 | 164.5468 | 5.0477 | 186.8688 | 5.1012 | 188.8492 | 5.4950 | 203.4256 |
| RC | 0.2635 | 9.7574 | 0.2108 | 7.8053 | 0.2686 | 9.9460 | 0.7968 | 29.4986 | 0.8385 | 31.0421 |
| ZC | 0.0084 | 0.03143 | 0.16863 | 6.2430 | 0.1827 | 6.7648 | 0.0471 | 1.7446 | 0.0486 | 1.8012 |
| D/T C | 3.2003 | 118.4755 | 3.2705 | 121.0752 | 3.3250 | 123.0934 | 3.0975 | 114.6719 | 3.1692 | 117.3250 |
| CF | 5.6604 | 209.5492 | 5.8396 | 216.1820 | 5.2996 | 196.1924 | 5.1186 | 189.4936 | 4.7538 | 175.9891 |
| P | 7.8930 | 292.2014 | 7.8930 | 292.2014 | 7.8930 | 292.2014 | 7.8930 | 292.2014 | 7.8930 | 292.2014 |
| EF | 14.6413 | 542.0221 | 14.6413 | 542.0221 | 14.6413 | 542.0221 | 14.6413 | 542.0221 | 14.6413 | 542.0221 |
| DF | 48.7257 | 1803.828 | 48.7257 | 1803.828 | 48.7257 | 1803.828 | 48.7257 | 1803.828 | 48.7257 | 1803.828 |
| D/S F | 0.8493 | 31.4413 | 0.8493 | 31.4413 | 0.8493 | 31.4413 | 0.8493 | 31.4413 | 0.8493 | 31.4413 |
| LS | 0.6105 | 22.6018 | 0.6105 | 22.6018 | 0.6105 | 22.6018 | 0.6105 | 22.6018 | 0.6105 | 22.6018 |
| GL | 0.00135 | 0.05029 | 0.00135 | 0.05029 | 0.00135 | 0.05029 | 0.00135 | 0.05029 | 0.00135 | 0.05029 |
| WL | 5.9257 | 219.3727 | 4.9651 | 183.8102 | 4.7756 | 176.7939 | 4.4997 | 166.5806 | 4.3457 | 160.8783 |
| Wmx | 2.8360 | 104.9899 | 2.9260 | 108.3220 | 2.9164 | 107.9668 | 3.3666 | 124.6337 | 2.9328 | 108.5735 |
| Wmn | 0.5355 | 19.8260 | 0.4455 | 16.49 | 0.4551 | 16.8491 | 0.00492 | 0.1823 | 0.4387 | 16.2424 |
| | 100 | 3702 | 100 | 3702 | 100 | 3702 | 100 | 3702 | 100 | 3702 |

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

Table 4.8: LULC Changes in terms of Area for Goa Region from 2010-2015.

| Year | 2010-11 | 2011-2012 | 2012-2013 | 2013-2014 | 2014-2015 | | | | | |
|-------|------------------------|------------------------|------------------------|------------------------|------------------------|------------|----------|-----------|----------|----------|
| Class | Area(%) | Area(%) | Area(%) | Area(%) | Area(%) | | | | | |
| | Area(km ²) | Area(km ²) | Area(km ²) | Area(km ²) | Area(km ²) | | | | | |
| BU | 5.2800 | 195.4663 | 5.1855 | 191.9676 | 5.2803 | 195.4789 | 5.2803 | 195.4789 | 5.2803 | 194.5767 |
| KC | 5.2897 | 195.8246 | 5.8385 | 216.1411 | 4.8079 | 177.9884 | 4.8248 | 178.6171 | 4.1459 | 203.4256 |
| RC | 0.3305 | 12.2376 | 0.1545 | 5.7211 | 0.1425 | 5.2779 | 0.5163 | 19.1156 | 1.1286 | 31.0421 |
| ZC | 0.048571 | 1.79808537 | 0.14206 | 5.25908534 | 0.31596 | 11.6969 | 0.11964 | 4.4291 | 0.4086 | 1.8012 |
| D/T C | 4.0971 | 151.6773 | 3.4008 | 125.9005 | 3.2088 | 118.7930 | 3.8994 | 144.3560 | 4.6497 | 117.3250 |
| CF | 4.7266 | 174.9801 | 4.8027 | 177.7966 | 5.5301 | 204.7271 | 4.7510 | 175.8854 | 3.5584 | 175.9891 |
| P | 7.8930 | 292.2014 | 7.9118 | 292.8961 | 7.8979 | 292.3837 | 7.8979 | 292.3837 | 7.8979 | 292.2014 |
| EF | 14.6413 | 542.0221 | 14.6413 | 542.0221 | 14.6413 | 542.0221 | 14.6413 | 542.0221 | 14.6413 | 542.0221 |
| DF | 48.7257 | 1803.8285 | 48.7258 | 1803.8285 | 48.7258 | 1803.8285 | 48.7257 | 1803.8285 | 48.7257 | 1803.828 |
| D/S F | 0.8493 | 31.4413 | 0.8493 | 31.4413 | 0.8493 | 31.4413 | 0.8493 | 31.4413 | 0.8493 | 31.4413 |
| LS | 0.6105 | 22.6018 | 0.6105 | 22.6018 | 0.6105 | 22.6018 | 0.6105 | 22.6018 | 0.6105 | 22.6018 |
| GL | 0.00135 | 0.05029 | 0.00136 | 0.05029 | 0.00136 | 0.05029 | 0.001358 | 0.05029 | 0.001358 | 0.050296 |
| WL | 4.1343 | 153.0541 | 0.0079 | 0.2923 | 4.6514 | 172.1981 | 4.5456 | 168.2781 | 4.7654 | 160.8783 |
| Wmx | 2.23 | 82.5547 | 4.3886 | 162.4658 | 1.6717 | 61.8893438 | 3.3363 | 123.5114 | 1.6435 | 108.5735 |
| Wmn | 1.1415 | 42.2612 | 3.3391 | 123.6152 | 1.6645 | 61.6221 | 0 | 0 | 1.6928 | 16.2424 |
| | 100 | 3702 | 100 | 3702 | 100 | 3702 | 100 | 3702 | 100 | 3702 |

Table 4.9: LULC Changes in terms of Area for Goa Region from 2015-2018.

| Year | 2015-16 | | 2016-2017 | | 2017-2018 | | 2005-2018 | |
|-------|----------|----------------|-----------|----------------|-----------|----------------|----------------|-----------------------|
| | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Area(%) | Area(km^2) | Area(km^2) | Total LULC Change (%) |
| BU | 5.4880 | 203.1673 | 5.2208 | 193.2737 | 5.4926 | 203.3387 | 1.2591 | |
| KC | 5.0118 | 185.5381 | 5.2634 | 194.8498 | 4.7872 | 177.2226 | 0.1646 | |
| RC | 0.08589 | 3.17968 | 0.199 | 7.3663 | 1.4412 | 53.3564 | 1.1777 | |
| ZC | 0.00313 | 0.1159 | 0.0358 | 1.3256 | 0.1039 | 3.8478 | 0.103 | |
| D/T C | 6.0781 | 225.0117 | 4.0804 | 151.0560 | 5.6208 | 208.0829 | 2.4205 | |
| CF | 3.1888 | 118.0500 | 4.6595 | 172.4937 | 2.3614 | 87.4194 | -3.299 | |
| P | 7.9251 | 293.3884 | 7.8878 | 292.007 | 7.9188 | 293.1574 | 0.0258 | |
| EF | 14.118 | 522.6501 | 14.751 | 546.0684 | 14.0724 | 520.9626 | -0.5688 | |
| DF | 48.7447 | 1804.5295 | 48.688 | 1802.4177 | 48.7287 | 1803.9368 | 0.0029 | |
| D/S F | 0.8526 | 31.5650 | 0.8519 | 31.5388 | 0.8587 | 31.7913 | 0.0094 | |
| LS | 0.6124 | 22.6711 | 0.6072 | 22.4795 | 0.6107 | 22.6082 | 0.0001 | |
| GL | 0.000939 | 0.03477 | 0.0011 | 0.04007 | 0.001043 | 0.03863 | -0.0003 | |
| WL | 4.5261 | 167.557 | 4.3669 | 161.6613 | 4.4702 | 165.4897 | -1.4555 | |
| Wmx | 1.3655 | 50.5542 | 1.4968 | 55.4109 | 2.5246 | 93.4617 | -0.3114 | |
| Wmn) | 1.9985 | 73.9865 | 1.8911 | 70.0099 | 1.00715 | 37.2850 | 0.4719 | |
| | 100 | 3702 | 100 | 3702 | 100 | 3702 | | |

+ denotes increase in the LULC class, - denotes decrease in the LULC class

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

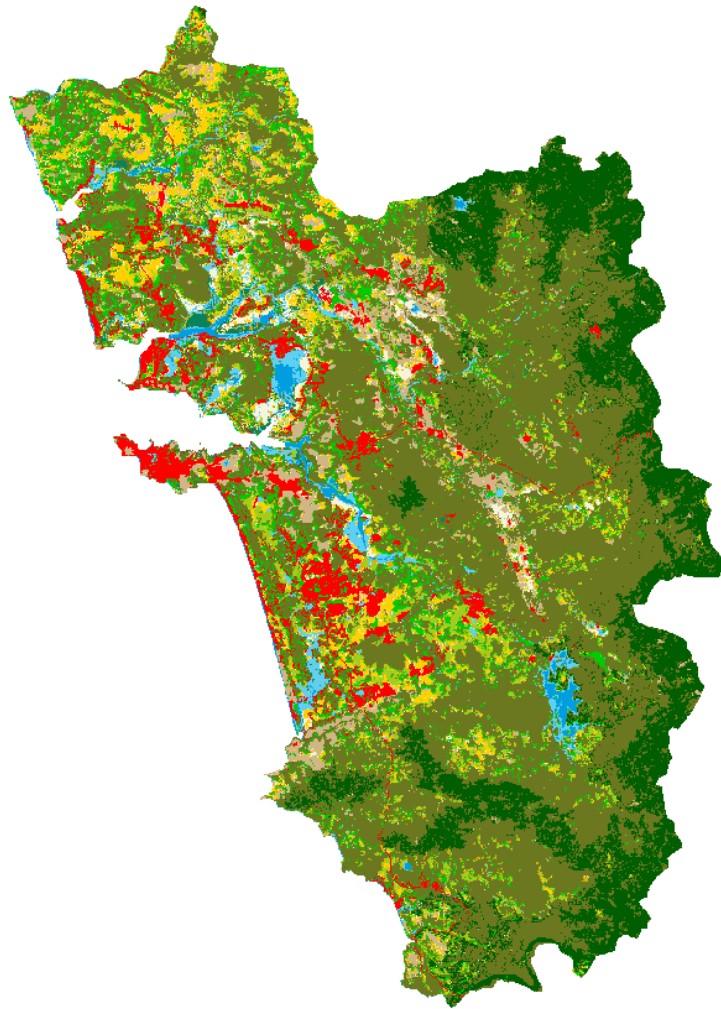
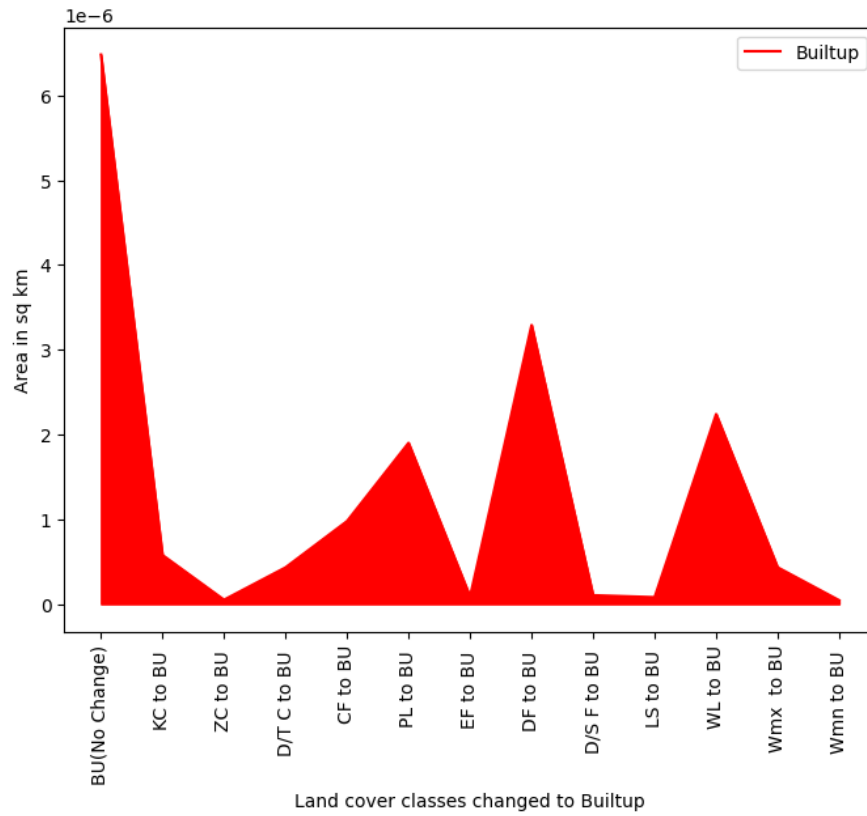
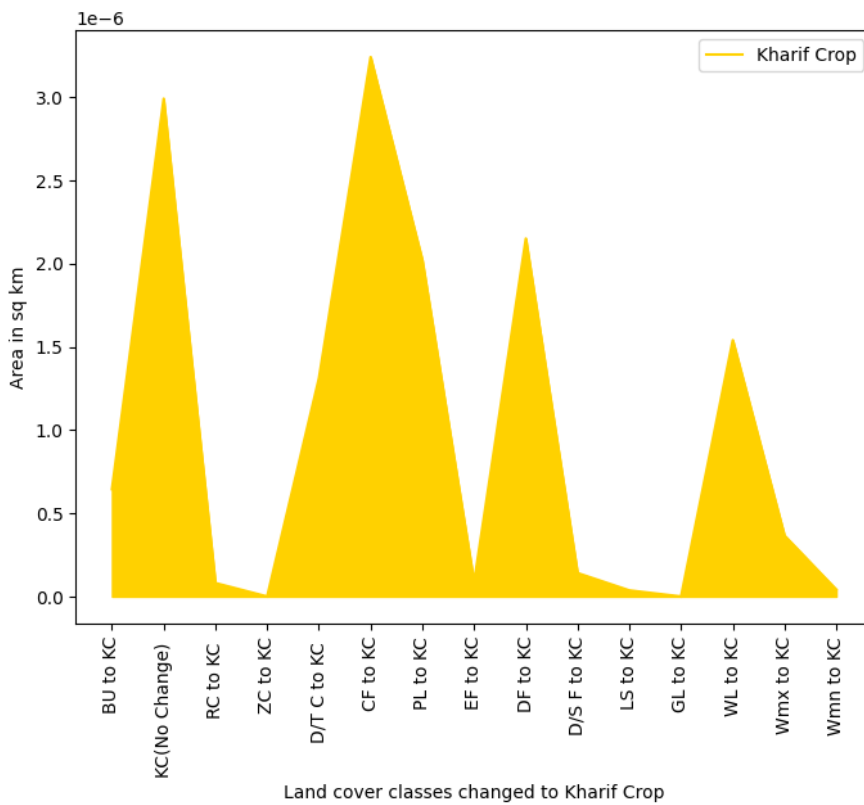


Figure 4.21: Multiclass Change Map for the Goa Region.



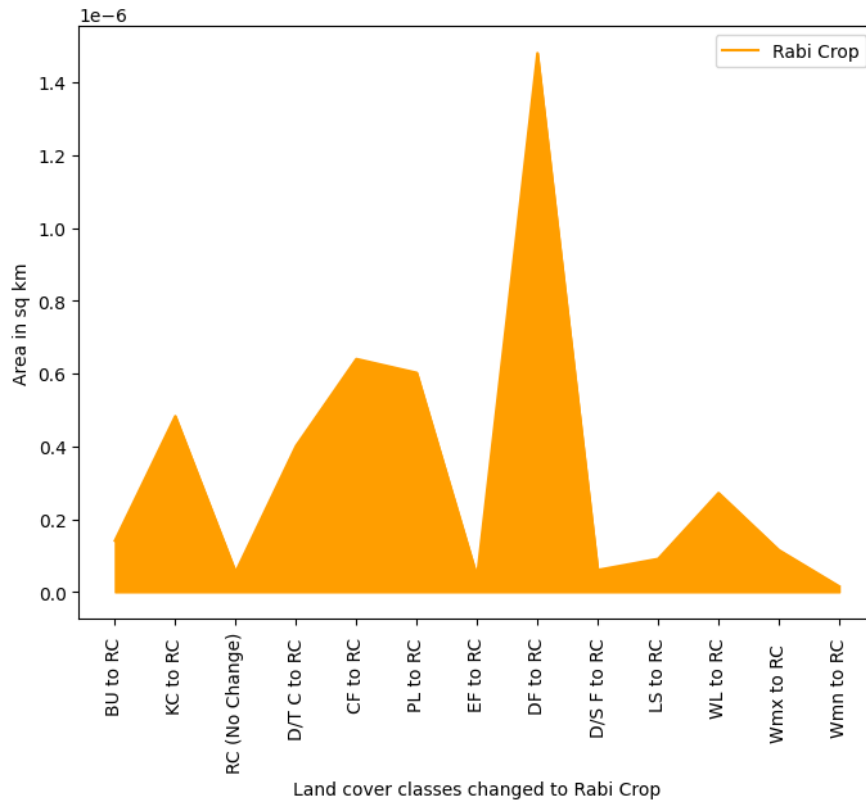
(a) Built-up



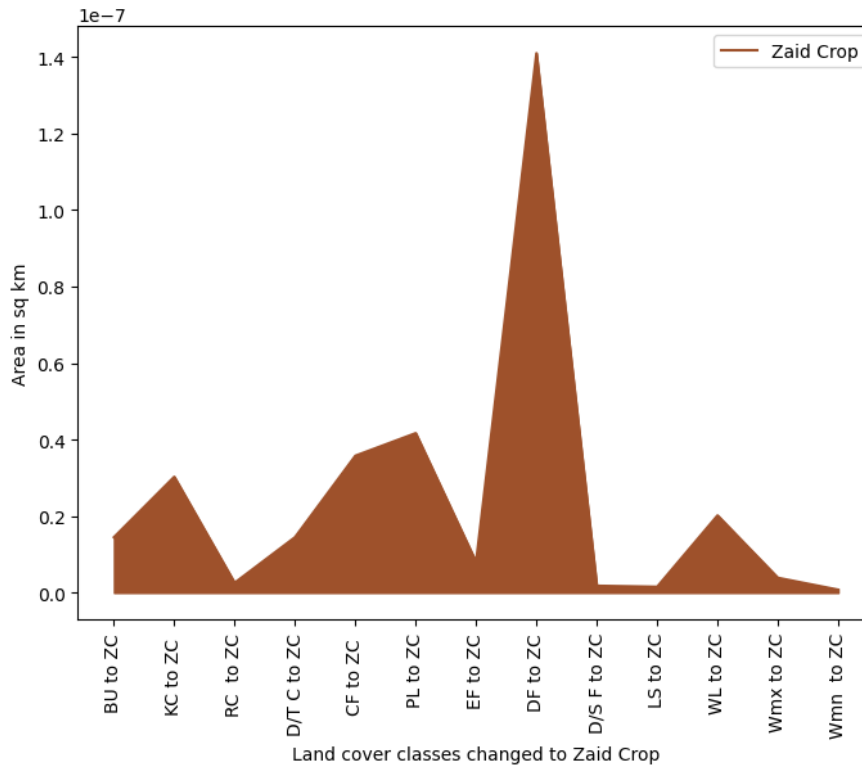
(b) Kharif Crop

Figure 4.22: Area Wise Changes in LULC Classes (Built-up and Kharif Crop).

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

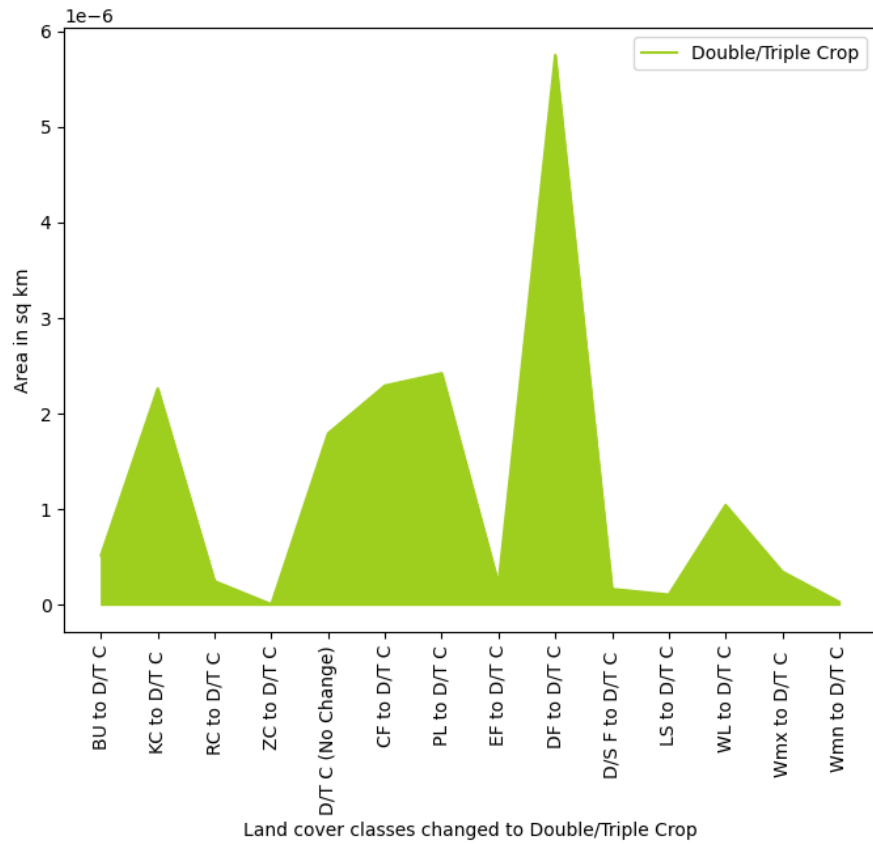


(a) Rabi Crop

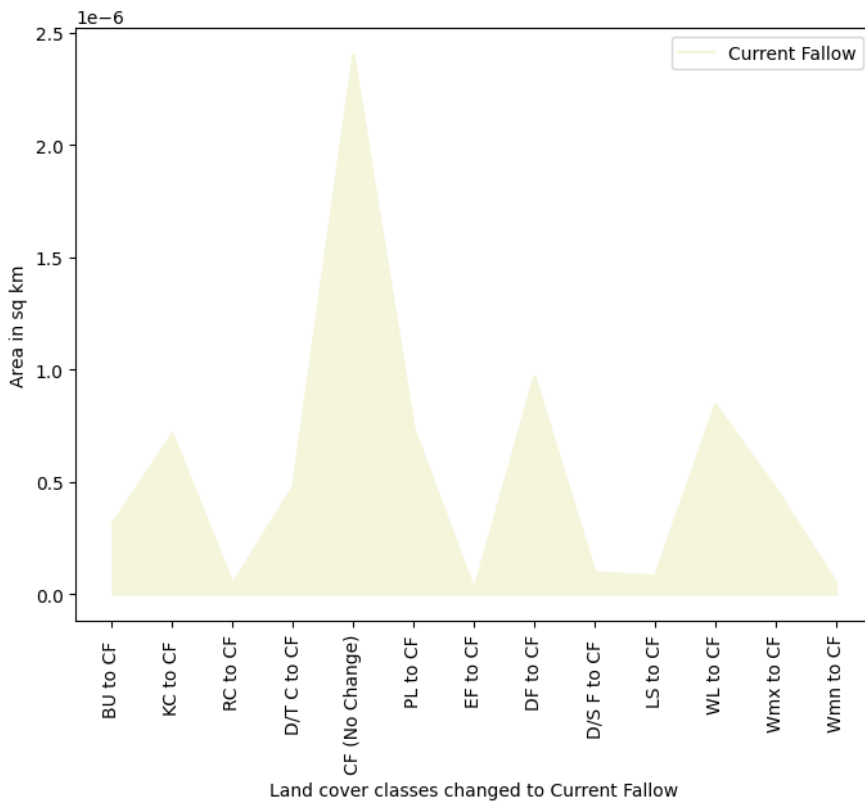


(b) Zaid Crop

Figure 4.23: Area Wise Changes in LULC Classes (Rabi Crop and Zaid Crop).



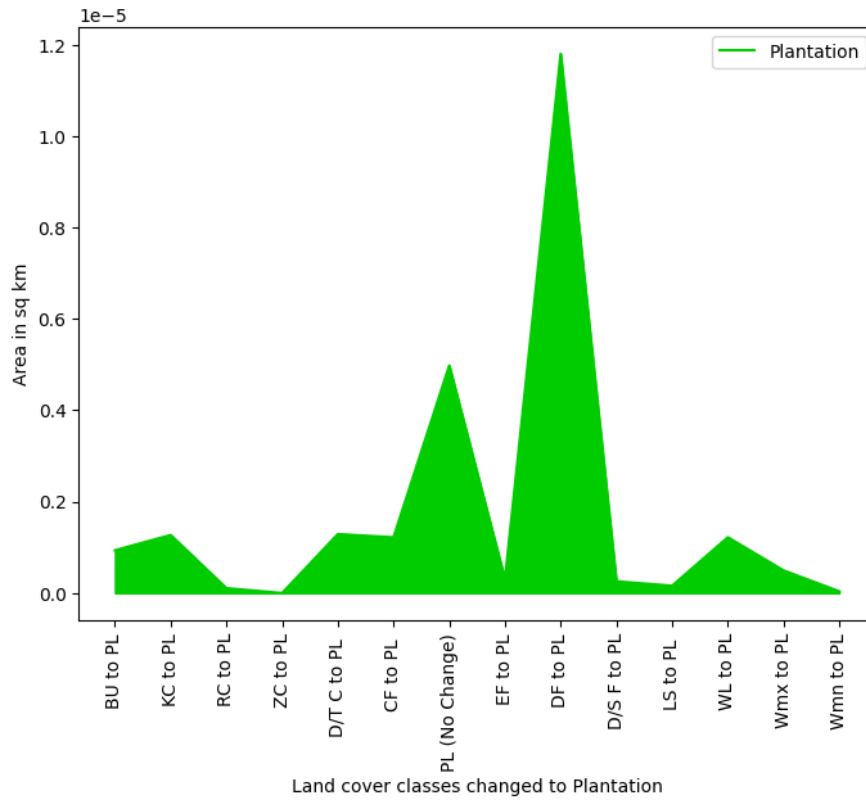
(a) Double/Triple Crop



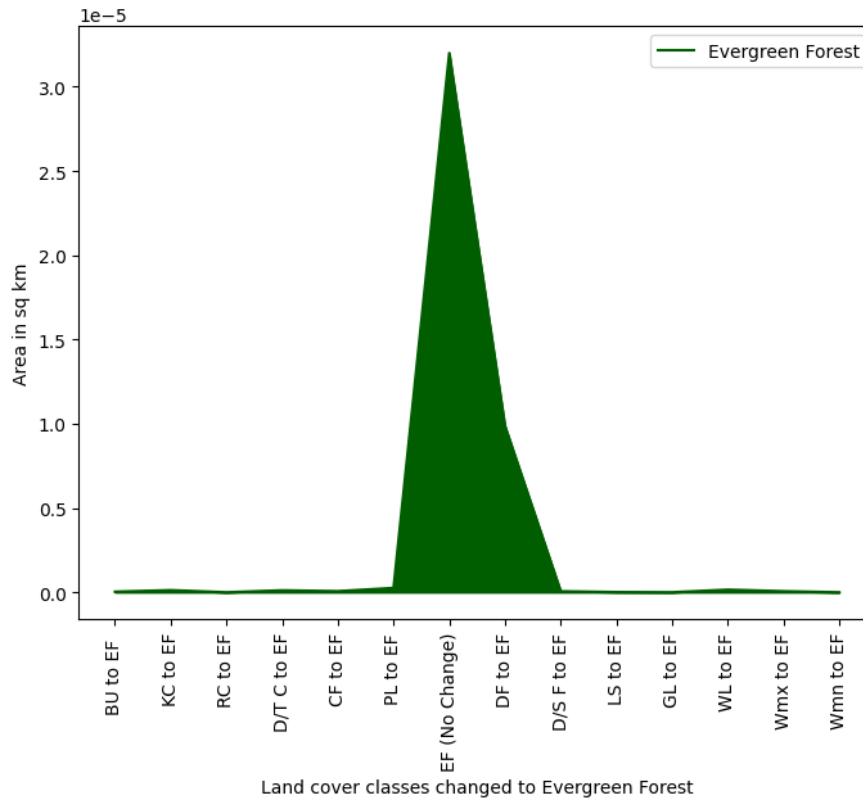
(b) Current Fallow

Figure 4.24: Area Wise Changes in LULC Classes (Double/Triple Crop and Current Fallow).

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

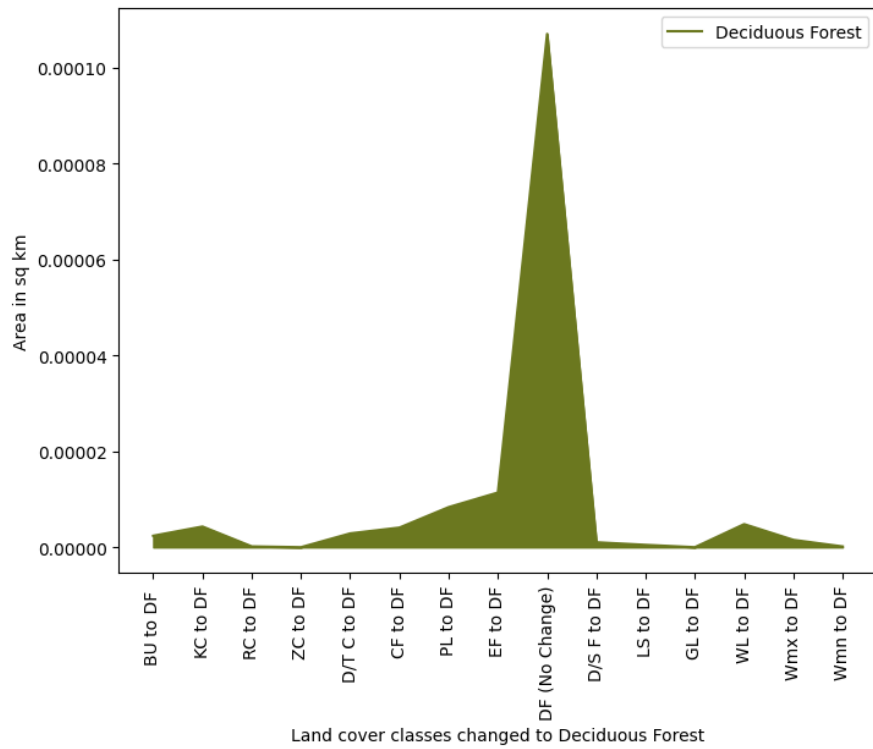


(a) Plantation

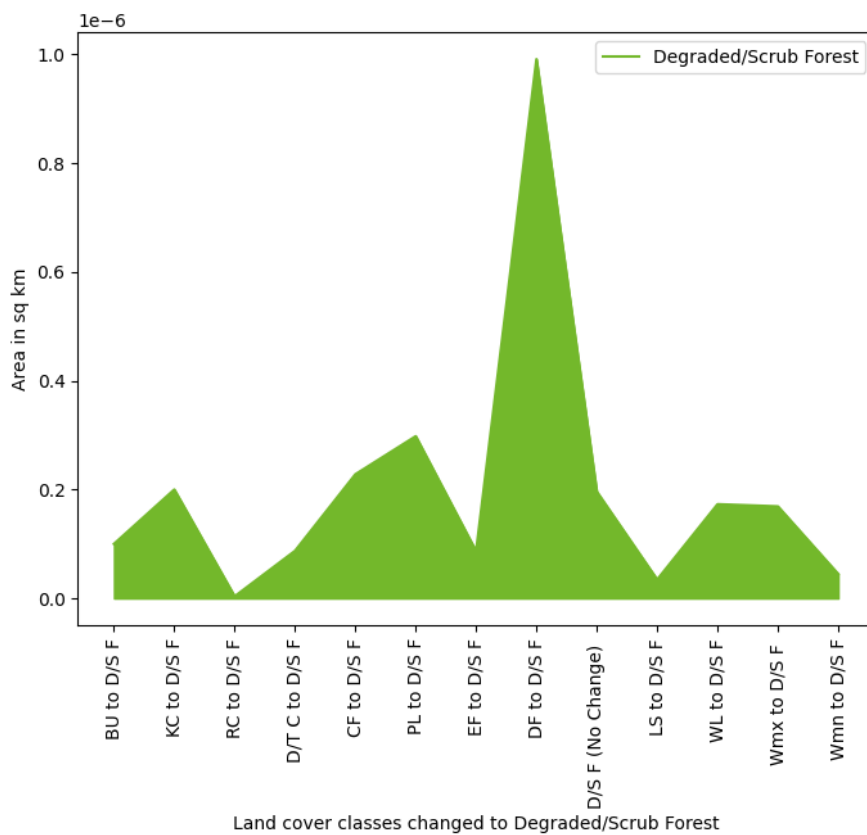


(b) Evergreen Forest

Figure 4.25: Area Wise Changes in LULC Classes (Plantation and Evergreen Forest).



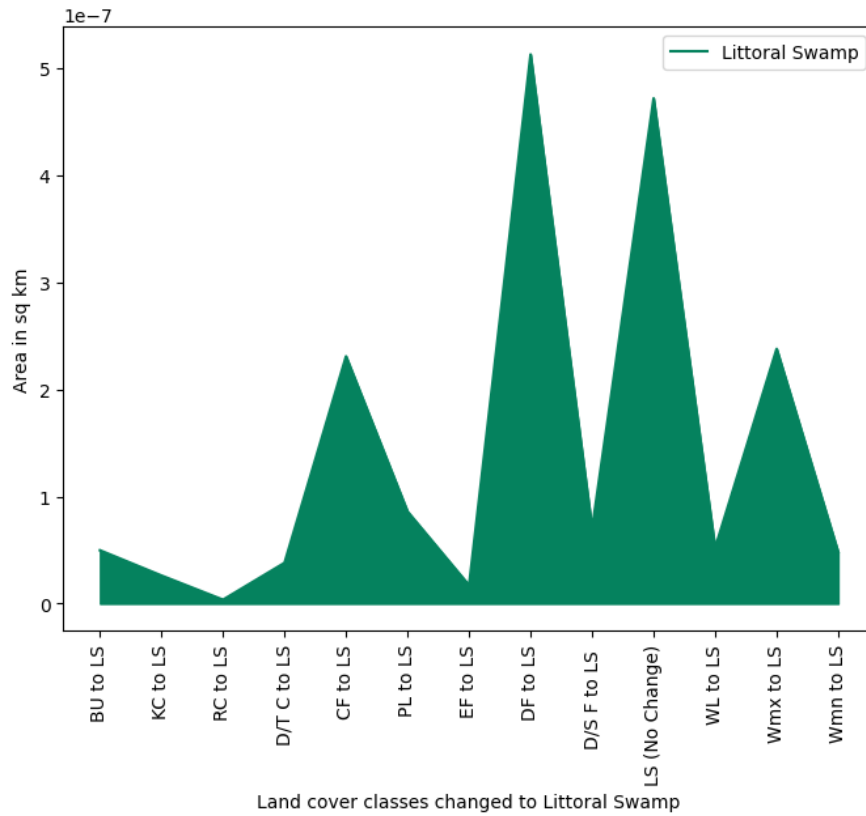
(a) Deciduous Forest



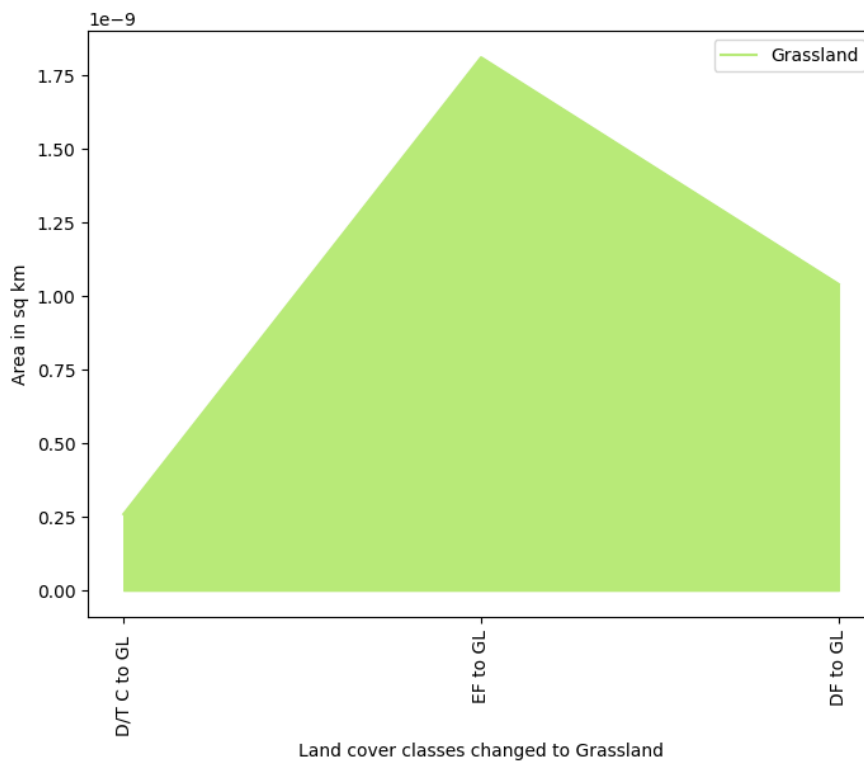
(b) Degraded/Scrub Forest

Figure 4.26: Area Wise Changes in LULC Classes (Deciduous Forest and Degraded/Scrub Forest).

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques

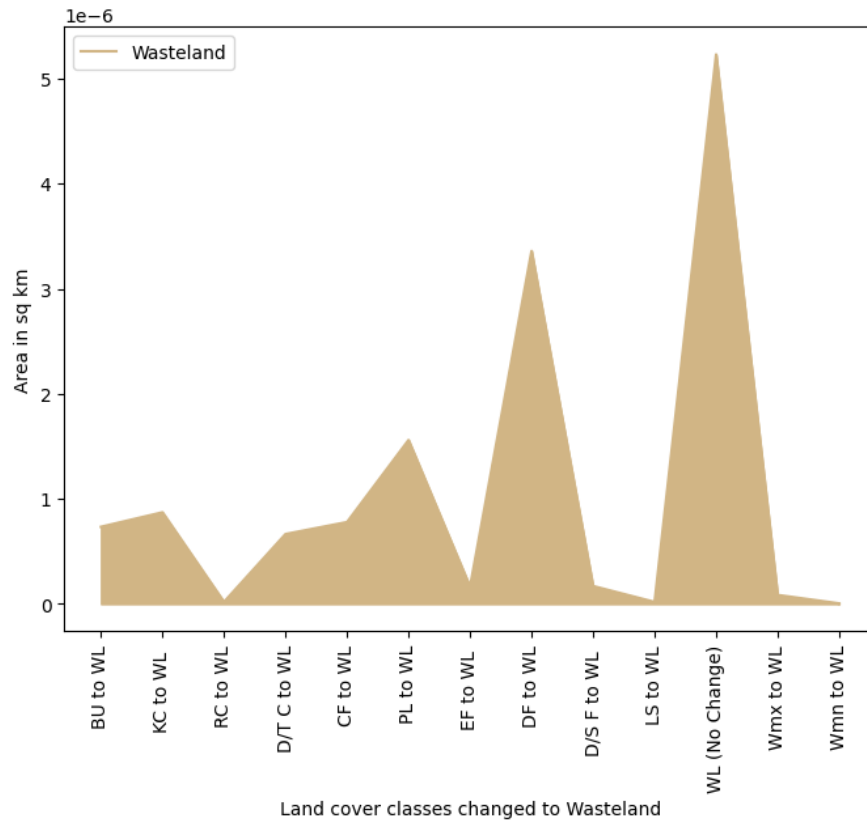


(a) Littoral Swamp

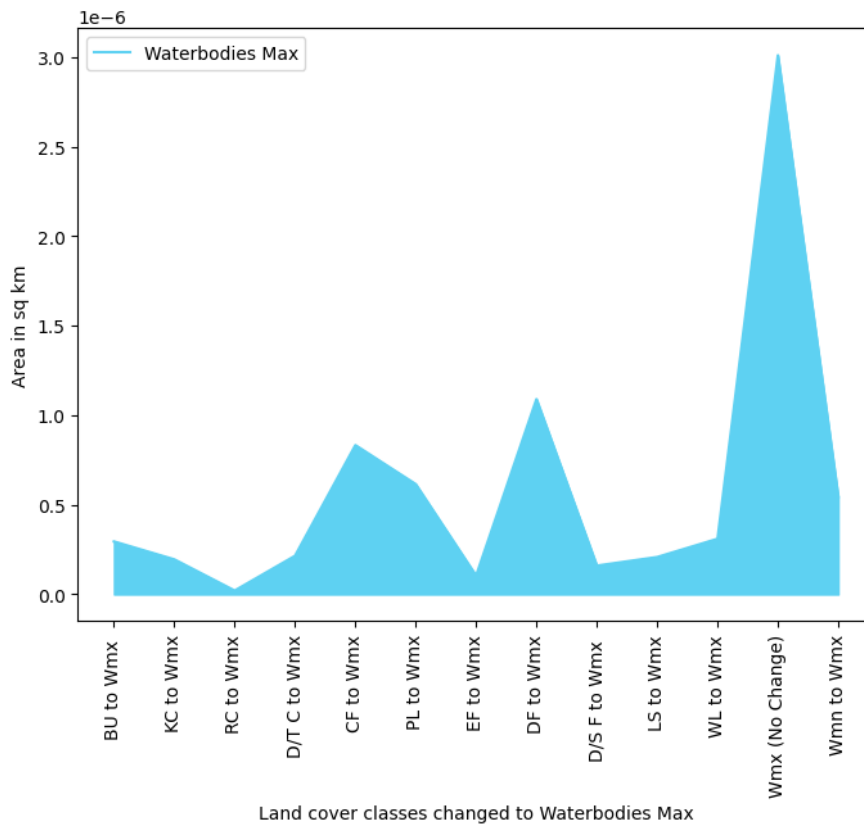


(b) Grassland

Figure 4.27: Area Wise Changes in LULC Classes (Littoral Swamp and Grassland).



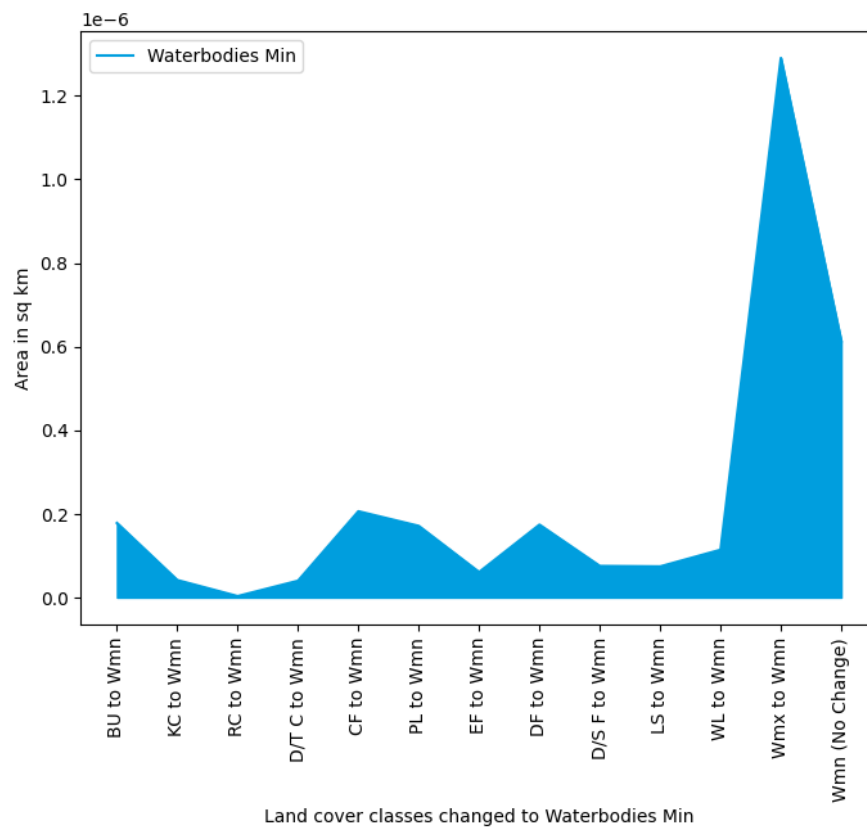
(a) Wasteland



(b) Waterbodies max

Figure 4.28: Area Wise Changes in LULG Classes (Wasteland and Waterbodies max).

4. A Deep Learning Approach for Multi-Class Change Detection in Land Use Land Cover Data using Supervised Techniques



(a) Waterbodies mn

Figure 4.29: Area Wise Changes in LULC Classes (Waterbodies min).

4.5 SUMMARY

This chapter presented a solution for LULCCD with improved classification accuracy in less computational time. The proposed deep encoder-decoder model analyzed spatial and temporal features for effective CD. All the existing spatiotemporal CD models lag in multiscale feature analysis. However, introducing a dedicated multiscale feature analysis module that extracts spatiotemporal features in different window sizes helped improve accuracy. Compared to existing models, the model uses a combination of spatiotemporal analysis with an attention mechanism for identifying essential elements for producing binary and multiclass change maps. These STEDSAN (Goa Region) and STEDDAN (Dakshina Kannada Region) networks helped to minimize the number of learnable parameters effectively and significantly reduce overall execution time.

The performance of the proposed model is evaluated against the existing DL techniques. The dataset for analyzing the model performance is heterogeneous, with vegetation, buildings, roads, and other classes. The model exhibited better CD performance for all types of classes. The model was able to produce higher accuracy in such challenging scenarios also. The model demonstrated excellent performance even in the presence of limited training samples.

CHAPTER 5

A MULTITEMPORAL TECHNIQUE FOR LAND COVER CLASSIFICATION IN IMAGE TIME SERIES

5.1 INTRODUCTION

Deep learning SITS classification presently relies on obtaining significant features from temporal patterns Lyu et al. (2018). CNN's approaches are considered the best models for identifying patterns or trends in the 2-dimensional domain (2D). They have also achieved remarkable classification accuracy in the one dimensional (1D) field, incorporating time-series and spectral data Liu and Shi (2020) Cheng et al. (2018a). Since SITS' temporal relationships are extensive and complicated, modeling the DL architecture to assess remote sensing time series continues to be an open challenge Wang et al. (2019a). Recent advances in DL techniques hold great potential for large-scale assessment of the land cover.

The design of innovative approaches that merge spatio-temporal information from SITS with DL techniques is highly recommended. As a result, the capability of multiple DL algorithms is evaluated for predicting land cover using time series data. The following three goals are specifically addressed:

1. With the proposed framework, the effectiveness of several deep learning architectures is evaluated for the classification of land cover using spatiotemporal features from time series of satellite images using a test dataset.
2. Accuracy assessment for the classification performance of every model is done

by building a feature extractor to identify spatial and temporal properties. Depending on the features, various classifiers are trained and experimented on the LCC dataset.

3. The proposed approach extracted the best land cover from the multitemporal remote sensing images achieving higher F1 scores than the GRU, TCNN, and attention-based models.

The chapter is organized as follows: Section 5.1 is for introduction, Section 5.2 briefs the design of the overall framework used in this research. Section 5.3 describes the proposed framework for land cover classification and an explanation of each module used for classification. A diagrammatic representation of the proposed model is included in this section. Section 5.4 describes the data and details of the study area. Section 5.5 presents the results and its detailed analysis using various evaluation parameters, and the chapter is concluded in Section 5.6.

5.2 OVERALL FRAMEWORK

The complete framework is being divided into multiple phases and the proposed work is as shown in figure 5.1. The following sequence of operations with different models are being followed for spatiotemporal assessment of time series images. In this chapter, land cover classification (LCC) is performed for SITS data. The focus was on DNN architectures that account for spatial, temporal, and spatio-temporal information: The models devised are Gated Recurrent Unit (GRU), Temporal Convolutional Neural Networks (TCNN), GRU+ TCNN, Attention model on TCNN and GRU. The proposed model is also tested in a partitioned manner with the following three possible combinations, namely Univariate+Multivariate (U+M), Univariate+Multivariate+pixel Coordinates (U+M+C), and Univariate+ Multivariate(LSTM) + Coordinates. The TiSeLac dataset has been trained using the following models:

5.2.1 Gated Recurrent Unit

Traditional recurrent neural networks (RNN) can only cope with forward to backward sequences. Because they are unable to learn future information, they may lose it. As a

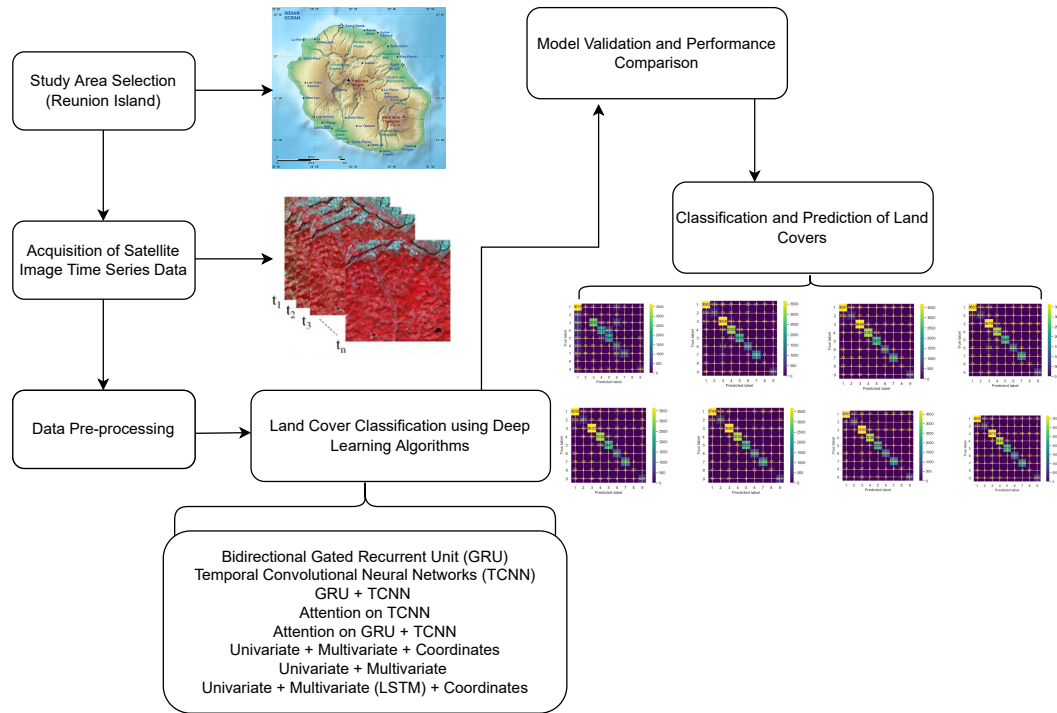


Figure 5.1: Full approach for Time Series Analysis of SITS.

result, the bidirectional recurrent neural network (Bi-RNN), which incorporates an RNN to deal with future input, was introduced by Penghua and Dingyi (2019). Bi-RNN is built by dividing a regular RNN into two directions, one moving clockwise and the other with reverse time. The same output layer is linked to both RNNs, as shown in figure 5.2. This architecture can provide detailed background data for the output layer intake sequence. The historical input data must penetrate both a forward and a reverse GRU, gathering the contextual data of a complete time series to simulate sequence-level LCC analysis using Bi-GRU.

5.2.2 Temporal CNN

The temporal CNN belongs to the category of RNNs, which aims to focus on the temporal dimension of the time series data with consideration of the spatial data by combining one dimensional (1D) FCN and causal convolutions Brock and Abdallah (2022). In this model, 1D convolutions are applied across the temporal dimension. Each dataset sample is reshaped to (number of channels \times number of days). A width filter equal to the

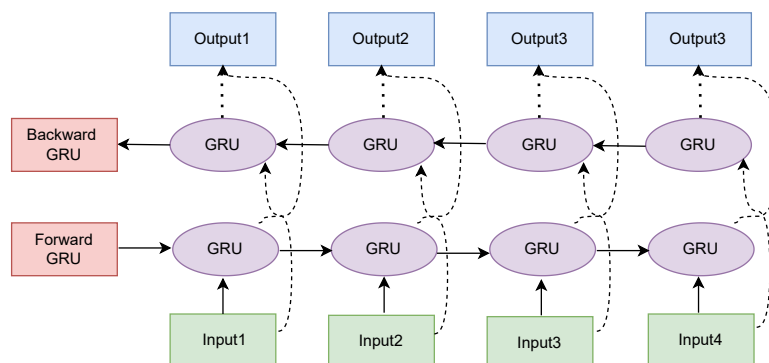


Figure 5.2: Bidirectional Gated Recurrent Unit.

number of channels moves across the dimension of days. The land cover analysis for the SITS is done by explicitly managing data flow through time. This model gives far better accuracy than the previous model, with a much shorter training time.

5.2.3 GRU+Temporal CNN

In this model, both GRU and Temporal CNNs are used separately on the dataset. The outputs of these are concatenated together and this concatenated tensor is passed through a few fully connected layers, terminating with a 9-way softmax. The accuracy of this model is slightly higher than the Temporal CNN model, however the training time per epoch is higher. Increasing the number of GRUs that are stacked together in this model doesn't help in improving the accuracy but does increase training time. Hence, it should be avoided.

5.2.4 Attention on Temporal CNN and GRU

Attention models are neural network input processing methods that let the network concentrate on individual components of a complicated input one at a time until the entire dataset has been classified. The idea is to divide challenging activities into manageable attentional chunks processed sequentially. A function called attention is used to translate a query to an output from an "s set" of key-value pairs. A scenario in which the end output, the fundamental values, and the query are all vectors. The weights assigned to each value are then stated by the query's compatibility function with the associated

key value. The result is then computed as a weighted sum of the values. The intuition behind using attention for our task is that the time series data of some days may be more crucial than others to predict land cover, and attention enables us to capture this Gar-not and Landrieu (2020). The figure 5.3 provides the architecture of TempCNN with attention mechanism. Attention has been applied to the Temporal CNN model and the GRU+Temporal CNN model.

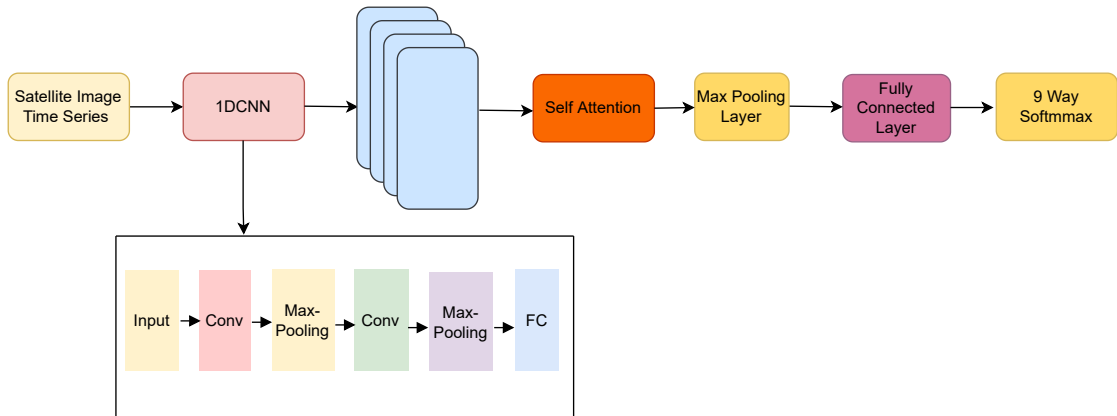


Figure 5.3: Attention on Temporal CNN.

5.3 PROPOSED FRAMEWORK

The architecture proposed in the figure 5.4 integrates three different models: the multi-variate model on the left, ten univariate models in the center, and a model for aggregation to position the information on the right. The description is as given below:

- **Multivariate model:** Uses a 1D convolution on the actual data. It consists of three convolutional layers with no presence of pooling layers in between with a filter size of 3 and a ReLU activation function.
- **Univariate Model:** This model makes use of 10 univariate models. Uses 1D convolutions individually for each feature and concatenates the results. It comprises two convolution layers, including max-pooling and flatten layers at both levels.

5. A Multitemporal Technique for Land Cover Classification in Image Time Series

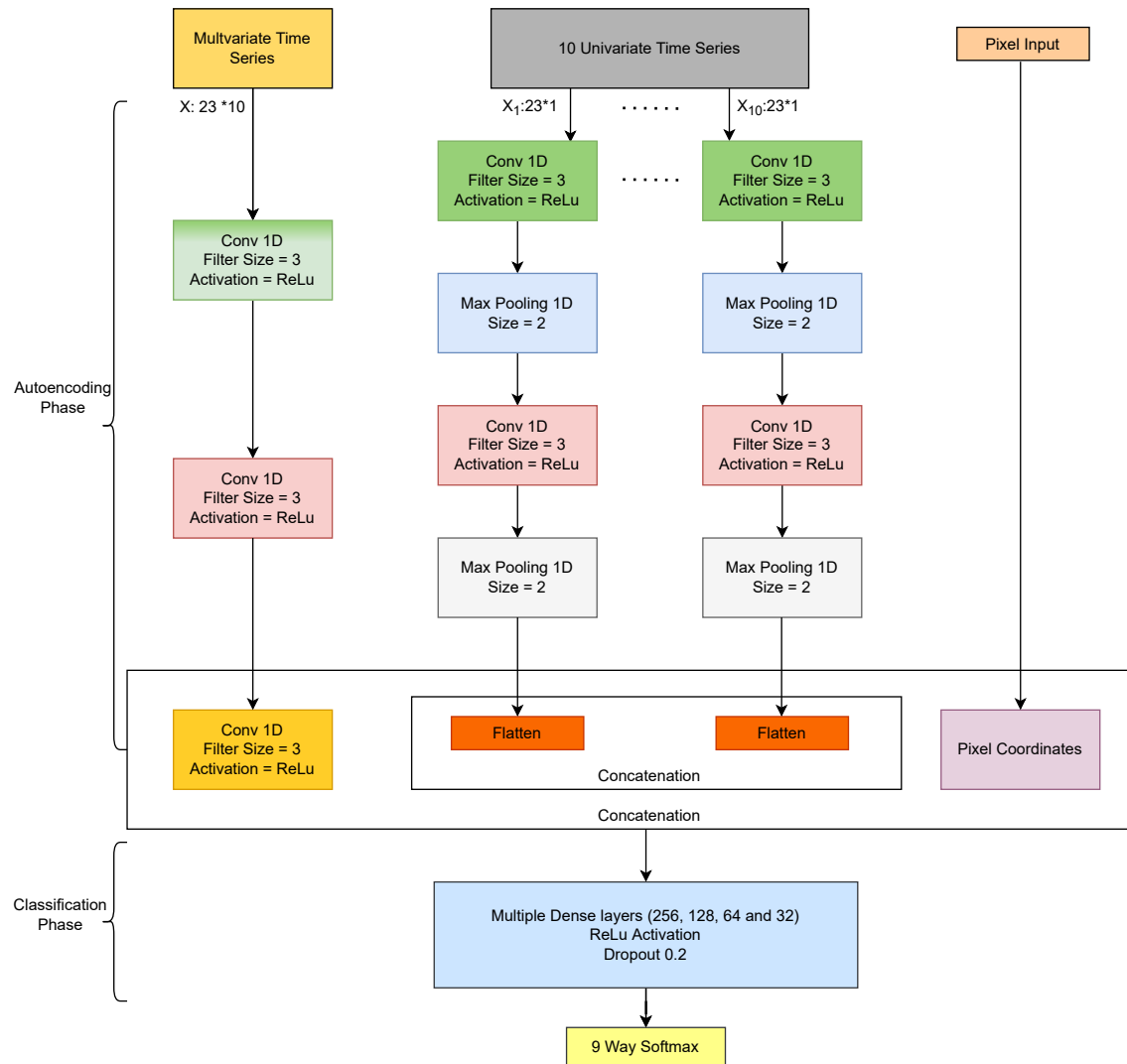


Figure 5.4: Proposed Framework of Univariate + Multivariate + Pixel Coordinates for Land Cover Classification.

Both layers' outcomes are concatenated to engineer the features at different levels.

- Pixel Coordinates: This model is utilized to pass through the preprocessed and scaled pixel coordinates to the final set of fully connected layers. Then, each of these feature extraction model outputs is concatenated to be classified with the usual fully-connected layers:

The concatenation of these models is performed for the task of LCC on the Tiselac dataset. The variants of the proposed framework are implemented using the Univariate + Multivariate + Coordinates component, Univariate + Multivariate component, and Univariate + Multivariate (LSTM) + Coordinates component.

5.4 STUDY AREA AND DATA DESCRIPTION

The dataset (TiseLac) comprises satellite images from an annual Landsat8 time series of 23 high-resolution images of dimensions of 2866×2633 pixels of the La Réunion Island at level 2A processing recorded in 2014. A dataset is made up of pixels: 230 columns, ten features, and 23 dates. Amongst the training and test datasets, 99,687 time series at a 30m spatial resolution are presented by the pixels from the 23 satellite images. The training set contains 81714 instances, whereas the testing set contains 17973 pixels taken from the 23 landsat8 images. Each pixel has ten features at each timestamp: seven surface reflectances, one for each independent multi-spectral band (OLI): Ultra Blue, Blue, Green, Red, NIR (Near Infrared), SWIR1 (Short Wave Infrared), and SWIR2. Each sample in the time-series dataset is arranged in temporal order, with features from 1 to 10 representing the initial timestamp and features from 220 to 230 representing the final timestamp. Table 5.1 shows the distribution of classes in the TiSeLaC dataset. Figure 5.5 depicts the details of the study region. This categorization addresses three significant difficulties that might be exploited and linked to land cover data. Each sample is assigned to one of 9 distinct classes. The essential perspective classes are retained, and spatial processing is used, supported by photo interpretation. The dataset was randomly sampled using pixels to create the fairest and most accurate ground truth possible. This classification aspired to accomplish two fun-

5. A Multitemporal Technique for Land Cover Classification in Image Time Series

damental exploitable challenges associated with the land cover data for the time-series images.

1. The rich band information comprises ten features with a description of each pixel.
2. Temporal data is showcased with the 23-time points amongst which band features are considered. The spatial information is associated with each pixel in the form of coordinates.

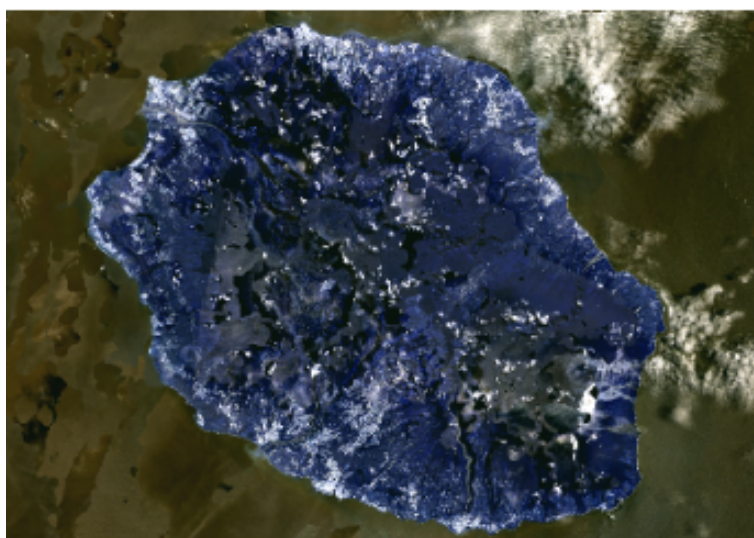


Figure 5.5: Reunion Island.

Table 5.1: Class distributions in the Training and Testing data for TiSeLaC Dataset.

| Class ID | Class Name | Instances Train | Instances Test |
|----------|-------------------------|-----------------|----------------|
| 1 | Urban areas | 16000 | 4000 |
| 2 | Other built-up surfaces | 3236 | 647 |
| 3 | Forests | 16000 | 4000 |
| 4 | Sparse Vegetation | 16000 | 3398 |
| 5 | Rocks and bare soil | 12942 | 2599 |
| 6 | Grassland | 5681 | 1136 |
| 7 | Sugarcane Crops | 7656 | 1531 |
| 8 | Other crops | 1600 | 154 |
| 9 | Water | 2599 | 519 |

The Table 5.1 shows that class proportions are equivalent between train and test splits and are a little unbalanced regarding minority classes like other crops and water. The various land cover classes are 1.Urban areas (UA), 2.Other builtup surfaces (OBA),3.Forests

(Fr), 4. Sparse Vegetation (SV), 5. Rocks and Bare Soil (RBS), 6. Grassland (GL), 7. Sugarcane Crops (SC), 8. Other Crops (OC) and 9. Water (Wr).

5.5 RESULTS AND DISCUSSION

5.5.1 Performance Evaluation Parameters

The performance of the various models for the LCC is evaluated using a test set. The metrics for evaluation are the accuracy, class-wise F1 scores, and macro and weighted average of F1 scores. The F1 score is the mean of precision and recall, given as $F1 = 2(P \cdot R) / (P + R)$. The capability of each model is computed by calculating the F1 score of every class for the LCC. The macro average is used to determine the general classification capacity of each model. The macro-average of the F1 scores is the average of all class-wise F1 scores, and it covers the class imbalance problem, providing attention to the rare essential classes. The weighted-average F1 score for the LCC is generated by computing the mean of all per-class F1 scores while considering each class's support. The dataset's number of actual class instances is referred to as support. The weight is effectively the fraction of each class's support relative to the total of all support values.

5.5.2 Comparative Methods

This section showcases the classification results using existing and proposed models. The LCC classification results are of temporal and spatio-temporal models. Table 5.2 and Table 5.3 shows the per-class accuracy, overall accuracy, per class F1 score, macro average, and weighted average F1 score for the eight DL models. As seen from Table 5.2, the basic models with spatial and temporal capability obtained comparable accuracy levels. The proposed architecture achieves higher accuracy levels than the spatiotemporal models, as shown in last three columns of Table 5.3. The proposed framework achieves an overall accuracy of 93% with macro average F1 scores of 87% and weighted average score of 93% with the overall support of 17973 for the classification task. As seen from the results, the proposed architecture has achieved similar accuracy and F1 scores, showcasing the model's effectiveness for classifying the land covers in time series data. The proposed framework's precision, recall, and F1 scores

are comparatively higher during the analysis process.

Bidirectional GRU In the Bidirectional GRU model, the prediction of samples is shown in the confusion matrix of figure 5.6 (a). As seen in figure 5.6 (a) the correct prediction shows the samples showcasing a higher recall rate for UA and Fr class and an average recall rate for SV and RBS class, and the forecast is slightly lower for the OBS, GL,SC, OC,and Wr classes. The F1 score is also calculated as shown in figure 5.7 (a), and it is zero for the other crops class, stating no prediction for this label. The overall accuracy for the Bidirectional GRU model is 69%. Macro average F1-score is 52%, weighted average F1 score is 69%. As seen it can be inferred from the tables and the matrix, the land cover prediction is not so accurate, being lower on the recall rate for each class showing average performance.

Temporal CNN The temporal CNN model uses 1D convolutions, and the prediction of samples is shown in the confusion matrix of figure 5.6 (b). The correct prediction, as seen in figure 5.6 (b), shows that the samples showcase a higher recall rate for UA,Fr,SV,RBS,GL,SC,and Wr and average recall rate for OBS, and the prediction is slightly lower for the OC class. The F1 score is also calculated for each category, as shown in figure 5.7 (b). The overall accuracy for the TempCNN model is 69%. Macro average F1 score is 81%, weighted Average F1 score is 89%. As seen it can be inferred from the tables and the matrix, the land cover prediction is slightly higher on the recall rate for almost seven classes for the DL model.

GRU+Temporal CNN The integration of GRU and temporal CNN model prediction of samples is shown in the confusion matrix of figure 5.6 (c). As seen in figure 5.6 (c), the correct prediction shows that the samples showcase a higher recall rate for UA, Fr,SV,RBS, GL,SC, and Wr average recall rate for OBS, and the forecast is slightly lower for the OC class. The F1 score is also calculated for each category, as shown in figure 5.7 (c). The overall accuracy for the GRU+TempCNN model is 91%. Macro average F1-score is 82%, weighted Average F1 score is 90%. As seen it can be inferred from the tables and the matrix, the integration of both models for the land cover prediction accuracy has increased by 1% only, with the recall rate being slightly higher for

almost seven classes for the DL model.

Attention on Temporal CNN The attention to the temporal CNN model using self-attention and the prediction of samples are shown in the confusion matrix of figure 5.6 (d). The correct forecast, as seen in figure 5.6 (d) , shows that the samples showcase a higher recall rate for UA, Fr,SV,RBS, GL,SC, and Wr, average recall rate for OBS, and the prediction is slightly lower for the OC class. The F1 score is also calculated for each category, as shown in figure 5.7 (d). The overall accuracy for the attention on the temporal CNN model is 91%. Macro average F1-score is 83%, weighted average F1 score is 91%. As seen it can be inferred from the tables and the matrix, the integration of both models for the land cover prediction accuracy has increased by 1% only, with the recall rate being slightly higher for almost seven classes for the DL model.

Attention on Temporal CNN +GRU The attention to temporal CNN with the GRU model and the prediction of samples are shown in the confusion matrix of figure 5.6 (e). The correct prediction, as seen in figure 5.6 (e), the samples showcase a higher recall rate for UA, Fr,SV,RBS, GL,SC, and Wr average recall rate for OBS, and the prediction is slightly lower for the OC class. The F1 score is also calculated for each class, as shown in figure 5.7 (e). The overall accuracy for the attention on the temporal CNN with the GRU model is 90%. Macro average F1-score is 82%, Weighted average F1 score is 90%. As seen it can be inferred from the tables and the matrix, the integration of both models for the land cover prediction accuracy has decreased by 1% only, with the recall rate being slightly higher for almost seven classes for the DL model.

5.5.3 Result Analysis with the Proposed Framework

5.5.3.1 Univariate+ Multivariate + Coordinates

In the Univariate+ Multivariate + Coordinates model, the prediction of samples is shown in the confusion matrix of figure 5.6 (f). As seen in figure 5.6 (f), the correct prediction shows that the samples showcase higher recall rates for UA, Fr,SV,RBS,GL,SC, and Wr ,average recall rates for the OBS and OC classes. The F1 score is also calculated for each category, as shown in figure 5.7 (f). The overall accuracy for the attention on this model is 93%. Macro average F1 score is 87%, weighted average F1 score is 93%. As

seen it can be inferred from the tables and the matrix, the integration of the three models for the prediction of the land cover accuracy has increased by 3% only, with the recall rate being higher for almost seven classes for the DL model.

5.5.3.2 Univariate+ Multivariate

The Univariate+ Multivariate model without the pixel coordinates the prediction of samples are as shown in the confusion matrix of figure 5.6 (g). As seen in figure 5.6 (g), the correct prediction shows the samples showcasing higher recall rates for UA, OBS, Fr,SV,RBS, GL,SC, and Wr and slightly average recall rates for the OC class. The F1 score is also calculated for each category, as shown in figure 5.7 (g). The overall accuracy for the attention on Univariate+ Multivariate model is 93%. Macro average F1 score is 87%, weighted average F1 score is 93%. As seen it can be inferred from the tables and the matrix the integration of the two models for the prediction of the land cover the accuracy is almost the same as the previous model except for showing higher recall rate on OBS class and with the recall rate being more higher for almost eight classes for the DL model.

5.5.3.3 Univariate + Multivariate (LSTM) + Coordinates

In the Univariate + Multivariate (LSTM) + Coordinates model with the inclusion of LSTM (RNN), the prediction of samples is shown in the confusion matrix of figure 5.6 (h). As seen in figure 5.6 (h), the correct prediction shows the samples showcasing higher recall rates for UA, Fr,SV,RBS, GL,SC, and Wr, with slightly average recall rates for the OBS and OC classes. The F1 score is also calculated for each class as shown in figure 5.7 (h) . The overall accuracy for the attention on Univariate+ Multivariate(LSTM)+ Coordinates model is 93%. Macro average F1 score is 87%, weighted average F1 score is 93%. As seen it can be inferred from the tables and the matrix the integration of the two models for the prediction of the land cover the accuracy is almost the same as the previous model except for showing higher recall rate on OBS class and with the recall rate being more higher for almost 8 classes for the DL model.

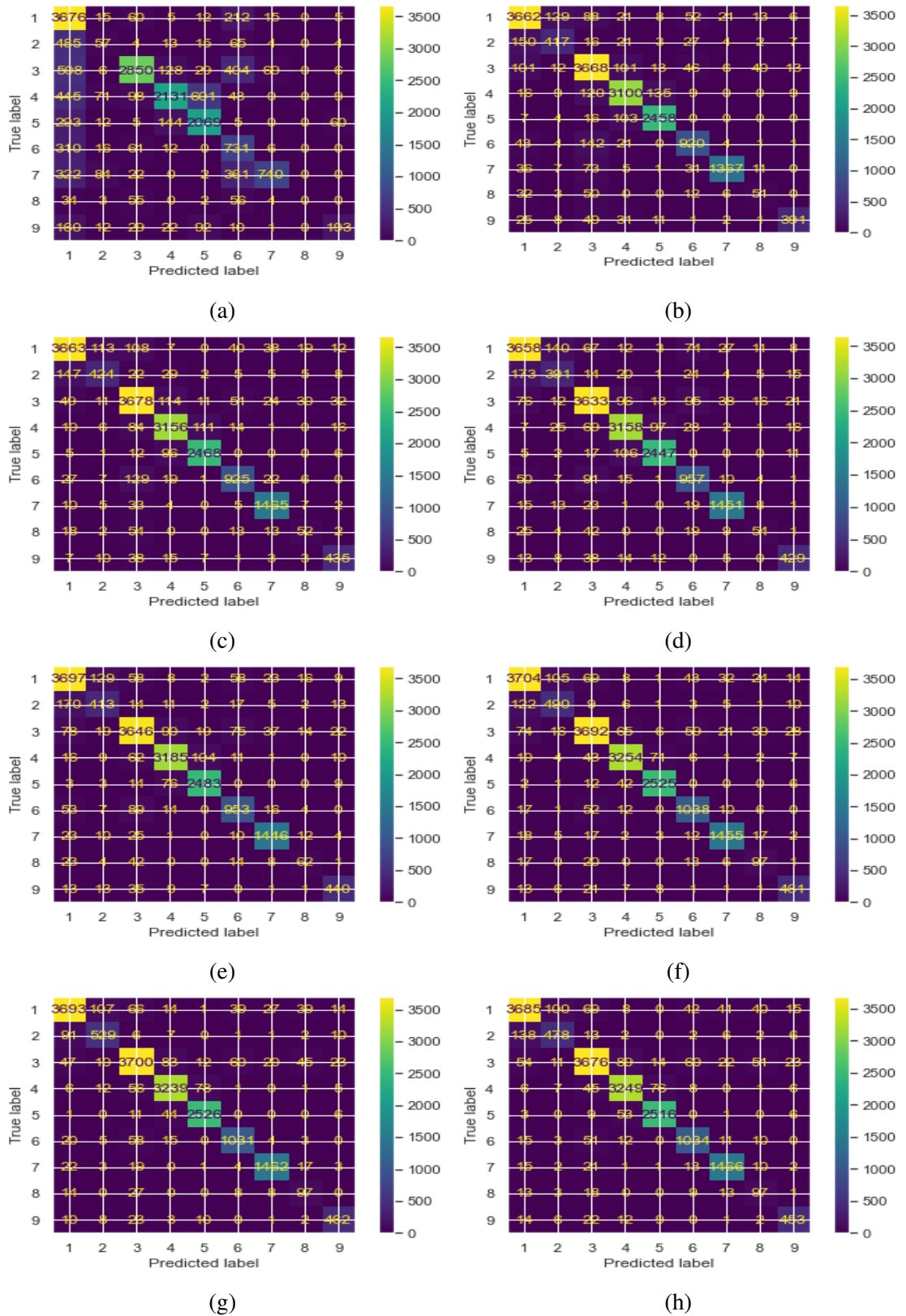


Figure 5.6: Confusion Matrix of Land Cover Classification for the Time Series Dataset using the DL Models: (a) Bidirectional GRU (b) Temporal CNN (c) GRU + Temporal CNN (d) Attention on Temporal CNN (e) Attention on Temporal CNN + GRU (f) Univariate + Multivariate + Coordinates (g) Univariate + Multivariate (h) Univariate + Multivariate (LSTM) + Coordinates.

5. A Multitemporal Technique for Land Cover Classification in Image Time Series

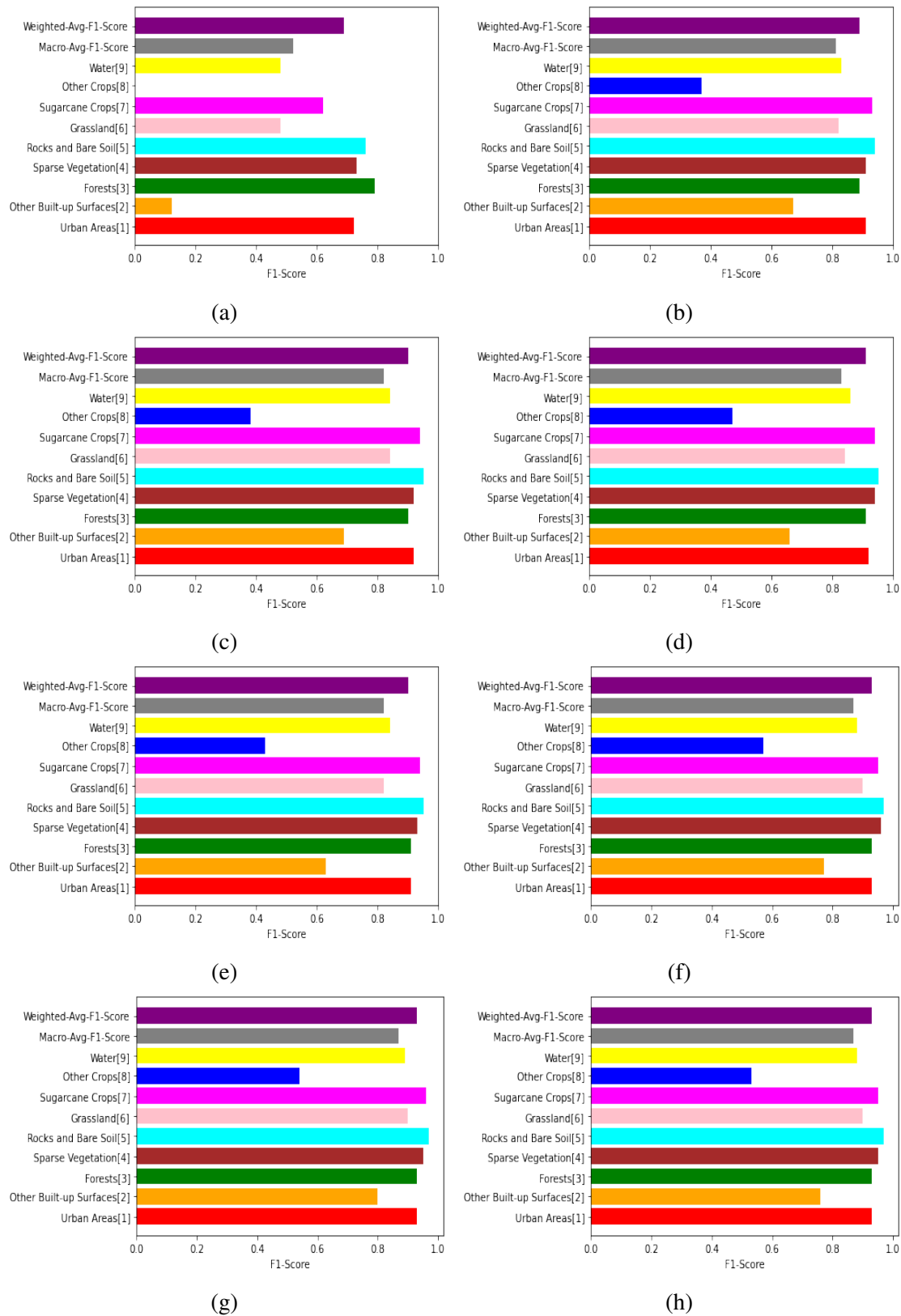


Figure 5.7: F1 Scores of Land Cover Classification for the Time Series Dataset using the DL Models: (a) Bidirectional GRU (b) Temporal CNN (c) GRU + Temporal CNN (d) Attention on Temporal CNN (e) Attention on Temporal CNN + GRU (f) Univariate + Multivariate + Coordinates (g) Univariate + Multivariate (h) Univariate + Multivariate (LSTM) + Coordinates.

Table 5.2: Precision, Recall and F1 score values for the GRU, Temporal CNN and Attention Models on Temporal CNN.

| Class ID | GRU | | | | | Temporal CNN | | | | | GRU+ Temporal CNN | | | | | Attention on Temp.CNN | | | | |
|--------------|------|------|------|-------|--|--------------|------|------|-------|--|-------------------|------|------|-------|--|-----------------------|------|------|-------|--|
| | P | R | F1 | S | | P | R | F1 | S | | P | R | F1 | S | | P | R | F1 | S | |
| 1 (UA) | 0.59 | 0.92 | 0.72 | 4000 | | 0.90 | 0.92 | 0.91 | 4000 | | 0.93 | 0.92 | 0.92 | 4000 | | 0.91 | 0.92 | 0.92 | 4000 | |
| 2 (OBS) | 0.21 | 0.09 | 0.12 | 647 | | 0.70 | 0.64 | 0.67 | 647 | | 0.73 | 0.66 | 0.69 | 647 | | 0.68 | 0.64 | 0.66 | 647 | |
| 3 (Fr) | 0.90 | 0.71 | 0.79 | 4000 | | 0.87 | 0.92 | 0.89 | 4000 | | 0.88 | 0.92 | 0.90 | 4000 | | 0.91 | 0.91 | 0.91 | 4000 | |
| 4 (SV) | 0.87 | 0.63 | 0.73 | 3398 | | 0.91 | 0.91 | 0.91 | 3398 | | 0.92 | 0.93 | 0.92 | 3398 | | 0.94 | 0.94 | 0.94 | 3398 | |
| 5 (RBS) | 0.73 | 0.80 | 0.76 | 2588 | | 0.93 | 0.95 | 0.94 | 2588 | | 0.95 | 0.95 | 0.95 | 2588 | | 0.95 | 0.96 | 0.95 | 2588 | |
| 6 (GL) | 0.39 | 0.64 | 0.48 | 1136 | | 0.84 | 0.81 | 0.82 | 1136 | | 0.88 | 0.81 | 0.84 | 1136 | | 0.84 | 0.84 | 0.84 | 1136 | |
| 7 (SC) | 0.88 | 0.48 | 0.62 | 1531 | | 0.97 | 0.89 | 0.93 | 1531 | | 0.93 | 0.96 | 0.94 | 1531 | | 0.94 | 0.94 | 0.94 | 1531 | |
| 8 (OC) | 0.00 | 0.00 | 0.00 | 154 | | 0.43 | 0.33 | 0.37 | 154 | | 0.43 | 0.34 | 0.38 | 154 | | 0.56 | 0.40 | 0.47 | 154 | |
| 9 (Wr) | 0.70 | 0.37 | 0.48 | 519 | | 0.92 | 0.75 | 0.83 | 519 | | 0.85 | 0.84 | 0.84 | 519 | | 0.87 | 0.85 | 0.86 | 519 | |
| accuracy | | | 0.69 | 17973 | | | | 0.89 | 17973 | | | | 0.91 | 17973 | | | | 0.91 | 17973 | |
| macro avg | 0.58 | 0.52 | 0.52 | 17973 | | 0.83 | 0.79 | 0.81 | 17973 | | 0.83 | 0.81 | 0.82 | 17973 | | 0.84 | 0.82 | 0.83 | 17973 | |
| weighted avg | 0.73 | 0.69 | 0.69 | 17973 | | 0.89 | 0.89 | 0.89 | 17973 | | 0.90 | 0.91 | 0.90 | 17973 | | 0.91 | 0.91 | 0.91 | 17973 | |

Table 5.3: Precision, Recall and F1 score values for the Attention on Temporal CNN+GRU, Univariate+Multivariate+Coordinates, Univariate+Multivariate and Univariate+Multivariate(LSTM)+Coordinates.

| Class ID | Attention on Temporal CNN + GRU | | | | | Univariate+Multivariate+Coordinates | | | | | Univariate+Multivariate | | | | | Univariate+Multivariate(LSTM)+Coordinates | | | | |
|---------------------|---------------------------------|------|------|-------|--|-------------------------------------|------|------|-------|--|-------------------------|------|------|-------|--|---|------|------|-------|--|
| | P | R | F1 | S | | P | R | F1 | S | | P | R | F1 | S | | P | R | F1 | S | |
| 1 (UA) | 0.91 | 0.91 | 0.91 | 4000 | | 0.93 | 0.93 | 0.93 | 4000 | | 0.95 | 0.92 | 0.93 | 4000 | | 0.93 | 0.92 | 0.93 | 4000 | |
| 2 (OBS) | 0.65 | 0.60 | 0.63 | 647 | | 0.78 | 0.76 | 0.77 | 647 | | 0.78 | 0.82 | 0.80 | 647 | | 0.78 | 0.74 | 0.76 | 647 | |
| 3 (Fr) | 0.91 | 0.91 | 0.91 | 4000 | | 0.94 | 0.92 | 0.93 | 4000 | | 0.93 | 0.93 | 0.93 | 4000 | | 0.94 | 0.92 | 0.93 | 4000 | |
| 4 (SV) | 0.92 | 0.93 | 0.93 | 3398 | | 0.96 | 0.96 | 0.96 | 3398 | | 0.95 | 0.95 | 0.95 | 3398 | | 0.95 | 0.96 | 0.95 | 3398 | |
| 5 (RBS) | 0.95 | 0.95 | 0.95 | 2588 | | 0.97 | 0.98 | 0.97 | 2588 | | 0.96 | 0.98 | 0.97 | 2588 | | 0.96 | 0.97 | 0.97 | 2588 | |
| 6 (GL) | 0.79 | 0.84 | 0.82 | 1136 | | 0.88 | 0.91 | 0.90 | 1136 | | 0.90 | 0.91 | 0.90 | 1136 | | 0.89 | 0.91 | 0.90 | 1136 | |
| 7 (SC) | 0.94 | 0.95 | 0.94 | 1531 | | 0.95 | 0.95 | 0.95 | 1531 | | 0.96 | 0.95 | 0.96 | 1531 | | 0.94 | 0.96 | 0.95 | 1531 | |
| 8 (OC) | 0.55 | 0.35 | 0.43 | 154 | | 0.52 | 0.63 | 0.57 | 154 | | 0.47 | 0.63 | 0.54 | 154 | | 0.46 | 0.63 | 0.53 | 154 | |
| 9 (Wr) | 0.85 | 0.83 | 0.84 | 519 | | 0.87 | 0.89 | 0.88 | 519 | | 0.88 | 0.89 | 0.89 | 519 | | 0.88 | 0.87 | 0.88 | 519 | |
| accuracy | | | 0.90 | 17973 | | | | 0.93 | 17973 | | | | 0.93 | 17973 | | | | 0.93 | 17973 | |
| macro avg | 0.83 | 0.81 | 0.82 | 17973 | | 0.87 | 0.88 | 0.87 | 17973 | | 0.87 | 0.89 | 0.87 | 17973 | | 0.86 | 0.88 | 0.87 | 17973 | |
| weighted avg | 0.90 | 0.90 | 0.90 | 17973 | | 0.93 | 0.93 | 0.93 | 17973 | | 0.93 | 0.93 | 0.93 | 17973 | | 0.93 | 0.93 | 0.93 | 17973 | |

5.6 SUMMARY

This chapter aimed to evaluate the effectiveness of various spatiotemporal deep learning models for determining land cover on a broad scale in Reunion Island. The proposed framework achieves higher accuracy levels to demonstrate that the LCC capabilities of CNN (spatial) and LSTM (temporal) algorithms are rather limited. Overall, the proposed framework with the inclusion of LSTM outperformed the GRU, TempCNN, and variations to these models with integration of attention mechanisms in the tropical environment. Overall, the univariate, multivariate, and pixel coordinate model produced the best results in the study area considered, driving the use of these models for determining the landcover in the tropical setting more encouraging by decreasing the work of employing human interpreters in identifying LULC over large areas.

CHAPTER 6

CONCLUSIONS AND FUTURE SCOPE

This chapter presented the findings from the research tasks completed for the Ph.D. and included in this thesis. It summarizes the results and discusses how the work will progress in the future.

The main topic of this dissertation was the application of medium and high-resolution images obtained by remote sensing instruments, particularly optical multitemporal and time-series data. Today, it is evident that there is a tendency towards an increase in the accessibility of remote sensing images. The thesis concentrated on creating new paradigms and cutting-edge methodologies to process images retrieved from lengthy time series effectively. CD methods were investigated in this context to glean valuable information from the enormous collection of these images. Moreover, techniques for pre-processing optical multispectral images have received attention for executing successive automatic analyses of the changes.

The focus of the investigation was the issue of effectively detecting minute changes in the land cover datasets occurring at different instants of time and proper rehabilitation of pixels for the generation of the final change map. The thesis was more specifically focused on (i) building novel CD methods and unsupervised/supervised processes for evaluating and enhancing CD outcomes that benefit from the availability of spatiotemporal data and (ii) filtering out unwanted pixels and restoring relevant pixels in optical images to lessen their impact on subsequent processing and change analysis in multi-temporal data. Several images taken over the same geographic area are utilized using

the proposed approaches.

The most recent automatic and unsupervised methods for CD are thoroughly analyzed in Chapter 2. In particular, the multitemporal CD approaches and bitemporal CD techniques, which each need a pair of images or a whole time series of satellite images as input, were both investigated and discussed in this chapter. The shortcomings and difficulties of cutting-edge techniques have been examined and discussed. A gap in the literature exists regarding the definition of unsupervised and automatic CD techniques based on all of the temporal data from extended image time series. No method uses the pairwise computed change maps temporal correlation in multi-temporal images. Additionally, only some methods proposed in the literature explicitly address the temporal uniformity of changes.

The dissertation presents three crucial advancements to the SOTA: The first contribution (Chapter 3) is a novel CD method for self-producing training data to detect changes in small regions or areas based on deep learning for medium and very high spatial resolution images. The lightweight deep learning framework is trained using the changed and unchanged pixels produced from the image patches of the bi-temporal images using superpixel segmentation and parallel FCM clustering. The intermediate pixels were utilized to extract the semantic information from the feature representations to classify and build the final change map. Experiments on the two datasets affirm the viability and usefulness of the proposed technique. The bitemporal image patches overcame the issue of limited training samples to train and test the network with self-supervision, which reduced the possibility of errors in generating the accurate final change map. The performance assessment of the proposed method is being evaluated using five current SOTA CD methods with the inclusion of deep learning techniques used for the comparative analysis.

The second contribution (Chapter 4) attempted to apply deep learning methods for LULCCD and analysis in multitemporal remote sensing images. To evaluate the bitemporal images spatially, a novel dual-branch encoder-decoder network is built with operations like downsampling with strided convolution for the encoder and upsampling with transpose convolution for the decoder. The analysis is more effective for each

LULC class by including an attention module(self, spatial, and channel) at each stage in the decoding network to analyze the nature, trend, and pattern changes in the multi-temporal data. From 2005 to 2018, the model successfully assessed nearly all changes in the 15 LULC classes. By qualitative comparisons with proposed CD methods that are now considered SOTA, the effectiveness of our methodology using was tested using a dataset of a small region like Dakshina Kannada. The experiments demonstrated that the model effectively accounted for multiclass changes in each LULC class to provide change maps. Since LULC serves as a critical element for climate change, environmental monitoring, etc., specific changes in the LULC classes may impact the environment and contribute to making effective planning related to land monitoring and assessment of nature.

In the third contribution (Chapter 5), utilizing the inherent properties of SITS, an investigation of the feasibility of using spatiotemporal deep learning models for LCC was assessed. This research considered eight deep learning models: GRU, temporal CNN, GRU + temporal CNN, attention on temporal CNN, attention on temporal CNN + GRU, multivariate + univariate + pixel coordinate with its three variants, and the inclusion of LSTM for multivariate data. It was found that, for the majority of land cover classes, the complementary spatiotemporal information extracted by the proposed model (univariate + multivariate + pixel coordinates) from the SITS significantly increased the accuracy of the model in the classification of the land cover classes in contrast to the GRU and temporal CNN, which are only made to focus on temporal features. The evaluated spatiotemporal models, exceptionally well-trained on regional data, could distinguish between the various land cover classes. When viewed in the context of the global assessment and the long-term goal for the sustainable development goals, the size and scope of this study are valuable for tracking significant changes in land cover classes and forest areas.

6.1 FUTURE SCOPE

Many deep learning-based CD applications have shown that these techniques have had considerable success in the remote sensing community's field of CD. Further studies

might be carried out to address existing problems and look into intriguing subjects brought up throughout the Ph.D. to expand the research activities completed in the thesis. Yet, there are numerous difficulties with the procedures, and these include the following:

- Numerous platforms and sensors for the timely acquisitions of remote sensing data are being developed, and with them come several complex challenges, including high dimensionality in the datasets (spatial resolution, complex features of multispectral and hyperspectral images), complicated data structures (preservation of nonlinearity and overlapping distributions which may occur in the data), and the nonlinear optimization difficulty (increased computational complexity). The complexity of multi-source data significantly aggravates the challenge of developing solid and discriminative representations from training data using deep learning techniques. These might be considered severe data processing difficulties for heterogeneous sources of remote sensing datasets based on the type of sensors deployed.
- Massive training samples, often acquired through labor and time intensive procedures like field surveys and human interpretation of remote sensing results, are necessary for supervised deep learning techniques. With few training samples, building a reliable model for deep learning-based methods is extremely difficult. Techniques for unsupervised deep learning must be developed for effective CD and analysis of the study area considered for evaluation by the researchers in future years.
- As discussed in chapter 2, numerous effective and precise deep-learning models and frameworks are proposed for this work. Researchers currently repeatedly suggest novel deep learning-based CD algorithms. Yet, selecting an effective one and ensuring its correctness for various purposes can be challenging. In real-world applications, it is essential to consider the robustness of deep learning algorithms proposed for remote sensing CD.

BIBLIOGRAPHY

- Abbas Khan, K., Zaman, K., Shoukry, A. M., Sharkawy, A., Gani, S., Ahmad, J., Khan, A. and Hishan, S. S. (2019). “Natural disasters and economic losses: controlling external migration, energy and environmental resources, water demand, and financial development for global prosperity.” *Environmental Science and Pollution Research*, 26, 14287–14299.
- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P. and Süsstrunk, S. (2012). “Slic superpixels compared to state-of-the-art superpixel methods.” *IEEE transactions on pattern analysis and machine intelligence*, 34(11), 2274–2282.
- Albert, A., Kaur, J. and Gonzalez, M. C. (2017). “Using convolutional networks and satellite imagery to identify patterns in urban environments at a large scale.” In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, 1357–1366.
- Aminikhanghahi, S. and Cook, D. J. (2017). “A survey of methods for time series change point detection.” *Knowledge and information systems*, 51(2), 339–367.
- Antrop, M. (2009). “Land-use changes affected by urban and industrial development.” *Land Use, Land Cover and Soil Sciences-Volume I: Land Cover, Land Use and the Global Change*, 169.
- Attioui, S. and Najah, S. (2021). “Unsupervised change detection method in sar images based on deep belief network using an improved fuzzy c-means clustering algorithm.” *IET Image Processing*.
- Bai, S., Kolter, J. Z. and Koltun, V. (2018). “An empirical evaluation of generic

BIBLIOGRAPHY

- convolutional and recurrent networks for sequence modeling.” *arXiv preprint arXiv:1803.01271*.
- Bai, Y., Zhao, Y., Shao, Y., Zhang, X. and Yuan, X. (2022). “Deep learning in different remote sensing image categories and applications: status and prospects.” *International Journal of Remote Sensing*, 43(5), 1800–1847.
- Bandara, W. G. C. and Patel, V. M. (2022). “A transformer-based siamese network for change detection.” *arXiv preprint arXiv:2201.01293*.
- Barrett, E. C. (2013). *Introduction to environmental remote sensing*, Routledge.
- Baudhuin, H., Lambot, A. and Schaus, P. “” change detection in satellite imagery using deep learning.” .
- Binkowski, M., Marti, G. and Donnat, P. (2018). “Autoregressive convolutional neural networks for asynchronous time series.” In *International Conference on Machine Learning*, PMLR, 580–589.
- Borovykh, A., Bohte, S. and Oosterlee, C. W. (2017). “Conditional time series forecasting with convolutional neural networks.” *arXiv preprint arXiv:1703.04691*.
- Bovolo, F., Marchesi, S. and Bruzzone, L. (2011). “A framework for automatic and unsupervised detection of multiple changes in multitemporal images.” *IEEE Transactions on Geoscience and Remote Sensing*, 50(6), 2196–2212.
- Brauwers, G. and Frasincar, F. (2021). “A general survey on attention mechanisms in deep learning.” *IEEE Transactions on Knowledge and Data Engineering*.
- Brock, J. and Abdallah, Z. S. (2022). “Investigating temporal convolutional neural networks for satellite image time series classification.” *arXiv preprint arXiv:2204.08461*.
- Bruzzone, L., Bovolo, F., Paris, C., Solano-Correa, Y. T., Zanetti, M. and Fernández-Prieto, D. (2017). “Analysis of multitemporal sentinel-2 images in the framework of the esa scientific exploitation of operational missions.” In *2017 9th International*

- Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, IEEE, 1–4.
- Bruzzone, L. and Cossu, R. (2003). “An adaptive approach to reducing registration noise effects in unsupervised change detection.” *IEEE Transactions on Geoscience and Remote Sensing*, 41(11), 2455–2465.
- Camilleri, S., De Giglio, M., Stecchi, F. and Pérez-Hurtado, A. (2017). “Land use and land cover change analysis in predominantly man-made coastal wetlands: towards a methodological framework.” *Wetlands Ecology and Management*, 25(1), 23–43.
- Campos-Taberner, M., García-Haro, F. J., Martínez, B., Izquierdo-Verdiguier, E., Atzberger, C., Camps-Valls, G. and Gilabert, M. A. (2020). “Understanding deep learning in land use classification based on sentinel-2 time series.” *Scientific reports*, 10(1), 1–12.
- Chen, Z., Zhou, Y., Wang, B., Xu, X., He, N., Jin, S. and Jin, S. (2022). “Egde-net: A building change detection method for high-resolution remote sensing imagery based on edge guidance and differential enhancement.” *ISPRS Journal of Photogrammetry and Remote Sensing*, 191, 203–222.
- Cheng, G., Yang, C., Yao, X., Guo, L. and Han, J. (2018a). “When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative cnns.” *IEEE transactions on geoscience and remote sensing*, 56(5), 2811–2821.
- Cheng, H., Wu, H., Zheng, J., Qi, K. and Liu, W. (2021). “A hierarchical self-attention augmented laplacian pyramid expanding network for change detection in high-resolution remote sensing images.” *ISPRS Journal of Photogrammetry and Remote Sensing*, 182, 52–66.
- Cheng, L., Zang, H., Ding, T., Sun, R., Wang, M., Wei, Z. and Sun, G. (2018b). “Ensemble recurrent neural network based probabilistic wind speed forecasting approach.” *Energies*, 11(8), 1958.

- Coppin, P., Lambin, E., Jonckheere, I. and Muys, B. (2002). “Digital change detection methods in natural ecosystem monitoring: A review.” *Analysis of multi-temporal remote sensing images*, 3–36.
- Coppin, P. R. and Bauer, M. E. (1994). “Processing of multitemporal landsat tm imagery to optimize extraction of forest cover change features.” *IEEE Transactions on Geoscience and remote Sensing*, 32(4), 918–927.
- Coppin, P. R. and Bauer, M. E. (1996). “Digital change detection in forest ecosystems with remote sensing imagery.” *Remote sensing reviews*, 13(3-4), 207–234.
- Cui, Z., Chen, W. and Chen, Y. (2016). “Multi-scale convolutional neural networks for time series classification.” *arXiv preprint arXiv:1603.06995*.
- Dai, X. and Khorram, S. (1998). “The effects of image misregistration on the accuracy of remotely sensed change detection.” *IEEE Transactions on Geoscience and Remote sensing*, 36(5), 1566–1577.
- Dai, X. and Khorram, S. (1999). “Data fusion using artificial neural networks: a case study on multitemporal change analysis.” *Computers, Environment and Urban Systems*, 23(1), 19–31.
- Deilami, B. R., Ahmad, B. B., Saffar, M. R. and Umar, H. Z. (2015). “Review of change detection techniques from remotely sensed images.” *Research Journal of Applied Sciences, Engineering and Technology*, 10(2), 221–229.
- Dong, H., Ma, W., Wu, Y., Zhang, J. and Jiao, L. (2020). “Self-supervised representation learning for remote sensing image change detection based on temporal prediction.” *Remote Sensing*, 12(11), 1868.
- Du, B., Ru, L., Wu, C. and Zhang, L. (2019). “Unsupervised deep slow feature analysis for change detection in multi-temporal remote sensing images.” *IEEE Transactions on Geoscience and Remote Sensing*, 57(12), 9976–9992.

- Du, H., Zhuang, Y., Dong, S., Li, C., Chen, H., Zhao, B. and Chen, L. (2021). “Bilateral semantic fusion siamese network for change detection from multitemporal optical remote sensing imagery.” *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5.
- Dvornek, N. C., Ventola, P., Pelphrey, K. A. and Duncan, J. S. (2017). “Identifying autism from resting-state fmri using long short-term memory networks.” In *International Workshop on Machine Learning in Medical Imaging*, Springer, 362–370.
- Ekim, B. and Sertel, E. (2021). “Deep neural network ensembles for remote sensing land cover and land use classification.” *International Journal of Digital Earth*, 1–14.
- Fan, C., Zhang, Y., Pan, Y., Li, X., Zhang, C., Yuan, R., Wu, D., Wang, W., Pei, J. and Huang, H. (2019). “Multi-horizon time series forecasting with temporal attention learning.” In *Proceedings of the 25th ACM SIGKDD International conference on knowledge discovery & data mining*, 2527–2535.
- Fang, H., Du, P., Wang, X., Lin, C. and Tang, P. (2021). “Unsupervised change detection based on weighted change vector analysis and improved markov random field for high spatial resolution imagery.” *IEEE Geoscience and Remote Sensing Letters*.
- Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L. and Muller, P.-A. (2018a). “Data augmentation using synthetic data for time series classification with deep residual networks.” *arXiv preprint arXiv:1808.02455*.
- Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L. and Muller, P.-A. (2018b). “Transfer learning for time series classification.” In *2018 IEEE international conference on big data (Big Data)*, IEEE, 1367–1376.
- Gao, F., Dong, J., Li, B. and Xu, Q. (2016). “Automatic change detection in synthetic aperture radar images based on pcanet.” *IEEE Geoscience and Remote Sensing Letters*, 13(12), 1792–1796.
- Garnot, V. S. F. and Landrieu, L. (2020). “Lightweight temporal self-attention for classifying satellite images time series.” In *International Workshop on Advanced Analytics and Learning on Temporal Data*, Springer, 171–181.

- Geng, J., Ma, X., Zhou, X. and Wang, H. (2019). “Saliency-guided deep neural networks for sar image change detection.” *IEEE Transactions on Geoscience and Remote Sensing*, 57(10), 7365–7377.
- Geng, Y. and Luo, X. (2018). “Cost-sensitive convolution based neural networks for imbalanced time-series classification.” *arXiv preprint arXiv:1801.04396*.
- Ghaderi, A., Sanandaji, B. M. and Ghaderi, F. (2017). “Deep forecast: Deep learning-based spatio-temporal forecasting.” *arXiv preprint arXiv:1707.08110*.
- Gharbia, R., Khalifa, N. E. M. and Hassanien, A. E. (2020). “Land cover classification using deep convolutional neural networks.” In *International Conference on Intelligent Systems Design and Applications*, Springer, 911–920.
- Glorot, X. and Bengio, Y. (2010). “Understanding the difficulty of training deep feed-forward neural networks.” In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, JMLR Workshop and Conference Proceedings, 249–256.
- Gong, M., Zhan, T., Zhang, P. and Miao, Q. (2017). “Superpixel-based difference representation learning for change detection in multispectral remote sensing images.” *IEEE Transactions on Geoscience and Remote Sensing*, 55(5), 2658–2673.
- Gong, Z., Chen, H., Yuan, B. and Yao, X. (2018). “Multiobjective learning in the model space for time series classification.” *IEEE transactions on cybernetics*, 49(3), 918–932.
- Grover, A., Kapoor, A. and Horvitz, E. (2015). “A deep hybrid model for weather forecasting.” In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, 379–386.
- Guo, Q., Zhang, J., Zhong, C. and Zhang, Y. (2021). “Unsupervised multiple change detection for multispectral images based on ammf and spatiospectral channel augmentation.” *IEEE Geoscience and Remote Sensing Letters*.

- Hall, F. G., Strebel, D. E., Nickeson, J. E. and Goetz, S. J. (1991). “Radiometric rectification: toward a common radiometric response among multitemporal, multisensor images.” *Remote sensing of environment*, 35(1), 11–27.
- Hansen, M. C., Potapov, P. V., Pickens, A. H., Tyukavina, A., Hernandez-Serna, A., Zalles, V., Turubanova, S., Kommareddy, I., Stehman, S. V., Song, X.-P. et al. (2022). “Global land use extent and dispersion within natural land cover using landsat data.” *Environmental Research Letters*, 17(3), 034050.
- Hatami, N., Gavet, Y. and Debayle, J. (2018). “Classification of time-series images using deep convolutional neural networks.” In *Tenth international conference on machine vision (ICMV 2017)*, volume 10696, SPIE, 242–249.
- He, P., Zhao, X., Shi, Y. and Cai, L. (2021). “Unsupervised change detection from remotely sensed images based on multi-scale visual saliency coarse-to-fine fusion.” *Remote Sensing*, 13(4), 630.
- He, T. and Wang, S. (2021). “Multi-spectral remote sensing land-cover classification based on deep learning methods.” *The Journal of Supercomputing*, 77(3), 2829–2843.
- Huang, H., Hu, X., Zhao, Y., Makkie, M., Dong, Q., Zhao, S., Guo, L. and Liu, T. (2017). “Modeling task fmri data via deep convolutional autoencoder.” *IEEE transactions on medical imaging*, 37(7), 1551–1561.
- Hussain, M., Chen, D., Cheng, A., Wei, H. and Stanley, D. (2013). “Change detection from remotely sensed images: From pixel-based to object-based approaches.” *ISPRS Journal of photogrammetry and remote sensing*, 80, 91–106.
- Im, J. and Jensen, J. R. (2005). “A change detection model based on neighborhood correlation image analysis and decision tree classification.” *Remote Sensing of Environment*, 99(3), 326–340.
- Ingram, K., Knapp, E. and Robinson, J. (1981). “Change detection technique development for improved urbanized area delineation.” *NASA, Comput. Sci. Corp., Springfield, MD, CSC/TM-81/6087*.

- Jakka, T. K., Reddy, Y. M. and Rao, B. P. (2019). “Gwdwt-fcm: change detection in sar images using adaptive discrete wavelet transform with fuzzy c-mean clustering.” *Journal of the Indian Society of Remote Sensing*, 47(3), 379–390.
- Janssen, L. L. and Vanderwel, F. J. (1994). “Accuracy assessment of satellite derived land-cover data: a review.” *Photogrammetric engineering and remote sensing;(United States)*, 60(4).
- Jensen, J. R. (2009). *Remote sensing of the environment: An earth resource perspective 2/e*, Pearson Education India.
- Jian, P., Chen, K. and Cheng, W. (2021). “Gan-based one-class classification for remote-sensing image change detection.” *IEEE Geoscience and Remote Sensing Letters*.
- Kalinicheva, E., Sublime, J. and Trocan, M. (2020). “Unsupervised satellite image time series clustering using object-based approaches and 3d convolutional autoencoder.” *Remote Sensing*, 12(11), 1816.
- Kingma, D. P. and Ba, J. (2014). “Adam: A method for stochastic optimization.” *arXiv preprint arXiv:1412.6980*.
- Kleynhans, W., Salmon, B. P., Olivier, J. C., Van den Bergh, F., Wessels, K. J., Grobler, T. L. and Steenkamp, K. C. (2012). “Land cover change detection using autocorrelation analysis on modis time-series data: Detection of new human settlements in the gauteng province of south africa.” *IEEE Journal of selected topics in applied earth observations and remote sensing*, 5(3), 777–783.
- Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2017). “Imagenet classification with deep convolutional neural networks.” *Communications of the ACM*, 60(6), 84–90.
- Lei, Y., Liu, X., Shi, J., Lei, C. and Wang, J. (2019). “Multiscale superpixel segmentation with deep features for change detection.” *Ieee Access*, 7, 36600–36616.
- Li, R., Shahn, Z., Li, J., Lu, M., Chakraborty, P., Sow, D., Ghalwash, M. and Lehman, L.-w. H. (2020). “G-net: a deep learning approach to g-computation for coun-

- terfactual outcome prediction under dynamic treatment regimes.” *arXiv preprint arXiv:2003.10551*.
- Li, S., Jin, X., Xuan, Y., Zhou, X., Chen, W., Wang, Y.-X. and Yan, X. (2019a). “Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting.” *Advances in neural information processing systems*, 32.
- Li, Y., Peng, C., Chen, Y., Jiao, L., Zhou, L. and Shang, R. (2019b). “A deep learning method for change detection in synthetic aperture radar images.” *IEEE Transactions on Geoscience and Remote Sensing*, 57(8), 5751–5763.
- Li, Y., Yu, R., Shahabi, C. and Liu, Y. (2017). “Diffusion convolutional recurrent neural network: Data-driven traffic forecasting.” *arXiv preprint arXiv:1707.01926*.
- Liang, J., Xu, J., Shen, H. and Fang, L. (2020). “Land-use classification via constrained extreme learning classifier based on cascaded deep convolutional neural networks.” *European Journal of Remote Sensing*, 53(1), 219–232.
- Liao, B., Zhang, J., Wu, C., McIlwraith, D., Chen, T., Yang, S., Guo, Y. and Wu, F. (2018). “Deep sequence learning with auxiliary information for traffic prediction.” In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 537–546.
- Lim, B. (2018). “Forecasting treatment responses over time using recurrent marginal structural networks.” *advances in neural information processing systems*, 31.
- Lim, B., Arık, S. Ö., Loeff, N. and Pfister, T. (2021). “Temporal fusion transformers for interpretable multi-horizon time series forecasting.” *International Journal of Forecasting*, 37(4), 1748–1764.
- Lim, B., Zohren, S. and Roberts, S. (2019). “Enhancing time-series momentum strategies using deep neural networks.” *The Journal of Financial Data Science*, 1(4), 19–38.

- Lim, B., Zohren, S. and Roberts, S. (2020). “Recurrent neural filters: Learning independent bayesian filtering steps for time series prediction.” In *2020 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 1–8.
- Liu, C.-L., Hsaio, W.-H. and Tu, Y.-C. (2018). “Time series classification with multivariate convolutional neural network.” *IEEE Transactions on Industrial Electronics*, 66(6), 4788–4797.
- Liu, H. and Zhou, Q. (2004). “Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison.” *International Journal of Remote Sensing*, 25(5), 1037–1050.
- Liu, S. and Shi, Q. (2020). “Local climate zone mapping as remote sensing scene classification using deep learning: A case study of metropolitan china.” *ISPRS Journal of Photogrammetry and Remote Sensing*, 164, 229–242.
- Lu, D., Mausel, P., Brondizio, E. and Moran, E. (2004). “Change detection techniques.” *International journal of remote sensing*, 25(12), 2365–2401.
- Lv, N., Chen, C., Qiu, T. and Sangaiah, A. K. (2018). “Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in sar images.” *IEEE transactions on industrial informatics*, 14(12), 5530–5538.
- Lv, Y., Duan, Y., Kang, W., Li, Z. and Wang, F.-Y. (2014). “Traffic flow prediction with big data: a deep learning approach.” *IEEE Transactions on Intelligent Transportation Systems*, 16(2), 865–873.
- Lv, Z., Liu, T., Benediktsson, J. A. and Falco, N. (2021). “Land cover change detection techniques: Very-high-resolution optical images: A review.” *IEEE Geoscience and Remote Sensing Magazine*, 10(1), 44–63.
- Lyu, H., Lu, H., Mou, L., Li, W., Wright, J., Li, X., Li, X., Zhu, X. X., Wang, J., Yu, L. et al. (2018). “Long-term annual mapping of four cities on different continents by applying a deep information learning method to landsat data.” *Remote Sensing*, 10(3), 471.

- Malhotra, P., TV, V., Vig, L., Agarwal, P. and Shroff, G. (2017). “Timenet: Pre-trained deep recurrent neural network for time series classification.” *arXiv preprint arXiv:1706.08838*.
- Martinez, C., Perrin, G., Ramasso, E. and Rombaut, M. (2018). “A deep reinforcement learning approach for early classification of time series.” In *2018 26th European Signal Processing Conference (EUSIPCO)*, IEEE, 2030–2034.
- Mehdiyev, N., Lahann, J., Emrich, A., Enke, D., Fettke, P. and Loos, P. (2017). “Time series classification using deep learning for process planning: A case from the process industry.” *Procedia Computer Science*, 114, 242–249.
- Meng, W., Wang, L., Du, A. and Li, Y. (2020). “Sar image change detection based on data optimization and self-supervised learning.” *IEEE Access*, 8, 217290–217305.
- Mittelman, R. (2015). “Time-series modeling with undecimated fully convolutional neural networks.” *arXiv preprint arXiv:1508.00317*.
- Mnih, V., Heess, N., Graves, A. et al. (2014). “Recurrent models of visual attention.” *Advances in neural information processing systems*, 27.
- MohanRajan, S. N., Loganathan, A. and Manoharan, P. (2020). “Survey on land use/land cover (lu/lc) change analysis in remote sensing and gis environment: Techniques and challenges.” *Environmental Science and Pollution Research*, 27(24), 29900–29926.
- Mulla, D. J. (2013). “Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps.” *Biosystems engineering*, 114(4), 358–371.
- Nedd, R., Light, K., Owens, M., James, N., Johnson, E. and Anandhi, A. (2021). “A synthesis of land use/land cover studies: Definitions, classification systems, meta-studies, challenges and knowledge gaps on a global landscape.” *Land*, 10(9), 994.
- Ngoy, K. I., Qi, F. and Shebitz, D. J. (2021). “Analyzing and predicting land use and land cover changes in new jersey using multi-layer perceptron–markov chain model.” *Earth*, 2(4), 845–870.

- NRSC, H. (2010). “National Land Use and Land Cover Mapping Using Multi-Temporal AWiFS Data.”).
- Ojima, D., Galvin, K. and Turner, B. (1994). “The global impact of land-use change.” *BioScience*, 44(5), 300–304.
- Olson, G. A., Cheriyyadat, A., Mali, P. and O’Hara, C. G. (2004). “Detecting and managing change in spatial data-land use and infrastructure change analysis and detection.” In *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*, volume 2, IEEE, 729–734.
- Pandeeswari, B., Sutha, J. and Parvathy, M. (2021). “A novel synthetic aperture radar image change detection system using radial basis function-based deep convolutional neural network.” *Journal of Ambient Intelligence and Humanized Computing*, 12, 897–910.
- Pandey, P. C., Koutsias, N., Petropoulos, G. P., Srivastava, P. K. and Ben Dor, E. (2021). “Land use/land cover in view of earth observation: data sources, input dimensions, and classifiers—a review of the state of the art.” *Geocarto International*, 36(9), 957–988.
- Panuju, D. R., Paull, D. J. and Griffin, A. L. (2020). “Change detection techniques based on multispectral images for investigating land cover dynamics.” *Remote Sensing*, 12(11), 1781.
- Peng, D., Bruzzone, L., Zhang, Y., Guan, H., Ding, H. and Huang, X. (2020). “Semicd-net: A semisupervised convolutional neural network for change detection in high resolution remote-sensing images.” *IEEE Transactions on Geoscience and Remote Sensing*.
- Peng, D., Bruzzone, L., Zhang, Y., Guan, H. and He, P. (2021). “Scdnet: A novel convolutional network for semantic change detection in high resolution optical remote sensing imagery.” *International Journal of Applied Earth Observation and Geoinformation*, 103, 102465.

- Penghua, Z. and Dingyi, Z. (2019). “Bidirectional-gru based on attention mechanism for aspect-level sentiment analysis.” In *Proceedings of the 2019 11th International Conference on Machine Learning and Computing*, 86–90.
- Petitjean, F., Inglada, J. and Gançarski, P. (2012). “Satellite image time series analysis under time warping.” *IEEE transactions on geoscience and remote sensing*, 50(8), 3081–3095.
- Radke, R. J., Andra, S., Al-Kofahi, O. and Roysam, B. (2005). “Image change detection algorithms: a systematic survey.” *IEEE transactions on image processing*, 14(3), 294–307.
- Rajesh, S., Nisia, T. G., Arivazhagan, S. and Abisekaraj, R. (2020). “Land cover/land use mapping of liss iv imagery using object-based convolutional neural network with deep features.” *Journal of the Indian Society of Remote Sensing*, 48(1), 145–154.
- Ramachandran, P., Parmar, N., Vaswani, A., Bello, I., Levskaya, A. and Shlens, J. (2019). “Stand-alone self-attention in vision models.” *arXiv preprint arXiv:1906.05909*.
- Rangapuram, S. S., Seeger, M. W., Gasthaus, J., Stella, L., Wang, Y. and Januschowski, T. (2018). “Deep state space models for time series forecasting.” *Advances in neural information processing systems*, 31.
- Robinove, C. J. (1982). “Computation with physical values from landsat digital data.” *Photogrammetric Engineering and Remote Sensing*, 48(5), 781–784.
- Rodrigues, F., Markou, I. and Pereira, F. C. (2019). “Combining time-series and textual data for taxi demand prediction in event areas: A deep learning approach.” *Information Fusion*, 49, 120–129.
- Rousset, G., Despinoy, M., Schindler, K. and Mangeas, M. (2021). “Assessment of deep learning techniques for land use land cover classification in southern new caledonia.” *Remote Sensing*, 13(12), 2257.

- Saha, S., Bovolo, F. and Bruzzone, L. (2020a). “Change detection in image time-series using unsupervised lstm.” *IEEE Geoscience and Remote Sensing Letters*.
- Saha, S., Mou, L., Zhu, X. X., Bovolo, F. and Bruzzone, L. (2020b). “Semisupervised change detection using graph convolutional network.” *IEEE Geoscience and Remote Sensing Letters*, 18(4), 607–611.
- Salinas, D., Bohlke-Schneider, M., Callot, L., Medico, R. and Gasthaus, J. (2019). “High-dimensional multivariate forecasting with low-rank gaussian copula processes.” *Advances in neural information processing systems*, 32.
- Santos, T. and Kern, R. (2016). “A literature survey of early time series classification and deep learning.” *Sami@ iknow*.
- Sawaya, K. E., Olmanson, L. G., Heinert, N. J., Brezonik, P. L. and Bauer, M. E. (2003). “Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery.” *Remote sensing of Environment*, 88(1-2), 144–156.
- Sefrin, O., Riese, F. M. and Keller, S. (2021). “Deep learning for land cover change detection.” *Remote Sensing*, 13(1), 78.
- Sen, R., Yu, H.-F. and Dhillon, I. S. (2019). “Think globally, act locally: A deep neural network approach to high-dimensional time series forecasting.” *Advances in neural information processing systems*, 32.
- Shi, J., Zhang, Z., Tan, C., Liu, X. and Lei, Y. (2021). “Unsupervised multiple change detection in remote sensing images via generative representation learning network.” *IEEE Geoscience and Remote Sensing Letters*.
- Shi, W., Zhang, M., Zhang, R., Chen, S. and Zhan, Z. (2020). “Change detection based on artificial intelligence: State-of-the-art and challenges.” *Remote Sensing*, 12(10), 1688.
- Shu, Y., Li, W., Yang, M., Cheng, P. and Han, S. (2021). “Patch-based change detection method for sar images with label updating strategy.” *Remote Sensing*, 13(7), 1236.

- Siddiqui, S. A., Mercier, D., Munir, M., Dengel, A. and Ahmed, S. (2019). “Tsviz: Demystification of deep learning models for time-series analysis.” *IEEE Access*, 7, 67027–67040.
- Simonyan, K., Vedaldi, A. and Zisserman, A. (2013). “Deep inside convolutional networks: Visualising image classification models and saliency maps.” *arXiv preprint arXiv:1312.6034*.
- Singh, A. (1989). “Review article digital change detection techniques using remotely-sensed data.” *International journal of remote sensing*, 10(6), 989–1003.
- Smyl, S. (2020). “A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting.” *International Journal of Forecasting*, 36(1), 75–85.
- Solano-Correa, Y. T., Bovolo, F., Bruzzone, L. and Fernández-Prieto, D. (2017). “Spatio-temporal evolution of crop fields in sentinel-2 satellite image time series.” In *2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, IEEE, 1–4.
- Solano-Correa, Y. T., Bovolo, F., Bruzzone, L. and Fernández-Prieto, D. (2018). “Automatic derivation of cropland phenological parameters by adaptive non-parametric regression of sentinel-2 ndvi time series.” In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 1946–1949.
- Song, W., Liu, L., Liu, M., Wang, W., Wang, X. and Song, Y. (2020). “Representation learning with deconvolution for multivariate time series classification and visualization.” In *International Conference of Pioneering Computer Scientists, Engineers and Educators*, Springer, 310–326.
- Soua, R., Koesdwiady, A. and Karray, F. (2016). “Big-data-generated traffic flow prediction using deep learning and dempster-shafer theory.” In *2016 International joint conference on neural networks (IJCNN)*, IEEE, 3195–3202.

- Tewkesbury, A. P., Comber, A. J., Tate, N. J., Lamb, A. and Fisher, P. F. (2015). “A critical synthesis of remotely sensed optical image change detection techniques.” *Remote Sensing of Environment*, 160, 1–14.
- Townshend, J. R., Justice, C. O., Gurney, C. and McManus, J. (1992). “The impact of misregistration on change detection.” *IEEE Transactions on Geoscience and remote sensing*, 30(5), 1054–1060.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. and Polosukhin, I. (2017). “Attention is all you need.” *Advances in neural information processing systems*, 30.
- Verbesselt, J., Hyndman, R., Newnham, G. and Culvenor, D. (2010). “Detecting trend and seasonal changes in satellite image time series.” *Remote sensing of Environment*, 114(1), 106–115.
- Verma, D. and Jana, A. (2019). “Lulc classification methodology based on simple convolutional neural network to map complex urban forms at finer scale: Evidence from mumbai.” *arXiv preprint arXiv:1909.09774*.
- Wang, D., Chen, X., Jiang, M., Du, S., Xu, B. and Wang, J. (2021). “Ads-net: An attention-based deeply supervised network for remote sensing image change detection.” *International Journal of Applied Earth Observation and Geoinformation*, 101, 102348.
- Wang, H., Zhao, X., Zhang, X., Wu, D. and Du, X. (2019a). “Long time series land cover classification in china from 1982 to 2015 based on bi-lstm deep learning.” *Remote Sensing*, 11(14), 1639.
- Wang, L., Wang, Z. and Liu, S. (2016a). “An effective multivariate time series classification approach using echo state network and adaptive differential evolution algorithm.” *Expert Systems with Applications*, 43, 237–249.
- Wang, S., Hua, G., Hao, G. and Xie, C. (2017). “A cycle deep belief network model for multivariate time series classification.” *Mathematical Problems in Engineering*, 2017.

- Wang, W., Chen, C., Wang, W., Rai, P. and Carin, L. (2016b). “Earliness-aware deep convolutional networks for early time series classification.” *arXiv preprint arXiv:1611.04578*.
- Wang, Y., Smola, A., Maddix, D., Gasthaus, J., Foster, D. and Januschowski, T. (2019b). “Deep factors for forecasting.” In *International conference on machine learning*, PMLR, 6607–6617.
- Wang, Z. and Oates, T. (2015). “Spatially encoding temporal correlations to classify temporal data using convolutional neural networks.” *arXiv preprint arXiv:1509.07481*.
- Wen, R. and Torkkola, K. (2019). “Deep generative quantile-copula models for probabilistic forecasting.” *arXiv preprint arXiv:1907.10697*.
- Wen, R., Torkkola, K., Narayanaswamy, B. and Madeka, D. (2017). “A multi-horizon quantile recurrent forecaster.” *arXiv preprint arXiv:1711.11053*.
- Wijaya, H., Kurniawati, H. and Hutama, S. (2018). “Industrialization impact on worker mobility and land use in peri urban area (case study of semarang district, indonesia).” In *IOP Conference Series: Earth and Environmental Science*, volume 123, IOP Publishing, 012037.
- Woo, S., Park, J., Lee, J.-Y. and Kweon, I. S. (2018). “Cbam: Convolutional block attention module.” In *Proceedings of the European conference on computer vision (ECCV)*, 3–19.
- Xie, D., Zhang, L. and Bai, L. (2017). “Deep learning in visual computing and signal processing.” *Applied Computational Intelligence and Soft Computing*, 2017.
- Yang, J., Nguyen, M. N., San, P. P., Li, X. L. and Krishnaswamy, S. (2015). “Deep convolutional neural networks on multichannel time series for human activity recognition.” In *Twenty-fourth international joint conference on artificial intelligence*.
- Yang, X. (2020). “An overview of the attention mechanisms in computer vision.” In *Journal of Physics: Conference Series*, volume 1693, IOP Publishing, 012173.

- Zhan, T., Gong, M., Jiang, X. and Zhao, W. (2021). “Transfer learning-based bilinear convolutional networks for unsupervised change detection.” *IEEE Geoscience and Remote Sensing Letters*.
- Zhang, C., Yue, P., Tapete, D., Jiang, L., Shangguan, B., Huang, L. and Liu, G. (2020). “A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images.” *ISPRS Journal of Photogrammetry and Remote Sensing*, 166, 183–200.
- Zhang, H., Lin, M., Yang, G. and Zhang, L. (2021a). “Escnet: An end-to-end superpixel-enhanced change detection network for very-high-resolution remote sensing images.” *IEEE Transactions on Neural Networks and Learning Systems*.
- Zhang, P., Gong, M., Zhang, H., Liu, J. and Ban, Y. (2018). “Unsupervised difference representation learning for detecting multiple types of changes in multitemporal remote sensing images.” *IEEE Transactions on Geoscience and Remote Sensing*, 57(4), 2277–2289.
- Zhang, W., Li, J., Zhang, F., Sun, J. and Zhang, K. (2021b). “Unsupervised change detection of multispectral images based on pca and low-rank prior.” *IEEE Geoscience and Remote Sensing Letters*.
- Zhao, W., Chen, X., Ge, X. and Chen, J. (2020). “Using adversarial network for multiple change detection in bitemporal remote sensing imagery.” *IEEE Geoscience and Remote Sensing Letters*.
- Zheng, Y., Liu, Q., Chen, E., Ge, Y. and Zhao, J. L. (2016). “Exploiting multi-channels deep convolutional neural networks for multivariate time series classification.” *Frontiers of Computer Science*, 10(1), 96–112.
- Zhu, Q., Guo, X., Deng, W., Shi, S., Guan, Q., Zhong, Y., Zhang, L. and Li, D. (2022). “Land-use/land-cover change detection based on a siamese global learning framework for high spatial resolution remote sensing imagery.” *ISPRS Journal of Photogrammetry and Remote Sensing*, 184, 63–78.

- Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F. and Fraundorfer, F. (2017).
“Deep learning in remote sensing: A comprehensive review and list of resources.”
IEEE Geoscience and Remote Sensing Magazine, 5(4), 8–36.

PUBLICATIONS

JOURNAL PAPERS

1. Nitesh Naik, K Chandrasekaran, Andrzej Stateczny, M Venkatesan and P Prabhavathy, (2022), Spatiotemporal Assessment of Satellite Image Time Series for Land Cover Classification Using Deep Learning Techniques: A Case Study of Reunion Island, France,. (DOI: <https://doi.org/10.3390/rs14205232>, URL: <https://www.mdpi.com/2072-4292/14/20/5232>) [IF:5.6, Q1 SCIE].
2. Nitesh Naik, K Chandrasekaran, M Venkatesan and P Prabhavathy, (2023) ,Dual Attention Guided Deep Encoder-Decoder Network for Change Analysis in Land use/Land cover for Dakshina Kannada District, Karnataka, India. (DOI: <https://doi.org/10.1007/s12665-022-10713-1>, URL: <https://link.springer.com/article/10.1007/s12665-022-10713-1>) [IF:3.1, Q1 SCI, SCIE].
3. Nitesh Naik, K Chandrasekaran, M Venkatesan and P Prabhavathy, Spatio-Temporal Analysis of Land Use/Land Cover Change Detection in Small Regions using Self-Supervised Lightweight Deep Learning, (DOI:<https://doi.org/10.1007/s00477-023-02554-6>,URL: <https://link.springer.com/article/10.1007/s00477-023-02554-6>) [IF:4.2, Q1 SCI, SCIE].
4. Nitesh Naik, K Chandrasekaran, M Venkatesan and P Prabhavathy, Assessment of Land use and Land Cover Change Detection and Prediction using Deep Learning Techniques for the southwestern coastal region, Goa, India. *Environmental Monitoring and Assessment SpringerNature* [IF:3.0, Q2 SCIE] (“Accepted”).

5. Bhavesh Verma, Nitesh Naik, K Chandrasekaran, M Venkatesan and Jeny Rajan, Forecasting Land-Use and Land-Cover Change using Hybrid CNN-LSTM Model, *IEEE Geoscience and Remote Sensing Letters* [IF:4.8, Q1 SCIE] (Under Review).

CONFERENCE PAPERS

1. Nitesh Naik., K.Chandrasekaran, M.Venkatesan and P Prabhavathy (2021) Deep Learning-Based Prediction, Classification, Clustering Models for Time Series Analysis: A Systematic Review. In: Vishal Goar, Manoj Kuri, Rajesh Kumar, Tomonobu Senjyu (eds) *Advances in Information Communication Technology and Computing AICTC 2021 Lecture Notes in Networks and Systems*, vol 392. Springer, Singapore. (DOI: https://doi.org/10.1007/978-981-19-0619-0_34, URL: https://link.springer.com/chapter/10.1007/978-981-19-0619-0_34)
2. Nitesh Naik., K.Chandrasekaran, M.Venkatesan and P Prabhavathy (2021) Attention-Based Bitemporal Image Deep Feature-Level Change Detection for High Resolution Imagery. In: Rajesh Doriya, Badal Soni, Anupam Shukla, Xiao-Zhi Gao (eds) *Machine Learning, Image Processing, Network Security and Data Sciences. MIND 2021. Lecture Notes in Networks and Systems*, vol 946. Springer, Cham. (DOI: https://doi.org/10.1007/978-981-19-5868-7_20, URL: https://link.springer.com/chapter/10.1007/978-981-19-5868-7_20) (**Best Paper Award in Machine Learning Track**)

BIODATA

Name: NITESH NAVNATH NAIK

Date of Birth: 17th July 1984

Gender: Male

Marital Status: Married

Father's Name: Navnath

Mother's Name: Navita

Address : House no.29/4
Head Land Sada
Bhatwadi (P.O)
Mormugao Goa-403804

E-mail: naiknitesh22gec05@gmail.com

Mobile: +91 9823562097

Qualification: B.E in Information Technology (College of Engineering, Goa)

M.Tech in Computer Science & Engineering
(National Institute of Technology, Karnataka)

Areas of Interest: Data Science and Machine Learning