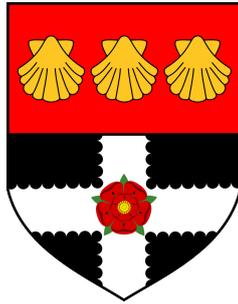


**From Images To Changes: Enhancing Synthetic Aperture Radar**

**Change Detection Via Registration and Deep Despeckling Model**



**Mohamed Ihmeida**

Department of Computer Sciences

University of Reading

This dissertation is submitted for the degree of

*Doctor of Philosophy*

April 2024



I would like to dedicate this thesis to my loving parents . . .



## **Declaration**

I declare that this is my own work, and the use of all material from other sources has been properly and fully acknowledged.

Mohamed Ihmeida

April 2024



## **Acknowledgements**

First, I am thankful to Allah Almighty, who blesses me with sound health and the ability to perform and complete my work. Without Allah's help, I could not do my work.

I would like to express my appreciation to my supervisor, Dr Muhammad Shahzad, for his support during publications and for following up with me by submission.

I would like to thank my monitoring committee sincerely, Dr Martin Lester and Dr Lily Sun, for their comments and feedback on my research throughout the PhD journey.

I would like to thank the Libyan Government and Libyan attaché for sponsoring this research.

I want to express my gratitude to my fellow PhD peer, Dr Chandresh Pravin, for accompanying me on this research journey. Our frequent and ongoing conversations about research and academic matters provided me with significant comfort. My appreciation for your support knows no bounds.

Finally, I would like to sincerely thank my parents for their prayers and support during my studies. Without their support, I would not do my PhD research. I would also like to thank my wife, children, brothers & sisters and friends for their continuous love and support.



## Abstract

Change Detection (CD) in Synthetic Aperture Radar (SAR) is an essential task in the field of Earth Observation (EO). It focuses on identifying the change for the same geographical region between two SAR images acquired at different times. SAR offers several advantages over optical sensors. For instance, spaceborne SAR sensors are able to provide day/night capability to map the globe in virtually all weather conditions. Moreover, SAR's microwave signals can pass through the cloud cover, allowing it to acquire data and generate images even in the presence of clouds, fog and dust. Despite these advantages of SAR, CD in SAR remains a highly challenging problem due to the misregistration of multi-temporal SAR images and speckle noise. Both these challenges adversely affect the performance of SAR-based CD techniques. In this thesis research, we have thoroughly discussed these challenges and have proposed novel solutions to improve the overall performance of SAR-based CD algorithms. For instance, we have proposed a deep neural network-based despeckling model (DM) that effectively suppresses speckle noise and enhances the performance of the existing CD methods. Specifically, the proposed despeckling methodology consists of two modules where the first despeckling module passes the input SAR image through a series of convolutional layers to suppress speckle noise and later feeds the resulting noise-reduced image to the subsequent change detection module.

For change detection, we initiate a preclassification step employing the logarithmic ratio operator and the hierarchical FCM algorithm. Subsequently, we utilise a layer attention module that exploits correlations among multi-layer convolutions. This module produces

robust cascaded feature representations learned by the network. These robust representations not only allow the proposed despeckling architecture to be resilient to multi-temporal SAR acquired from one SAR imaging process (i.e., the same number of SAR images looks before and after the change) but also enable it to deal with any combination of single or multi-look images acquired prior and after the change. In addition to this despeckling model, we have also developed a robust loss function that effectively suppresses the speckle noise, thereby improving the change detection accuracy. Both the despeckling model and the proposed noise-tolerant loss function are evaluated extensively on three public real SAR datasets, achieving superior performance compared to existing state-of-the-art SAR CD methods in all benchmark datasets.

# Table of contents

<b>List of figures</b>	<b>xv</b>
<b>List of tables</b>	<b>xvii</b>
<b>Nomenclature</b>	<b>xix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background of Research Problem . . . . .	1
1.2 Motivation and Research Questions . . . . .	3
1.3 Aims and Objectives . . . . .	4
1.4 Thesis Overview . . . . .	5
1.5 Publications and Main Contributions . . . . .	6
<b>2 Literature Review</b>	<b>9</b>
2.1 Holistic Overview . . . . .	10
2.2 Change Detection Framework . . . . .	12
2.3 Image Registration . . . . .	13
2.3.1 Image Registration Classification . . . . .	14
2.3.2 Classical Image Registration Techniques . . . . .	14
2.3.3 Conventional Image Registration Methods . . . . .	16
2.3.4 Deep Learning Based Image Registration . . . . .	20

2.4	Noise Removal from Remote Sensing Images . . . . .	25
2.4.1	Denoising Optical Images . . . . .	26
2.4.2	Despeckling SAR Images . . . . .	26
2.5	Remote Sensing Change Detection . . . . .	28
2.5.1	Optical Images Change Detection . . . . .	28
2.5.2	SAR Images Change Detection . . . . .	30
2.5.3	Hybrid Images Change Detection . . . . .	33
2.6	Summary . . . . .	34
<b>3</b>	<b>Systematic Investigation of Image Registration for SAR Change Detection</b>	<b>37</b>
3.1	Overview . . . . .	37
3.2	Methodology . . . . .	38
3.2.1	Pre-processing . . . . .	38
3.2.2	Change Detection . . . . .	42
3.3	Datasets . . . . .	42
3.3.1	Suez Canal Dataset . . . . .	43
3.3.2	Dry River Dataset . . . . .	43
3.4	Experimental Results & Discussion . . . . .	43
3.4.1	Evaluation Metrics . . . . .	44
3.4.2	Image Registration Results . . . . .	45
3.4.3	Change Detection Results . . . . .	48
3.5	Summary . . . . .	50
<b>4</b>	<b>Enhanced SAR Change Detection Using Deep Despeckling Model</b>	<b>53</b>
4.1	Overview . . . . .	53
4.2	Proposed Despeckling Model . . . . .	54
4.2.1	Network Architecture . . . . .	54

---

4.2.2	Loss Function . . . . .	55
4.3	Datasets . . . . .	57
4.3.1	Synthetic SAR Images . . . . .	58
4.3.2	Real SAR Images . . . . .	58
4.4	Experimental Results and Discussion . . . . .	60
4.4.1	Despeckling Metrics . . . . .	60
4.4.2	Change Detection Metrics . . . . .	62
4.4.3	Despeckling Results . . . . .	63
4.4.4	Varying Loss Functions . . . . .	66
4.4.5	Performance Analysis . . . . .	67
4.4.6	Change Detection Results . . . . .	69
4.5	Summary . . . . .	75
<b>5</b>	<b>Deep Learning Based on Change Detection for SAR Images</b>	<b>79</b>
5.1	Overview . . . . .	79
5.2	Methodology . . . . .	80
5.2.1	Baseline Architecture . . . . .	81
5.2.2	Enhanced Loss Function . . . . .	81
5.2.3	Enhanced Architecture . . . . .	82
5.3	Experimental Results & Discussion . . . . .	82
5.3.1	Datasets . . . . .	83
5.3.2	Ablation Study . . . . .	84
5.3.3	Results & Discussions . . . . .	87
5.4	Summary . . . . .	90
<b>6</b>	<b>Conclusion &amp; Future Work</b>	<b>93</b>
6.1	Summary of Research Findings . . . . .	93

6.2	Contribution to Knowledge . . . . .	95
6.3	Limitations and Future Work . . . . .	97
	<b>References</b>	<b>101</b>
	<b>Appendix A Published Papers</b>	<b>133</b>
	<b>Appendix B Source code</b>	<b>167</b>

# List of figures

2.1	Change detection problem . . . . .	11
2.2	Typical change detection results for different remote sensing datasets [1–3].	12
2.3	Framework of change detection problem . . . . .	13
2.4	Image registration steps with highlighted NNDR, Putative matches and Good Matches . . . . .	16
3.1	Change Detection process . . . . .	38
3.2	Nearest Neighbour Distance Ratio . . . . .	40
3.3	Image Registration Procedure . . . . .	41
3.4	Image Registration process for Dry River Dataset . . . . .	47
3.5	Image Registration process for Suez Canal Dataset . . . . .	48
3.6	The Registration Result for both SAR datasets . . . . .	49
3.7	CM results for Dry River dataset. . . . .	51
3.8	CM results for Suez Canal dataset. . . . .	52
4.1	The Methodology structure . . . . .	55
4.2	Despeckling proposed network for better change detection accuracy . . . .	56
4.3	Sample of BSD500 dataset that used to train the model . . . . .	58
4.4	Change detection datasets used to test the despeckling model . . . . .	59
4.5	Visualised results of Farmland dataset with different despeckling methods. .	65

4.6	Visualised results of Ottawa dataset with different despeckling methods. . .	66
4.7	Visualised results of Yellow River dataset with different despeckling methods.	66
4.8	Relationship between DM and F1 score for Farmland dataset . . . . .	68
4.9	Relationship between DM and F1 score for Yellow River dataset . . . . .	68
4.10	Relationship between DM and F1 score for Ottawa dataset . . . . .	69
4.11	Visualised results of Farmland dataset with different despeckling methods .	71
4.12	Visualised results of Yellow River dataset with different despeckling methods	74
4.13	Visualised results of Ottawa dataset with different despeckling methods . .	76
5.1	Proposed Change Detection Method . . . . .	84
5.2	Visualised Proposed loss function results of Farmland with different loss function . . . . .	85
5.3	Visualised Proposed loss function results of Yellow River with different loss function . . . . .	86
5.4	Visualised Proposed loss function results of Ottawa with different loss functions	86
5.5	Visualised results of change detection methods . . . . .	87

# List of tables

2.1	Comparison between optical and SAR change detection. . . . .	33
3.1	Multi-temporal remote sensing SAR images specifications. $H$ is height, and $W$ is width . . . . .	43
3.2	Quantitative Comparison of the Six Image Registration Algorithms. . . . .	46
3.3	Comparison of RMSE. . . . .	49
4.1	Proposed Method Network Configuration. . . . .	55
4.2	Despeckling Metrics. . . . .	61
4.3	Confusion Metrics. . . . .	62
4.4	Quantitative evaluation of despeckling methods on samples of synthetic testing SAR datasets . . . . .	64
4.5	Quantitative evaluation of filters on Multi-temporal SAR datasets . . . . .	65
4.6	Relationship between different DM loss functions and F1 Score for DDNet. . . . .	67
4.7	Relationship between different DM loss functions and F1 Score for LANTNet. . . . .	67
4.8	Quantitative evaluation on Farmland change detection based on different despeckling filters. . . . .	72
4.9	Quantitative evaluation on Ottawa change detection based on different de- speckling filters . . . . .	77

4.10	Quantitative evaluation on Yellow River change detection based on different despeckling filters. . . . .	78
5.1	Enhanced Architecture Where $lr$ learning rate and Epochs is the number of epochs in the experiment . . . . .	83
5.2	Multi-temporal SAR images specifications. $H$ is height, and $W$ is width . . .	84
5.3	Ablation Studies of the Proposed Method for LANTNet. . . . .	85
5.4	Change Detection Results of Farmland Dataset. . . . .	88
5.5	Change Detection Results of Yellow River Dataset. . . . .	89
5.6	Change Detection Results of Ottawa Dataset. . . . .	90

# Nomenclature

## Acronyms / Abbreviations

AIR Automatic Image Registration

AKAZE Accelerated-KAZE

BRIEF Binary Robust Independent Elementary Features

BRISK Binary Robust Invariant Scalable Keypoints

BRISK Robust Invariant Scalable Key points

CD Change Detection

CE Cross Entropy

CM Change Map

CNN Conventional Neural Network

DDNet Dual Domain Neural Network

DI Difference Image

DL Deep Learning

DM Despeckling Model

DNN Deep Neural Network

DoG Difference-of-Gaussians

EM Electromagnetic

FCM Fuzzy C-Means

FCN Fully Convolutional Networks

GAN Generative Adversarial Networks

GCP Ground Control point

GT Ground Truth

KL Kullback-Leibler Divergence

KNN K-nearest neighbours

LANTNet Layer Attention-based Noise-Tolerant Network

LoG Laplacian-of-Gaussian

MAE Mean Absolute Error

NNDR Nearest Neighbour Distance Ratio

NR-ELM Neighbourhood Ratio Extreme Learning Machine

ORB Oriented FAST and rotated BRIEF

PCAK Principle Component Analysis *K*-means

PCA Principal Component Analysis

PM Putative Matching

RANSAC Random sample consensus

RMSE Root Mean Square Error

RS Remote Sensing

SAR Synthetic Aperture Radar

SIFT Scale Invariant Feature Transform

SURF Speeded Up Robust Features

SVM Support Vector Machine

VHR Very High-Resolution



# Chapter 1

## Introduction

### 1.1 Background of Research Problem

Image-based change detection is one of the significant fields in remote sensing (RS) and computer vision. RS change detection (CD) here refers to finding differences in an area from remotely sensed images captured for the same geographical region at different periods [4, 5]. Change detection has been widely used in many applications, such as deforestation monitoring in the agricultural sector. Civilian applications of change detection include monitoring urban area development and city extension. In the military, it is used in gathering information about new military installations, movement of the enemy's military forces, and damage assessment. In the areas of climate change, it has been used to monitor deforestation and disasters [5, 6].

The RS CD process is vital in Earth Observation (EO) because it endeavours to distinguish the changed and unchanged pixels of multi-temporal EO images covering the same geographical region but at different times. The multi-temporal images (images at different times) should be co-registered to determine the correct position for each pixel in both multi-temporal images before being inputted into the change detection method. The more accurate the image registration process, the better the change detection performance [7–10].

Once co-registered, the change map (a result of the change detection algorithm) can be easily obtained using classical change detection methods by computing a difference image (DI), simply the intensity difference between the two images.

However, change detection in EO is nontrivial owing to inherent challenges such as co-registration errors, illumination variations, viewpoint, shadows, atmospheric effects (e.g., presence of clouds, fog, etc.), and varying sensor characteristics. Moreover, surface reflectance from incoherent objects (such as vegetation) can adversely affect the performance of optical CD algorithms. Synthetic aperture radar offers distinct advantages over optical sensors for CD in EO because it is not affected by weather conditions, provides penetration through clouds and vegetation, and offers sensitivity to small changes, making it capable of detecting changes that may be missed by optical CD methods. This technique allows us to remotely map the reflectivity of objects or environments with high spatial resolution through the transmission and reception of electromagnetic signals in the microwave spectrum, which easily penetrates through clouds and provides all-weather day/night sensing capability, making it suitable for applications related to disaster assessment (such as flooding and earthquake) [11, 5, 12]. However, SAR data suffers from speckle noise caused by random interference between the coherent returns issued from the many scatterers on the earth's surface. This speckle noise is the main challenge that affects CD accuracy. Eliminating the noise will increase CD accuracy. Existing change detection methods struggle with this noise, and despeckling it in the pre-processing step is necessary for this case [5, 12].

Therefore, a deep understanding of speckle filters is essential to improve the accuracy of existing change detection methods. Moreover, a change detection method should be robust to the noise and identify the difference between relevant and irrelevant information in the input images to detect the change. Although this challenge with change detection is not easy, deep learning has recently been used to solve remote sensing and computer vision challenges and achieve state-of-the-art results in these problems, such as image classification and object

detection. However, there are still many challenging problems for deep learning to solve. One of these challenges is change detection. The traditional CD methods have misclassified changed and unchanged objects or pixels over two periods (multi-temporal). Deep learning is expected to provide higher accuracy and low computational cost for SAR change detection.

## 1.2 Motivation and Research Questions

The advancement in satellite imaging technology, characterized by enhanced resolution and capabilities, particularly the ability to collect data without physical contact, has sparked increased interest among researchers in Earth's surface change detection. SAR change detection aims to determine the difference between multi-temporal images for an exact location [4, 2, 13]. Change detection techniques are broadly employed in several applications, such as disaster assessment [14], environmental monitoring [15], land management [16], and urban change analysis [17, 12]. SAR images are advantageous over optical images because they can provide information about disasters in the darkness and all weather conditions. However, SAR images suffer from speckle noise. This kind of noise heavily affects the performance of change detection techniques. Research studies are trying to resolve this challenge by despeckling noise or using deep learning SAR change detection methods. These approaches still struggle with speckle noise, mainly when different noise levels exist between pre-change and post-change SAR images. Addressing this challenge will enable us to provide a clear change map image that can help us assess the disasters quickly. Moreover, it can, in turn, support local governments in making effective and timely decisions to prevent or mitigate material losses and lives.

The motivation for this research arises from the need to address the existing gaps in the field of SAR change detection. Current research has encountered a speckle noise challenge in the SAR change detection task. This study aims to fill this void with two approaches: First, does despeckling SAR images before changing detection methods enhance CD performance?

Second, we explore different deep-learning approaches for SAR change detection. We specifically try to address the question of whether deep learning can deal with speckle noise without despeckling techniques as a pre-processing step. Finally, we investigate different image registration-based feature detection algorithms to reduce the change detection error by misregistering the two SAR images. By addressing these questions, we anticipate that our research will advance our understanding of SAR change detection tasks.

### **1.3 Aims and Objectives**

To explore and improve the performance of SAR change detection methods, we aim to address the issue of speckle noise that adversely affects the accuracy of these methods. In order to achieve this goal, the following specific objectives have been identified:

- Review the literature and investigate different image registration techniques in remote sensing image registration to reduce misregistration errors and improve change detection accuracy.
- Evaluate the matching and registration process in existing methods for change detection.
- Review the literature and conduct experiments to evaluate the state-of-the-art algorithm in SAR change detection.
- Develop a despeckling method based on deep learning to improve the accuracy of change detection methods by reducing speckle noise.
- Evaluate and compare the performance of the proposed despeckling method with the state-of-the-art despeckling methods.
- Develop a loss function for deep learning to reduce the influence of speckle noise and improve the change detection performance.

## 1.4 Thesis Overview

The thesis is broken into six chapters. These chapters are listed as the following:

**Chapter One: Introduction:** This chapter concentrates on the research background, motivations and research questions, aims and objectives, problems, the structure of the current thesis and publications and main contributions.

**Chapter Two: Literature Review:** This chapter discusses change detection and its types: classical and deep learning change detection techniques. Image registration is discussed as follows: classical and deep learning image registration techniques are used in producing change detection. The image registration section also provides details about feature detectors and then introduces the six famous feature detection and description methods. Feature matching, outlier rejection, and performance evaluation are also described in Chapter 2. Moreover, image denoising for optical images and image despeckling for SAR images are reviewed. Remote sensing change detection is discussed in detail in optical image change detection, SAR image change detection, and hybrid image change detection. Finally, the summary of this chapter.

**Chapter Three: Systematic Investigation of Image Registration for SAR Change Detection:** This chapter investigates the importance of image registration techniques in enhancing change detection performance. Two SAR datasets are used for registration and change detection tasks. The two datasets have large water areas, and changing areas are small. Therefore, different pre-processing have been investigated for better change detection accuracy. Four change detection methods have been compared with the SAR datasets. We rely on the subject image assessment in this chapter because of the absence of a ground truth label.

**Chapter Four: Enhanced Change Detection Performance Based on Deep Despeckling of SAR Images:** This chapter introduces our proposed despeckling model to enhance the performance of the change detection methods. The proposed model is then compared to different despeckling methods, including the state of the art methods. Two types of SAR datasets have been used to train and test our proposed method in this chapter. Subsequently, intensive experiments are performed over three real SAR datasets for change detection tasks using four change detection methods, including the current state-of-the-art algorithms. Following the results and discussion are presented. At the end of this chapter is the summary.

**Chapter Five: Deep Learning-Based Change Detection for SAR Images:** This chapter explores deep learning networks that have been used in change detection. A proposed deep learning method has been introduced, and robust loss function and better setting parameters have been discussed for the proposed methods. The experiments have been done using three SAR CD datasets. The results and discussion section compares our method with the state of art methods.

**Chapter Six: Conclusion and Future Work:** This chapter offers a conclusion for the research and outlines the contributions that this research has made. Any limitations faced during this research are also presented. It also lists the future work that follows the research done in this thesis.

## 1.5 Publications and Main Contributions

In this research, we have three main contributions.

- **Chapter Three contribution and publication:**

We investigated six image registration algorithms with different remote sensing datasets, including SAR images. We have found that no algorithm can perform best on all SAR

datasets. In addition, we use an Otsu with a 5x5 Gaussian filter to remove the unwanted information under the water surface in the Suez dataset. This pre-processing step assists in identifying the small change for the ships in the image at  $T_1$  (pre-change) and image at  $T_2$  (post-change). We published a Survey paper on this work entitled "Image Registration Techniques and Applications: Comparative Study on Remote Sensing Imagery" at the DeSE conference. This paper concludes that no singular algorithm demonstrates superior performance across all datasets. However, after assessing the average performance across the three datasets, it becomes evident that SIFT exhibits the most robust performance [7].

- **Chapter Four contribution and publication:**

We developed a despeckling model (DM) based on a convolutional neural network (CNN). It is trained on synthetic SAR datasets and tested on synthetic and real SAR CD datasets. DM is able to suppress the speckle noise and improve the performance of the change detection methods. A conference paper was published based on this work for BCD-SGAI 2023. It is entitled "Deep Despeckling of SAR Images to Improve Change Detection Performance" [18].

- **Chapter Five contribution and publication:**

Our proposed method adapts from LANTNet. We have proposed a tolerant loss function that is more resistant to speckle noise. It improves the CD performance. It is important to mention that combining the despeckling model and the CD loss function has further enhanced the CD performance. This work was accepted by the IEEE Access journal. The paper entitled "Enhanced Change Detection Performance Based on Deep Despeckling of Synthetic Aperture Radar Images" [19].



# Chapter 2

## Literature Review

Remote sensing images have been developed and utilised across various research domains, focusing on change detection. The application of remote sensing for change detection is crucial for discerning alterations between two images captured at different times within the same geographical area [13, 2]. This process holds significance in diverse sectors, including deforestation monitoring and damage assessment [4]. Remote sensing images serve as a valuable resource for observing the Earth's surface. In the context of denoising, optical satellite images offer researchers frequent data for Earth's surface monitoring. However, this data is susceptible to interference from darkness, clouds, and atmospheric conditions, necessitating the collection of images during daylight and in favourable weather conditions [13].

In comparison, SAR images can work at night and are not affected by weather conditions, darkness and clouds [20]. This characteristic gives SAR data an advantage over other RS data. The limitation of SAR data is that it is influenced by speckle noise [21, 22]. Despeckling this noise can improve the change detection result [23, 2, 24]. This chapter is structured as follows: first, image registration, which is an essential step for change detection, is discussed, and then image despeckling is also the fundamental step for change detection. The change detection techniques are debated. Finally, the summary of this chapter is provided.

## 2.1 Holistic Overview

Remote sensing change detection aims to identify the change between two multi-temporal images for the same geographical region at different times [25, 4, 26]. It provides valuable insights for various purposes, such as monitoring deforestation [4], detecting targets [27], promoting agricultural progress [28], managing land [16], and analysing urban changes [17]. Moreover, the CD algorithms help to extract vital information to assess the change, especially in case of natural disasters (e.g., earthquakes, floods, droughts, and hurricanes [14, 29, 30]), which in turn supports the local governments to make an effective and timely decision to prevent or mitigate material losses and lives.

Change detection is a quantitative analysis technique used to identify and measure surface changes in phenomena or objects over two distinct time periods [31]. It is a fundamental technology within the field of earth observation. It aims to differentiate between changed and unchanged pixels in multi-temporal remote sensing images captured from the same geographic area but at different times [32, 33]. The primary objective of a change detection system is to assign a binary label to each pixel based on a pair or series of co-registered images. A positive label indicates that the corresponding pixel represents an area that has changed, while a null label denotes an area that has remained unchanged (refer to Figures 2.1 and 2.2) [34]. Change detection is a powerful tool for various applications, including video surveillance, urban area mapping, and other forms of multi-temporal analysis.

Let  $I_1$  and  $I_2$  be two co-registered images that cover the same geographical region and have the same size  $W * H$  at different times,  $T_1$  and  $T_2$ , respectively [4, 35]. The same sensor captures both images. In the context of change detection, it is customary to convert both images to grayscale prior to analysis, ensuring consistency in their representation throughout the process.

$$I_1 : X * Y \rightarrow \mathbf{N} \quad (2.1)$$

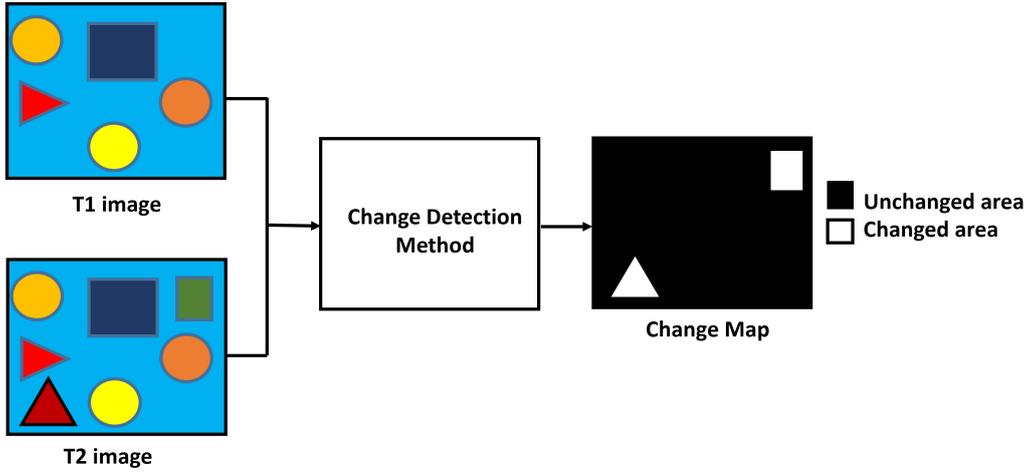


Fig. 2.1 Change detection problem

$$I_2 : X * Y \rightarrow \mathbf{N} \quad (2.2)$$

Where  $X = [1, w]$  and  $Y = [1, H]$ .

$I_1(x, y)$  and  $I_2(x, y)$  are the pixel Intensity values at coordinates  $(x, y)$ . In typical change detection methods, a change map can be obtained by a difference image operation based on differencing or log-ratio functions. Both DI operations are defined in equations 2.3 and 2.4.

$$DI(x, y) = |I_1(x, y) - I_2(x, y)| \quad (2.3)$$

$$DI(x, y) = \left| \log \left( \frac{I_2(x, y) + 1}{I_1(x, y) + 1} \right) \right| \quad (2.4)$$

The final analysis of the difference image (DI result) is to obtain the change map. Obtaining an accurate CM relies on several steps, including image registration, image denoising or despeckling (depending on the data type, in optical image denoising is used, whereas, in SAR images, image despeckling is used) and CD algorithms performance. These steps are part of the change detection framework. They will be discussed in section 2.2.

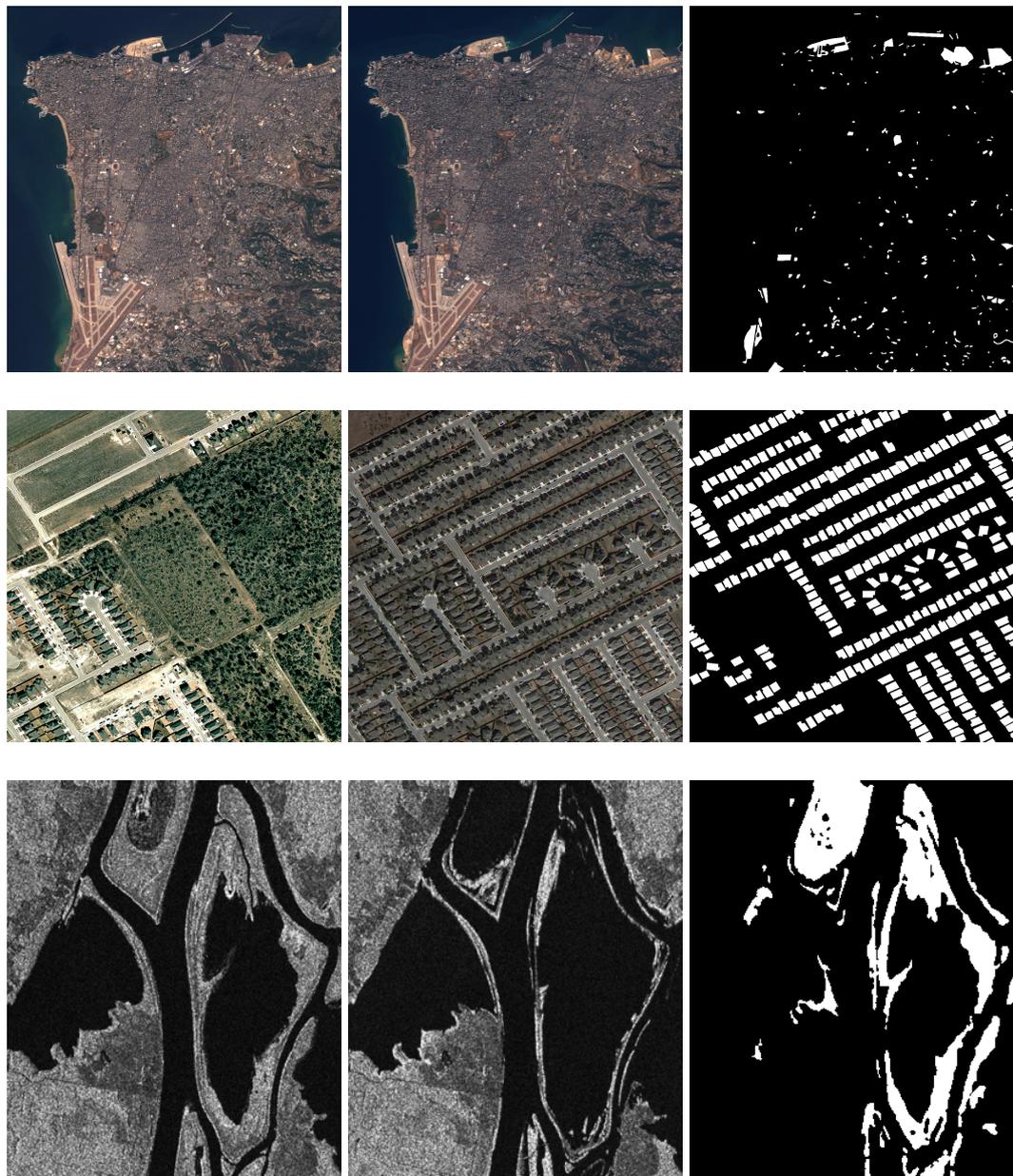
(a) Image at ( $T_1$ )(b) Image at ( $T_2$ )(c) Change map ( $CM$ )

Fig. 2.2 Typical change detection results for different remote sensing datasets [1–3]. Columns: (1<sup>st</sup> column) image at  $T_1$  before the change, (2<sup>nd</sup> column) image at  $T_2$  after the change, whereas (3<sup>rd</sup> column) the change between both images at  $T_1$  and  $T_2$ .

## 2.2 Change Detection Framework

Change detection identifies the difference in two multi-temporal images for the same location [36, 37]. Figure 2.3 demonstrates the change detection framework. Firstly, the process

starts with two RS images for an exact location captured at different times. These images are more likely to be affected by atmospheric interference or speckle noise; this depends on the images' type. These noises need to be removed before the image registration to reduce the error in the change detection result. In some cases, image registration has been executed before the image denoising/ despeckling process, or there will still be some noise after the denoising process. Most available datasets have co-registered images; therefore, Figure 2.3 starts with image registration and then the image denoising (despeckling) process. In addition, image registration is performed to align both images to exact coordinates. These processes are completed in the pre-processing step. Subsequently, the change detection technique highlights and classifies images into changed and unchanged pixels to constitute a change map. Finally, the ground truth image is compared with the change map to compute the performance evaluation. The following section discusses image registration.

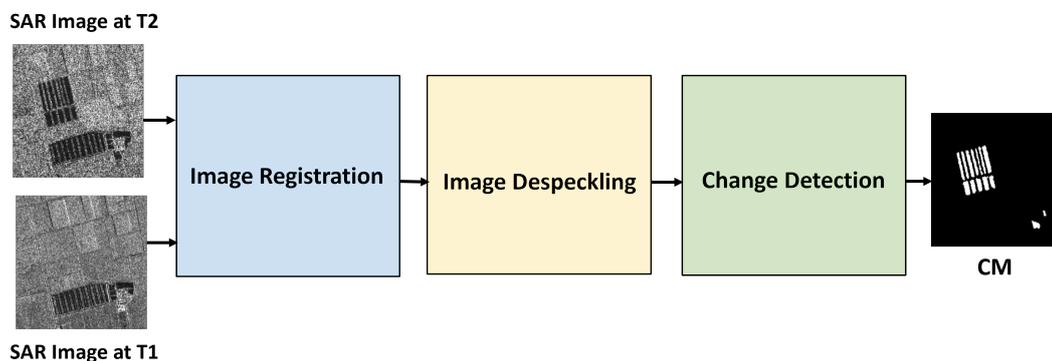


Fig. 2.3 Framework of change detection problem

## 2.3 Image Registration

Image registration is fundamental for many remote sensing and computer vision applications. It is the first step in change detection, image fusion, image stitching, and many more [38, 39]. This section studies image registration with possible applied scenarios and techniques used for feature detection and descriptions, feature matching, and outlier rejection.

### 2.3.1 Image Registration Classification

Image registration scenarios for remote sensing can be classified below based on how images are taken.

- **Images taken from different viewpoints (Viewpoints Registration):** It is used to integrate information from one moving sensor or multiple sensors from different viewpoints into the same object to make a 3D model. Landmark navigation and plant exploration are examples of applications that take advantage of this kind of registration [40, 41, 38]. Image stitching is one application of viewpoint registration, too.
- **Multi-temporal Images:** They are utilised for change detection and land resource surveys that include monitoring agricultural and land cover features extracted from data captured from one or more sensors over time [42, 40, 39, 38].
- **Different sensors (Multi-modal Registration):** This application is essential for integrating complementary information from different sensors. It benefits, for instance, the land cover uses such as yield estimate in agriculture, flood monitoring and detection of illegal crops. The fusion of different remote sensing data illustrates countless promises in assisting the decision-making process in several previous applications [42, 39, 38].

### 2.3.2 Classical Image Registration Techniques

Classical image registration can be categorised into manual and automatic methods.

- **Manual Image Registration:**

In the last two decades, traditional image geo-referencing started with manual image registration. A human does this process to allocate the ground control point (GCP),

called a tie point, in reference to sensed images. This step is equivalent to the feature detection. The next step is to match this GCP in both images by human operators. This process is facing many challenges [40], including i) time cost (especially for remotely sensed images because RS images usually cover a large size), ii) finding and allocating precise features is more complex in RS than in medical images, and iii) it requires expert operators. Based on these challenges, there are requirements for semi-automatic or automatic registration approaches.

- **Automatic Image Registration (AIR):**

AIR can be categorised into Area-Based Methods and Feature-Based Methods. The area-based (Intensity-Based) method is usually used when an image reflects a relatively smooth surface/scene short of essential features. It is popularly used for medical image registration. However, it is very time-consuming and is influenced by image noise. In contrast, feature-based methods, such as lines, corners, contours, and edges, aim to identify a corresponding region in both reference and sensed images. It is less computationally expensive and more resistant to noise. Therefore, it is commonly adapted for remote sensing image registration [42, 40, 38]. Our study follows the feature-based approach, which contains several steps: feature detection and description, feature matching, outlier removal, homography and image resampling, which is the process of geometrically transforming digital images. In this case, it will transfer the sensed image to the reference image coordination system. Figure 2.4 shows the process of feature-based image registration. Traditional methods, such as Scale Invariant Feature Transform (SIFT) [43], Speeded-Up Robust Features (SURF) [44], Oriented FAST and Rotated BRIEF (ORB) [45], KAZE [46], Accelerated-KAZE (AKAZE) [47] and Binary Robust Invariant Scalable Keypoints (BRISK) [48] are famous feature detectors and descriptors popularly used in image registration. Feature matching methods find corresponding features from two feature descriptors: reference

and sensed images. Nearest neighbour distance ratio (NNDR) is used to reduce initial false matches from feature matching. NNDR is pre-defined in detail in Equation 3.1 and 3.2. RANSAC is used to remove further outliers from putative matches [49]. The good matches after RANSAC are used in transformation matrix estimation [49].

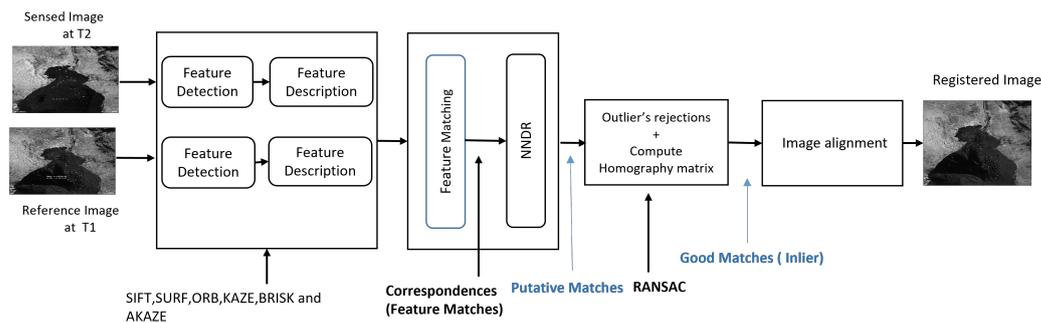


Fig. 2.4 Image registration steps with highlighted NNDR, Putative matches and Good Matches

### 2.3.3 Conventional Image Registration Methods

Conventional image registration techniques refer to well-established, standard approaches for aligning images through geometric transformations based on identified common features. The key steps involve feature detection, matching, outlier rejection and image resampling. Salient features like edges, corners, or distinct image patches that are common between the images are first detected [40]. Correspondences between these features are then established using mathematical methods such as keypoint matching algorithms, a classical technique in computer vision used to identify and match an object that exists in two images if it has a similar enough appearance in both images. With a set of matched control points, the transformation parameters (e.g. rotation, scale, translation) can be estimated to align one image to the other spatially [50]. This technique allows one image to be warped and overlaid on the reference image. Resampling of the pixel values is conducted to ensure accurate alignment. Conventional registration methods work well when point correspondences are

robust. However, they may fail for images with significant differences in appearance or geometry. This section discusses the critical steps of conventional feature-based image registration because it provides better results for remote sensing than area-based image registration [51, 52].

- Feature Detection Algorithms

Feature detection algorithms detect features (sometimes called keypoints) and provide a feature descriptor, making it easy to find the relevant feature in other images. Figure 2.4 shows that the first step in image registration is feature detection, which establishes a feature descriptor. The six feature detection algorithms are briefly described below. The input images are called reference and sensed images. Lowe [43] introduced the Scale Invariant Feature Transform (SIFT) to solve corner detecting problems with scaling invariance. The main stages of computation in SIFT lead to generating a set of image features. SIFT detector is based on the Difference-of-Gaussians operator, which approximates Laplacian-of-Gaussian. Feature points are detected by searching local maxima using Difference-of-Gaussians at various scales of subject images. The description method extracts a  $16 \times 16$  neighbourhood around each detected feature and further segments the region into sub-blocks, rendering 128 values. Bay et al. [44] presented Speeded-Up Robust Features (SURF), which relies on Gaussian scale-space analysis of images as SIFT. It uses different detectors and descriptors to speed up the computation, which is the disadvantage of the SIFT algorithm. Hessian matrix [53] has been used as a detector. Blob-like structures can be detected at locations where the determinant of the Hessian matrix is maximum. The Hessian matrix is a mathematical construct that describes the local curvature or second-order derivatives of a function, such as the intensity values in an image.

In 2011, Rublee et al. [45] developed Oriented FAST and Rotated BRIEF (ORB) to deal with the high computational burden of SIFT and SURF, especially for real-time

applications, for example, visual odometry. ORB consists of the FAST (Features from Accelerated Segment Test) detector and BRIEF (Binary Robust Independent Elementary Features) descriptor, performing well with low computational cost. This combination has provided a fast algorithm for real-time applications requiring corner feature extraction. FAST [54] is a sufficient corner feature detection for a real-time application that matches visual features, such as Parallel Tracking and Mapping [45]. In summary, FAST is several times faster than any existing corner detectors. However, it is affected by a high noise level [45]. The BRIEF descriptor is a simple binary test between pixels in a smoothed image patch [55]. Although BRIEF's performance is similar to SIFT in many cases, such as robustness to lighting, blur and perspective distortion, it performs poorly with rotation. Therefore, rBRIEF has been developed in [45] that significantly enhances rotated images compared with BRIEF. In summary, ORB is much faster than SIFT and SURF. However, it failed to address the scale invariance issue. Leutenegger et al. [48] introduced Binary Robust Invariant Scalable Keypoints (BRISK) that uses Adaptive and Generic Accelerated Segment Test (AGAST) [56] as a corner detector and filter them with FAST corner score when looking for maxima in a scale-space pyramid. The feature descriptor is built by classifying the characteristic direction of every feature to achieve rotation invariance. The descriptor is constructed as a binary string to achieve illumination invariance. BRISK was introduced to provide solutions for high-performance algorithms such as SIFT and SURF. It was also reported to be faster than SIFT and SURF.

In 2012, Alcantarilla et al. [46] introduced KAZE, which takes advantage of non-linear scale-space through non-linear diffusion filtering. This method blurs images locally adaptive to feature points, reducing noise and retaining regions' boundaries in images under processing. Scale normalised determinant of Hessian Matrix is used as KAZE detector. It is calculated at multiple scale levels. The maxima of detector response are

chosen as feature points using a moving window. Feature descriptor presents rotation invariance property by finding dominant orientation in the circular neighbourhood around every feature detector. KAZE is more computationally expensive than SURF, but the KAZE experiments in [46] have proved it is less computationally expensive than SIFT due to the computation in the non-linear scale-space [46, 57]. Alcantarill et al. [47] developed Accelerated-KAZE (AKAZE). It is similar to KAZE based on non-linear diffusion filtering. However, it solves the computational burden of creating a non-linear scale space using a mathematically well-organised framework, Fast Explicit Diffusion, embedded in a pyramidal framework to accelerate feature detection in non-linear scale spaces significantly. The AKAZE detector contains a determinant of the Hessian Matrix. The rotation invariance quality has been improved by Scharr filters [47]. The maxima of detector responses in spatial locations are chosen as feature points. The feature descriptor of AKAZE is based on a highly efficient Modified Local Difference Binary. By experiments, although AKAZE is computationally less expensive than KAZE, SIFT, and SURF, it is more expensive than BRISK and ORB [47].

- Feature Matching

Feature matching is essential for many problems in computer vision, such as object recognition, 3D reconstruction from multiple images, image registration, and motion tracking [39, 38]. The two main types of matching strategies are Brute Force and K-nearest neighbours (KNN). Brute Force is classified based on the type of feature descriptors. It takes the descriptor of one feature in the first set (reference image) and matches all other features in the second set (sensed image) using some distance calculation; the closest one is returned. The floating-point descriptor is used in SIFT, SURF and KAZE, whereas the binary descriptor is used in ORB, BRISK and AKAZE. The brute force algorithm specifies two parameters between features: distance metric

and crosscheck boolean. A crosscheck boolean is used to validate if the two features match. This method classifies the algorithms based on the descriptor types. For example, the  $L_1$ -norm function (also called Least Absolute Deviations) is used for float-point descriptors such as SIFT, SURF and KAZE use, and the second type of descriptors is Hamming distance is used for ORB, BRISK and AKAZE [40]. Whereas KNN with a pre-defined value,  $K$  is an approach that may produce many matching points. Therefore, there is a need to identify suitable matches using NNDR, a threshold introduced by Lowe to reduce the number of false matches (Outlier) [43]. The matching value after NNDR is called putative matching. After applying NNDR, the RANSAC algorithm removes the rest of the outliers.

- **Outlier Rejection and Image Alignment**

Random Sample Consensus (RANSAC) is a famous algorithm for rejecting outlier points, also known as false matching points [58]. RANSAC is introduced as a general and straightforward approach applicable to various problems. It is utilised to reject many outliers and estimate the homography matrix using suitable matches. The homography matrix transfers the second image (sensed image) to the first (reference image) for image alignment.

### **2.3.4 Deep Learning Based Image Registration**

Deep learning has made significant achievements in computer vision [59, 60], speech processing [61, 62], and image processing [63, 64]. Within deep learning, popular models such as deep belief networks (DBNs) [65], auto-encoders (AEs) [66], and convolutional neural networks (CNNs) [67] are commonly used. These models share a similar structure, consisting of multiple layers that progressively abstract features from lower to higher levels through non-linear transformations. Some deep models focus on capturing data distribution characteristics by minimising reconstruction errors [68], while others leverage stochastic

gradient descent in the backpropagation algorithm to learn semantic features [69]. Deep learning has demonstrated its superiority and robustness when applied to remote sensing images [70–72]. Liu et al. [73] applied deep Boltzmann machines to object detection. Cheng et al. [74] developed a rotation-invariant CNN for object detection in very high-resolution (VHR) optical remote sensing images, where the object and its rotated counterparts exhibit similar features. Subsequently, they proposed the rotation-invariant and Fisher discriminative CNN (RIFD-CNN) for object detection [75]. The RIFD-CNN ensures rotation invariance and enhances feature discrimination by minimising within-class scatter and maximising between-class separation using Fisher discrimination. Scott et al. [76] employed deep CNNs for landcover classification. Zhao and Du [77] introduced a deep-learning approach for hyperspectral image classification.

Deep learning techniques have also been applied to image registration and patch matching [78, 79]. Wu et al. [80] proposed a stacked auto-encoder for unsupervised deep feature extraction in medical image registration. Han et al. [81] introduced a Siamese network to match image patches, extracting patch-pair features through identical CNNs. Alternatively, Zagoruyko et al. [82] evaluated features extracted by identical or diverse CNNs or jointly learned from paired image patches. The results demonstrated the clear superiority of the joint processing approach. Deep learning image registration can have various applications. This research classifies them as remote sensing image registration and non-remote sensing image registration.

- **Remote Sensing Image Registration**

Remote sensing image registration based on deep learning can be classified into supervised and unsupervised approaches [78]. The supervised deep learning methods predict geometric transformation parameters from the reference and sensed images [83]. Miao et al. [84, 85] also adopted a supervised approach, converting the registration problem into a regression problem to learn parameters (such as affine transformation

matrix or matching point displacement) for image registration. DeTone et al. [86] also introduced a VGG-style projection transformation parameter regression model to assess homography between natural images. Moreover, DeTone et al. [87] developed a deep learning network called SuperPoint for feature point detection and description, yielding a more extensive feature point set than traditional methods. However, a notable limitation of SuperPoint is that the positions of interest points are uncertain and must be obtained through simulation, making them unsuitable for real-world scenarios. Specifically, using a supervised learning method based on a regression network faces challenges in simultaneously optimising rotation and scale terms of different dimensions, thus restricting the accuracy of supervised learning registration. Unsupervised deep learning matching approaches employ a spatial transformation network to warp the sensed image and align it with the reference image coordinates [88]. Notably, these methods do not rely on any human annotations for training. The optimisation of geometric constraints between the two images is achieved through a similarity loss function. Balakrishnan et al. [89] introduced an unsupervised image registration framework based on a CNN. This framework involved formulating the mapping between registration pairs and deformation fields. The sensed image was then distorted using a spatial transformation network, with the target loss function measuring the gray similarity between the distorted image and the reference image. Dalca et al. [90] proposed a probabilistic model and derived an unsupervised learning inference algorithm by leveraging the latest advancements in CNNs. Furthermore, De Vos et al. [91] designed flexible ConvNets for both affine image registration and deformable image registration. This involved stacking multiple ConvNets into a larger architecture, allowing for image registration at varying levels of detail, from coarse to fine.

- **Non Remote Sensing Image Registration**

Medical image registration using deep learning can be classified into five approaches: deep similarity metrics, Supervised End-to-End registration, Deep Reinforcement Learning, and Unsupervised End-to-End registration [92, 93]. The deep similarity metrics approach involves training various types of Deep Neural Networks to understand visual similarity metrics. This training used a substantial collection of paired and annotated ground-truth data. As a result of this training process, the networks became proficient in accurately and meaningfully capturing the structural differences between inputted pairs of images/patches. Particularly, they excelled in handling deformable transformations and various modalities, which traditional similarity metrics struggled with. Two main drawbacks are associated with this approach: it relies heavily on access to a large set of paired and annotated ground-truth data to train the network. Unfortunately, obtaining such datasets is often challenging in medical applications. Additionally, the approach still relies on conventional iterative-based methods, which are slow and impractical for clinical use [94, 95].

Supervised End-to-End Registration emerged as a significant breakthrough for the community. It successfully eliminated the computational burden and time constraints associated with traditional iterative-based registration methods. It has revolutionised the field by enabling the registration process to be conducted in a single step, making real-time clinical use feasible. Initially introduced in 2016 by [84] for rigid registration and later expanded in 2017 by Sokooti et al. [96] to include deformable registration, it has quickly become the dominant category and continues to be actively researched. Nevertheless, a significant limitation of this paradigm is its reliance on a large set of paired annotated ground-truth data for training the network, which poses a substantial obstacle in developing approaches within this category. In recent years, Deep Reinforcement Learning, also known as Agent-Based Registration, has learned to iteratively generate the final transformation, aiming to maximise positive envigenerate

the final transformation iteratively. Unlike the Deep similarity metrics approach, conventional similarity measures like normalised mutual information or Local Cross-Correlation are regularly utilised. However, a significant challenge in developing such paradigms is the agents' inability to interact effectively with the extensive state space introduced by deformable registration. It's worth noting that all the proposed approaches within this paradigm have been designed for rigid registration. Moreover, the success of this approach still relies on using ground-truth data to train the agents.

To address this issue, the Unsupervised End-to-End Registration paradigm was introduced. In this approach, various DNNs can be trained without relying on ground-truth data to build the regression model and determine the transformation parameters in a single step. Instead of using an extensive ground-truth dataset, data augmentation techniques are applied to a limited number of input samples as seeds. The learning process is guided by a traditional similarity measure (or a combination of them) used as a loss function. While this approach has achieved considerable success in unimodal registration, multi-modal registration poses a more significant challenge. The main difficulty lies in the inefficiency of the multi-modal similarity measures used as loss functions during the network's learning process. Consequently, networks trained on such measures inherit this inefficiency. The key to advancing this category lies in developing more efficient and powerful novel similarity measures in the future. It's worth noting that both supervised and unsupervised approaches have their limitations and drawbacks.

Weakly-semi-supervised End-to-End registration emerged in 2017 [92], effectively addressing the limitations of supervised and unsupervised approaches while incorporating their strengths. Some weak-semi-supervised end-to-end registration methods are label-driven, meaning they can implicitly learn to identify paired landmarks in inputted

images for the registration process based on just a few fully-annotated ground-truth samples. Other approaches adopt a dual-supervised approach, leveraging both similarity measures (similar to unsupervised methods) and a small number of ground-truth samples to fine-tune the network. Remarkably, it has been demonstrated that even with just a few ground-truth samples, transfer learning from other body organs or modalities is entirely feasible for medical image registration [97]. Additionally, certain methods utilise Generative Adversarial Networks (GANs) as their underlying technique, where the competitive interplay between the generator and discriminator benefits from a few ground-truth samples to construct and refine the model. This weakly/semi-supervised paradigm has proven to be highly practical and promising, attracting significant research focus for the future. It effectively combines the advantages of both supervised and unsupervised methods, making it a compelling and valuable direction for further exploration.

To summarise this subsection, we highlight the use of image registration in different fields and here we try to describe the main two subfields in the area of image registration. Our goal in this research is to focus on SAR change detection, which is a subfield of remote sensing. Therefore, we discuss remote sensing in general. In the following subsection, noise removal from remote-sensing images is discussed.

## **2.4 Noise Removal from Remote Sensing Images**

The noise removal process is an essential step for remote sensing change detection. It helps with noise reduction, which subsequently improves the change detection algorithms' performance. In this section, the noise removal process has been classified into four classes: traditional denoising optical images, deep learning denoising optical images, traditional

Despeckling SAR images and deep learning despeckling SAR images. These classes are discussed in the following sections.

### **2.4.1 Denoising Optical Images**

Denoising is primarily employed to diminish noise while retaining fine details within an image. In optical images, the pixel intensities frequently encounter additive Gaussian noise, characterised by a Gaussian distribution [13]. To counter this noise, linear filters are commonly employed. For example, Elad et al. [98] introduced a classical denoising algorithm that denoises through learning dictionaries and sparse representation. Yan et al. [99] presented non-local dictionary learning that utilises the multi-resolution structure and sparsity characteristics of wavelets in every decomposition level of the wavelet transform to denoise the images. Cui et al. [100] employed joint bilateral filtering for image denoising, effectively minimising noise while preserving edge details without excessively smoothing the image. This approach enhances the traditional bilateral filter by incorporating colour distance calculation based on the colour differences between image pixels.

### **2.4.2 Despeckling SAR Images**

Several approaches have been proposed to address speckle noise. For instance, pioneering work in the despeckling of SAR images was presented by Lee [21]. Later, Lee [101] refined [102] to remove noisy edge boundaries in SAR images by enhancing the edge representation using local statistics (average and variance) within a 7x7 window. A drawback of this approach is its reliance on a fixed mask size [11]. Kuan [103] proposed an adaptive speckle-noise smoothing filter that can handle different noise types without prior knowledge of the original statistics of the image. However, it tends to oversmooth image details and has high computational complexity [11]. Lope et al. [104] then proposed an Enhanced Lee filter and comprehensively analysed well-known filters by experimenting with varying the

local coefficients of despeckled SAR images. Their approach allows the preservation of fine details, such as texture and edge information, in the heterogeneous regions of the observed SAR image. Zhu et al. [105] further improved despeckling performance by combining an enhanced Lee filter with a median filter.

Recently, deep learning methods have been employed in despeckling SAR images. Ravani et al. [106] introduced a practical deep-learning method for despeckling synthetic aperture radar images. Notably, this approach bypasses the need for ground truth despeckled images. The experimentation involved assessing results on both the ImageNet dataset and real SAR images. The outcomes were validated through qualitative and quantitative measures, showcasing advancements compared to contemporary methods. Gu et al. [107] proposed a robust, deep, fully convolutional architecture tailored for despeckling multi-source SAR images. The architecture incorporates convolution and deconvolution layers to establish a non-linear transformation between noisy and clean SAR images. With skip connections integrated, the framework retains image details and embraces a residual learning approach. Initial outcomes using simulated and authentic SAR images exhibit improved despeckling performance and computational efficiency compared to existing techniques. Denis et al. [108] delved into a comprehensive analysis of ongoing advancements in image despeckling methods, specifically comparing patch-based nonlocal filtering and deep learning strategies. While these two approaches have distinct properties, their combination remains unexplored. This study aims to extract the maximum benefits from each approach while addressing inherent challenges [108]. Vitale et al. [109] proposed a novel cost function considering spatial and statistical characteristics. This approach serves a dual purpose: to balance the trade-off between spatial resolution and noise suppression and to identify an optimal cost function for solo learning-based despeckling. Evaluations conducted on both real and simulated data demonstrate compelling performance outcomes. Gleich and Sipos [110] were involved in the despeckling of SAR data using deep convolutional structures.

Preliminary findings showcased promising results using synthetic and real-world images. Cozzolino et al. [111] scrutinised the integration of deep learning techniques to enhance nonlocal despeckling. Employing a non-iterative, nonlocal means despeckling approach, the authors used weights generated by a well-designed deep CNN for each estimation window. Comparative evaluations of synthetic and real SAR data revealed the superiority of this approach over traditional nonlocal methods.

## **2.5 Remote Sensing Change Detection**

Change detection is the process of identifying differences in the state of an object or phenomenon by observing it at different times. Two types of remote sensing data, optical and synthetic aperture radar images, are discussed in this section. Our research concentrates on SAR change detection images. Therefore, we briefly discuss optical change detection in the next section.

### **2.5.1 Optical Images Change Detection**

Change detection involves the identification of discrepancies in an object or phenomenon's condition by observing it at distinct points in time [112, 113]. Optical change detection images are employed in tasks like monitoring urban expansion, assessing disasters, and overseeing environmental preservation and habitat monitoring. As depicted in Figure 2.3, the change detection process comprises data acquisition, data pre-processing, the application of a change detection algorithm, and evaluating accuracy [13]. Conventional change detection methods can be categorised into algebra-based, transformation-based, classification-based, and clustering-based approaches [5]. Algebraic methods, encompassing techniques like image differencing [114, 115], image ratioing [116], and change vector analysis [117], typically derive insights from algebraic operations on corresponding areas of multi-temporal

data. Transformation-based techniques discern alterations by converting registered images into feature spaces. Widely-used transformations include principal component analysis (PCA) [118] and tasselled cap transformation [119]. Classification-oriented change detection algorithms determine changing regions through classifications, including post-classification comparisons [120] and direct classification of bitemporal data [121]. Clustering-based algorithms yield change maps by clustering bitemporal data into changed and unchanged areas. Prominent clustering algorithms include K-means [122] and fuzzy c-means (FCM) [123].

In recent years, numerous DL algorithms have been introduced [124–127, 5, 128]. Early change detection networks entail classification networks that input small image patches and output corresponding categories [82, 129]. With the emergence of fully convolutional networks (FCN) [130], fully convolutional change detection networks have gained favour [131, 132]. In contrast to traditional change detection techniques, DL-based algorithms possess more hyperparameters, greater resilience to input data variations, and improved generalisation capabilities [131, 126, 128].

Zhan et al. [133] proposed a supervised change detection method for optical aerial images grounded in the deep Siamese network. Lyu et al. [134] proffered a supervised CD approach centred around a recurrent neural network. Geng et al. [135] introduced a supervised binary CD approach based on contractive autoencoders. Specific methods endeavour to decrease the demand for labelled training samples by employing a pre-classification scheme to deduce an initial coarse change map. Zhang et al. [136] delineated a binary CD method leveraging this initial map to spot improbable pairs, subsequently employed to instruct a mapping neural network. Xu et al. [137] developed a binary CD method deploying an autoencoder to comprehend the correspondence between pre-change and post-change images. Nonetheless, these approaches possess limited applicability due to the necessity of training or fine-tuning the model for specific datasets. Saha et al. [138] proposed unsupervised context-sensitive framework deep change vector analysis for CD in multi-temporal VHR

images that exploit convolutional neural network features to obtain deep features that can model spatial relationships among neighbouring pixels.

In general, Optical image change detection, such as VHR images, provides excellent visual resolution to humans to identify change during disasters. However, as mentioned earlier, it struggles during dark, fog, and clouds. Therefore, we concentrate our research on SAR images change detection, which is discussed in the next section.

### 2.5.2 SAR Images Change Detection

Synthetic aperture radar offers distinct advantages over optical sensors for CD in EO because it is not affected by weather conditions, provides penetration through clouds and vegetation, and shows sensitivity to small changes, making it capable of detecting changes that optical CD methods may miss. This technique allows us to remotely map the reflectivity of objects or environments with high spatial resolution through the emission and reception of electromagnetic signals in the microwave spectrum, which enables ease of penetration through clouds and provides all-weather day/night sensing capability, making it suitable for applications related to disaster assessment such as flooding and earthquake [11].

Typically, optical CD methods rely mainly on supervised machine learning approaches [12, 139, 140]. However, owing to the lack of annotated SAR datasets, the majority of SAR CD approaches primarily rely on unsupervised learning [141–143]. Several methods for unsupervised SAR CD have been proposed in the literature. For instance, Celik [23] proposed a simple unsupervised CD method using principal component analysis and  $k$ -means where change detection was achieved by partitioning the feature vector space into two clusters. Krinidis et al. [144] proposed fuzzy local information C-means (FLICM) to improve the clustering quality and aim to be robust to noise and preserve the image details. Gong et al. [145] also proposed fuzzy c-means (FCM), a reformulated FLICM to cluster image pixels into changed and unchanged. The aforementioned methods are performed under speckle-free

images. These approaches perform fairly well. However, SAR data suffer from speckle noise, which arises owing to the coherent nature of SAR imaging, which causes interference patterns in the received signals. This speckle noise makes information extraction from SAR images challenging and, consequently, adversely affects change detection accuracy [146–148].

SAR change detection has been widely used in many applications such as urban extension [36], agricultural monitoring [149], target detection [150] disaster monitoring [41] and assessment [151]. Typically, owing to the lack of annotated SAR datasets, most researchers rely on unsupervised methods [25, 152, 153, 3] to address SAR CD. However, the problem is highly challenging owing to the presence of speckle noise, which negatively impacts SAR images and reduces the change detection accuracy [146, 147, 152]. For this purpose, many researchers have formulated SAR CD in three sequential steps: image pre-processing, difference image generation, and classification [154]. The image pre-processing stage includes despeckling (denoising) and image registration. Image despeckling aims to reduce the impact of speckle noise and enhance SAR image quality. However, oversmoothing usually occurs when doing so, which may result in the loss of geometric details. After despeckling, the latter image registration aids in aligning multi-temporal images in the same reference coordinate system, enabling accurate change detection [7, 8]. To generate difference image, various methods have been proposed in the literature, including image differencing (also known as subtracting) [4], log ratio [155], neighbourhood-based ratio [156], Gauss-ratio operator [157] and mean- and log-ratio difference [158]. Finally, the classification of DI typically includes thresholding and clustering [159].

Some approaches use the clustered DI image (preclassification result) to subsequently train a classifier model and then combine the information from the preclassification and classifier results to generate a change map. For instance, Gao et al. [160] computed the preclassification result by computing a DI via log-ratio and fuzzy c-means clustering and later trained the PCANet model (classifier) to obtain the initial classification, which was

fused with the preclassification results to obtain the final change map. Similarly, Gao et al. [161] proposed an approach that employs a neighbourhood-based ratio to generate the difference image and then adopts an extreme learning machine (ELM) to model the high probability pixel based on the difference image, which is later used with the initial change map to yield the final change map. Wang et al. [162] employed a semi-supervised Laplacian support vector machine (SVM) to differentiate between changed and unchanged regions. To initialise the SVM, a pseudo-training set is generated using saliency similarity detection. This pseudo-training set consists of labelled changed and unchanged pixels. The Laplacian SVM effectively utilises the prior information from the available labelled samples and incorporates unlabelled samples to improve its discriminatory capabilities. Lv et al. [147] presented feature learning utilising a stacked contractive autoencoder to extract temporal change features from superpixels while effectively suppressing noise. Li et al. [163] proposed a Gamma correction and fuzzy local information c-means clustering model to reduce the impact of speckle noise and improve the performance. Liu et al. [78] introduced a locally restricted CNN for SAR change detection. They enhanced the original CNN architecture by incorporating a local spatial constraint, thereby improving CD performance.

Recently, a few approaches have aimed to explicitly suppress the inherent speckle noise to improve the SAR CD performance. For example, Qu et al. [2] proposed DDNet, a method that leverages features extracted from both the spatial and frequency domains to mitigate the impact of speckle noise. Gao et al. [25] also presented a Siamese adaptive fusion network for SAR image change detection, which focused on extracting high-level semantic features from multi-temporal SAR images while effectively suppressing speckle noise. Meng et al. [164] introduced a layer attention module that leverages the correlation among multiple convolutional layers and designs a loss function that minimises the influence of speckle noise, thereby enhancing the change detection performance. A limitation of these approaches is their inability to tackle different types of speckle noise effectively in images prior and

after the change, for example, single-look prior image and multi-look post-change image, which makes it difficult for SAR CD methods to perform well due to varying speckle-noise characteristics [5]. Table 2.1 summarises the difference between optical and SAR change detection.

Table 2.1 Comparison between optical and SAR change detection.

Feature	Optical Change Detection	SAR Change Detection
Data Source	Visible and Infrared	Microwave
Operational Limitations	Affected by Cloud	Operate in All Weather Conditions
Day/Night Capability	Daylight Only	Operate Day or Night
Spatial Resolution	Higher	Lower (Varies Depending on Sensor)
Temporal Resolution	Frequent	Less Frequent
Sensitivity to Surface Changes	Limited in Certain Conditions	High Sensitivity Penetrates Vegetation
Availability of Datasets	Widely Available	Limited
Learning techniques	supervised and Unsupervised	Unsupervised manner
Applications	Land Cover Monitoring, Agriculture	Disaster Monitoring, Deforestation

### 2.5.3 Hybrid Images Change Detection

Hybrid change detection is also called multisensor images CD or heterogeneous images CD. It combines the pre-change and post-change images from optical and SAR images. Several architectures have been introduced for conducting change detection tasks with pre-change and post-change images captured by the same sensor [165, 124]. However, only a few methods exist that can effectively address the CD challenge using multi-temporal data obtained from different sensors [166, 136]. This is due to the complications introduced by dissimilarities in spatial resolution [136] and spectral attributes among the sensor data [167]. An initial approach to address multisensor inputs is to individually generate classification or segmentation maps for each multisensory image, followed by a comparison to extract regions of change [166]. However, such post-classification strategies are susceptible to errors. Another popular strategy is to project pre-change and post-change images into a shared feature space, allowing for comparability in this new domain [168, 138]. Various

techniques achieve this projection, such as employing generative adversarial networks [138] and homogeneous pixel transformation [168]. Alternatively, learning a mapping function between the pre-change and post-change images is another approach [136]. Similar to the methods employed in single-sensor CD, symmetric and Siamese deep neural networks have also been adopted for multisensor CD. In work by Zhao et al. [169], an approximately symmetric deep neural network was leveraged to align images within the same feature space. Wang et al. [170] introduced a deep convolutional neural network-based Siamese architecture, incorporating a hybrid convolutional feature extraction module for processing multisensor data. Lastly, Ebel et al. [171] introduced a novel Siamese architecture to fuse SAR and optical observations for multi-model change detection. The work in this thesis concentrates on the SAR dataset from the same sensor discussed in section 2.5.2. These datasets will be discussed in detail in the experimental Chapters 3, 4 and 5.

## 2.6 Summary

This chapter focused on various aspects of image processing, particularly on image registration and denoising techniques. This section highlighted the importance and need for image registration in remote sensing and non-remote sensing data-denoising techniques for optical images and despeckling methods for synthetic aperture radar images. Traditional algorithms and the integration of deep learning for improved performance in these areas were also explored. The discussion started with denoising optical images, highlighting the prevalence of additive Gaussian noise and introducing denoising techniques like dictionary learning, nonlocal dictionary learning using wavelets, and joint bilateral filtering. Moving on to SAR images, the chapter covered despeckling methods. Early approaches by Lee and others [21, 101, 102] were presented, focusing on addressing the challenges of speckle noise. Recent advancements in deep learning-based approaches for despeckling were explored, showcasing their ability to outperform traditional methods.

The chapter also discussed remote sensing change detection, explaining the concept of identifying differences in objects over time. Traditional change detection methods, categorised into algebra-based, transformation-based, classification-based, and clustering-based approaches, were discussed. The emergence of deep learning algorithms in change detection was highlighted, emphasising their robustness and improved generalisation. Lastly, the chapter introduced hybrid change detection, combining data from different sensors for enhanced accuracy. Various strategies for comparing multisensory data, such as shared feature spaces and deep neural networks, were explored.

This chapter provided a comprehensive overview of denoising, speckle noise reduction, and change detection techniques in optical and SAR images. The integration of traditional methods and advanced deep learning approaches demonstrated their potential to enhance the accuracy and effectiveness of remote sensing applications. This research focused on SAR data for its advantages over optical images. Therefore, the experimental Chapters 3, 4 and 5 will focus on SAR image registration for change detection, as well as SAR change detection techniques based on despeckling and deep learning.



# Chapter 3

## Systematic Investigation of Image Registration for SAR Change Detection

### 3.1 Overview

Image registration is a fundamental process in many computer vision applications. Typically, it aligns reference and sensed images to the same coordinate system [7, 172]. In practical terms, image registration becomes necessary in order to minimise the disparity between the coordinate systems used in both images. It is used in change detection, image fusion, and image stitching [173, 8]. The motivation of this chapter is to illustrate the importance of image registration in change detection and to elucidate its process, subsequently providing input to change detection algorithms. For this purpose, six image registration algorithms have been benchmarked with two multi-temporal SAR datasets. However, the two multi-temporal SAR datasets have water surfaces with some unwanted information, which misleads the change detection algorithms by highlighting unchanged areas in the change map. To overcome this issue, we use the Otsu method that removes unwanted information under the water surface after the registration process, which increases the change detection performance.

The remaining sections are structured as follows: pre-processing (includes image registration methods and the Otsu method) and change detection algorithms. Subsequently, the datasets section covers non-registered SAR datasets. Next is the experimental results and discussion section, which discusses the environment setup, evaluation metrics, results, and discussion. Finally, the summary of this chapter is concluded.

## 3.2 Methodology

This section discusses the methodology in three stages: first, image registration to align the two SAR images. Secondly, Otsu's thresholding [174] is employed to separate the registered and reference images into two classes, foreground and background, based on the grayscale intensity values of their pixels. This step helps eliminate unwanted information under the water's surface. Finally, the change detection stage briefly discusses four change detection algorithms. Figure 3.1 presents how the change detection task depends on the image registration process.

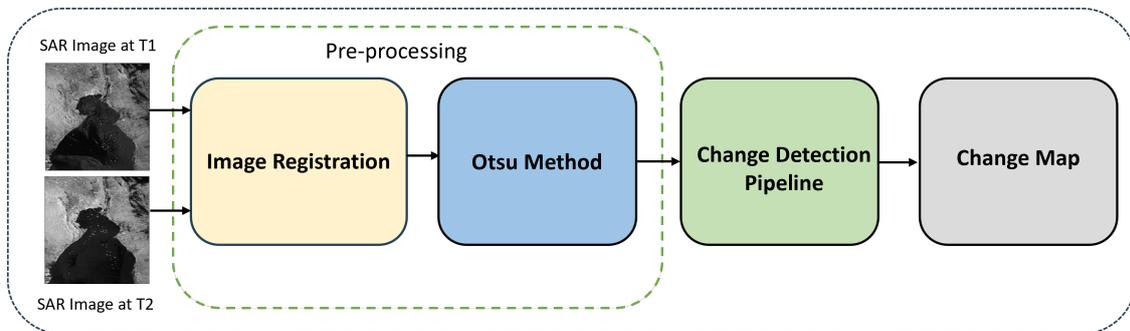


Fig. 3.1 Change Detection process

### 3.2.1 Pre-processing

This section describes two parts of the pre-processing phase: Image Registration and the Otsu method.

### 1. Image Registration

Image registration is the alignment of two multi-temporal images to one image. It is a significant step for change detection. The change detection task runs depending on registered multi-temporal SAR images. The registration aims to align two multi-temporal images to the same coordinate system. If the multi-temporal images were not registered correctly, it would provide an inaccurate change detection result. Therefore, an accurate image registration result is required to perform the change detection task. For this purpose, six famous image registration algorithms have been discussed in the previous chapter and are used in this chapter for the registration process. These algorithms have been performed and compared to choose the best algorithm for each dataset based on the highest precision result. Based on the following registration, change detection is performed.

Figure 3.3 demonstrates the existing image registration steps with highlighted NNDR, Putative matches and Good Matches. This process contains four steps to produce the registered image. The first step is to obtain feature detection and description in both input images by the six image registration algorithms. The second step is feature matching, which matches features in both images by Brute Force and K-nearest neighbours. NNDR is pre-defined in equation 3.1 [175] and Figure 3.2. In this work, we have applied brute force and KNN, then set the NNDR threshold value to 0.7 based on our experimentation. This value is aimed at removing 90% of the false matches while discarding less than 5% of the correct matches between descriptors in reference and sensed images.

$$\frac{\|D_R - D_{S1}\|}{\|D_R - D_{S2}\|} < T_{ratio} \quad (3.1)$$

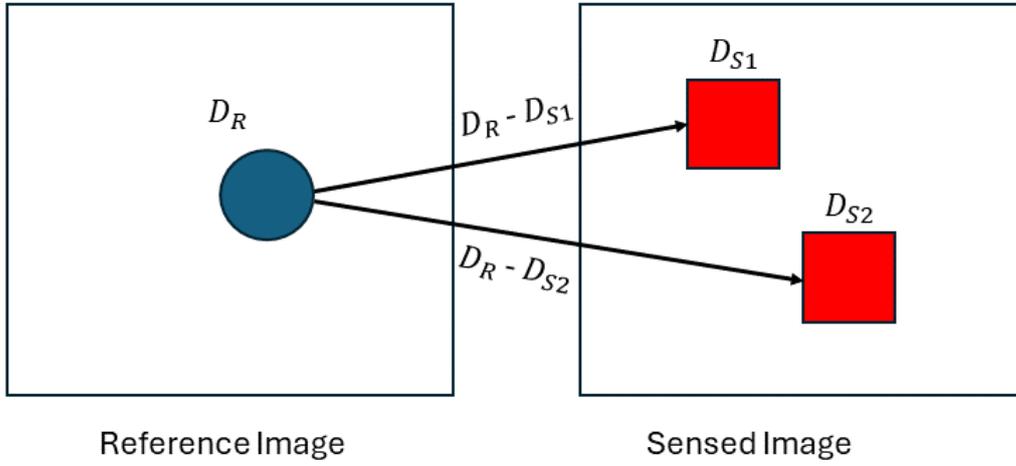


Fig. 3.2 Nearest Neighbour Distance Ratio

where  $D_R$  is the reference image feature descriptor,  $D_{S1}$  and  $D_{S2}$  are the first and second closest descriptors to  $D_R$  in the sensed image respectively.  $\|D_R - D_{S1}\|$  is the distance of  $D_R$  to  $D_{S1}$ ,  $\|D_R - D_{S2}\|$  is the distance of  $D_R$  to  $D_{S2}$ , and  $T_{ratio}$  is threshold, set as 0.7 because it provides best matching result with our datasets.

The matching value after applying NNDR is called Putative Matching (PM). Following this step, the RANSAC algorithm is executed to remove the rest of the outlier points and compute the homography matrix using the inlier points. The inlier points are good matches after removing the rest of the outlier points from the putative matching. The homography matrix transfers the sensed image  $I_2$  at  $T_2$  to the reference image coordinate system  $I_1$  at  $T_1$ . Finally, the image warping is done by aligning both images to produce the registered image. Once the registration process is finished, the reference and registered images will be fed to the change detection method to generate a change detection map. The registration process is applied to the six image registration algorithms: SIFT, SURF, ORB(5000), KAZE, BRISK and AKAZE. The best way to evaluate image registration results with the absence of ground truth is to reduce the shift between the reference image and the registered image. This process can be done

by calculating the root mean square error (RMSE) [176]. The evaluation metrics and datasets will be discussed in 3.4 and 3.3.

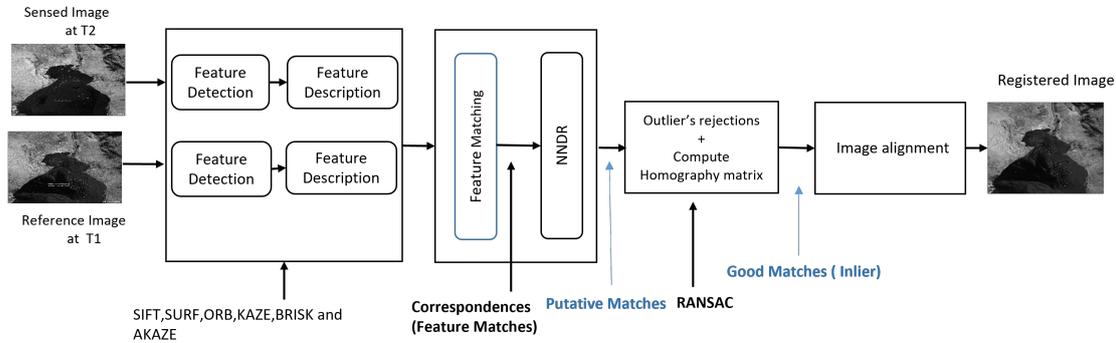


Fig. 3.3 Image Registration Procedure

## 2. Otsu Method

This subsection focuses on detecting changes between two SAR images, specifically within the Suez Canal dataset. This dataset contains small ships in both multi-temporal images and unwanted data submerged underwater, which can affect change detection accuracy. To address this issue, we explore various thresholding methods to separate the image background from the foreground. The first method is Otsu's thresholding [174], which initially constructs a histogram of pixel intensities in a grayscale image and then normalises it to represent probabilities. It tests different threshold values to segment the image into foreground and background classes at each threshold. It calculates the variance between these classes for each separation, aiming to maximise it. The method selects the threshold that maximises interclass variance, effectively pinpointing the most distinguishable point between foreground and background. This threshold is then used for image segmentation, proving valuable for automatic thresholding and segmentation, mainly when clear intensity differences exist between object and background regions. The second method involves applying a 5x5 Gaussian filter to reduce noise and smooth SAR images, followed by Otsu's thresholding. Both methods will be evaluated, and the results will be discussed in Section 3.4.

### 3.2.2 Change Detection

In this section, we investigate the performance of four change detection methods using two SAR datasets that are described in Section 3.3. The change detection methods are PCA- $k$ -means (PCAK) [23], NR-ELM [161], DDNet [2] and LANTNet [164] where the DDNet and LANTNet are the current state-of-the-art CD methods. As we discussed in the literature review section 2.5.2, PCA- $k$ -means (PCAK) proposed in [23] to perform a simple unsupervised CD method using principal component analysis and  $k$ -means where change detection was achieved by partitioning the feature vector space into two clusters. NR-ELM introduced in [161] to execute an approach that employs a neighbourhood-based ratio to generate the difference image and then adopts an extreme learning machine (ELM) to model the high probability pixel based on the difference image, which is later used with the initial change map to yield the final change map. DDNet presented in [2] to obtain features from spatial and frequency domains to minimise the speckle noise. LANTNet presented in [164] proposed a robust loss function and a layer attention-based noise-tolerant network (LANTNet) that benefits from feature correlations among multi-convolutional layers and suppresses the impact of noisy labels.

## 3.3 Datasets

The datasets used in this chapter are three Co-registered SAR datasets. These datasets will be used in chapters five and six. However, in terms of demonstrating the importance of image registration for the change detection task. In this chapter, the datasets need to meet two conditions: first, multi-temporal SAR images that have the same size and second, cover the same geographical region. Two public data meet these requirements, which are the Suez Canal [177, 7] and Dry River [178] datasets. These data are sufficient to be used for image

registration and change detection algorithms as they are utilised in an unsupervised manner. These data are described below:

### 3.3.1 Suez Canal Dataset

The two SAR images were captured by the Copernicus Sentinel-1 mission on the 21<sup>st</sup> and 25<sup>th</sup> of March 2021. On 21<sup>st</sup> of March, it illustrates the regular maritime traffic in the canal, with vessels visible every 2 to 3 km. On 25<sup>th</sup> March, a 400 m ship blocked the canal. Both images are publicly available on the European Space Agency (ESA) Website [177].

### 3.3.2 Dry River Dataset

These multi-temporal SAR images were captured at different times for the same area, showing the river's drought and vegetation in the second image. There are no known data on when both images were captured. However, these images are used for image registration tasks in [178].

Table 3.1 Multi-temporal remote sensing SAR images specifications.  
*H* is height, and *W* is width

Dataset	Name	Image H*W in Pixels
1	Dry1 and Dry2	531*787
2	Suez21 and Suez25	956*1140

## 3.4 Experimental Results & Discussion

Experiments were conducted on two sets of data, which are introduced in detail in section 3.3. Python 3.7 with OpenCV version 3.4.2.17 were used to perform the experiments. The platform of the experiments was Google Colaboratory Pro (Colab Pro) environment with a Tesla GPU P100-PCI-E-16 GB RAM 147.15 GB Disk.

### 3.4.1 Evaluation Metrics

Three evaluation metrics are used to evaluate the image matching algorithms: the number of inlier points provided in Table 3.2, Precision and computational cost.

- The number of Inlier points (Good matches) is the number of matching points after applying the RANSAC algorithm.
- Computational cost is vital to examine which algorithm can be used for real-time applications.
- Precision (the inlier ratio) defines the number of correct matches out of the putative matches as shown in Figure 3.3, and it is expressed in Equation 3.2.

$$Precision = \frac{Good\ Matches}{Putative\ Matches} * 100\% \quad (3.2)$$

- Registration Accuracy is evaluated by the root-mean-square error (RMSE) criterion [179–181]. A total of N corresponding point pairs  $(x_i, y_i), (x''_i, y''_i)$  are randomly selected from the reference and registered images. The point pairs are arbitrarily chosen and refined to reduce the residual to as low a level as possible. Hence, those point pairs are used as the reference to test the precision of model parameters. The *RMSE* is computed according to:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - x''_i)^2 + (y_i - y''_i)^2} \quad (3.3)$$

Where  $(x''_i, y''_i)$  denotes the transformed coordinates of  $(x'_i, y'_i)$ .

It is essential to mention that the two SAR datasets do not have a ground truth to evaluate the change detection performance, so we rely on subjective image assessment.

### 3.4.2 Image Registration Results

The image registration process goes through four steps: image matching, putative matching, obtaining inlier points, rejecting outlier points, and finally, using inlier points to obtain the registered image. This process has been illustrated for both datasets in Figures 3.5 and 3.4. The image matching step is excluded from both figures because it does not show sufficient information; just matching lines cover whole images. At the same time, Putative matching is shown in the first column. The red dots in this column are the number of features detected and rejected by the ratio threshold. In contrast, the blue lines are putative matching between two images. The second column is the outlier points rejected by RANSAC algorithm matching. Subsequently, Inlier points appeared in green lines that were used to generate a homography matrix to obtain the registered image. From subjective assessment of the Image registration process for Suez Canal and Dry River datasets in Figures 3.5 and 3.4, it can be seen that all the diagonal matching between image at ( $T_1$ ) and image at ( $T_2$ ) is wrong matching. It can be directly classified as outliers. In contrast, all the straight matching with 180 degrees are likely to be good matching. This notice can be observed in the first three columns. The fourth column shows a registered image by each algorithm.

The six famous image-matching algorithms have been experimented with different multi-temporal SAR datasets, as discussed in the previous section. The image registration result has been demonstrated in Table 3.2. The evaluation metric that is used to evaluate this process is precision, and it is illustrated in Section 3.4.1. It can be noticed that ORB(50000) has a high precision percentage of 99.38% with the dry river dataset, and AKAZE performs better precision with the Suez Canal dataset with 87.05%. It is the lowest computational cost in both datasets. Therefore, it is recommended for real-time applications. This result proves that no algorithm performs better in both datasets. It depends on the characteristics of the dataset. For instance, all the algorithms perform better in the dry river dataset than the Suez

Canal dataset because the Suez Canal has more illumination, and water covers about half of the images.

Nonetheless, all algorithms detect more features on the land sides of suez canal data than dry river data, as seen in Table 3.2 in the FD1 and FD2 columns. Moreover, there are more outliers in the Suez Canal than in the dry river dataset, as shown in Figures 3.5 and 3.4. From the results above, the Suez Canal is registered with the AKAZE algorithm, whereas dry river data is registered with the ORB(50000) algorithm because they provide a high precision percentage and lowest RMSE values. The final result is demonstrated in Figure 3.6 for the two datasets and will be used for the change detection task for its high precision and smallest RMSE, as shown in Table 3.3.

Table 3.2 Quantitative Comparison of the Six Image Registration Algorithms.

The **Bold font** is the highest value in precision and the smallest Image Matching Time. FD 1: Number of features Detected in Image 1 and FD 2: Number of features Detected in Image 2.

PM: Putative Matches. IMT: Image Matching Time (s). P%: Precision.

Algorithms	FD1	FD2	PM	Inliers	Outliers	P%	IMT(s)
Image Pair 1: Dry River							
SIFT	3494	2302	381	369	12	96.85	0.27
SURF	4259	3794	213	203	10	95.31	0.28
ORB(50000)	19065	14828	967	961	6	<b>99.38</b>	2.52
BRISK	6514	5184	590	586	4	99.32	0.34
KAZE	1513	1069	218	214	4	98.17	0.04
AKAZE	1214	827	169	166	3	98.22	<b>0.02</b>
Image Pair 2: Suez21 and Suez25							
SIFT	7605	7958	366	298	68	81.42	13
SURF	11683	14043	377	289	88	76.66	7.13
ORB(50000)	41690	42272	2592	2068	524	79.78	3.1
BRISK	35861	38253	1423	1013	410	71.19	15.58
KAZE	4896	4490	1141	816	325	71.52	4.1
AKAZE	4347	4291	811	763	48	<b>87.05</b>	<b>2.62</b>

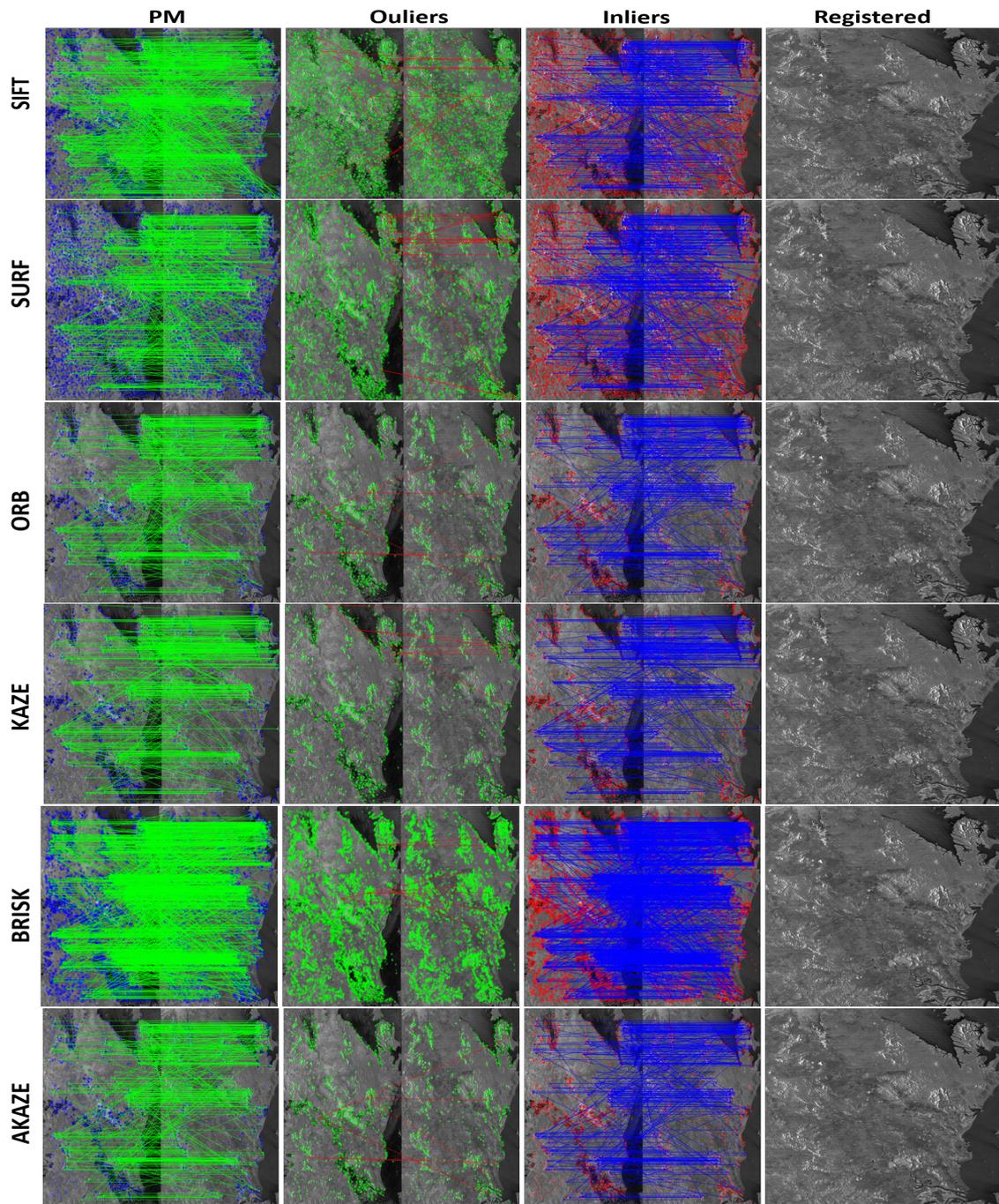


Fig. 3.4 Image Registration process for Dry River Dataset  
PM is putative Matching, Outlier is the rejected points by RANSAC, and Inliers are the good points that are used for the registration



Fig. 3.5 Image Registration process for Suez Canal Dataset

PM is putative Matching, Outlier is the rejected points by RANSAC, and Inliers are the good points that are used for the registration

### 3.4.3 Change Detection Results

To validate the effectiveness of the pre-processing method, we compared the results of change detection methods (PCA- $k$ -means (PCK) [23], NR-ELM [161], DDNet [2] and

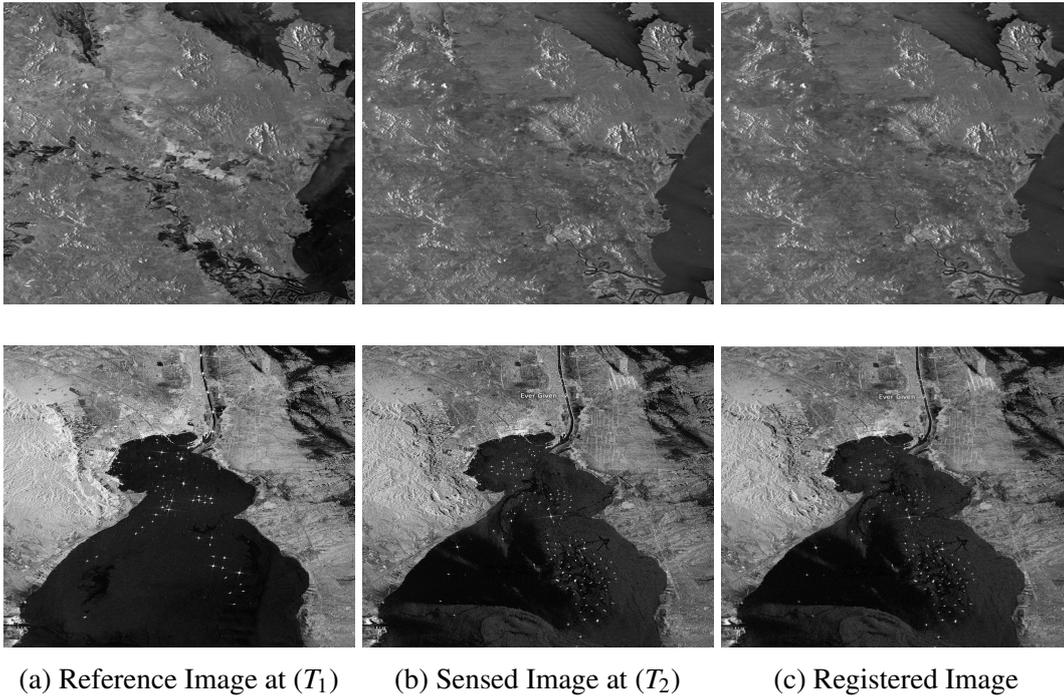


Fig. 3.6 The Registration Result for both SAR datasets  
 Rows: (1<sup>st</sup> row) Dry River dataset, and (2<sup>nd</sup> row) Suez Canal dataset

Table 3.3 Comparison of RMSE.

Method	SIFT	SURF	ORB(50000)	KAZE	BRISK	AKAZE
RMSE (Dry River)	2.97	3.32	<b>2.93</b>	3.46	3.35	3.49
RMSE (Suez)	4.49	5.42	5.48	6.54	7.09	<b>2.28</b>

LANTNet [164]) with and without the Otsu method using two real SAR datasets. Figures 3.7 and 3.8 demonstrate that applying the 5\*5 Gaussian filter and then Otsu thresholding considerably enhanced the change map results compared with CM without the Otsu for all the change detection methods.

In the dry river dataset, CD methods without the Otsu method yield unsatisfactory results. They highlight the water surface area in the whole image as a changed area. The Otsu algorithm helps change detection methods to exclude certain water areas that are common between the two images. However, it highlights many unchanged areas to be changed in

the change map. Otsu algorithm with Gaussian filter helps CD methods produce better CD images than the previous two methods as presented in Figure 3.7. In the Suez Canal dataset, the CD method without the processing step highlights the whole water surface and ignores the ships. Otsu algorithm provides better accuracy compared to change detection methods without Otsu. However, based on this method for the four CD algorithms, the result classifies a wide land area as changed areas, especially in PCAK and NR-ELM. The CM image is much better with DDNet and LANTNet. Finally, Otsu with Gaussian filter assists the CD methods in providing the best CM images for the CD methods by removing the unwanted information under the water surface and highlighting the small ships in CM images. LANTNet and DDNet provide better results compared to PCAK and NR-ELM methods as illustrated in Figure 3.8.

### 3.5 Summary

This chapter discussed the systematic investigation of image registration for SAR change detection. The process involved four steps: image matching, putative matching, obtaining inlier and rejecting outlier points, and obtaining the registered image. This process was illustrated for both the Suez Canal and Dry River datasets, emphasising the importance of each step. Image-matching lines were excluded from the illustration as they covered entire images. Putative matching, outliers, and inlier points were visually represented, leading to the generation of a homography matrix for image registration. Various image-matching algorithms were tested with precision as the evaluation metric, showing that different algorithms performed better depending on the dataset characteristics. The Suez Canal dataset was registered with the AKAZE algorithm, while the Dry River data was registered with ORB(50000) due to high precision and low RMSE (Root Mean Square Error) values. Change detection results were also presented, demonstrating the effectiveness of Otsu thresholding and Gaussian filtering in enhancing change maps for different change detection

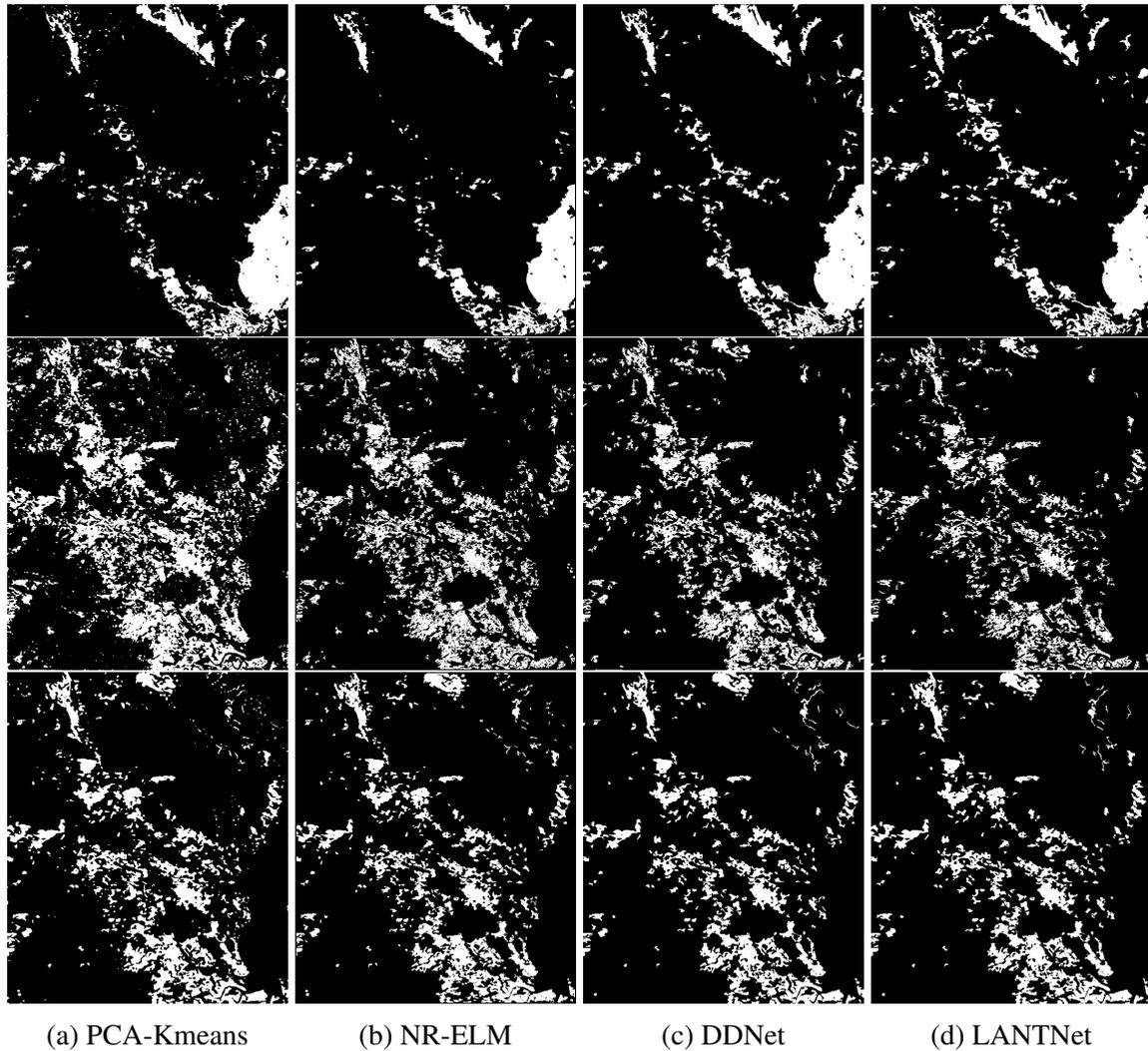


Fig. 3.7 CM results for Dry River dataset.

Rows: (1<sup>st</sup> row) CM without Otsu, (2<sup>nd</sup> row) suez canal data CM with Otsu thresholding, and (3<sup>rd</sup> row) CM with Otsu thresholding after applying 5\*5 Gaussian filter

methods. To summarise, the registered and reference images needed to be despeckled if they had speckle noise. However, both datasets seemed to be clear and did not need despeckling. The next chapter discussed despeckling noise for better change detection accuracy. A novel despeckling method was developed based on the state of the art and evaluated with comparisons to different despeckling methods.

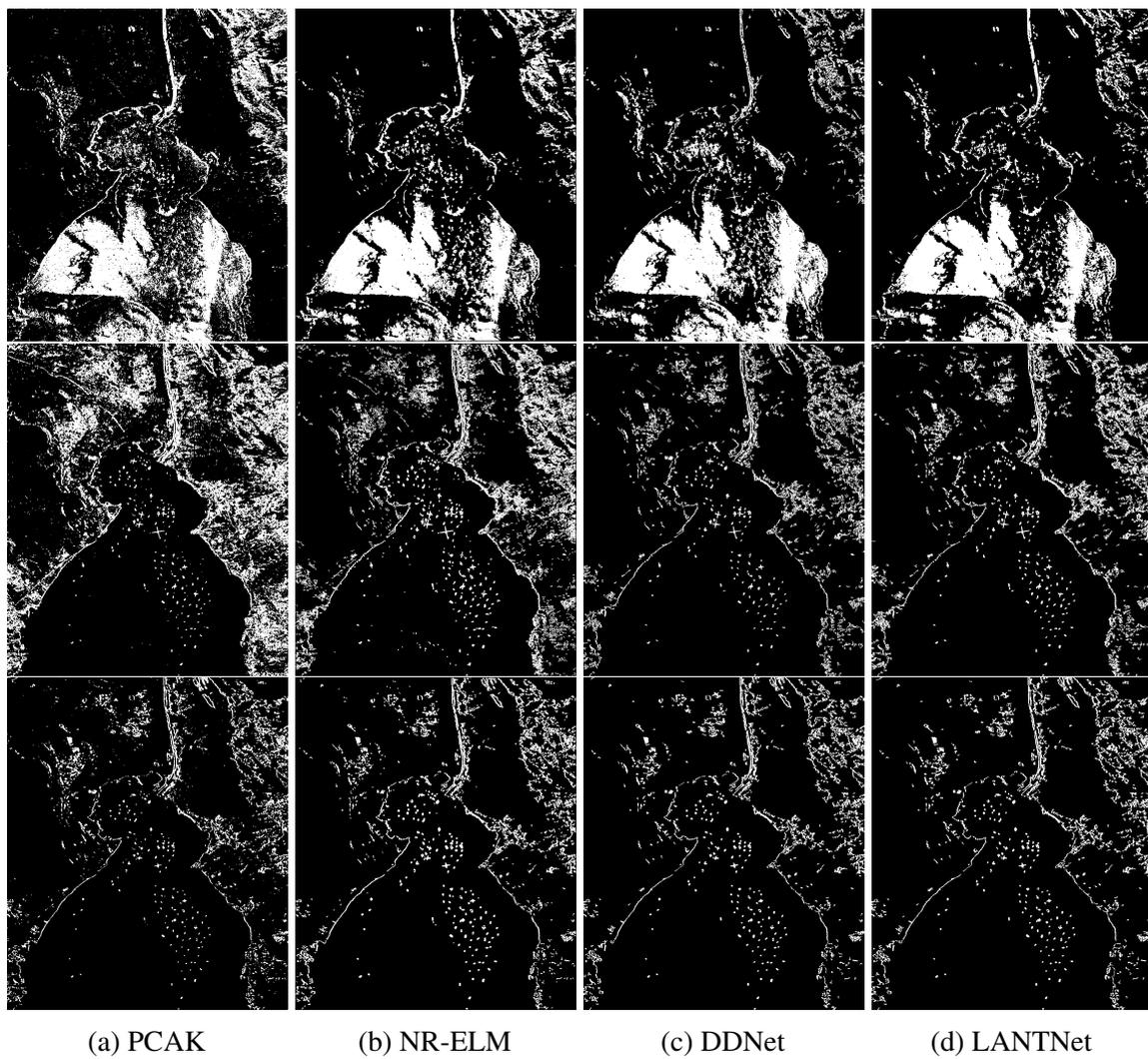


Fig. 3.8 CM results for Suez Canal dataset.

Rows: (1<sup>st</sup> row) CM without Otsu, (2<sup>nd</sup> row) suez canal data CM with Otsu thresholding, and (3<sup>rd</sup> row) is CM with Otsu thresholding after applying 5\*5 Gaussian filter

# Chapter 4

## Enhanced SAR Change Detection Using Deep Despeckling Model

### 4.1 Overview

SAR CD identifies the change between two multi-temporal SAR images for the same geographical region. The main challenge for the CD task is speckle noise. It is caused by interference patterns that arise when coherent electromagnetic waves reflect off rough surfaces or scatter from a volume of small particles. In images, speckle noise appears as a granular pattern that can reduce image quality and change detection performance. Various techniques have been developed to reduce or remove speckle noise, including filtering [182], wavelet-based methods [183], and statistical modelling [184]. In the context of CD, many CD methods are introduced to deal with this challenge. For example, Qu et al. [2] also presented DDNet to obtain features from spatial and frequency domains to minimise the speckle noise. Meng et al. [164] proposed a robust loss function and a LANTNet that benefits from feature correlations among multi-convolutional layers and suppresses the impact of noisy labels. Despite the robustness of modern deep learning-based methods against various noise types, they struggle to suppress speckle noise, effectively limiting their change detection capabilities.

Additionally, speckle noise levels differ between single-look and multi-look SAR imaging processes [5], further degrading change detection algorithms' performance. This research introduces a resilient deep convolutional neural network-based Despeckling Model (DM) that can suppress speckle noise and improve the performance of state-of-the-art SAR CD methods.

This chapter is structured to combine two subfields of research: image despeckling and SAR change detection. We hypothesised that more despeckling for SAR images before the CD task would improve the CD performance. In other words, the highest despeckling metrics for our SAR change detection images will provide the highest change detection accuracy. To investigate this hypothesis by despeckling SAR change detection datasets and computing the despeckling metrics for despeckling methods. Subsequently, CD methods will be performed on the despeckled SAR CD datasets. Finally, the CD metrics will be computed and compared to despeckling metrics to check the hypothesis.

## **4.2 Proposed Despeckling Model**

The proposed methodology is a despeckling model (DM) that passes the input SAR image through a series of convolutional layers to suppress speckle noise and later feeds the resulting noise-reduced image to the subsequent change detection method. In the following, we discuss them in detail, where we first present the proposed despeckling architecture, despeckling loss function, and adaptations we have made to the baseline change detection approach by proposing a noise-resilient loss function.

### **4.2.1 Network Architecture**

The proposed method is inspired by ID-CNN method [185]. It consists of ten convolutional layers that include batch normalisation and ReLU activation functions. Each layer has 64

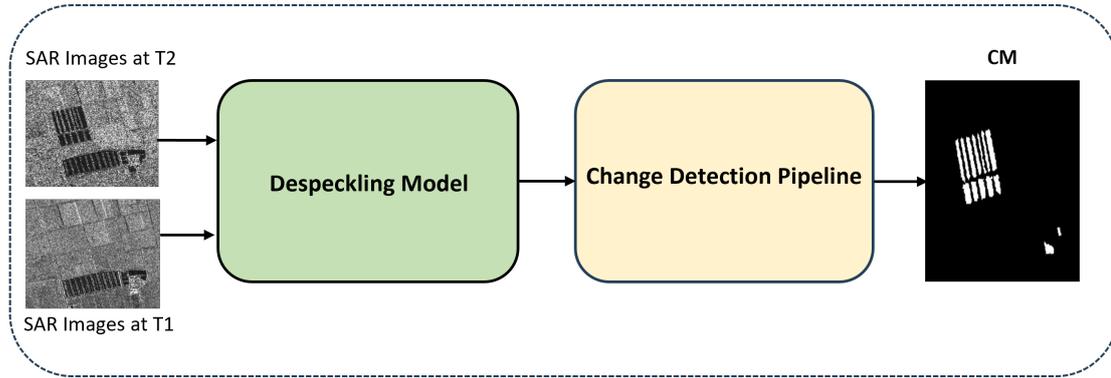


Fig. 4.1 The Methodology structure

filters with a stride of one, and zero padding is used to ensure that the output of each layer has the exact dimensions as the input image, except the last one has one filter. At the network's end, a hyperbolic tangent is stacked to work as a nonlinear function. The details of the proposed method architecture are described in table 4.1. In this context,  $L_1$  and  $L_{10}$  refer to a series of Conv-ReLU layers. Similarly, the layers between  $L_2$  and  $L_9$  represent Conv-BN and ReLU layers as described in figure 4.2.

Table 4.1 Proposed Method Network Configuration.

-	Layer	Filter Size	Filters	Output Size
$L_1$	Conv + ReLU	3*3*1	64	256*256*64
$L_2 - L_9$	Conv + BN + ReLU	3*3*64	64	256 *256* 64
$L_{10}$	Conv + ReLU	3*3*64	1	256 *256*1

### 4.2.2 Loss Function

The loss function is essential and part and parcel of the learning process. The utilisation of loss functions is crucial for learning, particularly in image reconstruction tasks that use convolutional neural networks (CNNs). Many studies have investigated various loss functions and their combinations to improve the learning process for tasks like super-resolution [186],

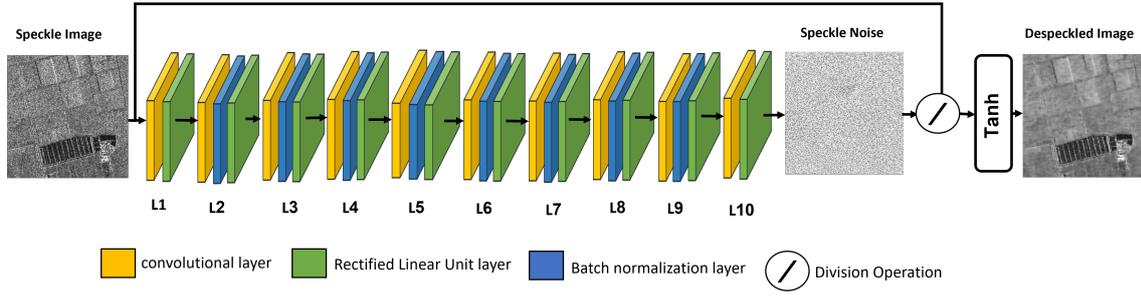


Fig. 4.2 Despeckling proposed network for better change detection accuracy

semantic segmentation [130], change detection [187], and style transfer [188]. In earlier studies on CNN-based image restoration, the optimisation process involved calculating the  $L_2$ -norm (Euclidean loss) or  $L_1$ -norm between the predicted (despeckled noise) and ground truth (clean image) images on a per-pixel basis.

Given an image pair  $X, Y$ , where  $Y$  is a speckle (noisy) image, and  $X$  is the corresponding ground truth image,  $X$  and  $Y$  have the same shape  $W \cdot H$ . The pre-pixel Euclidean loss is defined in equation 4.1.

$$L_E(\theta) = \frac{1}{W \cdot H} \sum_{w=1}^W \sum_{h=1}^H \|X^{(w,h)} - \hat{X}^{(w,h)}\|^2 \quad (4.1)$$

Where  $\hat{X}$  is the despeckled output image, and  $\theta$  is the learning network (parameters) for generating the despeckled image.

Although the Euclidean loss has effectively solved numerous image restoration problems, it frequently produces artifacts in the resulting estimated image [189]. To address this problem, Wang et al [185] has integrated a supplementary total variation (TV) loss into the loss function to promote smoother outcomes. This method removes the artefacts but smooths the images, which causes the loss of some information and reduces the performance of change detection, for instance. To overcome this, I have utilised a structural similarity index (SSIM), which image quality assessment techniques rely on quantifying errors between a reference (clean) and a sample image (despeckled) [190]. We use it as supplementary to

Euclidean loss, which can remove the artifacts and maintain the necessary information. This causes an increase in the change detection performance. SSIM is described in equation 4.7. Because the SSIM value is limited between 0 - 1. Therefore, there is a need for a predefined weight  $\lambda_{SSIM}$  for the loss function to increase and adjust the SSIM value and control the importance of the SSIM loss function.

$$SSIM(X, \hat{X}) = \frac{(2\mu_X\mu_{\hat{X}} + C_1) \cdot (2\sigma_{X\hat{X}} + C_2)}{(\mu_X^2 + \mu_{\hat{X}}^2 + C_1) \cdot (\sigma_X^2 + \sigma_{\hat{X}}^2 + C_2)} \quad (4.2)$$

Where  $X$  and  $\hat{X}$  are the reference (noise-free) and despeckled images, respectively,  $\mu_X$  and  $\mu_{\hat{X}}$  are the mean values of  $X$  and  $\hat{X}$  respectively. Similarly,  $\sigma_X$  and  $\sigma_{\hat{X}}$  are the standard deviations of  $X$  and  $\hat{X}$  respectively. While  $\sigma_{X\hat{X}}$  is the covariance between  $X$  and  $\hat{X}$ . Finally,  $C_1$  and  $C_2$  are constants set to be 0.01 and 0.03 respectively [190].

The total loss is thus calculated as follows:

$$Total - loss = L_E(\theta) + \lambda_{SSIM} \cdot SSIM \quad (4.3)$$

Where  $L_T$  is the total loss and  $\lambda_{SSIM}$  represents the weighting of the auxiliary SSIM in the loss. Based on the experimental results shown in Tables 4.6 and 4.7, we chose a value of  $\lambda_{SSIM}$  for 5.

## 4.3 Datasets

This chapter used two types of datasets to train and examine the despeckling model. The first is the Berkeley Segmentation Dataset 500, which is widely employed to generate synthetic SAR images. In addition, real SAR images (for the purpose of change detection purpose) were employed to assess the model's performance. Both datasets are described in detail In the following subsections:

### 4.3.1 Synthetic SAR Images

The Berkeley Segmentation Dataset 500 (BSD-500) was initially developed to evaluate the segmentation of natural edges, including object contours, object interior and background boundaries [191]. It included 500 natural images with carefully manually annotated boundaries and edges of natural objects collected from multiple users. This dataset has been widely used to generate synthetic SAR images for the purpose of despeckling [192–194]. Inspired by these studies, we have used BCD-500 to train our despeckling model.



Fig. 4.3 Sample of BSD500 dataset that used to train the model

### 4.3.2 Real SAR Images

For the purpose of change detection, we employed three real SAR image datasets that are multi-temporal and have been co-registered and corrected geometrically.

- *Farmland and Yellow River Datasets:*

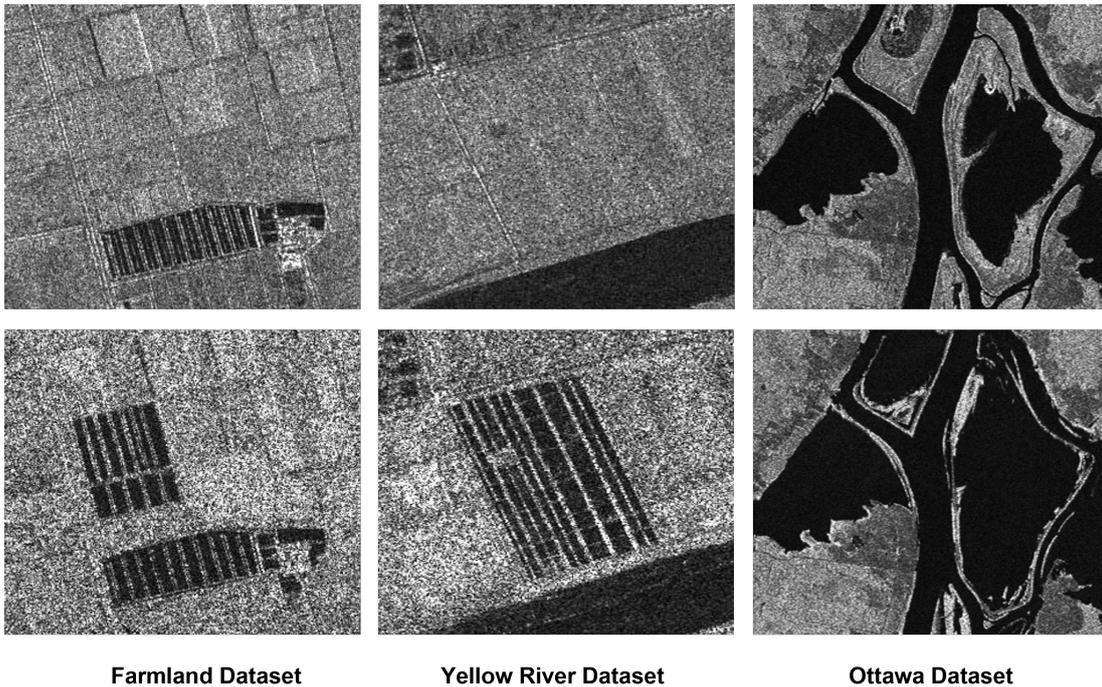


Fig. 4.4 Change detection datasets used to test the despeckling model

RADARSAT-2 captured the images for both datasets in the region of the Yellow River Estuary in China on 18<sup>th</sup> June 2008 (pre-change) and 19<sup>th</sup> June 2009 (post-change). The pre-change images are single-look, whereas the post-change images have been acquired via a multi-look (four) imaging process. The single-look pre-change image is significantly influenced by speckle noise compared to the four-look post-change image [25]. The disparity between the single and four looks in these two SAR datasets poses a significant challenge for change detection methods.

- *Ottawa Dataset:*

The images for this dataset were also captured by RADARSAT-2 in May 1997 (pre-change) and August 1997 (post-change) in the areas affected by floods [1–3]. Because of the single imaging process, both the pre-and post-change images are less affected by noise in this dataset.

## 4.4 Experimental Results and Discussion

This section has been divided into two subsections. Firstly, we discuss the metrics of despeckling and change detection. Subsequently, the performance investigation of different loss functions will be discussed. Performance Investigation of the Despeckling Model will also be discussed. The results and discussion have two subsections for despeckling and change detection results.

### 4.4.1 Despeckling Metrics

The main challenge in SAR image despeckling is reducing speckle noise while still preserving the image's fine details. One major issue is the lack of a "ground truth" for the desired speckle-free reflectivity. Another important consideration is the relationship between the quality and reliability of the despeckled SAR images. Despeckling techniques can be evaluated by analysing the degradation in homogeneous regions and the preservation of fine details in heterogeneous areas. The quality of the despeckled SAR image can be assessed through visual inspection or by using performance metrics such as with-reference and without-reference indexes. When a reference image is available, it becomes easier to compare and improve upon the results of the despeckling process. Some of the major performance metrics using a reference image are the peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) [195], Edge Preservation Index (EPI) [196] and UIQI [197], while in without-reference indexes, the Equivalent Number of Looks (ENL) is used. ENL gives information on the speckle reduction ability of an algorithm, and it is given by the ratio between the square of the mean to the Standard deviation. The highest ENL shows the best despeckling performance. The below table is summarised the despeckling metrics [198, 11].

$$ENL = \frac{\mu_I^2}{\sigma_I^2} \quad (4.4)$$

Table 4.2 Despeckling Metrics.

-	Full reference (pair)	Non reference (single)
1	SSIM	$\mu$
2	PSNR	$\sigma$
3	UIQI	ENL

$$\mu_I = \frac{1}{N * M} \sum_{N=1, M=1}^{N, M} I(i, j) \quad (4.5)$$

$$\sigma_I^2 = \frac{1}{N * M} \sum_{N=1, M=1}^{N, M} (I(i, j) - \mu)^2 \quad (4.6)$$

Where  $I$  is the despeckled image,  $\mu$  is the mean,  $\sigma$  is the standard deviation, and  $\sigma^2$  is the variance.

$$SSIM(X, \hat{X}) = \frac{(2\mu_X \mu_{\hat{X}} + C_1) \cdot (2\sigma_{X\hat{X}} + C_2)}{(\mu_X^2 + \mu_{\hat{X}}^2 + C_1) \cdot (\sigma_X^2 + \sigma_{\hat{X}}^2 + C_2)} \quad (4.7)$$

Where  $X$  and  $\hat{X}$  are the reference (noise-free) and despeckled images, respectively,  $\mu_X$  and  $\mu_{\hat{X}}$  are the mean values of  $X$  and  $\hat{X}$  respectively. Similarly,  $\sigma_X$  and  $\sigma_{\hat{X}}$  are the standard deviations of  $X$  and  $\hat{X}$  respectively. While  $\sigma_{X\hat{X}}$  is the covariance between  $X$  and  $\hat{X}$ . Finally,  $C_1$  and  $C_2$  are constants set to be 0.01 and 0.03 respectively [190].

$$PSNR = 10 \log_{10} \left( \frac{MAX^2}{MSE} \right) \quad (4.8)$$

Where MAX is the maximum possible pixel value of the image, and MSE is the Mean Squared Error between the original and compressed images.

$$UIQI = \frac{\sigma_{X\hat{X}}}{\sigma_X \sigma_{\hat{X}}} \cdot \frac{2\mu_X \mu_{\hat{X}}}{(\mu_X)^2 + (\mu_{\hat{X}})^2} \cdot \frac{2\sigma_X \sigma_{\hat{X}}}{\sigma_X^2 + \sigma_{\hat{X}}^2} \quad (4.9)$$

The first component is the correlation coefficient between the x and y images, quantifying the extent of their linear correlation. The second component, with a value range of [0,1], gauges the proximity of the mean luminance levels between x and y. The third component assesses the similarity in contrast between the images. The range of values for the *UIQI* index spans from -1 to 1, with the optimal value of 1 achieved only when the images are identical.

#### 4.4.2 Change Detection Metrics

Quantitative evaluation indices including false positives (FP), false negatives (FN), true positive (TP), true negative (TN), precision (P), recall (R), overall accuracy (OA) (also called observed by change) and F1 score (F1) [199, 200]. All these metrics are represented in the Formulas 4.10 to 4.13. Clearly, a greater P value leads to a decrease in false alarms, while a higher R-value indicates a reduced rate of incorrect detections. The OA measures the proportion of accurately detected pixels in the image. However, when the number of altered pixels is only a small part of the entire image, relying on these three metrics could result in overestimating the outcome. To avoid this, the F1 score is used as it addresses the limitations of P and R and provides a more comprehensive evaluation of the performance. It is important to note that larger F1 values indicate better overall performance [201].

Table 4.3 Confusion Metrics.

GT	GT	Metrics
0	0	TN
0	1	FP
1	0	FN
1	1	TP

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

$$P = \frac{TP}{(TP + FP)} \quad (4.11)$$

$$OA = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (4.12)$$

$$F1 = \frac{(2 \cdot P \cdot R)}{(P + R)} \quad (4.13)$$

### 4.4.3 Despeckling Results

To evaluate the despeckling performance of the proposed despeckling model on both synthetic and real SAR datasets, we compare the performance of the DM with the following four despeckling methods: Lee [21], Enhanced Lee [104], SAR2SAR [202] and ID-CNN [185]. Note that [202] and [185] are the most recent state-of-the-art image despeckling algorithms. In the following, we divide the despeckling result into results on synthetic SAR images and results on real SAR images.

#### 1. Results on Synthetic SAR Images

We have randomly chosen four pairs of images (noise-free and despeckle images): the desert, riverside, Buildings, and Rivers classes were respectively set up with synthetic SAR images. To assess the despeckling performance of different methods on synthetic SAR, we use three fundamental metrics: PSNR, SSIM and UIQI. These metrics show how proficient each method operates within the context of these data pairs. Table 4.4 represents the performance of despeckling methods on synthetic testing SAR datasets. In general, Enhanced Lee consistently achieves the highest scores in terms of PSNR and UIQI, indicating its superior performance in reducing noise and preserving image quality. On the other hand, SAR2SAR excels in SSIM scores, highlighting its ability to maintain structural similarity with the original data. However, the despeckling model

achieves average performance in all three metrics. This divergence in performance can be attributed to the distinct approaches and algorithms employed by these despeckling methods.

Table 4.4 Quantitative evaluation of despeckling methods on samples of synthetic testing SAR datasets

Testing data	Metrics	Noisy	Lee	Enhanced Lee	SAR2SAR	ID-CNN	DM
Pair 1	PSNR $\uparrow$	8.77	8.53	<b>14.64</b>	13.07	9.12	9.65
	SSIM $\uparrow$	0.08	0.02	0.31	<b>0.65</b>	0.05	0.07
	UIQI $\uparrow$	0.67	0.65	<b>0.87</b>	0.85	0.70	0.72
Pair 2	PSNR $\uparrow$	7.97	11.54	<b>14.64</b>	12.02	8.90	14.14
	SSIM $\uparrow$	0.09	0.15	0.30	<b>0.34</b>	0.11	0.23
	UIQI $\uparrow$	0.76	0.87	<b>0.93</b>	0.87	0.79	0.92
Pair 3	PSNR $\uparrow$	7.65	11.68	<b>15.78</b>	14.25	9.20	9.94
	SSIM $\uparrow$	0.02	0.05	0.23	<b>0.54</b>	0.03	0.03
	UIQI $\uparrow$	0.72	0.87	<b>0.93</b>	0.90	0.79	0.82
Pair 4	PSNR $\uparrow$	8.63	12.95	<b>19.04</b>	18.77	10.14	11.12
	SSIM $\uparrow$	0.03	0.07	0.27	<b>0.61</b>	0.05	0.06
	UIQI $\uparrow$	0.74	0.89	<b>0.96</b>	0.95	0.80	0.83

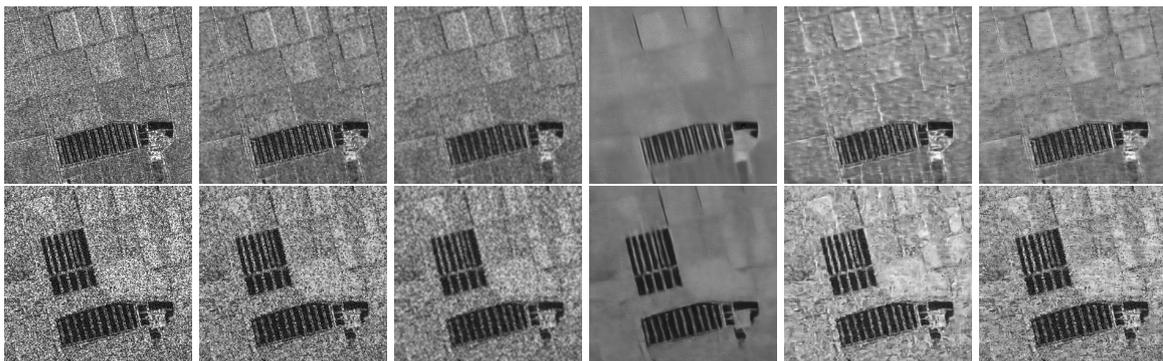
## 2. Results on Real SAR Images

In this section, we have three genuine SAR datasets designated for change detection purposes at our disposal. These datasets, however, lack a reference, noise-free (ground truth) image for direct comparison with the despeckled output. Consequently, there arises the necessity to employ non-reference image quality assessment metrics, such as Equivalent Number of Looks, to evaluate and contrast the performance of our despeckling model against other established despeckling methods. As illustrated in Figures 4.5, 4.6 and 4.7, these figures depict the performance of various despeckling techniques applied to synthetic SAR testing datasets. Notably, the SAR2SAR algorithm tends to oversmooth the three real SAR datasets, resulting in the highest ENL values, as indicated in Table 4.5. Conversely, Lee and Enhanced methods gently despeckle

Table 4.5 Quantitative evaluation of filters on Multi-temporal SAR datasets

Metric	Image	Noisy	Lee	Enhanced Lee	SAR2SAR	ID-CNN	DM
$\mu \uparrow$	Farmland1	118.43	118.23	118.3	121.22	<b>134.54</b>	130.91
	Farmland2	122.47	122.24	122.23	104.61	<b>149.35</b>	135.37
	Ottawa1	60.88	60.63	60.48	66.49	<b>71.93</b>	66.83
	Ottawa2	71.55	70.96	71.25	76.26	<b>88.75</b>	75.44
	Yellow River1	101.70	101.48	101.70	112.52	<b>118.49</b>	118.08
	Yellow River2	105.57	105.15	105.35	94.29	<b>115.32</b>	109.94
$\sigma \downarrow$	Farmland1	44.85	28.35	26.07	<b>21.02</b>	30.09	28.73
	Farmland2	67.43	53.59	38.78	<b>27.64</b>	46.16	49.88
	Ottawa1	55.83	<b>51.82</b>	49.43	56.32	61.39	57.69
	Ottawa2	54.77	49.48	<b>48.17</b>	50.6	62.29	52.39
	Yellow River1	43.26	33.54	<b>31.58</b>	34.53	39.15	36.63
	Yellow River2	68.23	58.00	46.20	<b>38.55</b>	51.68	53.99
ENL $\uparrow$	Farmland1	6.97	17.4	20.59	<b>33.27</b>	19.99	20.76
	Farmland2	3.3	5.2	9.94	<b>14.32</b>	10.47	7.37
	Ottawa1	1.19	1.37	<b>1.50</b>	1.39	1.37	1.34
	Ottawa2	1.70	2.1	2.19	<b>2.27</b>	2.03	2.07
	Yellow River1	5.53	9.15	10.37	<b>10.62</b>	9.16	10.39
	Yellow River2	2.39	3.29	5.20	<b>5.98</b>	4.98	4.15

the SAR datasets, so ENL values are small. At the same time, our custom despeckling model effectively eliminates noise, resulting in a significant increase in the ENL scores.



(a) Farm (T2) (b) Lee (c) Enh Lee (d) S2S (e) ID-CNN (f) Proposed

Fig. 4.5 Visualised results of Farmland dataset with different despeckling methods. The 1<sup>st</sup> row represents Farmland data at  $T_1$  and 2<sup>nd</sup> row represents Farmland at  $T_2$

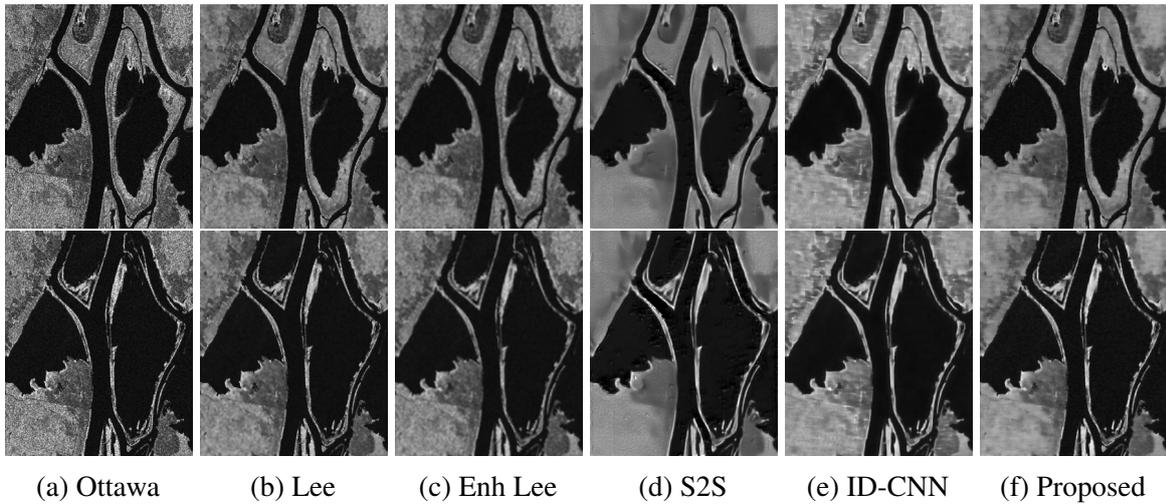


Fig. 4.6 Visualised results of Ottawa dataset with different despeckling methods.

The 1<sup>st</sup> row represents Ottawa data at  $T_1$  and 2<sup>nd</sup> row represents Ottawa at  $T_2$

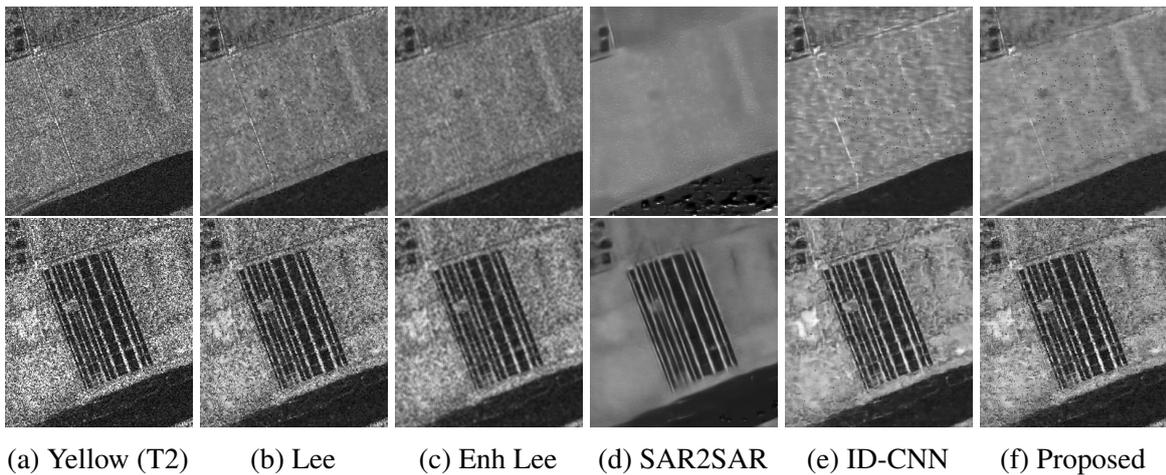


Fig. 4.7 Visualised results of Yellow River dataset with different despeckling methods.

The 1<sup>st</sup> row represents Yellow River data at  $T_1$  and 2<sup>nd</sup> row represents Yellow River at  $T_2$

#### 4.4.4 Varying Loss Functions

In this section, we explore the impact of various loss functions in despeckling methods on the performance of recent change detection methods. For each loss function, we train the despeckling model using the respective loss function and subsequently apply change detection methods, including DDNet and LANTNet. The DM loss function that yields

Table 4.6 Relationship between different DM loss functions and F1 Score for DDNet.

Loss Function Method	Farmland	Yellow River	Ottawa
MSE + $L_{TV}$ [185]	88.82	77.91	94.62
MSE+10 · SSIM	89.1	81.5	94.73
MSE+5 · SSIM	<b>89.70</b>	<b>90.79</b>	<b>94.87</b>

Table 4.7 Relationship between different DM loss functions and F1 Score for LANTNet.

Loss Function Method	Farmland	Yellow River	Ottawa
MSE + $L_{TV}$ [185]	87.60	79.16	93.90
MSE+10 · SSIM	88.50	80.43	94.20
MSE+5 · SSIM	<b>89.20</b>	<b>91.72</b>	<b>94.88</b>

the highest change detection accuracy is selected. 4.6 and 4.7 show the change detection performance with different DM loss functions, with the highest F1 Score highlighted in bold.

#### 4.4.5 Performance Analysis

To validate the effectiveness of the despeckling model, we compared the results of change detection methods with and without the despeckling model using three real SAR datasets. Figures 4.8, 4.9 and 4.10 demonstrates that the proposed despeckling model considerably enhanced the F1 score for existing (including state-of-the-art) change detection methods. In all these experiments, we empirically set the  $\lambda_{SSIM}$  to be 5 in the loss objective (4.3) as a tradeoff between despeckling and change detection performance, as we investigated in Tables 4.6 and 4.7. It is evident that the performance of the CD methods improves once we pass them through the proposed despeckling model in three SAR datasets. However, in Figure 4.10, the NR-ELM algorithm with DM obtained a lower F1 because Ottawa dataset is less affected by the speckle noise. This is why we see a higher F1 score even with all other methods without DM. Secondly, Compared to other methods, the NR-ELM is more resistant to speckle noise because of the inherent despeckling process encoded within its architecture.

Therefore, the decline in the F1 score when we include the DM module is due to the fact that an additional despeckling process oversmooths the input image, subsequently decreasing the F1 score. These results will be explained in more detail in section 4.4.3.

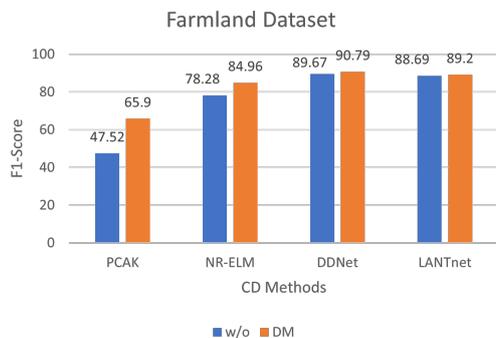


Fig. 4.8 Relationship between DM and F1 score for Farmland dataset

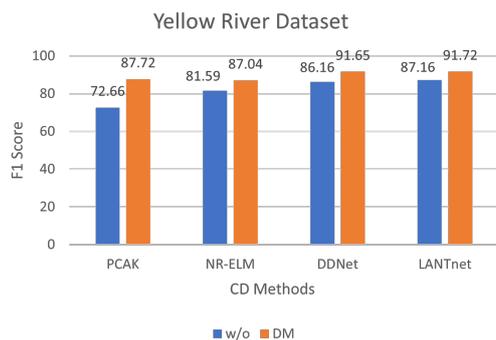


Fig. 4.9 Relationship between DM and F1 score for Yellow River dataset

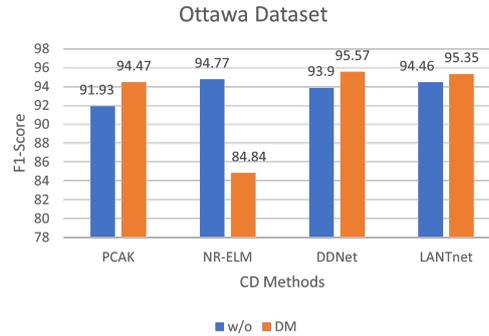


Fig. 4.10 Relationship between DM and F1 score for Ottawa dataset

#### 4.4.6 Change Detection Results

To evaluate the impact of the proposed despeckling model on change detection purposes, we compare the effectiveness of the proposed DM with other existing despeckling methods such as Lee [21], Enhanced Lee [104], SAR2SAR [202] and ID-CNN [185] on three real SAR datasets. Subsequently, we feed the despeckled SAR images to four aforementioned change detection methods, namely PCA- $k$ -means (PCAK) [23], NR-ELM [161], DDNet [2] and LANTNet [164]. PCAK employs principle component analysis for feature extraction and utilises the  $k$ -means clustering algorithm for classification. NR-ELM incorporates the neighbourhood ratio for feature extraction using the difference image, followed by classification using an extreme learning machine. DDNet is a dual-domain network that exploits spatial and frequency domain features to mitigate speckle noise. LANTNet is a layer attention-based noise-tolerant network that leverages the correlation between convolutional layers. Both DDNet and LANTNet are currently state-of-the-art change detection methods.

- **Results of Farmland dataset**

From Figure 4.11, it can be observed that the change map generated by PCAK misclassifies many unchanged pixels compared to GT. The Enhanced Lee filter significantly improves the results for PCAK, increasing the accuracy from 47.44% to 79.44%, while the proposed DM achieves 65.90%. It is worth mentioning that, Farmland

dataset is heavily influenced by speckle noise, and change detection algorithms usually perform poorly compared to Ottawa dataset, which is less affected by speckle noise. Simply applying PCAK, a simple CD method, without despeckling, results in poor performance, as shown in Table 4.8. Another reason for this poor performance, in addition to the speckle noise, is because the pre- and post-change images in Farmland dataset are different looks, i.e., single and multi-looks before and after the change with varying noise levels. Using the despeckling process somewhat takes this into account and improves the performance, as seen in Table 4.8, where all despeckling methods consistently improve the results with PCAK. Specifically, the Enhanced Lee performs the best here because it is well suited for stronger speckle noise and helps PCAK to significantly smooth the image, while DM is designed to support and generically enhance the overall CD performance. NR-ELM produces better results with less noise but misses some changed pixels. The DM filter improves NR-ELM's performance from 78.28% to 84.96%.

Furthermore, DDNet performed better than PCAK and NR-ELM did. The DM enhances the F1 score for DDNet from 86.67% to 89.70%, i.e., it demonstrates higher accuracy than PCAK and NR-ELM, although slightly lower than DDNet, while DM improves the accuracy of LANTNet from 88.69% to 89.20%. The proposed method improved performance after incorporating the DM module, increasing accuracy from 89.91% to 91.28%. Notably, the despeckled data using the SAR2SAR filter performed poorly and yielded lower results than the original methods without the despeckling model. It is evident that the DM outperforms other despeckling methods in terms of the F1 score for the purpose of change detection. Moreover, it consistently outperforms other change detection methods without a DM. In other words, the DM suppresses speckle noise even when two Farmland image pairs have different looks, such as single-look (pre-change) and four-look (post-change). This type of suppression is

reflected positively in the performance of the change detection methods as shown in Table 4.8.

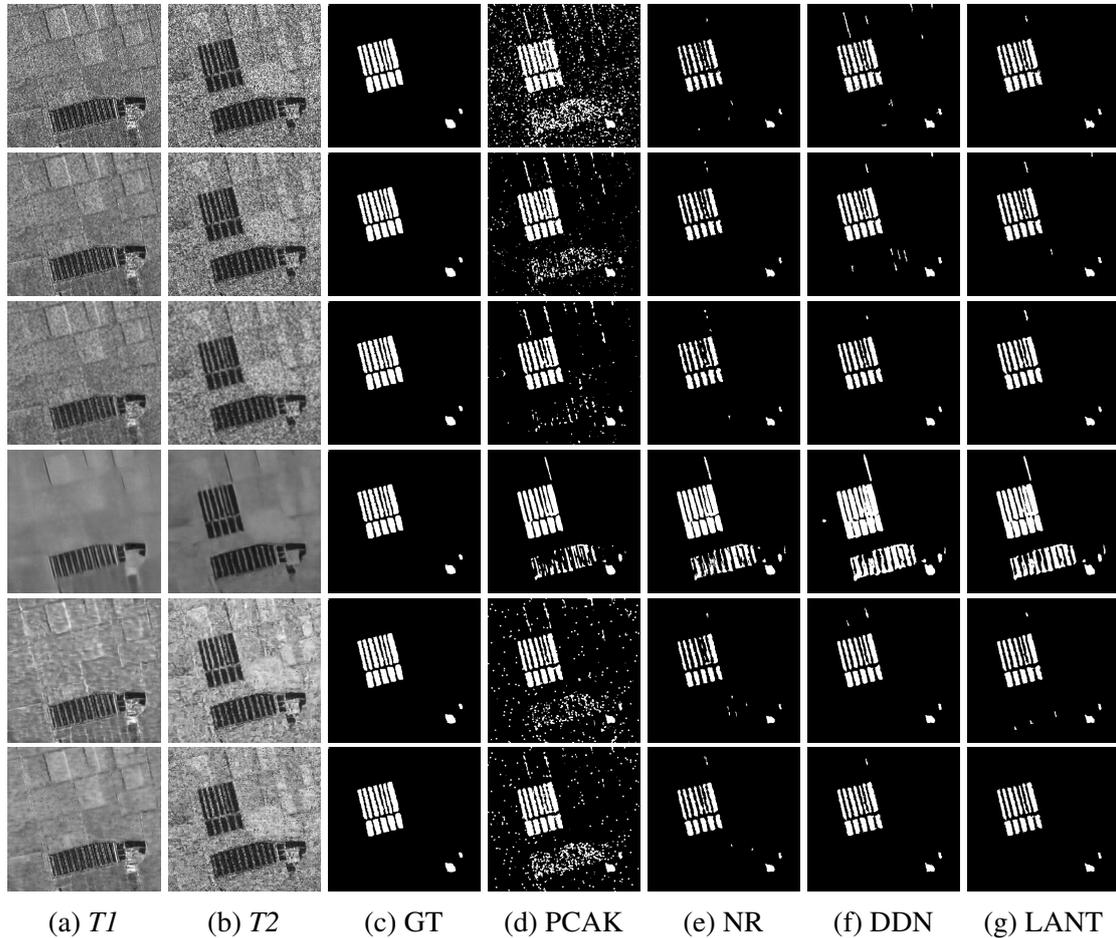


Fig. 4.11 Visualised results of Farmland dataset with different despeckling methods  
 Rows: (1<sup>st</sup> row) Farmland without despeckling (w/o), (2<sup>nd</sup> row) Farmland despeckled with lee, (3<sup>rd</sup> row) Farmland despeckled with enhanced lee, (4<sup>th</sup> row) Farmland despeckled with SAR2SAR, (5<sup>th</sup> row) Farmland despeckled with IDCNN, (6<sup>th</sup> row) Farmland despeckled with proposed DM. Columns: Farmland image captured at (a)  $T_1$  and (b)  $T_2$ . (c) refers to the ground truth (GT) image. Results obtained by methods (d) PCAK [23], (e) NR-ELM [161], (f) DDNet [2], and (g) LANTNet [164].

#### • Results of Yellow River dataset

In Figure 4.12, it is noticeable that the change map generated by PCAK misclassifies many unchanged pixels as changed ones compared with the GT. The Lee filter reduces speckle noise and improves the CM. The DM performs as the best filter, effectively

Table 4.8 Quantitative evaluation on Farmland change detection based on different despeckling filters.

Here, w/o means it is the original method without despeckling, DM is our proposed despeckling model while S2S is SAR2SAR, Enh refers to Enhanced Lee, OA is Overall Accuracy, NR is NR-ELM, DDN is DDNet and LANT is LANTNet

Methods	Metrics	w/o	Lee [21]	Enh [104]	S2S [202]	ID-CNN [185]	DM
PCAK [23]	Recall↑	90.04	90.87	67.32	95.69	85.45	90.76
	Precision↑	32.27	57.73	96.89	66.51	54.35	51.74
	OA↑	88.22	95.53	97.94	96.89	94.89	94.44
	F1↑	47.52	70.60	<b>79.44</b>	78.48	66.44	65.90
NR [161]	Recall↑	65.20	68.82	66.52	97.50	66.64	75.39
	Precision↑	97.92	99.13	98.51	57.17	97.23	97.33
	OA↑	97.86	98.12	97.96	95.52	97.91	98.42
	F1↑	78.28	81.24	79.42	72.08	79.08	<b>84.96</b>
DDN [2]	Recall↑	82.26	86.58	78.25	99.26	81.52	82.81
	Precision↑	91.59	92.76	98.21	48.66	97.57	97.85
	OA↑	98.50	98.81	98.63	93.76	98.79	98.87
	F1↑	86.67	89.57	87.11	65.30	88.82	<b>89.70</b>
LANT [164]	Recall↑	81.35	80.51	81.76	98.46	79.87	81.18
	Precision↑	97.50	96.14	96.27	52.74	96.98	98.98
	OA↑	98.77	98.66	98.73	94.69	98.66	98.84
	F1↑	88.69	87.64	88.42	68.69	87.60	<b>89.20</b>

suppressing noise and significantly improving the F1 score from 72.66% to 87.7% for the PCAK method. NR-ELM produces better results with less noise but misses some changed pixels, whereas the DM filter enhances NR-ELM's performance from 81.59% to 87.04%.

Furthermore, DDNet outperformed PCAK and NR-ELM results. The DM considerably enhanced the F1 score from DDNet from 86.65% to 90.79%. LANTNet achieves higher accuracy than PCAK and NR-ELM. DM has enhanced the F1-score for LANTNet from 88.44% to 91.1%. After applying the proposed DM, the proposed method's performance has improved from 88.44% to 91.83%. It is worth mentioning that the despeckled data using the SAR2SAR filter does not perform well and yields lower

results compared to the original methods without despeckling, such as DDNet and LANTNet. It is evident that DM achieves a superior F1 score for change detection methods compared to other despeckling methods due to the ability to efficiently cope with the single-look pre-change and multi-look post-change SAR images via robust loss function.

- **Results of Ottawa dataset**

Compared to previous datasets, the Ottawa dataset is less affected by speckle noise. This is evident from the achieved better change detection results of 91.93% using the PCAK method without any despeckling process on Ottawa dataset compared to the previous two datasets. Including the proposed DM further improves the F1 score value from 91.93% to 94.47%. NR-ELM [161] provides better results compared to PCAK, Lee slightly improves the F1 score from 94.15% to 94.77%, whereas DM reduces the performance to 84.84% as shown in Figure 4.10 and Table 4.13. The proposed DM with the NR-ELM degrades the performance because of oversmoothing. This is because NR-ELM has an inherent despeckling process encoded within its architecture. Moreover, this is also the case for other despeckling methods except the Lee method, which does not degrade (but slightly improve) the performance. A possible reason for this could be because, in comparison, Lee [21] is the least strong despeckling method and, therefore, does not result in much oversmoothing, which degrades the performance.

DDNet performed better than PCAK and NR-ELM, and the proposed DM improves the F1 score for DDNet from 93.90% to 94.87%. LANTNet produces better accuracy than PCAK, NR-ELM and DDNet. Its accuracy has further improved by the proposed DM from 94.46% to 94.88%. With the proposed loss objective, the performance slightly improves from 94.46% to 94.50%, which is further enhanced from 94.50% to 95.79% when used in conjunction with the DM as shown in Figure 4.13 and Table 4.8. It can

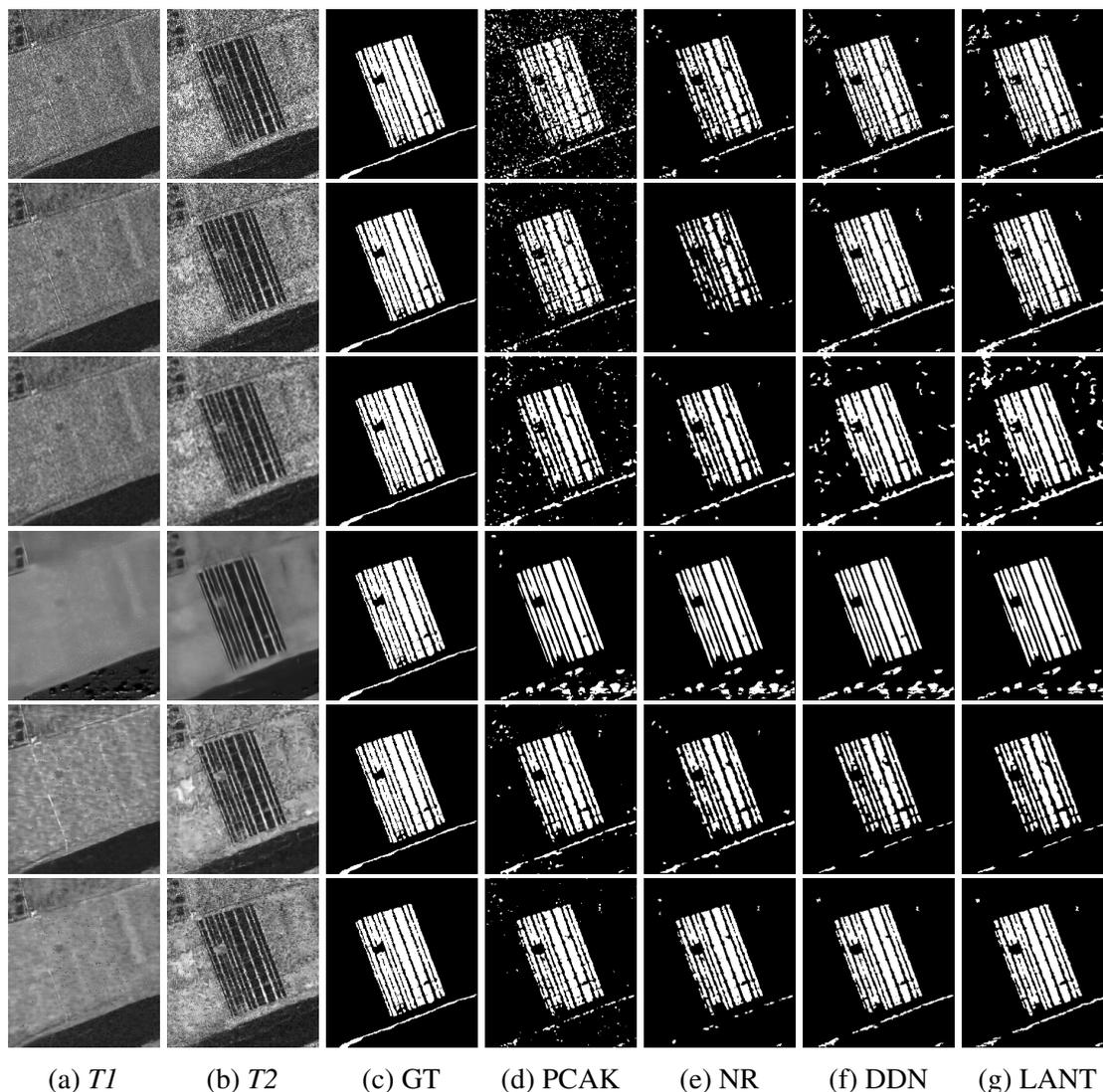


Fig. 4.12 Visualised results of Yellow River dataset with different despeckling methods Rows: (1<sup>st</sup> row) Farmland without despeckling (w/o), (2<sup>nd</sup> row) Farmland despeckled with lee, (3<sup>rd</sup> row) Farmland despeckled with enhanced lee, (4<sup>th</sup> row) Farmland despeckled with SAR2SAR, (5<sup>th</sup> row) Farmland despeckled with IDCNN, (6<sup>th</sup> row) Farmland despeckled with proposed DM. Columns: Farmland image captured at (a)  $T_1$  and (b)  $T_2$ . (c) refers to the ground truth (GT) image. Results obtained by methods (d) PCAK [23], (e) NR-ELM [161], (f) DDNet [2], and (g) LANTNet [164].

be observed from the Ottawa dataset results that the CD methods without despeckling already perform well because the data is less affected by noise. Nevertheless, with DM, the performance of these CD methods was further improved.

In summary, addressing our hypothesis, a high Equivalent Number of Looks score in SAR image processing signifies effective noise reduction, resulting in a less noisy and more uniform image. However, this heightened noise reduction can introduce a tradeoff with change detection sensitivity. As observed in our previous results, a high ENL score often corresponds to an oversmoothing effect on SAR images, leading to the potential loss of subtle or small changes in the image. This loss is attributed to the smoothing or averaging processes employed for noise reduction, which can blur or merge neighbouring pixels, making it challenging to distinguish actual changes from noise. Consequently, this tradeoff results in a lower F1 score. The challenge lies in striking a balance between noise reduction and maintaining change detection sensitivity, a common dilemma in SAR image processing. Our findings show that our proposed despeckling model does not necessarily attain the highest ENL score. Still, it excels in achieving a superior F1 score when compared to other despeckling methods. This superiority can be attributed to DM's ability to handle single-look pre-change and multi-look post-change SAR images efficiently.

## 4.5 Summary

This chapter centred around the goal of enhancing the accuracy of change detection in SAR imagery by addressing the significant issue of speckle noise. Speckle noise, characterised by granular patterns in SAR images, posed a major challenge to achieving accurate change detection. Various despeckling methods were developed to mitigate this noise, including filtering and statistical modelling. In response to this challenge, we introduced a novel despeckling model (DM) based on deep convolutional neural networks to enhance SAR change detection performance.

The methodology detailed in this chapter was designed to test the hypothesis that intensifying the despeckling of SAR images before conducting change detection could lead to more

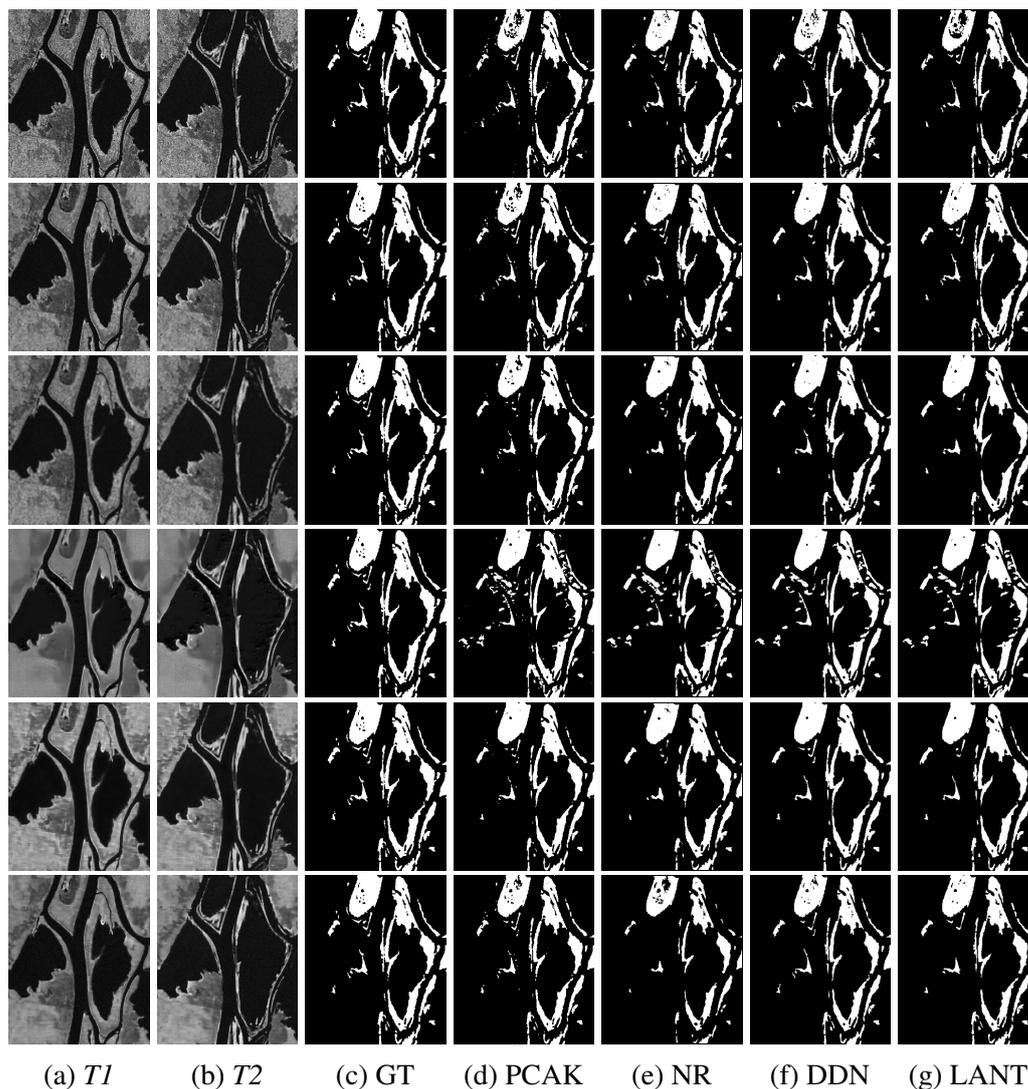


Fig. 4.13 Visualised results of Ottawa dataset with different despeckling methods  
 Rows: (1<sup>st</sup> row) Farmland without despeckling (w/o), (2<sup>nd</sup> row) Farmland despeckled with lee, (3<sup>rd</sup> row) Farmland despeckled with enhanced lee, (4<sup>th</sup> row) Farmland despeckled with SAR2SAR, (5<sup>th</sup> row) Farmland despeckled with IDCNN, (6<sup>th</sup> row) Farmland despeckled with proposed DM. Columns: Farmland image captured at (a)  $T_1$  and (b)  $T_2$ . (c) refers to the ground truth (GT) image. Results obtained by methods (d) PCAK [23], (e) NR-ELM [161], (f) DDNet [2], and (g) LANTNet [164].

accurate results. The structured approach involved despeckling SAR datasets, computing despeckling metrics for different methods, applying change detection to the despeckled datasets, and comparing the change detection metrics with despeckling metrics to validate

Table 4.9 Quantitative evaluation on Ottawa change detection based on different despeckling filters

Here, w/o means it is the original method without despeckling, DM is our proposed despeckling model while S2S is SAR2SAR, Enh refers to Enhanced Lee, OA is Overall Accuracy, NR is NR-ELM, DDN is DDNet, and LATNT is LANTNet.

Methods	Metrics	w/o	Lee [21]	Enh [104]	S2S [202]	ID-CNN [185]	DM
PCAK [23]	Recall $\uparrow$	88.16	91.58	88.74	88.85	92.01	91.00
	Precision $\uparrow$	96.05	96.28	97.74	82.14	96.85	98.20
	OA $\uparrow$	97.55	98.11	97.89	95.18	98.26	98.31
	F1 $\uparrow$	91.93	93.87	93.02	85.36	94.37	<b>94.47</b>
NR [161]	Recall $\uparrow$	93.14	94.79	87.67	88.68	92.07	73.92
	Precision $\uparrow$	95.19	94.74	94.59	80.56	92.65	99.53
	OA $\uparrow$	98.17	98.34	97.25	94.82	97.59	95.82
	F1 $\uparrow$	94.15	<b>94.77</b>	91.00	84.42	92.36	84.84
DDN [2]	Recall $\uparrow$	92.70	93.66	93.66	90.78	94.51	91.71
	Precision $\uparrow$	95.12	96.06	96.06	82.91	94.73	98.26
	OA $\uparrow$	98.09	98.39	98.39	95.58	98.30	98.43
	F1 $\uparrow$	93.90	94.84	94.85	86.67	94.62	<b>94.87</b>
LANT [164]	Recall $\uparrow$	91.8	94.67	90.73	89.91	92.62	91.66
	Precision $\uparrow$	97.30	94.48	95.11	82.49	95.23	98.33
	OA $\uparrow$	98.3	98.28	97.80	95.39	98.1	98.44
	F1 $\uparrow$	94.46	94.57	92.87	86.04	93.90	<b>94.88</b>

the hypothesis. The proposed despeckling model played a pivotal role in this methodology by effectively reducing speckle noise and improving overall image quality.

To validate its findings, we utilised two types of datasets: synthetic SAR images from the Berkeley Segmentation Dataset 500 and real SAR images from datasets such as Farmland, Yellow River, and Ottawa. The experimental results demonstrated that the proposed DM significantly enhanced change detection accuracy compared to other despeckling techniques. It struck a balance between noise reduction and preserving sensitivity to changes in SAR images, resulting in substantially improved change detection performance.

The work in this chapter focused on improving SAR change detection by addressing the challenge of speckle noise. It introduced the Despeckling Model, outlined a systematic

Table 4.10 Quantitative evaluation on Yellow River change detection based on different despeckling filters.

Here w/o means it is the original method without despeckling, DM is our proposed despeckling model while S2S is SAR2SAR, Enh refers to Enhanced Lee, OA is Overall Accuracy, NR is NR-ELM, DDN is DDNet, LANT is LANTNet

Methods	Metrics	w/o	Lee [21]	Enh [104]	S2S [202]	ID-CNN [185]	DM
PCA-K [23]	Recall↑	74.96	78.40	74.52	81.80	81.37	82.59
	Precision↑	70.50	87.80	82.75	83.31	92.79	93.53
	OA↑	89.80	94.12	92.58	93.74	95.49	95.82
	F1↑	72.66	82.84	78.42	82.55	86.70	<b>87.72</b>
NR [161]	Recall↑	72.18	48.35	70.19	78.30	79.76	79.32
	Precision↑	93.83	99.72	92.22	85.53	95.08	96.42
	OA↑	94.11	90.63	93.54	93.68	95.59	95.73
	F1↑	81.59	65.13	79.71	81.76	86.75	<b>87.04</b>
DDN [2]	Recall↑	83.46	86.32	82.86	80.46	64.06	86.58
	Precision↑	90.09	91.41	81.89	85.00	90.40	95.44
	OA↑	95.35	96.06	93.59	93.90	93.43	96.83
	F1↑	86.65	88.79	82.37	82.67	77.91	<b>90.79</b>
LANT [164]	Recall↑	82.44	84.00	83.03	79.84	65.93	87.51
	Precision↑	92.45	91.18	71.49	87.83	99.04	94.99
	OA↑	95.61	95.64	90.94	94.35	93.72	96.91
	F1↑	87.16	87.44	76.83	83.64	79.16	<b>91.1</b>

methodology, and demonstrated its effectiveness in enhancing change detection performance. This research contributed to the field of SAR image processing by providing a robust solution to mitigate the disruptive effects of speckle noise on change detection outcomes. The next chapter will concentrate on Deep learning based on change detection for SAR images.

# Chapter 5

## Deep Learning Based on Change Detection for SAR Images

### 5.1 Overview

Change detection is one of the important tasks in computer vision. It is responsible for highlighting the differences in an area from multi-temporal satellite images captured for the same geographical region at different periods [23, 6]. Some of the change detection techniques have been discussed in Chapter 2. The most challenging aspect of SAR change detection methods is speckle noise [23, 24, 2]. Chapter 3 debates the importance of image registration for the change detection problem. Moreover, Chapter 4 highlights the importance of reducing speckle noise to improve the change detection map and F1 Score.

This chapter will explore and develop deep learning in change detection to reduce speckle noise and improve the F1 score. Deep learning, such as conventional neural networks, has been employed to solve computer vision challenges and achieve state-of-the-art results in these problems, such as image classification, semantic segmentation, and object detection [203, 130, 204, 205]. Therefore, this chapter aims to produce a convolutional neural network to enhance change detection performance. The results of this method will be com-

pared to the state-of-the-art techniques in change detection problems, such as DDNET and LANTNet.

These two objectives will be discussed in detail in section 5.3. The remaining sections are structured as follows: related work briefly discusses change detection techniques. Then, the methodology will discuss the proposed network. Subsequently, three co-registered SAR datasets have been discussed. Next is the experimental results and discussion section. This section discusses the environment setup, evaluation metrics, results, and discussion. Finally, here is the summary of this chapter.

## 5.2 Methodology

Existing unsupervised change detection methods utilise clustering algorithms such as hierarchical Fuzzy C-Means [206] and Fuzzy C-Means (FCM) [207] to generate pseudo-labels with a high probability for network training. While this method solves the need for label data, errors commonly affect network performance. In addition, the attention mechanism is utilised to emphasise the essential parts of the input while disregarding irrelevant information, but it often neglects the correlations among multiple convolution layers. Meng et al. [164] proposed a layer attention module to weigh features from different layers based on the learned correlation matrix to address this limitation. LANTNet is a module that effectively combines spatial information from low-level layers with semantic information from high-level layers, emphasising informative layers and suppressing redundant ones. The process involves matrix multiplication to assign adaptive weights to the input feature groups, followed by calculating the attention matrix using a softmax operation. The weighted feature matrix is then multiplied by the attention matrix, reshaped, and combined with the original input to produce the final output. The change map is generated through a series of convolution and fully connected layers. The trained network can classify all pixels from the multitemporal SAR images to obtain the final change map as shown in Figure 5.1.

### 5.2.1 Baseline Architecture

In the pre-classification module, we employ the logarithmic ratio operator and the hierarchical FCM algorithm, and you can find more detailed information about this module in [160]. After pre-classification, we extract image patches ( $R \times R$  pixels) from  $I_1$ ,  $I_2$ , and the difference image, each labelled accordingly. These patches are then combined to create new image patches ( $3 \times R \times R$  pixels), which serve as our training data. The convolutional stem consists of four convolution layers. Initially, a  $1 \times 1$  convolution is used to extract a shallow feature denoted as  $F_0$ . Subsequently, we use three  $3 \times 3$  convolution layers to extract intermediate features  $F_i$ , where  $i$  ranges from 1 to 3. For the shallow feature  $F_0$ , we set the channel count to 16, and for the intermediate features  $F_1$ ,  $F_2$ , and  $F_3$ , we set the channel count to 32. We then apply a  $1 \times 1$  convolution to expand the channel dimensions of  $F_0$  to 32. Following this,  $F_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  are merged to create a feature group, which is then input into the layer attention module.

### 5.2.2 Enhanced Loss Function

In this section, we adapt the training strategy and propose a loss function that is more noise-resistant to speckle noise. However, this loss function does not provide satisfactory performance. To this end, we first designed a robust loss function that is more resistant to speckle noise. The loss function combines MSE and Kullback-Leibler Divergence ( $KL$ ). The loss function is expressed as follows:

$$L_{\text{MSE}}(X, \hat{X}) = \|X - \hat{X}\|^2 \quad (5.1)$$

$$L_{\text{KL}}(X, \hat{X}) = \hat{X} \cdot (\log \hat{X} - X) \quad (5.2)$$

$$L_T = \alpha \cdot L_{MSE} + \beta \cdot L_{KL} \quad (5.3)$$

where  $\alpha$  and  $\beta$  are two weighting hyperparameters that control the relative significance of the individual components within the model, enabling the precise adjustments to enhance the performance based on the assigned values of  $\alpha$  and  $\beta$ . Based on our empirical study,  $\alpha$  and  $\beta$  were set to 0.9 and 0.1 to trade off noise robustness and convergence efficiency. The KL acts similarly to CE, with the difference that CE penalises the network based on its predictions. In contrast, KL mainly evaluates the disparity between the probability distribution predicted by the network and the distribution of the reference ground truth. Therefore, we argue that combining MSE and KL can provide a better change detection performance and suppress speckle noise (see Section 5.3.2). In the following section, we present the results of our proposed methodology along with the training details.

### 5.2.3 Enhanced Architecture

In this section, we have made enhancements to LANTNet's parameters to improve its change detection performance further. Specifically, we have adjusted the number of training epochs, increasing it to 70 epochs to allow for more comprehensive model training. Furthermore, we have fine-tuned the learning rate, setting it at 0.001, facilitating a more precise convergence of the model during the training process. The chosen of these parameters was based on the experiments in Table 5.1. These parameter adjustments are instrumental in optimising LANTNet's effectiveness in detecting changes within SAR imagery.

## 5.3 Experimental Results & Discussion

Experiments were conducted on three co-registered SAR datasets, which are introduced in detail in this section. Python 3.7 was used to perform the experiment. The platform of

Table 5.1 Enhanced Architecture

Where  $lr$  learning rate and Epochs is the number of epochs in the experiment

Parameters	Farmland	Yellow River	Ottawa
Lr 0.01 Epochs 50	82.14	83.85	93.42
Lr 0.001 Epochs 50	85.48	86.91	94.94
Lr 0.01 Epochs 60	88.25	87.15	93.79
Lr 0.001 Epochs 60	89.67	87.06	94.80
Lr 0.01 Epochs 70	88.22	87.33	95.35
<b>Lr 0.001 Epochs 70</b>	<b>89.91</b>	<b>88.44</b>	<b>95.35</b>

the experiments was Tesla GPU P100-PCIE-16 GB RAM 147.15 GB Disk. The result and discussion will be discussed in the next subsection.

### 5.3.1 Datasets

Three co-registered SAR change detection datasets were discussed in the previous chapter. Below is the description for SAR change detection datasets:

- Farmland and Yellow River Datasets

These two datasets were captured by RADARSAT-2 in the region of Yellow River Estuary in China in June 2008 and June 2009, respectively. Hence, the influence of speckle noise on the image captured in 2008 is much greater than that of the one acquired in 2009

- Ottawa dataset It captured two images of a flooding event that caused a rise of lakes near Ottawa, Canada. It was taken by the RADARSAT-1 satellite SAR sensor in July and August 1997. The size of each image is  $290 \times 350$  pixels. [3, 2, 1]. Table 5.2 presents the specifications of datasets used in this chapter.

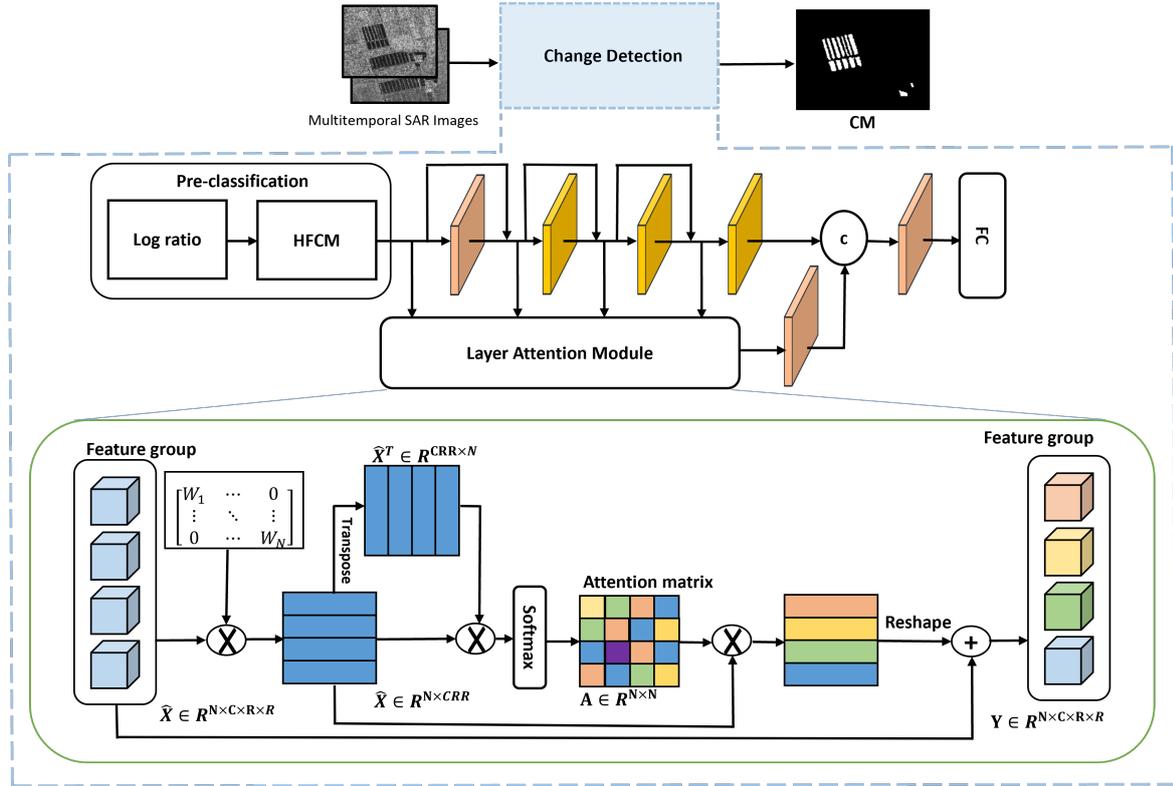


Fig. 5.1 Proposed Change Detection Method

### 5.3.2 Ablation Study

This section examines how well our new loss function performs in LANTNet. We do this by comparing it to other commonly used loss functions in similar applications. To make sure our evaluation covers a wide range of scenarios, we use three different SAR change detection datasets. By comparing these results, we aim to demonstrate how our proposed loss function improves change detection accuracy when used in LANTNet as presented in Table 5.3 and

Table 5.2 Multi-temporal SAR images specifications.

$H$  is height, and  $W$  is width

Dataset	Name	Type of Image	Image H*W in Pixels
1	Farmland	SAR	306*291
2	Yellow River	SAR	257*289
3	Ottawa	SAR	290*350

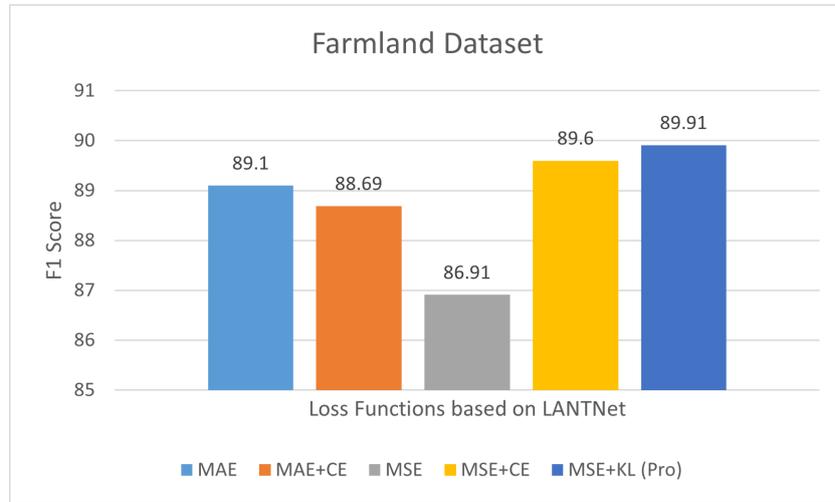


Fig. 5.2 Visualised Proposed loss function results of Farmland with different loss function

Table 5.3 Ablation Studies of the Proposed Method for LANTNet.

Loss Function Method	Farmland	Yellow River	Ottawa
MAE	89.1	87.24	94.24
MAE+CE [164]	88.69	87.16	94.46
MSE	86.91	86.22	94.75
MSE+CE	89.60	88.23	94.54
MSE+KL(Pro)	<b>89.91</b>	<b>88.44</b>	<b>95.35</b>

Figures 5.2, 5.3 and 5.4. This analysis will help us understand our loss function's strengths, showing its resistance against speckle noise in the SAR change detection task.

The proposed loss function has undergone rigorous comparisons with existing loss functions across three distinct SAR change detection datasets. In this comprehensive evaluation, the proposed loss function consistently outperformed all other counterparts, as evidenced by the results depicted in Table 5.3 and Figures 5.2, 5.3 and 5.4. Consequently, we have chosen to employ this superior loss function within our novel change detection method. By doing so, we aim to mitigate the adverse effects of speckle noise and enhance the overall performance of our change detection system.

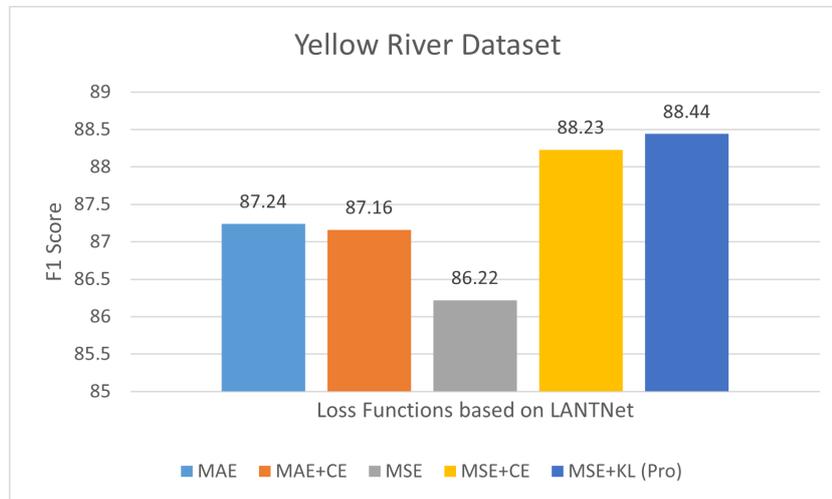


Fig. 5.3 Visualised Proposed loss function results of Yellow River with different loss function

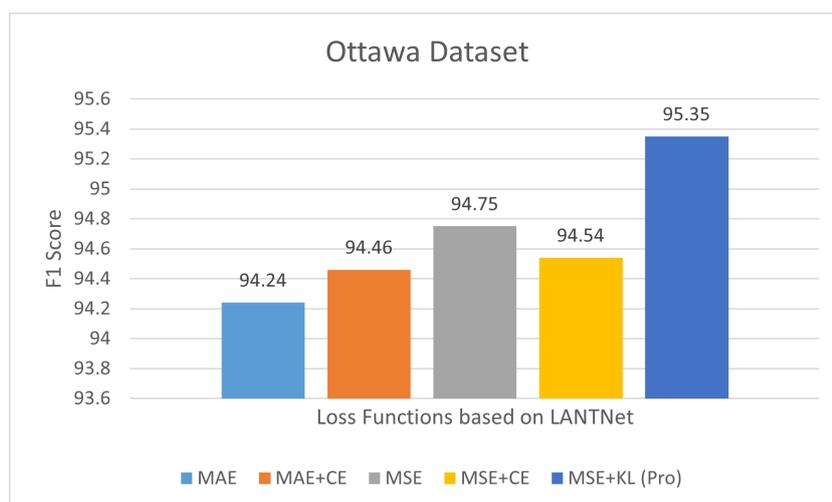


Fig. 5.4 Visualised Proposed loss function results of Ottawa with different loss functions

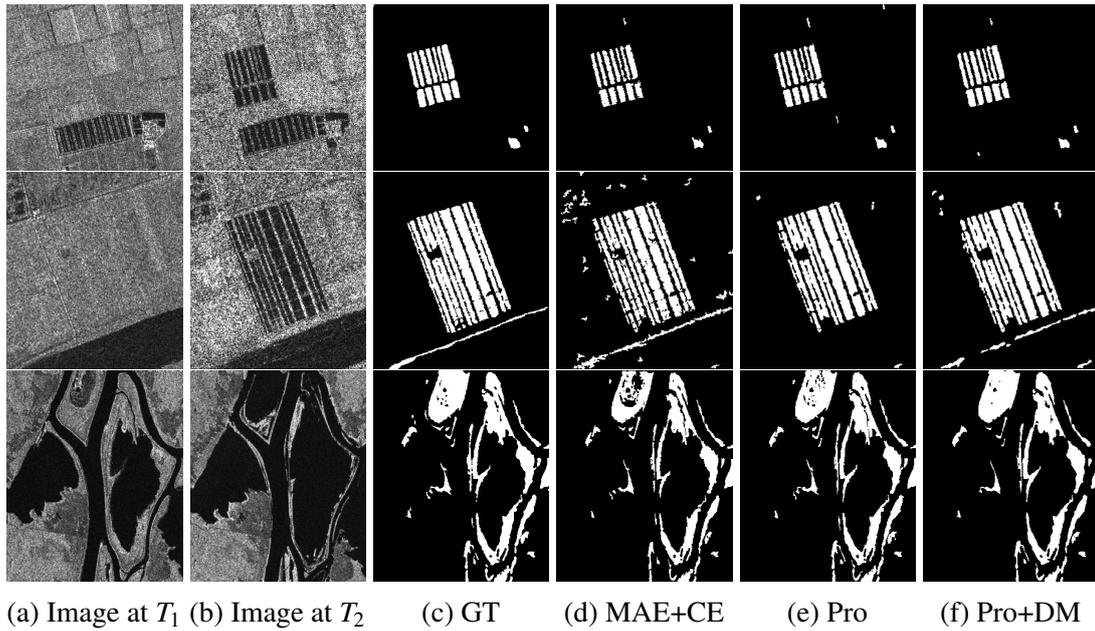


Fig. 5.5 Visualised results of change detection methods on the (1<sup>st</sup> row) Farmland dataset, the (2<sup>nd</sup> row) Yellow River, and the (3<sup>rd</sup> row) Ottawa dataset. Image captured at (a) and (b). (c) GT, (d) loss function MAE+ CE and (e) results by the proposed method MSE +KL(Pro).

### 5.3.3 Results & Discussions

To assess the effectiveness of our proposed approach, we conduct a comparative analysis against several established change detection methods, including PCA- $k$ -means (PCKA)[23], NR-ELM[161], DDNet [2], LANTNet [164] and Pro + DM. PCKA relies on principal component analysis for feature extraction and employs the PCKA-means clustering algorithm for classification. NR-ELM incorporates neighbourhood ratio for feature extraction, utilising the difference image, and subsequently applies an extreme learning machine for classification. DDNet is a dual-domain network that harnesses spatial and frequency domain features to mitigate speckle noise. LANTNet, on the other hand, is a noise-tolerant network that incorporates layer attention mechanisms to exploit correlations between convolutional layers. Pro + DM refers to the proposed method after despeckling SAR CD datasets with the despeckling model we introduce in Chapter 4.

- **Results on the Farmland Dataset:**

From Figure 5.5 and Table 5.4, it can be observed that the change map generated by PCAK misclassifies many unchanged pixels as changed ones compared with the GT which increase FP value and produces insufficient CD performance 47.52%. NR-ELM produces better results. This can be seen from the small value of FP 7. This leads to yielding the F1 score to 78.28%. Furthermore, DDNet outperformed PCAK and NR-ELM results. DDNet obtains 86.67% score. This happens because FP and FN are very small values compared to the TP values. LANTNet gets better performance than the DDNet method. The proposed method outperforms all previous methods because the proposed loss function is more resistant to speckle noise. It achieves 89.91% accuracy. Finally, (Pro + DM), which applies the despeckling model before the proposed method, can reduce speckle noise and considerably enhance CD performance from 88.91% to 91.28%.

Table 5.4 Change Detection Results of Farmland Dataset.

$R$  here is Recall,  $P$  is Precision, Pro is the proposed method, and Pro + DM is the proposed method with the despeckling model

Methods	TP	TN	FP	FN	R	P	OA	F1 Score <sup>↑</sup>
PCAK [23]	4745	73817	9959	525	90.04	32.27	95.53	47.52
NR [161]	3436	83703	73	1834	65.20	97.92	97.86	78.28
DDNet [2]	4335	83378	398	935	82.26	91.59	98.81	86.67
LANTNet [164]	4287	83666	110	983	81.35	97.50	98.77	88.69
Pro	4460	83594	180	810	84.08	97.67	98.89	<b>89.91</b>
Pro + DM	4390	83709	67	880	86.45	96.67	99.02	<b>91.28</b>

- **Results on the Yellow River Dataset:**

From Table 5.5, distinct trends emerge in the performance of various change detection methodologies. Firstly, while PCAK yields reasonably satisfactory results, NR-ELM exhibits superior accuracy at 81.59%, owing to its enhanced resistance to speckle

noise compared to PCA. DDNet surpasses PCAK and NR-ELM by achieving an impressive score of 86.65%, primarily due to notably reduced instances of false positives and false negatives. Exceeding the performance of DDNet, LANTNet attains a score of 87.16%, signifying its superior change detection capabilities. Our proposed method emerges as the top-performing approach, achieving an accuracy rate of 88.44%, primarily attributable to its innovative loss function specifically designed to mitigate the influence of speckle noise. To further enhance performance, we introduce (Pro + DM), which mitigates speckle noise and significantly augments change detection accuracy to an impressive 91.83%.

Table 5.5 Change Detection Results of Yellow River Dataset.

$R$  here is Recall,  $P$  is Precision, Pro is the proposed method, and Pro + DM is the proposed method with the despeckling model

Methods	TP	TN	FP	FN	R	P	OA	F1 Score $\uparrow$
PCAK [23]	10068	56628	42213	3364	74.96	70.50	89.80	72.66
NR [161]	9695	60204	637	3737	72.18	93.83	94.11	81.59
DDNet [2]	11210	59608	1233	2222	83.46	90.09	95.35	86.65
LANTNet [164]	11074	59936	905	2358	82.44	92.45	95.64	87.16
Pro	11293	60028	813	2139	84.08	93.28	96.03	<b>88.44</b>
Pro + DM	11861	59920	921	1571	89.53	94.25	97.12	<b>91.83</b>

- **Results on the Ottawa Dataset:**

As discussed in Chapter 4, the Ottawa dataset is less affected by speckle noise. This is evident from the achieved better change detection results of 91.93% using the PCAK method. NR-ELM provides better results compared to PCAK, with an F1 score of 94.15% as shown in Figure 5.5 and Table 5.6. DDNet performed better than PCAK but slightly less than NR-ELM, with an F1 score for DDNet of 93.90%. LANTNet produces better accuracy than PCAK, NR-ELM and DDNet, with an F1 score of 94.46%. The proposed method obtains the highest performance than the previous CD methods, and its accuracy has further improved by the proposed DM from 94.50%

to 95.79%. It can be observed from the Ottawa dataset results that the CD methods without despeckling already perform well because the data is less affected by noise. Nevertheless, with DM, the performance of these CD methods was further improved.

Table 5.6 Change Detection Results of Ottawa Dataset.

$R$  here is Recall,  $P$  is Precision, Pro is the proposed method, and Pro + DM is the proposed method with the despeckling model

Methods	TP	TN	FP	FN	R	P	OA	F1 Score $\uparrow$
PCAK [23]	14148	84869	582	1901	88.16	96.05	97.55	91.93
NR [161]	14948	84695	756	1101	93.14	95.19	98.17	94.15
DDNet [2]	14878	84688	763	1171	92.70	95.12	98.09	93.90
LANTNet [164]	14733	85042	409	1316	91.8	97.30	98.3	94.46
Pro	14717	85069	382	1332	91.70	97.47	98.31	<b>94.50</b>
Pro + DM	15042	85137	314	1007	93.73	97.96	98.70	<b>95.79</b>

## 5.4 Summary

In this chapter, we explored the improved change detection performance achieved through the novel inclusion of a specialised loss function. This innovation demonstrated substantial advantages over current state-of-the-art change detection methods. Furthermore, we provided a detailed account of the methodology employed to enhance LANTNet’s training regime. Through extensive experimentation with different hyperparameters, we aimed to augment its proficiency in change detection further. Specifically, we extended the number of training epochs to 70, fostering a more comprehensive model training process. Additionally, we fine-tuned the learning rate to 0.001, enhancing model convergence during training. These parameter adjustments played a pivotal role in optimising LANTNet’s effectiveness in identifying changes within SAR imagery.

We utilised three co-registered SAR change detection datasets, including the Farmland, Yellow River, and Ottawa datasets. Our experiments were conducted using a Google Co-

laboratory Pro environment with a Tesla GPU P100-PCIE, 16 GB RAM, and 147.15 GB Disk. We evaluated the performance of different loss functions, including MAE, MAE+CE, MSE, MSE+CE, and MSE+KL(Pro), within the LANTNet framework. This analysis aimed to demonstrate the effectiveness of the proposed loss function in improving change detection accuracy, particularly in scenarios affected by speckle noise. We compared our approach to established change detection methods, such as PCA-k-means (PCK), NR-ELM, DDNet, LANTNet, and Pro + DM. The results showed that our method consistently outperformed these methods, achieving higher accuracy rates, especially when dealing with speckle noise-affected data.

This chapter provided insights into the enhancements made to LANTNet's parameters and presented a thorough evaluation of the performance of different loss functions. We demonstrated the effectiveness of our proposed approach in improving change detection accuracy, with a particular focus on datasets affected by speckle noise. The next chapter will discuss the conclusion and future work.



# Chapter 6

## Conclusion & Future Work

### 6.1 Summary of Research Findings

Change detection, which is based on synthetic aperture radar, is a critical topic in remote sensing and computer vision areas. Change detection has gained more attention after the success of deep learning methods in providing better accuracy compared to existing methods. This research has concentrated on SAR images for change detection because they have an advantage over optical change detection (CD); they can work in darkness and in difficult weather. With this advantage, SAR CD can assist in assessing disasters during darkness and bad weather. SAR CD identifies changes in two multi-temporal SAR images for the same geographical region. We aim to systematically investigate and enhance the performance of SAR change detection methods, aiming to address the issue of speckle noise that adversely affects the accuracy of these methods. This aim can be addressed by the following questions: first, can the enhancement of image registration accuracy lead to an improvement in change detection performance? Second, does despeckling SAR images prior to change detection methods enhance CD performance? Finally, We explore different deep-learning approaches for SAR change detection. We specifically try to address whether deep learning can deal with speckle noise without despeckling techniques as a pre-processing step.

By answering these questions in this thesis, we will advance our understanding of SAR change detection tasks. In Chapter 3, we investigated the importance and need of image registration for SAR CD by comparing several commonly used image registration-based feature detection algorithms. Subsequently, we applied Ostu thresholding with a Gaussian filter to remove the influence of water surface areas that reduced the capacity of change detection methods. In Chapter 4, we proposed a solution for the main challenge of SAR CD, which is speckle noise. Speckle noise influences SAR images and reduces the performance of change detection methods. We developed a convolutional neural network to despeckle speckle noise in SAR CD datasets before feeding them to change detection methods. Our proposed despeckling model outperforms other despeckling methods with regard to the  $F1$  score for the purpose of change detection using three SAR change detection datasets. Moreover, we investigated our hypothesis that the more despeckling SAR CD datasets, the better CD performance ( $F1$ ) was obtained by comparing the ENL for all despeckling methods and the  $F1$  score for change detection methods. We found that this hypothesis is not completely accurate because as we increase the despeckling of SAR change detection on certain datasets, the image becomes oversmoothed, which results in the loss of essential information in SAR images. This ultimately results in a degradation of the overall change detection performance.

In Chapter 5, we addressed the final research question by enhancing the performance of LANTNet [164]. Our significant contribution involves introducing a novel loss function that combines two existing ones, rendering it more resilient to speckle noise. This innovation substantially improved change detection performance, surpassing the capabilities of current state-of-the-art methods. Furthermore, these enhancements required the fine-tuning of critical parameters, including extending the training regime by increasing the number of training epochs to 70 and optimising the learning rate to 0.001. These adjustments played a pivotal role in maximising the effectiveness of LANTNet for detecting changes in SAR imagery. We

systematically evaluated various loss functions, including  $MAE$ ,  $MAE + CE$ ,  $MSE$ ,  $MSE + CE$ , and  $MSE + KL(Pro)$ , within the LANTNet framework. This comprehensive analysis aimed to showcase the superior performance of our proposed loss function, particularly in scenarios affected by speckle noise.

Through benchmarking our method against established change detection approaches, including PCA- $k$ -means (PCK), NR-ELM, DDNet, LANTNet, and Pro + DM, our findings consistently demonstrated the superior accuracy of our approach. Notably, our method exhibited exceptional accuracy, especially when handling speckle noise-affected data. In summary, this chapter discussed the enhancements made to LANTNet's parameters and offered a comprehensive evaluation of the performance of diverse loss functions. Our thesis highlights the effectiveness of our proposed methodology in significantly improving change detection accuracy, with a particular emphasis on datasets corrupted with speckle noise.

## 6.2 Contribution to Knowledge

This thesis has made significant contributions to the field of Synthetic Aperture Radar change detection, offering novel insights and advancements in several key areas:

- ***Enhanced SAR Change Detection Techniques:***

This study has introduced and validated several techniques aimed at enhancing the accuracy of SAR change detection. By systematically exploring image registration-based feature detection algorithms, we have shed light on the critical role of accurate image alignment in improving change detection results. Additionally, our proposed despeckling model (DM) stands out as an innovative solution to address speckle noise, a persistent challenge in SAR imagery. Our research demonstrates that DM outperforms existing despeckling methods, underscoring its potential to enhance change detection accuracy significantly.

- ***Deep Learning Approaches:***

In response to the growing interest in deep learning, particularly convolutional neural networks, this research has delved into the applicability of deep learning techniques for SAR change detection. By exploring the capacity of deep learning models to handle speckle noise without the need for a custom pre-processing step, we have advanced the understanding in relation to the capabilities and limitations of these models in SAR applications. Our findings provide valuable insights for researchers and practitioners seeking to harness the power of deep learning in SAR change detection tasks.

- ***Comparative Analysis:***

This thesis has detailed extensive benchmarking and comparative analyses, systematically evaluating various change detection methods and loss functions. By benchmarking our proposed method against established methods, including PCA-k-means (PCAK), NR-ELM, DDNet, LANTNet, and Pro + DM, we have contributed valuable performance metrics and insights. Our research consistently demonstrates the superior accuracy of our method, particularly when dealing with data corrupted by speckle noise, offering a reliable benchmark for future research and applications.

- ***Disaster Assessment and Beyond:***

Beyond the realm of academia, our research has practical implications for disaster assessment and management. The capability of SAR change detection to operate during darkness and adverse weather conditions makes it a valuable tool for disaster assessment. By improving the accuracy of change detection methods, especially in scenarios impacted by speckle noise, our work contributes to more precise disaster assessment, thereby enhancing the resilience of communities in the face of natural or man-made disasters.

This thesis advances the state of knowledge in SAR change detection by providing valuable techniques, insights, and benchmarking results. The contributions made extend beyond the academic realm, offering practical benefits in disaster assessment and management and providing a solid foundation for further research in this critical field.

## 6.3 Limitations and Future Work

Encountering challenges and setbacks is an inherent part of any research journey. This section delves into the difficulties encountered during this research and their impact on the study's progression. Furthermore, this segment explores potential enhancements that have arisen from the research, offering valuable insights for future studies to consider and implement.

### Limitations

- *Limited Generalizability:*

While our proposed despeckling model (DM) has shown promise in enhancing SAR change detection, its performance may vary depending on specific dataset characteristics and applications. Further research is needed to evaluate its generalizability on a broader range of SAR datasets and scenarios.

- *Deep Learning Complexity:*

While we have explored the potential of deep learning for SAR change detection, the computational complexity of deep learning models, particularly when dealing with large SAR images, remains a challenge. Future work should focus on optimising deep learning architectures for efficient processing of SAR data.

- *Dataset Oversmoothing:*

Our investigation into despeckling revealed that excessive pre-processing can lead to oversmoothed images, resulting in a loss of essential information. Balancing the degree of despeckling with preserving critical details remains a challenge that requires further attention.

## Future Work

- ***Multi-Modal Data Fusion:***

Investigate the integration of SAR data with other remote sensing modalities, such as optical imagery or LiDAR data. The fusion of multiple data sources can provide a more comprehensive understanding of changes on the Earth's surface.

- ***Advanced Deep Learning Architectures:***

Explore more advanced deep learning architectures, including recurrent neural networks (RNNs) and transformers, to further enhance the capabilities of SAR change detection models. Experiment with state-of-the-art architectures to improve accuracy and robustness.

- ***Open Datasets and Benchmarks:***

Contribute to the development of open SAR change detection datasets and benchmarks. Standardized datasets can facilitate fair comparisons between different methods and encourage collaboration within the research community.

- ***Real-time Change Detection:***

Develop real-time SAR change detection systems that can process and analyse SAR data as it becomes available. This could have applications in disaster monitoring and response, where timely information is crucial.

- ***Automated Change Interpretation:***

Investigate methods for automatically interpreting the detected changes. This could involve categorising changes into different classes (e.g., natural disasters, urban expansion) or estimating quantitative attributes of changes (e.g., volume of deforestation).



# References

- [1] Yunhao Gao, Feng Gao, Junyu Dong, and Heng-Chao Li. Sar image change detection based on multiscale capsule network. *IEEE Geoscience and Remote Sensing Letters*, 18(3):484–488, 2020. doi: 10.1109/LGRS.2020.2977838.
- [2] Xiaofan Qu, Feng Gao, Junyu Dong, Qian Du, and Heng-Chao Li. Change detection in synthetic aperture radar images using a dual-domain network. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2021. doi: 10.1109/LGRS.2021.3073900.
- [3] Junjie Wang, Feng Gao, Junyu Dong, Shan Zhang, and Qian Du. Change detection from synthetic aperture radar images via graph-based knowledge supplement network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:1823–1836, 2022. doi: 10.1109/JSTARS.2022.3146167.
- [4] Lazhar Khelifi and Max Mignotte. Deep learning for change detection in remote sensing images: Comprehensive review and meta-analysis. *IEEE Access*, 8:126385–126400, 2020. doi: 10.1109/ACCESS.2020.3008036.
- [5] Wenzhong Shi, Min Zhang, Rui Zhang, Shanxiong Chen, and Zhao Zhan. Change detection based on artificial intelligence: State-of-the-art and challenges. *Remote Sensing*, 12(10), 2020. ISSN 2072-4292. doi: 10.3390/rs12101688. URL <https://www.mdpi.com/2072-4292/12/10/1688>.
- [6] Yanan You, Jingyi Cao, and Wenli Zhou. A survey of change detection methods based on remote sensing images for multi-source and multi-objective scenarios. *Remote*

- Sensing*, 12(15), 2020. ISSN 2072-4292. doi: 10.3390/rs12152460. URL <https://www.mdpi.com/2072-4292/12/15/2460>.
- [7] Mohamed Ihmeida and Hong Wei. Image registration techniques and applications: Comparative study on remote sensing imagery. In *2021 14th International Conference on Developments in eSystems Engineering (DeSE)*, pages 142–148, 2021. doi: 10.1109/DeSE54285.2021.9719538.
- [8] Lei Ma, Yu Liu, Xueliang Zhang, Yuanxin Ye, Gaofei Yin, and Brian Alan Johnson. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152:166–177, 2019. ISSN 0924-2716. doi: <https://doi.org/10.1016/j.isprsjprs.2019.04.015>. URL <https://www.sciencedirect.com/science/article/pii/S0924271619301108>.
- [9] Jun Zhang, Wenping Ma, Yue Wu, and Licheng Jiao. Multimodal remote sensing image registration based on image transfer and local features. *IEEE Geoscience and Remote Sensing Letters*, 16(8):1210–1214, 2019. doi: 10.1109/LGRS.2019.2896341.
- [10] Xiaolong Dai and Siamak Khorram. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote sensing*, 36(5):1566–1577, 1998. doi: 10.1109/36.718860.
- [11] Prabhishkek Singh, Manoj Diwakar, Achyut Shankar, Raj Shree, and Manoj Kumar. A review on sar image and its despeckling. *Archives of Computational Methods in Engineering*, 28:4633–4653, 2021. doi: 10.1007/s11831-021-09548-z.
- [12] Hao Chen and Zhenwei Shi. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. *Remote Sensing*, 12(10), 2020. ISSN 2072-4292. doi: 10.3390/rs12101662. URL <https://www.mdpi.com/2072-4292/12/10/1662>.
- [13] Anju Asokan and JJESI Anitha. Change detection techniques for remote sensing applications: a survey. *Earth Science Informatics*, 12(2):143–160, 2019. doi: 10.1007/s12145-019-00380-5.

- [14] Francesca Bovolo and Lorenzo Bruzzone. A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment. *IEEE Transactions on Geoscience and Remote Sensing*, 45(6):1658–1670, 2007. doi: 10.1109/TGRS.2007.895835. URL <https://doi.org/10.1109/TGRS.2007.895835>.
- [15] K. Nackaerts B. Muys P. Coppin Corresponding Author, I. Jonckheere and E. Lambin. Review article digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, 25(9):1565–1596, 2004. doi: 10.1080/0143116031000101675.
- [16] Jan Feranec, Gerard Hazeu, Susan Christensen, and Gabriel Jaffrain. Corine land cover change detection in europe (case studies of the netherlands and slovakia). *Land use policy*, 24(1):234–247, 2007. doi: 10.1016/j.landusepol.2006.02.002.
- [17] Cláudia M Viana, Sandra Oliveira, Sérgio C Oliveira, and Jorge Rocha. Land use/land cover change detection and urban sprawl analysis. In *Spatial modeling in GIS and R for earth and environmental sciences*, pages 621–651. Elsevier, 2019. doi: 10.3390/IJGI5020015. URL <https://doi.org/10.3390/ijgi5020015>.
- [18] Mohamed Ihmeida and Muhammad Shahzad. Deep despeckling of SAR images to improve change detection performance. In Max Bramer and Frederic T. Stahl, editors, *Artificial Intelligence XL - 43rd SGAI International Conference on Artificial Intelligence, AI 2023, Cambridge, UK, December 12-14, 2023, Proceedings*, volume 14381 of *Lecture Notes in Computer Science*, pages 115–126. Springer, 2023. doi: 10.1007/978-3-031-47994-6\_9. URL [https://doi.org/10.1007/978-3-031-47994-6\\_9](https://doi.org/10.1007/978-3-031-47994-6_9).
- [19] Mohamed Ihmeida and Muhammad Shahzad. Enhanced change detection performance based on deep despeckling of synthetic aperture radar images. *IEEE Access*, 11:95734–95746, 2023. doi: 10.1109/ACCESS.2023.3307208. URL <https://doi.org/10.1109/ACCESS.2023.3307208>.

- [20] Andreas Danklmayer, Björn Döring, Marco Schwerdt, and Madhu Chandra. Assessment of atmospheric propagation effects in SAR images. *IEEE Trans. Geosci. Remote. Sens.*, 47(10):3507–3518, 2009. doi: 10.1109/TGRS.2009.2022271. URL <https://doi.org/10.1109/TGRS.2009.2022271>.
- [21] Jong-Sen Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2(2):165–168, 1980. doi: 10.1109/TPAMI.1980.4766994. URL <https://doi.org/10.1109/TPAMI.1980.4766994>.
- [22] Farnaam Samadi, Gholamreza Akbarizadeh, and Hooman Kaabi. Change detection in SAR images using deep belief network: a new training approach based on morphological images. *IET Image Process.*, 13(12):2255–2264, 2019. doi: 10.1049/IET-IPR.2018.6248. URL <https://doi.org/10.1049/iet-ipr.2018.6248>.
- [23] Turgay Çelik. Unsupervised change detection in satellite images using principal component analysis and k -means clustering. *IEEE Geosci. Remote. Sens. Lett.*, 6(4):772–776, 2009. doi: 10.1109/LGRS.2009.2025059. URL <https://doi.org/10.1109/LGRS.2009.2025059>.
- [24] Aiyeola Sikiru Yommy, Rongke Liu, and Shuang Wu. Sar image despeckling using refined lee filter. In *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics*, volume 2, pages 260–265. IEEE, 2015. doi: 10.1109/IHMSC.2015.236.
- [25] Yunhao Gao, Feng Gao, Junyu Dong, Qian Du, and Heng-Chao Li. Synthetic aperture radar image change detection via siamese adaptive fusion network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:10748–10760, 2021. doi: 10.1109/JSTARS.2021.3120381.
- [26] Guillaume Quin, Beatrice Pinel-Puysegur, Jean-Marie Nicolas, and Philippe Loreaux. Mimosa: An automatic change detection method for sar time series. *IEEE Transactions on Geoscience and Remote Sensing*, 52(9):5349–5363, 2013. doi: 10.1109/TGRS.2013.2288271.

- [27] Xiangzhi Bai and Fugen Zhou. Analysis of new top-hat transformation and the application for infrared dim small target detection. *Pattern Recognition*, 43(6):2145–2156, 2010. doi: 10.1016/j.patcog.2009.12.023. URL <https://doi.org/10.1016/j.patcog.2009.12.023>.
- [28] Richard J Radke, Srinivas Andra, Omar Al-Kofahi, and Badrinath Roysam. Image change detection algorithms: a systematic survey. *IEEE transactions on image processing*, 14(3):294–307, 2005. doi: 10.1109/TIP.2004.838698.
- [29] Muwei Jian, Kin-Man Lam, Junyu Dong, and Linlin Shen. Visual-patch-attention-aware saliency detection. *IEEE Trans. Cybern.*, 45(8):1575–1586, 2015. doi: 10.1109/TCYB.2014.2356200. URL <https://doi.org/10.1109/TCYB.2014.2356200>.
- [30] Mingyang Zhang, Maoguo Gong, Haibo He, and Shengqi Zhu. Symmetric all convolutional neural-network-based unsupervised feature extraction for hyperspectral images classification. *IEEE Trans. Cybern.*, 52(5):2981–2993, 2022. doi: 10.1109/TCYB.2020.3020540. URL <https://doi.org/10.1109/TCYB.2020.3020540>.
- [31] Ashbindu Singh. Review article digital change detection techniques using remotely-sensed data. *International journal of remote sensing*, 10(6):989–1003, 1989. doi: 10.1080/01431168908903939. URL <https://doi.org/10.1080/01431168908903939>.
- [32] Sicong Liu, Lorenzo Bruzzone, Francesca Bovolo, and Peijun Du. Hierarchical unsupervised change detection in multitemporal hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1):244–260, 2014. doi: 10.1109/TGRS.2014.2321277. URL <https://doi.org/10.1109/TGRS.2014.2321277>.
- [33] Gang Yang, Heng-Chao Li, Wei-Ye Wang, Wen Yang, and William J Emery. Unsupervised change detection based on a unified framework for weighted collaborative representation with rddl and fuzzy clustering. *IEEE Transactions on Geoscience and Remote Sensing*, 57(11):8890–8903, 2019. doi: 10.1109/TGRS.2019.2923643. URL <https://doi.org/10.1109/TGRS.2019.2923643>.

- [34] Ruochen Liu, Zhihong Cheng, Langlang Zhang, and Jianxia Li. Remote sensing image change detection based on information transmission and attention mechanism. *IEEE Access*, 7:156349–156359, 2019. doi: 10.1109/ACCESS.2019.2947286. URL <https://doi.org/10.1109/ACCESS.2019.2947286>.
- [35] Caihong Mu, Chengzhou Li, Yi Liu, Menghua Sun, Licheng Jiao, and Rong Qu. Change detection in sar images based on the salient map guidance and an accelerated genetic algorithm. In *2017 IEEE Congress on Evolutionary Computation (CEC)*, pages 1150–1157. IEEE, 2017. doi: 10.1109/CEC.2017.7969436. URL <https://doi.org/10.1109/CEC.2017.7969436>.
- [36] Qian Shi, Mengxi Liu, Shengchen Li, Xiaoping Liu, Fei Wang, and Liangpei Zhang. A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection. *IEEE transactions on geoscience and remote sensing*, 60:1–16, 2021.
- [37] Xiaofeng Zhang, Shuli Cheng, Liejun Wang, and Haojin Li. Asymmetric cross-attention hierarchical network based on cnn and transformer for bitemporal remote sensing images change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–15, 2023. doi: 10.1109/TGRS.2023.3245674. URL <https://doi.org/10.1109/TGRS.2023.3245674>.
- [38] Barbara Zitová and Jan Flusser. Image registration methods: a survey. *Image Vis. Comput.*, 21(11):977–1000, 2003. doi: 10.1016/S0262-8856(03)00137-9. URL [https://doi.org/10.1016/S0262-8856\(03\)00137-9](https://doi.org/10.1016/S0262-8856(03)00137-9).
- [39] Jacqueline Le Moigne, Nathan S. Netanyahu, and Roger D. Eastman, editors. *Image Registration for Remote Sensing*. Cambridge University Press, 2011. ISBN 9780521516112. URL <http://www.cambridge.org/aus/catalogue/catalogue.asp?isbn=9780521516112>.
- [40] Kavitha Kuppala, Sandhya Banda, and Thirumala Rao Barige. An overview of deep learning methods for image registration with focus on feature-based approaches. *Inter-*

- national Journal of Image and Data Fusion*, 11(2):113–135, 2020. ISSN 19479824. doi: 10.1080/19479832.2019.1707720.
- [41] Ebrahim Hamidi, Brad G. Peter, David F. Muñoz, Hamed Moftakhari, and Hamid Moradkhani. Fast flood extent monitoring with sar change detection using google earth engine. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–19, 2023. doi: 10.1109/TGRS.2023.3240097.
- [42] Arthur A Goshtasby. *Theory And Applications of Image Registration*. John Wiley Sons, 2017.
- [43] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.*, 60(2):91–110, 2004. doi: 10.1023/B:VISI.0000029664.99615.94. URL <https://doi.org/10.1023/B:VISI.0000029664.99615.94>.
- [44] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. SURF: speeded up robust features. In Ales Leonardis, Horst Bischof, and Axel Pinz, editors, *Computer Vision - ECCV 2006, 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006, Proceedings, Part I*, volume 3951 of *Lecture Notes in Computer Science*, pages 404–417. Springer, 2006. doi: 10.1007/11744023\_32. URL [https://doi.org/10.1007/11744023\\_32](https://doi.org/10.1007/11744023_32).
- [45] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary R. Bradski. ORB: an efficient alternative to SIFT or SURF. In Dimitris N. Metaxas, Long Quan, Alberto Sanfeliu, and Luc Van Gool, editors, *IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6-13, 2011*, pages 2564–2571. IEEE Computer Society, 2011. doi: 10.1109/ICCV.2011.6126544. URL <https://doi.org/10.1109/ICCV.2011.6126544>.
- [46] Pablo Fernández Alcantarilla, Adrien Bartoli, and Andrew J. Davison. KAZE features. In Andrew W. Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid, editors, *Computer Vision - ECCV 2012 - 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VI*, volume

- 7577 of *Lecture Notes in Computer Science*, pages 214–227. Springer, 2012. doi: 10.1007/978-3-642-33783-3\_16. URL [https://doi.org/10.1007/978-3-642-33783-3\\_16](https://doi.org/10.1007/978-3-642-33783-3_16).
- [47] Pablo F. Alcantarilla, Jesús Nuevo, and Adrien Bartoli. Fast explicit diffusion for accelerated features in nonlinear scale spaces. In *BMVC 2013 - Electronic Proceedings of the British Machine Vision Conference 2013*, 2013. doi: 10.5244/C.27.13.
- [48] Stefan Leutenegger, Margarita Chli, and Roland Siegwart. BRISK: binary robust invariant scalable keypoints. In Dimitris N. Metaxas, Long Quan, Alberto Sanfeliu, and Luc Van Gool, editors, *IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6-13, 2011*, pages 2548–2555. IEEE Computer Society, 2011. doi: 10.1109/ICCV.2011.6126542. URL <https://doi.org/10.1109/ICCV.2011.6126542>.
- [49] Jiayi Ma, Xingyu Jiang, Aoxiang Fan, Junjun Jiang, and Junchi Yan. Image Matching from Handcrafted to Deep Features: A Survey. *International Journal of Computer Vision*, 129(1):23–79, 2021. ISSN 15731405. doi: 10.1007/s11263-020-01359-2.
- [50] Xiaolong Dai and Siamak Khorram. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Trans. Geosci. Remote. Sens.*, 37(5):2351–2362, 1999. doi: 10.1109/36.789634. URL <https://doi.org/10.1109/36.789634>.
- [51] Francisco PM Oliveira and Joao Manuel RS Tavares. Medical image registration: a review. *Computer methods in biomechanics and biomedical engineering*, 17(2):73–93, 2014.
- [52] Anchal Kumawat and Sucheta Panda. A robust edge detection algorithm based on feature-based image registration (FBIR) using improved canny with fuzzy logic (ICWFL). *Vis. Comput.*, 38(11):3681–3702, 2022. doi: 10.1007/S00371-021-02196-1. URL <https://doi.org/10.1007/s00371-021-02196-1>.
- [53] Ruan Lakemond, Clinton Fookes, and Sridha Sridharan. Affine adaptation of local image features using the hessian matrix. In Stefano Tubaro and Jean-Luc Dugelay,

- editors, *Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2009, 2-4 September 2009, Genova, Italy*, pages 496–501. IEEE Computer Society, 2009. doi: 10.1109/AVSS.2009.8. URL <https://doi.org/10.1109/AVSS.2009.8>.
- [54] Deepak Geetha Viswanathan. Features from Accelerated Segment Test (FAST) Deepak Geetha Viswanathan 1. 2009. URL <https://api.semanticscholar.org/CorpusID:17031649>.
- [55] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. BRIEF: Binary robust independent elementary features. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6314 LNCS(PART 4):778–792, 2010. ISSN 16113349. doi: 10.1007/978-3-642-15561-1\_56.
- [56] Elmar Mair, Gregory D Hager, Darius Burschka, Michael Suppa, and Gerhard Hirzinger. Adaptive and generic corner detection based on the accelerated segment test. In *European conference on Computer vision*, pages 183–196. Springer, 2010. doi: 10.1007/978-3-642-15552-9\_14. URL [https://doi.org/10.1007/978-3-642-15552-9\\_14](https://doi.org/10.1007/978-3-642-15552-9_14).
- [57] Jared Heinly, Enrique Dunn, and Jan-Michael Frahm. Comparative evaluation of binary features. In *European Conference on Computer Vision*, pages 759–773. Springer, 2012. doi: 10.1007/978-3-642-33709-3\_54. URL [https://doi.org/10.1007/978-3-642-33709-3\\_54](https://doi.org/10.1007/978-3-642-33709-3_54).
- [58] Martin A Fischler and Robert C Bolles. Random Sample Paradigm for Model Consensus: A Apphcatlons to Image Fitting with Analysis and Automated Cartography. *Graphics and Image Processing*, 24(6):381–395, 1981. ISSN 00010782.
- [59] Wu-Chih Hu, Liang-Bi Chen, Bo-Kai Huang, and Hong-Ming Lin. A computer vision-based intelligent fish feeding system using deep learning techniques for aquaculture. *IEEE Sensors Journal*, 22(7):7185–7194, 2022. doi: 10.1109/JSEN.2022.3151777.

- [60] Andre Esteva, Katherine Chou, Serena Yeung, Nikhil Naik, Ali Madani, Ali Mottaghi, Yun Liu, Eric Topol, Jeff Dean, and Richard Socher. Deep learning-enabled medical computer vision. *NPJ digital medicine*, 4(1):5, 2021. doi: 10.1038/S41746-020-00376-2. URL <https://doi.org/10.1038/s41746-020-00376-2>.
- [61] Young Hoon Jung, Trung Xuan Pham, Dias Issa, Hee Seung Wang, Jae Hee Lee, Mingi Chung, Bo-Yeon Lee, Gwangsu Kim, Chang D Yoo, and Keon Jae Lee. Deep learning-based noise robust flexible piezoelectric acoustic sensors for speech processing. *Nano Energy*, 101:107610, 2022. doi: <https://doi.org/10.1016/j.nanoen.2022.107610>. URL <https://www.sciencedirect.com/science/article/pii/S2211285522006887>.
- [62] Reinhold Haeb-Umbach, Shinji Watanabe, Tomohiro Nakatani, Michiel Bacchiani, Bjorn Hoffmeister, Michael L Seltzer, Heiga Zen, and Mehrez Souden. Speech processing for digital home assistants: Combining signal processing with deep-learning techniques. *IEEE Signal processing magazine*, 36(6):111–124, 2019. doi: 10.1109/MSP.2019.2918706. URL <https://doi.org/10.1109/MSP.2019.2918706>.
- [63] Andreas Maier, Christopher Syben, Tobias Lasser, and Christian Riess. A gentle introduction to deep learning in medical image processing. *Zeitschrift für Medizinische Physik*, 29(2):86–101, 2019. doi: <https://doi.org/10.1016/j.zemedi.2018.12.003>. URL <http://arxiv.org/abs/1810.05401>.
- [64] Jiawei Zhang, Chen Li, Md Mamunur Rahaman, Yudong Yao, Pingli Ma, Jinghua Zhang, Xin Zhao, Tao Jiang, and Marcin Grzegorzec. A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches. *Artificial Intelligence Review*, pages 1–70, 2022. doi: 10.1007/S10462-021-10082-4. URL <https://doi.org/10.1007/s10462-021-10082-4>.
- [65] Geoffrey E Hinton. Deep belief networks. *Scholarpedia*, 4(5):5947, 2009.
- [66] Geoffrey E. Hinton, Alex Krizhevsky, and Sida D. Wang. Transforming auto-encoders. In Timo Honkela, Wlodzislaw Duch, Mark A. Girolami, and Samuel Kaski, editors, *Artificial Neural Networks and Machine Learning - ICANN 2011 -*

- 21st International Conference on Artificial Neural Networks, Espoo, Finland, June 14-17, 2011, Proceedings, Part I*, volume 6791 of *Lecture Notes in Computer Science*, pages 44–51. Springer, 2011. doi: 10.1007/978-3-642-21735-7\_6. URL [https://doi.org/10.1007/978-3-642-21735-7\\_6](https://doi.org/10.1007/978-3-642-21735-7_6).
- [67] Zewen Li, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Trans. Neural Networks Learn. Syst.*, 33(12):6999–7019, 2022. doi: 10.1109/TNNLS.2021.3084827. URL <https://doi.org/10.1109/TNNLS.2021.3084827>.
- [68] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, Pierre-Antoine Manzagol, and Léon Bottou. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of machine learning research*, 11(12), 2010. doi: 10.5555/1756006.1953039. URL <https://dl.acm.org/doi/10.5555/1756006.1953039>.
- [69] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986. doi: 10.1007/978-3-031-42505-9\_8. URL [https://doi.org/10.1007/978-3-031-42505-9\\_8](https://doi.org/10.1007/978-3-031-42505-9_8).
- [70] Qiangqiang Yuan, Huanfeng Shen, Tongwen Li, Zhiwei Li, Shuwen Li, Yun Jiang, Hongzhang Xu, Weiwei Tan, Qianqian Yang, Jiwen Wang, et al. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sensing of Environment*, 241:111716, 2020. doi: <https://doi.org/10.1016/j.rse.2020.111716>. URL <https://www.sciencedirect.com/science/article/pii/S0034425720300857>.
- [71] Liangpei Zhang, Lefei Zhang, and Bo Du. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and remote sensing magazine*, 4(2):22–40, 2016. doi: 10.1109/MGRS.2016.2540798.
- [72] Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui-Song Xia, Liangpei Zhang, Feng Xu, and Friedrich Fraundorfer. Deep learning in remote sensing: A comprehensive review

- and list of resources. *IEEE geoscience and remote sensing magazine*, 5(4):8–36, 2017. doi: 10.1109/MGRS.2017.2762307.
- [73] Wei Liu, Rongrong Ji, and Shaozi Li. Towards 3d object detection with bimodal deep boltzmann machines over RGBD imagery. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, pages 3013–3021. IEEE Computer Society, 2015. doi: 10.1109/CVPR.2015.7298920. URL <https://doi.org/10.1109/CVPR.2015.7298920>.
- [74] Gong Cheng, Peicheng Zhou, Xiwen Yao, Chao Yao, Yanbang Zhang, and Junwei Han. Object detection in vhr optical remote sensing images via learning rotation-invariant hog feature. In *2016 4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA)*, pages 433–436. IEEE, 2016. doi: 10.1109/EORSA.2016.7552845.
- [75] Gong Cheng, Peicheng Zhou, and Junwei Han. RIFD-CNN: rotation-invariant and fisher discriminative convolutional neural networks for object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 2884–2893. IEEE Computer Society, 2016. doi: 10.1109/CVPR.2016.315. URL <https://doi.org/10.1109/CVPR.2016.315>.
- [76] Grant J. Scott, Matthew R. England, William A. Starms, Richard A. Marcum, and Curt H. Davis. Training deep convolutional neural networks for land-cover classification of high-resolution imagery. *IEEE Geosci. Remote. Sens. Lett.*, 14(4):549–553, 2017. doi: 10.1109/LGRS.2017.2657778. URL <https://doi.org/10.1109/LGRS.2017.2657778>.
- [77] Wenzhi Zhao and Shihong Du. Spectral-spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach. *IEEE Trans. Geosci. Remote. Sens.*, 54(8):4544–4554, 2016. doi: 10.1109/TGRS.2016.2543748. URL <https://doi.org/10.1109/TGRS.2016.2543748>.

- [78] Liangzhi Li, Ling Han, Mingtao Ding, Hongye Cao, and Huijuan Hu. A deep learning semantic template matching framework for remote sensing image registration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 181:205–217, 2021. doi: <https://doi.org/10.1016/j.isprsjprs.2021.09.012>. URL <https://www.sciencedirect.com/science/article/pii/S0924271621002446>.
- [79] Yue Wu, Jun-Wei Liu, Chenzhuo Zhu, Zhuangfei Bai, Qiguang Miao, Wenping Ma, and Maoguo Gong. Computational intelligence in remote sensing image registration: A survey. *Int. J. Autom. Comput.*, 18(1):1–17, 2021. doi: [10.1007/S11633-020-1248-X](https://doi.org/10.1007/S11633-020-1248-X). URL <https://doi.org/10.1007/s11633-020-1248-x>.
- [80] Guorong Wu, Minjeong Kim, Qian Wang, Brent C. Munsell, and Dinggang Shen. Correction to "scalable high-performance image registration framework by unsupervised deep feature representations learning". *IEEE Trans. Biomed. Eng.*, 64(1):250, 2017. doi: [10.1109/TBME.2016.2633139](https://doi.org/10.1109/TBME.2016.2633139). URL <https://doi.org/10.1109/TBME.2016.2633139>.
- [81] Xufeng Han, Thomas Leung, Yangqing Jia, Rahul Sukthankar, and Alexander C Berg. Matchnet: Unifying feature and metric learning for patch-based matching. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3279–3286, 2015. doi: [10.1109/CVPR.2015.7298948](https://doi.org/10.1109/CVPR.2015.7298948).
- [82] Sergey Zagoruyko and Nikos Komodakis. Learning to compare image patches via convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4353–4361, 2015. URL <https://api.semanticscholar.org/CorpusID:215827033>.
- [83] Grant Haskins, Uwe Kruger, and Pingkun Yan. Deep learning in medical image registration: a survey. *Mach. Vis. Appl.*, 31(1-2):8, 2020. doi: [10.1007/S00138-020-01060-X](https://doi.org/10.1007/S00138-020-01060-X). URL <https://doi.org/10.1007/s00138-020-01060-x>.

- [84] Shun Miao, Z. Jane Wang, and Rui Liao. A CNN regression approach for real-time 2d/3d registration. *IEEE Trans. Medical Imaging*, 35(5):1352–1363, 2016. doi: 10.1109/TMI.2016.2521800. URL <https://doi.org/10.1109/TMI.2016.2521800>.
- [85] Shun Miao, Z Jane Wang, Yefeng Zheng, and Rui Liao. Real-time 2d/3d registration via cnn regression. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 1430–1434. IEEE, 2016. doi: 10.1109/ISBI.2016.7493536. URL <https://doi.org/10.1109/ISBI.2016.7493536>.
- [86] Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. Deep image homography estimation. *CoRR*, abs/1606.03798, 2016. URL <http://arxiv.org/abs/1606.03798>.
- [87] Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. Superpoint: Self-supervised interest point detection and description. In *2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2018, Salt Lake City, UT, USA, June 18-22, 2018*, pages 224–236. Computer Vision Foundation / IEEE Computer Society, 2018. doi: 10.1109/CVPRW.2018.00060. URL [http://openaccess.thecvf.com/content\\_cvpr\\_2018\\_workshops/w9/html/DeTone\\_SuperPoint\\_Self-Supervised\\_Interest\\_CVPR\\_2018\\_paper.html](http://openaccess.thecvf.com/content_cvpr_2018_workshops/w9/html/DeTone_SuperPoint_Self-Supervised_Interest_CVPR_2018_paper.html).
- [88] Max Jaderberg, Karen Simonyan, Andrew Zisserman, and Koray Kavukcuoglu. Spatial transformer networks. pages 2017–2025, 2015. URL <https://proceedings.neurips.cc/paper/2015/hash/33ceb07bf4eeb3da587e268d663aba1a-Abstract.html>.
- [89] Guha Balakrishnan, Amy Zhao, Mert R. Sabuncu, John V. Guttag, and Adrian V. Dalca. Voxelmorph: A learning framework for deformable medical image registration. *IEEE Trans. Medical Imaging*, 38(8):1788–1800, 2019. doi: 10.1109/TMI.2019.2897538. URL <https://doi.org/10.1109/TMI.2019.2897538>.
- [90] Adrian V. Dalca, Guha Balakrishnan, John V. Guttag, and Mert R. Sabuncu. Unsupervised learning for fast probabilistic diffeomorphic registration. In Alejandro F. Frangi, Julia A. Schnabel, Christos Davatzikos, Carlos Alberola-López, and Gabor Fichtinger, editors, *Medical Image Computing and Computer Assisted Intervention*

- *MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I*, volume 11070 of *Lecture Notes in Computer Science*, pages 729–738. Springer, 2018. doi: 10.1007/978-3-030-00928-1\_82. URL [https://doi.org/10.1007/978-3-030-00928-1\\_82](https://doi.org/10.1007/978-3-030-00928-1_82).
- [91] Bob D. de Vos, Floris F. Berendsen, Max A. Viergever, Hessam Sokooti, Marius Staring, and Ivana Isgum. A deep learning framework for unsupervised affine and deformable image registration. *Medical Image Anal.*, 52:128–143, 2019. doi: 10.1016/J.MEDIA.2018.11.010. URL <https://doi.org/10.1016/j.media.2018.11.010>.
- [92] Hamid Reza Boveiri, Raouf Khayami, Reza Javidan, and Ali Reza MehdiZadeh. Medical image registration using deep neural networks: A comprehensive review. *Comput. Electr. Eng.*, 87:106767, 2020. doi: 10.1016/J.COMPELECENG.2020.106767. URL <https://doi.org/10.1016/j.compeleceng.2020.106767>.
- [93] Yabo Fu, Yang Lei, Tonghe Wang, Walter J Curran, Tian Liu, and Xiaofeng Yang. Deep learning in medical image registration: a review. *Physics in Medicine & Biology*, 65(20):20TR01, 2020.
- [94] Liya Zhao and Kebin Jia. Deep adaptive log-demons: Diffeomorphic image registration with very large deformations. *Comput. Math. Methods Medicine*, 2015: 836202:1–836202:16, 2015. doi: 10.1155/2015/836202. URL <https://doi.org/10.1155/2015/836202>.
- [95] Max Blendowski and Mattias P. Heinrich. Combining mrf-based deformable registration and deep binary 3d-cnn descriptors for large lung motion estimation in COPD patients. *Int. J. Comput. Assist. Radiol. Surg.*, 14(1):43–52, 2019. doi: 10.1007/S11548-018-1888-2. URL <https://doi.org/10.1007/s11548-018-1888-2>.
- [96] Hessam Sokooti, Bob D. de Vos, Floris F. Berendsen, Boudewijn P. F. Lelieveldt, Ivana Isgum, and Marius Staring. Nonrigid image registration using multi-scale 3d convolutional neural networks. In Maxime Descoteaux, Lena Maier-Hein, Alfred M. Franz, Pierre Jannin, D. Louis Collins, and Simon Duchesne, editors,

- Medical Image Computing and Computer Assisted Intervention - MICCAI 2017 - 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I*, volume 10433 of *Lecture Notes in Computer Science*, pages 232–239. Springer, 2017. doi: 10.1007/978-3-319-66182-7\_27. URL [https://doi.org/10.1007/978-3-319-66182-7\\_27](https://doi.org/10.1007/978-3-319-66182-7_27).
- [97] Padmavathi Kora, Chui Ping Ooi, Oliver Faust, U Raghavendra, Anjan Gudigar, Wai Yee Chan, K Meenakshi, K Swaraja, Pawel Plawiak, and U Rajendra Acharya. Transfer learning techniques for medical image analysis: A review. *Biocybernetics and Biomedical Engineering*, 42(1):79–107, 2022.
- [98] Michael Elad and Michal Aharon. Image denoising via learned dictionaries and sparse representation. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 1, pages 895–900, 2006. doi: 10.1109/CVPR.2006.142.
- [99] Ruomei Yan, Ling Shao, and Yan Liu. Nonlocal hierarchical dictionary learning using wavelets for image denoising. *IEEE Trans. Image Process.*, 22(12):4689–4698, 2013. doi: 10.1109/TIP.2013.2277813. URL <https://doi.org/10.1109/TIP.2013.2277813>.
- [100] Binge Cui, Xiudan Ma, Xiaoyun Xie, Guangbo Ren, and Yi Ma. Classification of visible and infrared hyperspectral images based on image segmentation and edge-preserving filtering. *Infrared Physics & Technology*, 81:79–88, 2017. doi: <https://doi.org/10.1016/j.infrared.2016.12.010>. URL <https://www.sciencedirect.com/science/article/pii/S1350449516305746>.
- [101] Jong-Sen Lee. Speckle analysis and smoothing of synthetic aperture radar images. *Computer graphics and image processing*, 17(1):24–32, 1981. doi: [https://doi.org/10.1016/S0146-664X\(81\)80005-6](https://doi.org/10.1016/S0146-664X(81)80005-6). URL <https://www.sciencedirect.com/science/article/pii/S0146664X81800056>.
- [102] Jong-Sen Lee. Refined filtering of image noise using local statistics. *Computer graphics and image processing*, 15(4):380–389, 1981. doi: <https://doi.org/10.1016/>

- S0146-664X(81)80018-4. URL <https://www.sciencedirect.com/science/article/pii/S0146664X81800184>.
- [103] Darwin T Kuan, Alexander A Sawchuk, Timothy C Strand, and Pierre Chavel. Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE transactions on pattern analysis and machine intelligence*, (2):165–177, 1985. doi: 10.1109/TPAMI.1985.4767641.
- [104] A. Lopes, R. Touzi, and E. Nezry. Adaptive speckle filters and scene heterogeneity. *IEEE Transactions on Geoscience and Remote Sensing*, 28(6):992–1000, 1990. doi: 10.1109/36.62623.
- [105] Junling Zhu, Jianguo Wen, and Yafeng Zhang. A new algorithm for sar image despeckling using an enhanced lee filter and median filter. In *2013 6th International congress on image and signal processing (CISP)*, volume 1, pages 224–228. IEEE, 2013. doi: 10.1109/CISP.2013.6743991.
- [106] Khilan Ravani, Shivam Saboo, and Jignesh S. Bhatt. A practical approach for SAR image despeckling using deep learning. In *2019 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2019, Yokohama, Japan, July 28 - August 2, 2019*, pages 2957–2960. IEEE, 2019. doi: 10.1109/IGARSS.2019.8897918. URL <https://doi.org/10.1109/IGARSS.2019.8897918>.
- [107] Feng Gu, Hong Zhang, Chao Wang, and Bo Zhang. Residual encoder-decoder network introduced for multisource sar image despeckling. In *2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA)*, pages 1–5. IEEE, 2017. doi: 10.1109/BIGSAR DATA.2017.8124932.
- [108] Loïc Denis, Charles-Alban Deledalle, and Florence Tupin. From patches to deep learning: Combining self-similarity and neural networks for sar image despeckling. In *2019 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2019, Yokohama, Japan, July 28 - August 2, 2019*, pages 5113–5116. IEEE, 2019. doi: 10.1109/IGARSS.2019.8898473. URL <https://doi.org/10.1109/IGARSS.2019.8898473>.

- [109] Sergio Vitale, Giampaolo Ferraioli, and Vito Pascazio. A new ratio image based CNN algorithm for SAR despeckling. In *2019 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2019, Yokohama, Japan, July 28 - August 2, 2019*, pages 9494–9497. IEEE, 2019. doi: 10.1109/IGARSS.2019.8899245. URL <https://doi.org/10.1109/IGARSS.2019.8899245>.
- [110] Dusan Gleich and Danijel Sipos. Deep despeckling of SAR images. In *2019 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2019, Yokohama, Japan, July 28 - August 2, 2019*, pages 1907–1910. IEEE, 2019. doi: 10.1109/IGARSS.2019.8899338. URL <https://doi.org/10.1109/IGARSS.2019.8899338>.
- [111] Davide Cozzolino, Luisa Verdoliva, Giuseppe Scarpa, and Giovanni Poggi. Nonlocal sar image despeckling by convolutional neural networks. In *2019 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2019, Yokohama, Japan, July 28 - August 2, 2019*, pages 5117–5120. IEEE, 2019. doi: 10.1109/IGARSS.2019.8897761. URL <https://doi.org/10.1109/IGARSS.2019.8897761>.
- [112] H Mahmoudzadeh. Digital change detection using remotely sensed data for monitoring green space destruction in tabriz. 2007.
- [113] Rongjun Qin, Jiaojiao Tian, and Peter Reinartz. 3d change detection—approaches and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122:41–56, 2016. doi: <https://doi.org/10.1016/j.isprsjprs.2016.09.013>. URL <https://www.sciencedirect.com/science/article/pii/S0924271616304026>.
- [114] NA Quarmby and JL Cushnie. Monitoring urban land cover changes at the urban fringe from spot hrv imagery in south-east england. *International Journal of Remote Sensing*, 10(6):953–963, 1989. doi: 10.1080/01431168908903937. URL <https://doi.org/10.1080/01431168908903937>.
- [115] Jinxia Zhu, Yanjun Su, Qinghua Guo, and Thomas C Harmon. Unsupervised object-based differencing for land-cover change detection. *Photogrammetric Engineering & Remote Sensing*, 83(3):225–236, 2017. doi: 10.14358/PERS.83.3.225.

- [116] Philip J Howarth and Gregory M Wickware. Procedures for change detection using landsat digital data. *International Journal of Remote Sensing*, 2(3):277–291, 1981. doi: 10.1080/01431168108948362. URL <https://doi.org/10.1080/01431168108948362>.
- [117] Sicong Liu, Lorenzo Bruzzone, Francesca Bovolo, Massimo Zanetti, and Peijun Du. Sequential spectral change vector analysis for iteratively discovering and detecting multiple changes in hyperspectral images. *IEEE transactions on geoscience and remote sensing*, 53(8):4363–4378, 2015.
- [118] JA Richards. Thematic mapping from multitemporal image data using the principal components transformation. *Remote Sensing of Environment*, 16(1):35–46, 1984. doi: [https://doi.org/10.1016/0034-4257\(84\)90025-7](https://doi.org/10.1016/0034-4257(84)90025-7). URL <https://www.sciencedirect.com/science/article/pii/0034425784900257>.
- [119] JIN Suming. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sensing of Environment*, 91(1):1–13, 2004. doi: <https://doi.org/10.1016/j.rse.2004.10.012>. URL <https://www.sciencedirect.com/science/article/pii/S0034425704003414>.
- [120] Ashish Ghosh, Niladri Shekhar Mishra, and Susmita Ghosh. Fuzzy clustering algorithms for unsupervised change detection in remote sensing images. *Inf. Sci.*, 181(4): 699–715, 2011. doi: 10.1016/J.INS.2010.10.016. URL <https://doi.org/10.1016/j.ins.2010.10.016>.
- [121] Jung-ho Im and John R Jensen. A change detection model based on neighborhood correlation image analysis and decision tree classification. *Remote Sensing of Environment*, 99(3):326–340, 2005. doi: <https://doi.org/10.1016/j.rse.2005.09.008>. URL <https://www.sciencedirect.com/science/article/pii/S0034425705002919>.
- [122] Junfu Liu, Keming Chen, Guangluan Xu, Xian Sun, Menglong Yan, Wenhui Diao, and Hongzhe Han. Convolutional neural network-based transfer learning for optical aerial images change detection. *IEEE Geosci. Remote. Sens. Lett.*, 17(1):127–131, 2020. doi: 10.1109/LGRS.2019.2916601. URL <https://doi.org/10.1109/LGRS.2019.2916601>.

- [123] Bin Cui, Yonghong Zhang, Li Yan, Jujie Wei, and Hongan Wu. An unsupervised SAR change detection method based on stochastic subspace ensemble learning. *Remote Sens.*, 11(11):1314, 2019. doi: 10.3390/RS11111314. URL <https://doi.org/10.3390/rs11111314>.
- [124] Rodrigo Caye Daudt, Bertr Le Saux, and Alexandre Boulch. Fully convolutional siamese networks for change detection. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 4063–4067. IEEE, 2018.
- [125] Daifeng Peng, Yongjun Zhang, and Haiyan Guan. End-to-end change detection for high resolution satellite images using improved unet++. *Remote Sens.*, 11(11):1382, 2019. doi: 10.3390/RS11111382. URL <https://doi.org/10.3390/rs11111382>.
- [126] Chenxiao Zhang, Peng Yue, Deodato Tapete, Liangcun Jiang, Boyi Shangguan, Li Huang, and Guangchao Liu. 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, 2020. doi: <https://doi.org/10.1016/j.isprsjprs.2020.06.003>. URL <https://www.sciencedirect.com/science/article/pii/S0924271620301532>.
- [127] Lin Zhang, Xiangyun Hu, Mi Zhang, Zhen Shu, and Hao Zhou. Object-level change detection with a dual correlation attention-guided detector. *ISPRS Journal of Photogrammetry and Remote Sensing*, 177:147–160, 2021. doi: <https://doi.org/10.1016/j.isprsjprs.2021.05.002>. URL <https://www.sciencedirect.com/science/article/pii/S0924271621001271>.
- [128] Xuan Hou, Yunpeng Bai, Ying Li, Changjing Shang, and Qiang Shen. High-resolution triplet network with dynamic multiscale feature for change detection on satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 177:103–115, 2021.
- [129] Shabnam Jabari and Yun Zhang. 6 urban change detection utilizing high-resolution optical images taken from different viewing angles and different platforms. *Urban Remote Sensing*, 2:125–148, 2018.

- [130] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.
- [131] Pablo F. Alcantarilla, Simon Stent, Germán Ros, Roberto Arroyo, and Riccardo Gherardi. Street-view change detection with deconvolutional networks. *Auton. Robots*, 42(7):1301–1322, 2018. doi: 10.1007/S10514-018-9734-5. URL <https://doi.org/10.1007/s10514-018-9734-5>.
- [132] Sheng Fang, Kaiyu Li, Jinyuan Shao, and Zhe Li. Snunet-cd: A densely connected siamese network for change detection of vhr images. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2021.
- [133] Yang Zhan, Kun Fu, Menglong Yan, Xian Sun, Hongqi Wang, and Xiaosong Qiu. Change detection based on deep siamese convolutional network for optical aerial images. *IEEE Geosci. Remote. Sens. Lett.*, 14(10):1845–1849, 2017. doi: 10.1109/LGRS.2017.2738149. URL <https://doi.org/10.1109/LGRS.2017.2738149>.
- [134] Haobo Lyu, Hui Lu, and Lichao Mou. Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote. Sens.*, 8(6):506, 2016. doi: 10.3390/RS8060506. URL <https://doi.org/10.3390/rs8060506>.
- [135] Jie Geng, Hongyu Wang, Jianchao Fan, and Xiaorui Ma. Change detection of sar images based on supervised contractive autoencoders and fuzzy clustering. In *2017 International workshop on remote sensing with intelligent processing (RSIP)*, pages 1–3. IEEE, 2017. doi: 10.1109/RSIP.2017.7958819.
- [136] Puzhao Zhang, Maoguo Gong, Linzhi Su, Jia Liu, and Zhizhou Li. Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116:24–41, 2016.
- [137] Yuan Xu, Shiming Xiang, Chunlei Huo, and Chunhong Pan. Change detection based on auto-encoder model for vhr images. In *MIPPR 2013: Pattern Recognition and*

- Computer Vision*, volume 8919, page 891902. SPIE, 2013. doi: 10.1117/12.2031104. URL <https://doi.org/10.1117/12.2031104>.
- [138] Sudipan Saha, Francesca Bovolo, and Lorenzo Bruzzone. Unsupervised multiple-change detection in vhr multisensor images via deep-learning based adaptation. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 5033–5036. IEEE, 2019.
- [139] G Meera Gandhi, BS Parthiban, Nagaraj Thummalu, and A Christy. Ndvi: Vegetation change detection using remote sensing and gis—a case study of vellore district. *Procedia computer science*, 57:1199–1210, 2015. doi: <https://doi.org/10.1016/j.procs.2015.07.415>. URL <https://www.sciencedirect.com/science/article/pii/S1877050915019444>.
- [140] Lu Xu, Weipeng Jing, Houbing Song, and Guangsheng Chen. High-resolution remote sensing image change detection combined with pixel-level and object-level. *IEEE Access*, 7:78909–78918, 2019. doi: 10.1109/ACCESS.2019.2922839. URL <https://doi.org/10.1109/ACCESS.2019.2922839>.
- [141] Yakoub Bazi, Lorenzo Bruzzone, and Farid Melgani. An unsupervised approach based on the generalized gaussian model to automatic change detection in multitemporal sar images. *IEEE Transactions on Geoscience and Remote Sensing*, 43(4):874–887, 2005. doi: 10.1109/TGRS.2004.842441.
- [142] Feng Gao, Xiao Wang, Yunhao Gao, Junyu Dong, and Shengke Wang. Sea ice change detection in SAR images based on convolutional-wavelet neural networks. *IEEE Geosci. Remote. Sens. Lett.*, 16(8):1240–1244, 2019. doi: 10.1109/LGRS.2019.2895656. URL <https://doi.org/10.1109/LGRS.2019.2895656>.
- [143] Heng-Chao Li, Turgay Çelik, Nathan Longbotham, and William J. Emery. Gabor feature based unsupervised change detection of multitemporal SAR images based on two-level clustering. *IEEE Geosci. Remote. Sens. Lett.*, 12(12):2458–2462, 2015. doi: 10.1109/LGRS.2015.2484220. URL <https://doi.org/10.1109/LGRS.2015.2484220>.

- [144] Stelios Krinidis and Vassilios Chatzis. A robust fuzzy local information c-means clustering algorithm. *IEEE Trans. Image Process.*, 19(5):1328–1337, 2010. doi: 10.1109/TIP.2010.2040763. URL <https://doi.org/10.1109/TIP.2010.2040763>.
- [145] Lorenzo Bruzzone and Diego Fernández-Prieto. An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images. *IEEE Trans. Image Process.*, 11(4):452–466, 2002. doi: 10.1109/TIP.2002.999678. URL <https://doi.org/10.1109/TIP.2002.999678>.
- [146] Huihui Dong, Wenping Ma, Yue Wu, Maoguo Gong, and Licheng Jiao. Local descriptor learning for change detection in synthetic aperture radar images via convolutional neural networks. *IEEE Access*, 7:15389–15403, 2019. doi: 10.1109/ACCESS.2018.2889326.
- [147] Ning Lv, Chen Chen, Tie Qiu, and Arun Kumar Sangaiah. Deep learning and super-pixel feature extraction based on contractive autoencoder for change detection in sar images. *IEEE Transactions on Industrial Informatics*, 14(12):5530–5538, 2018. doi: 10.1109/TII.2018.2873492.
- [148] Sudipan Saha, Muhammad Shahzad, Patrick Ebel, and Xiao Xiang Zhu. Supervised change detection using prechange optical-sar and postchange sar data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:8170–8178, 2022. doi: 10.1109/JSTARS.2022.3206898.
- [149] Robert E Kennedy, Philip A Townsend, John E Gross, Warren B Cohen, Paul Bolstad, YQ Wang, and Phyllis Adams. Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects. *Remote sensing of environment*, 113(7):1382–1396, 2009. doi: <https://doi.org/10.1016/j.rse.2008.07.018>. URL <https://www.sciencedirect.com/science/article/pii/S0034425709000601>.
- [150] Fan Hao, Zongfang Ma, Hong-peng Tian, Hao Wang, and Di Wu. Semi-supervised label propagation for multi-source remote sensing image change detection. *Comput.*

- Geosci.*, 170:105249, 2023. doi: 10.1016/J.CAGEO.2022.105249. URL <https://doi.org/10.1016/j.cageo.2022.105249>.
- [151] Fugui Wang and Y Jun Xu. Comparison of remote sensing change detection techniques for assessing hurricane damage to forests. *Environmental monitoring and assessment*, 162:311–326, 2010. doi: 10.1007/s10661-009-0798-8. URL <https://doi.org/10.1007/s10661-009-0798-8>.
- [152] Yangyang Li, Cheng Peng, Yanqiao Chen, Licheng Jiao, Linhao Zhou, and Ronghua Shang. A deep learning method for change detection in synthetic aperture radar images. *IEEE Transactions on Geoscience and Remote Sensing*, 57(8):5751–5763, 2019.
- [153] Dinghua Xue, Tao Lei, Xiaohong Jia, Xingwu Wang, Tao Chen, and Asoke K. Nandi. Unsupervised change detection using multiscale and multiresolution gaussian-mixture-model guided by saliency enhancement. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.*, 14:1796–1809, 2021. doi: 10.1109/JSTARS.2020.3046838. URL <https://doi.org/10.1109/JSTARS.2020.3046838>.
- [154] Lorenzo Bruzzone and Diego Fernández-Prieto. An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images. *IEEE Trans. Image Process.*, 11(4):452–466, 2002. doi: 10.1109/TIP.2002.999678. URL <https://doi.org/10.1109/TIP.2002.999678>.
- [155] RJ Dekker. Speckle filtering in satellite sar change detection imagery. *International Journal of Remote Sensing*, 19(6):1133–1146, 1998. doi: 10.1080/014311698215649. URL <https://doi.org/10.1080/014311698215649>.
- [156] Milad Majidi, Salman Ahmadi, and Reza Shah-Hosseini. A saliency-guided neighbourhood ratio model for automatic change detection of sar images. *International Journal of Remote Sensing*, 41(24):9606–9627, 2020. doi: 10.1080/01431161.2020.1826066. URL <https://doi.org/10.1080/01431161.2020.1826066>.

- [157] Biao Hou, Qian Wei, Yaoguo Zheng, and Shuang Wang. Unsupervised change detection in sar image based on gauss-log ratio image fusion and compressed projection. *IEEE journal of selected topics in applied earth observations and remote sensing*, 7(8):3297–3317, 2014. doi: 10.1109/JSTARS.2014.2328344.
- [158] Wenhua Zhang, Licheng Jiao, Fang Liu, Shuyuan Yang, and Jia Liu. Adaptive contourlet fusion clustering for sar image change detection. *IEEE Transactions on Image Processing*, 31:2295–2308, 2022. doi: 10.1109/TIP.2022.3154922.
- [159] Yunhao Gao, Feng Gao, Junyu Dong, Qian Du, and Heng-Chao Li. Synthetic aperture radar image change detection via siamese adaptive fusion network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:10748–10760, 2021. doi: 10.1109/JSTARS.2021.3120381.
- [160] Feng Gao, Junyu Dong, Bo Li, and Qizhi Xu. Automatic change detection in synthetic aperture radar images based on pcanet. *IEEE geoscience and remote sensing letters*, 13(12):1792–1796, 2016. doi: 10.1109/LGRS.2016.2611001.
- [161] Feng Gao, Junyu Dong, Bo Li, Qizhi Xu, and Cui Xie. Change detection from synthetic aperture radar images based on neighborhood-based ratio and extreme learning machine. *Journal of Applied Remote Sensing*, 10(4):046019, 2016. doi: 10.1117/1.JRS.10.046019.
- [162] Shaona Wang, Shuyuan Yang, and Licheng Jiao. Saliency-guided change detection for sar imagery using a semi-supervised laplacian svm. *Remote sensing letters*, 7(11):1043–1052, 2016.
- [163] Liangliang Li, Hongbing Ma, and Zhenhong Jia. Gamma correction-based automatic unsupervised change detection in sar images via flicm model. *Journal of the Indian Society of Remote Sensing*, pages 1–12, 2023. doi: 10.1007/s12524-023-01674-4. URL <https://doi.org/10.1007/s12524-023-01674-4>.

- [164] Desen Meng, Feng Gao, Junyu Dong, Qian Du, and Heng-Chao Li. Synthetic aperture radar image change detection via layer attention-based noise-tolerant network. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2022.
- [165] Yun Lin, Shutao Li, Leyuan Fang, and Pedram Ghamisi. Multispectral change detection with bilinear convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 17(10):1757–1761, 2019. doi: 10.1109/LGRS.2019.2953754.
- [166] Ling Wan, Yuming Xiang, and Hongjian You. A post-classification comparison method for sar and optical images change detection. *IEEE Geoscience and Remote Sensing Letters*, 16(7):1026–1030, 2019. doi: 10.1109/LGRS.2019.2892432.
- [167] Michele Volpi, Gustau Camps-Valls, and Devis Tuia. Spectral alignment of multi-temporal cross-sensor images with automated kernel canonical correlation analysis. *ISPRS journal of photogrammetry and remote sensing*, 107:50–63, 2015. doi: <https://doi.org/10.1016/j.isprsjprs.2015.02.005>.
- [168] Zhunga Liu, Gang Li, Gregoire Mercier, You He, and Quan Pan. Change detection in heterogenous remote sensing images via homogeneous pixel transformation. *IEEE Transactions on Image Processing*, 27(4):1822–1834, 2017. doi: 10.1109/TIP.2017.2784560.
- [169] Wei Zhao, Zhirui Wang, Maoguo Gong, and Jia Liu. Discriminative feature learning for unsupervised change detection in heterogeneous images based on a coupled neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12):7066–7080, 2017. doi: 10.1109/TGRS.2017.2739800.
- [170] Moyang Wang, Kun Tan, Xiuping Jia, Xue Wang, and Yu Chen. A deep siamese network with hybrid convolutional feature extraction module for change detection based on multi-sensor remote sensing images. *Remote Sensing*, 12(2):205, 2020. doi: 10.3390/rs12020205.
- [171] Patrick Ebel, Sudipan Saha, and Xiao Xiang Zhu. Fusing multi-modal data for supervised change detection. *The international archives of the photogrammetry*,

- remote sensing and spatial information sciences*, 43:243–249, 2021. doi: 10.5194/isprs-archives-XLIII-B3-2021-243-2021.
- [172] Michael E Zelinski, John R Henderson, and Elizabeth L Held. Image registration and change detection for artifact detection in remote sensing imagery. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIV*, volume 10644, pages 310–327. SPIE, 2018. doi: 10.1117/12.2303934. URL <https://doi.org/10.1117/12.2303934>.
- [173] Rongbo Fan, Bochuan Hou, Jinbao Liu, Jianhua Yang, and Zenglin Hong. Registration of multiresolution remote sensing images based on l2-siamese model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:237–248, 2020. doi: 10.1109/JSTARS.2020.3038922.
- [174] Jamileh Yousefi. Image binarization using otsu thresholding algorithm. *Ontario, Canada: University of Guelph*, 10, 2011. doi: 10.13140/RG.2.1.4758.9284.
- [175] SILA BAŞ. Point-based matching of oblique images acquired from airplane and uav platforms. Master’s thesis, Hacettepe University, Available: [openaccess.hacettepe.edu.tr](https://openaccess.hacettepe.edu.tr), 2020.
- [176] Tianfeng Chai and Roland R Draxler. Root mean square error (rmse) or mean absolute error (mae)?—arguments against avoiding rmse in the literature. *Geoscientific model development*, 7(3):1247–1250, 2014. doi: 10.5194/gmd-7-1247-2014.
- [177] ESA. Suez canal traffic jam seen from space, 2021. Available: [esa.int](https://esa.int), Accessed on: 30 Mar 2021.
- [178] Mohammadreza Pourfard, Tahmineh Hosseinian, Roghayeh Saeidi, Sayed Ahmad Motamedi, Mohammad Javad Abdollahifard, Reza Mansoori, and Reza Safabakhsh. Kaze-sar: Sar image registration using kaze detector and modified surf descriptor for tackling speckle noise. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–12, 2021. doi: 10.1109/TGRS.2021.3084411.

- [179] Maoguo Gong, Shengmeng Zhao, Licheng Jiao, Dayong Tian, and Shuang Wang. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE Trans. Geosci. Remote. Sens.*, 52(7):4328–4338, 2014. doi: 10.1109/TGRS.2013.2281391. URL <https://doi.org/10.1109/TGRS.2013.2281391>.
- [180] Benny Kupfer, Nathan S. Netanyahu, and Ilan Shimshoni. An efficient sift-based mode-seeking algorithm for sub-pixel registration of remotely sensed images. *IEEE Geosci. Remote. Sens. Lett.*, 12(2):379–383, 2015. doi: 10.1109/LGRS.2014.2343471. URL <https://doi.org/10.1109/LGRS.2014.2343471>.
- [181] Wenping Ma, Zelian Wen, Yue Wu, Licheng Jiao, Maoguo Gong, Yafei Zheng, and Liang Liu. Remote sensing image registration with modified SIFT and enhanced feature matching. *IEEE Geosci. Remote. Sens. Lett.*, 14(1):3–7, 2017. doi: 10.1109/LGRS.2016.2600858. URL <https://doi.org/10.1109/LGRS.2016.2600858>.
- [182] WG Zhang, F Liu, and LC Jiao. Sar image despeckling via bilateral filtering. *Electronics letters*, 45(15):781–783, 2009. doi: 10.1049/el.2009.1591.
- [183] Mario Mastriani. New wavelet-based superresolution algorithm for speckle reduction in SAR images. *CoRR*, abs/1608.00270, 2016. URL <http://arxiv.org/abs/1608.00270>.
- [184] Maryam Amirmazlaghani and Hamidreza Amindavar. A novel sparse method for despeckling SAR images. *IEEE Trans. Geosci. Remote. Sens.*, 50(12):5024–5032, 2012. doi: 10.1109/TGRS.2012.2195321. URL <https://doi.org/10.1109/TGRS.2012.2195321>.
- [185] Puyang Wang, He Zhang, and Vishal M Patel. Sar image despeckling using a convolutional neural network. *IEEE Signal Processing Letters*, 24(12):1763–1767, 2017.
- [186] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In David J. Fleet, Tomás Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV*, volume

- 8692 of *Lecture Notes in Computer Science*, pages 184–199. Springer, 2014. doi: 10.1007/978-3-319-10593-2\_13. URL [https://doi.org/10.1007/978-3-319-10593-2\\_13](https://doi.org/10.1007/978-3-319-10593-2_13).
- [187] Yi Liu, Chao Pang, Zongqian Zhan, Xiaomeng Zhang, and Xue Yang. Building change detection for remote sensing images using a dual-task constrained deep siamese convolutional network model. *IEEE Geosci. Remote. Sens. Lett.*, 18(5):811–815, 2021. doi: 10.1109/LGRS.2020.2988032. URL <https://doi.org/10.1109/LGRS.2020.2988032>.
- [188] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 2414–2423. IEEE Computer Society, 2016. doi: 10.1109/CVPR.2016.265. URL <https://doi.org/10.1109/CVPR.2016.265>.
- [189] Michal Dereziński and Manfred K. K Warmuth. The limits of squared euclidean distance regularization. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014. URL [https://proceedings.neurips.cc/paper\\_files/paper/2014/file/95151403b0db4f75bfd8da0b393af853-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2014/file/95151403b0db4f75bfd8da0b393af853-Paper.pdf).
- [190] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.*, 13(4):600–612, 2004. doi: 10.1109/TIP.2003.819861. URL <https://doi.org/10.1109/TIP.2003.819861>.
- [191] Pablo Arbeláez, Michael Maire, Charless C. Fowlkes, and Jitendra Malik. Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(5):898–916, 2011. doi: 10.1109/TPAMI.2010.161. URL <https://doi.org/10.1109/TPAMI.2010.161>.
- [192] Fawwaz Ulaby, M Craig Dobson, and José Luis Álvarez-Pérez. *Handbook of radar scattering statistics for terrain*. Artech House, 2019.

- [193] Heng-Chao Li, Wen Hong, Yirong Wu, and Ping-Zhi Fan. Bayesian wavelet shrinkage with heterogeneity-adaptive threshold for SAR image despeckling based on generalized gamma distribution. *IEEE Trans. Geosci. Remote. Sens.*, 51(4-2):2388–2402, 2013. doi: 10.1109/TGRS.2012.2211366. URL <https://doi.org/10.1109/TGRS.2012.2211366>.
- [194] Sathit Intajag and Sakreya Chitwong. Speckle noise estimation with generalized gamma distribution. In *2006 SICE-ICASE International Joint Conference*, pages 1164–1167. IEEE, 2006. doi: 10.1109/SICE.2006.315296.
- [195] Guang-Ting Li, Chun-Le Wang, Ping-Ping Huang, and Wei-Dong Yu. Sar image despeckling using a space-domain filter with alterable window. *IEEE Geoscience and Remote Sensing Letters*, 10(2):263–267, 2012. doi: 10.1109/LGRS.2012.2200875.
- [196] Gerardo Di Martino, Mariana Poderico, Giovanni Poggi, Daniele Riccio, and Luisa Verdoliva. Benchmarking framework for SAR despeckling. *IEEE Trans. Geosci. Remote. Sens.*, 52(3):1596–1615, 2014. doi: 10.1109/TGRS.2013.2252907. URL <https://doi.org/10.1109/TGRS.2013.2252907>.
- [197] Zhou Wang and Alan C. Bovik. A universal image quality index. *IEEE Signal Process. Lett.*, 9(3):81–84, 2002. doi: 10.1109/97.995823. URL <https://doi.org/10.1109/97.995823>.
- [198] Xiaoshuang Ma, Penghai Wu, Yanlan Wu, and Huanfeng Shen. A review on recent developments in fully polarimetric SAR image despeckling. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.*, 11(3):743–758, 2018. doi: 10.1109/JSTARS.2017.2768059. URL <https://doi.org/10.1109/JSTARS.2017.2768059>.
- [199] Giles M Foody. Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment*, 239:111630, 2020. doi: <https://doi.org/10.1016/j.rse.2019.111630>. URL <https://www.sciencedirect.com/science/article/pii/S0034425719306509>.

- [200] Xu Tang, Huayu Zhang, Lichao Mou, Fang Liu, Xiangrong Zhang, Xiao Xiang Zhu, and Licheng Jiao. An unsupervised remote sensing change detection method based on multiscale graph convolutional network and metric learning. *IEEE Trans. Geosci. Remote. Sens.*, 60:1–15, 2022. doi: 10.1109/TGRS.2021.3106381. URL <https://doi.org/10.1109/TGRS.2021.3106381>.
- [201] Yuxing Chen and Lorenzo Bruzzone. Self-supervised remote sensing images change detection at pixel-level. *arXiv preprint arXiv:2105.08501*, 2021.
- [202] Emanuele Dalsasso, Loïc Denis, and Florence Tupin. SAR2SAR: A semi-supervised despeckling algorithm for SAR images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.*, 14:4321–4329, 2021. doi: 10.1109/JSTARS.2021.3071864. URL <https://doi.org/10.1109/JSTARS.2021.3071864>.
- [203] Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, and José García Rodríguez. A review on deep learning techniques applied to semantic segmentation. *CoRR*, abs/1704.06857, 2017. URL <http://arxiv.org/abs/1704.06857>.
- [204] Luis Perez and Jason Wang. The effectiveness of data augmentation in image classification using deep learning. *CoRR*, abs/1712.04621, 2017. URL <http://arxiv.org/abs/1712.04621>.
- [205] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu. Object detection with deep learning: A review. *IEEE Trans. Neural Networks Learn. Syst.*, 30(11): 3212–3232, 2019. doi: 10.1109/TNNLS.2018.2876865. URL <https://doi.org/10.1109/TNNLS.2018.2876865>.
- [206] Yuhui Zheng, Byeungwoo Jeon, Danhua Xu, Q. M. Jonathan Wu, and Hui Zhang. Image segmentation by generalized hierarchical fuzzy c-means algorithm. *J. Intell. Fuzzy Syst.*, 28(2):961–973, 2015. doi: 10.3233/IFS-141378. URL <https://doi.org/10.3233/IFS-141378>.
- [207] James C Bezdek, Robert Ehrlich, and William Full. Fcm: The fuzzy c-means clustering algorithm. *Computers & geosciences*, 10(2-3):191–203, 1984. doi: <https://doi.org/10.3233/IFS-141378>.

1016/0098-3004(84)90020-7. URL [https://www.sciencedirect.com/science/article/pii/0098300484900207](https://www.sciencedirect.com/science/article/pii/S0098300484900207).

# **Appendix A**

## **Published Papers**

# Image Registration Techniques and Applications: Comparative Study on Remote Sensing Imagery

Mohamed Ihmeida  
Department of Computer Science  
University of Reading  
Reading RG6 6AY, UK  
tw821373@pgr.reading.ac.uk

Hong Wei  
Department of Computer Science  
University of Reading  
Reading RG6 6AY, UK  
h.wei@reading.ac.uk

**Abstract**—Image registration is a crucial task in many computer vision applications. It is the process of matching and aligning two or more images of a scene. These images can be captured from different viewpoints, different sensors, or different times. Feature based image registration has four main steps: feature detection and description, feature matching, outliers rejection and computing homography and image re-sampling. Computational cost and registration accuracy of feature-based image registration mainly depend on the robustness of feature detection and description methods. Therefore, choosing an optimal feature detection and description method is vital in image registration applications. This research illustrates a comparison between popular image registration algorithms; Scale-invariant feature transform (SIFT), Speeded Up Robust Features (SURF), Oriented FAST and Rotated BRIEF (ORB), KAZE, Binary Robust Invariant Scalable Keypoints (BRISK) and Accelerated-KAZE (AKAZE) in different scenarios: rotation (0 to 360 degrees), scaling (25% to 600%) and multitemporal. The remote sensing images that are used in the experiments are Radar images, Aerial images, and Unmanned Aerial Vehicle (UAV) images. Nearest Neighbour Distance Ratio (NNDR) is performed in the feature matching, whereas RANSAC is applied to reject the outliers matching. The results of the experiments show that SIFT outperforms other algorithms, showing strong stability and high precision in all scenarios. As for real-time application, ORB performs well, and it is the fastest algorithm for all scenarios and then AKAZE as the second fastest one.

**Keywords**—SIFT, SURF, ORB, BRISK, KAZE, AKAZE, RANSAC, Nearest Neighbour Distance Ratio, Image Matching, Image Registration

## I. INTRODUCTION

Image registration is a process to align the features of two or more images to one image. It is important in many applications such as image fusion, image stitching and change detection. Some popular feature detection and description algorithms are used in the image registration process. This paper focuses on six of them: SIFT, SURF, ORB, BRISK, KAZE, and AKAZE. In a real application, there is always a need to identify which algorithm performs the best in real scenarios. Many researchers have compared these algorithms by using the Affine Covariant Regions dataset (also known as Oxford dataset) or Iguazu dataset images that are used in KAZE experiments [1] [2]. These researchers have examined algorithms such as SIFT, SURF, ORB, BRISK, KAZE and

AKAZE [3], [4] [5] [6]. Previous studies have compared feature detection and description algorithms from different aspects. Micolajczyk and Schmid in [2], introduced the Affine Covariant Regions dataset and compared the performance of local feature descriptors such as SIFT, PCA-SIFT, GLOH and Cross-correlation using a variety of datasets that include rotation, zoom+ rotation, viewpoint change, image blur, JPEG compression and light change (a.k.a Oxford dataset). In contrast, rotation and scale change assessed in this research was based only on 30-45 degrees rotation and 200% to 250% scaling. Gauglitz, Hobias and Turk [7] assessed various feature detectors such as Harris Corner Detector, FAST, Hessian, Difference of Gaussians and several descriptors such as SIFT, SURF, Image Patch and some others for visual tracking. However, no rotation, and scaling evaluation were conducted. Pusztai and Hajder [8] presented a quantitative comparison of several feature detectors available in OpenCV 3 (SIFT, SURF, AKAZE, KAZE, MSER, ORB, FAST, GFTT, AGAST and BRISK). Tareen and Saleem [6] evaluated the feature detection and description algorithms (SIFT, SURF, ORB, KAZE, BRISK, AKAZE) using Oxford Dataset, other image pairs from OpenCV, VLFeat and vision toolbox of MATLAB. It provides a quantitative comparison of feature detectors and descriptors. However, a precision evaluation metric has not been used to evaluate the algorithms. Sharma and Jain [9] evaluated the feature detection and description algorithms (SIFT, SURF, ORB and AKAZE) for image registration and stitching using building dataset image pair, apartment dataset image pair and MATLAB dataset image pair. Based on their results, AKAZE was reported as the most accurate and the second fastest algorithms for image matching and stitching. Based on our knowledge no research has compared these algorithms with special attention on remote sensing data to obtain the answers to; which algorithm performs the best in precision with remote sensing datasets in rotation, scale change, and different viewpoints? Moreover, which algorithm performs well for real-time applications? This research aims to answer these questions in a case study with various remotely sensed images. The rest of the paper is organised as follows. Section II presents an image registration paradigm with applications and processes. Section III demonstrates experimental results with discussions. Finally, Section IV concludes the study.

## II. IMAGE REGISTRATION PARADIGM

Image registration is a fundamental task for many computer vision applications. It is the first step in image fusion, image stitching, change detection and many more [10] [11]. This section studies image registration with possible applied scenarios, and techniques used for feature detection and descriptions, feature matching, and outlier rejection.

### A. Image Registration Scenarios

Image registration scenarios for remote sensing can be classified as below, based on how images are taken.

- **Images taken from different viewpoints (Viewpoints Registration):** it is used to integrate information from one moving sensor or multiple sensors from different viewpoints to the same object to make a 3D model. Landmark navigation and plant exploration are examples of applications that obtain advantages from this kind of registration [10] [12] [11]. Image stitching is one application of viewpoints registration too.
- **Images taken from different times (Temporal Registration or Multi-temporal Images):** it is utilised for change detection, and land resource survey that includes monitoring agricultural and land cover features extracted from data captured from one or more sensors over time [10] [12] [13] [11].
- **Different sensors (Multi-modal Registration):** this application is vital for integrating complementary information from different sensors. It benefits, for instance, the land cover uses such as yield estimate in agriculture, flood monitoring and detection of illegal crops. The fusion of different remote sensing data illustrates countless promises in assisting the decision-making process in several previous applications [10] [12] [13] [11].

### B. Classical Image Registration

Classical image registration can be categorised into manual and automatic methods:

- **Manual Image Registration:** in the last two decades, traditional image geo-referencing started with a manual image registration process. A human does this process to allocate the ground control point (GCP) called Tie points in both reference and sensed images. This step is equivalent to feature detection. The next step is to match this GCP in both images by human operators. This process is facing many challenges [12]; including, i) time cost (especially for remote sensed (RS) images because RS images usually cover a large size), ii) finding and allocating precise features are more difficult in RS than in medical images, and iii) it requires expert operators. Based on these challenges, there are requirements for semi-automatic or automatic registration approaches.
- **Automatic Image Registration (AIR):** it can be categorised into Area Based Methods and Feature-Based Methods. Area Based Method (also called Intensity-Based Method) is usually used when an image reflects

a relatively smooth surface/scene short of essential features. It is popularly used for medical image registration. However, it is very time-consuming and is influenced by image noise. Whereas the feature-based method uses features such as lines, corners, contour and edges to find a corresponding region in a reference and sensed images. It is less computationally expensive and more resistant to noise. Therefore, it is commonly adapted for remote sensing image registration [10] [12] [13]. Our study follows the feature based approach, which contains several steps: feature detection and description, feature matching, outliers removal, homography and image re-sampling. Figure 1 shows the process of feature-based image registration. Traditional methods such as SIFT, SURF, ORB, KAZE, BRISK and AKAZE are several famous feature detectors and descriptors popularly used in image registration. Feature matching methods find corresponding features from two feature descriptors of reference and sensed images. Nearest Neighbour Distance Ratio (NNDR) is used to reduce initially false matches from feature matching. RANSAC is used to remove further outliers from putative matches [14]. The good matches after RANSAC are used in transformation matrix estimation [14].

### C. Feature Detection Algorithms

Figure 1 shows that feature detection from which a feature descriptor is established is the first step in image registration. The six feature detection algorithms are briefly described below. Feature detection algorithms detect features (also called keypoints) and provide a feature descriptor, making it easy to find the relevant feature in other images. The input images are called reference and sensed images.

- **SIFT:** Scale Invariant Feature Transform (SIFT) [15] is to solve corner detecting problems with scaling invariance. The main stages of computation in SIFT lead to generating a set of images features. SIFT detector is based on Difference-of-Gaussians (DoG) operator which approximates Laplacian-of-Gaussian (LoG). Feature points are detected by searching local maxima using DoG at various scales of subject images. The description method extracts a  $16 \times 16$  neighbourhood around each detected feature and further segments the region into sub-blocks rendering a total of 128 values. Equation (1) [15] illustrates the convolution of difference of two Gaussians (computed at a different scale) convolution with image  $I(x, y)$

$$D(x, y, \delta) = (G(x, y, k\delta) - G(x, y, \delta)) * I(x, y) \quad (1)$$

Where  $G(x, y, \delta)$  is the Gaussian function with different scales of  $\delta$ ,  $k$  is a constant used for scale change in the Gaussian function.

- **SURF:** Speeded-Up Robust Features (SURF) [16] relies on Gaussian scale-space analysis of images as SIFT. It uses different detectors, descriptors to speed up the computation that is the disadvantage of SIFT algorithm.

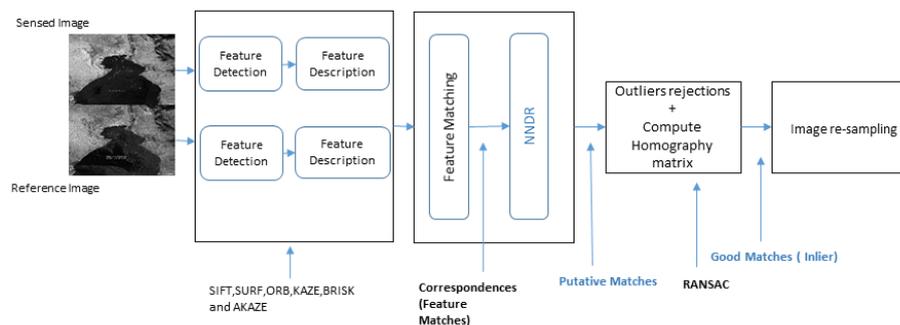


Fig. 1: Image registration steps with highlighted NNDR, Putative matches and Good Matches

Hessian matrix has been used as a detector. Blob-like structures can be detected at locations where the determinant of the Hessian matrix is maximum. Hessian matrix  $H(X, \delta)$  at point  $X = (x, y)$  in an image  $I$  at scale  $\delta$  is described as in equation (2) [16].

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix} \quad (2)$$

where  $L_{xx}(X, \sigma)$  is convolution of Gaussian second order derivative with  $I$  in point  $x$ , and similarly for  $L_{xy}(X, \sigma)$  and  $L_{yy}(X, \sigma)$ .

- **ORB**: Oriented FAST and Rotated BRIEF (ORB) [4] was developed to deal with the high computational burden of SIFT and SURF, especially for real-time applications, for example, visual odometry. ORB consists of FAST (Features from Accelerated Segment Test) detector and BRIEF (Binary Robust Independent Elementary Features) descriptor; both are performed well with low computational cost. This combination has provided a fast algorithm to deal with real-time applications which require corner feature extraction. FAST [17] is a perfect corner feature detection for a real-time application that matches visual features, such as Parallel Tracking and Mapping [4].

In summary, FAST is several times faster than any existing corner detectors. However, it is affected by a high level of noise [4]. The BRIEF descriptor is a simple binary test between pixels in smoothed image patch [18]. Although BRIEF's performance is similar to SIFT in many cases, such as robustness to lighting, blur and perspective distortion, it performs poorly with rotation. Therefore, rBRIEF has been developed in [4] that has significant enhancement in dealing with rotated images compared with BRIEF. In summary, ORB is much faster than SIFT and SURF. However, it has not been addressed the scale invariance in the experiment.

- **KAZE**: It takes advantage of non-linear scale-space through non-linear diffusion filtering. This method blurs images locally adaptive to feature points, reducing noise and retaining regions' boundaries in images under processing. Scale normalised determinant of Hessian Matrix is used as KAZE detector. It is calculated at multiple scale levels. The maxima of detector response are chosen as feature points using a moving window. Feature descriptor presents rotation invariance property by finding dominant orientation in the circular neighbourhood around every feature detector. KAZE is more computationally expensive than SURF, but the KAZE experiments in [1] have proved it is less computationally expensive than SIFT due to the computation in the non-linear scale-space [19] [1].
- **BRISK**: Binary Robust Invariant Scalable Keypoints (BRISK) [20] is an algorithm that is used Adaptive and Generic Accelerated Segment Test (AGAST) [21] as a corner detector and filters them with FAST corner score when looking for maxima in a scale-space pyramid. The feature descriptor is built on classifying the characteristic direction of every feature for attaining rotation invariance. The descriptor is constructed as a binary string to achieve illumination invariance. BRISK is introduced to provide solutions for high-performance algorithms such as SIFT and SURF. It was also reported to be faster than SIFT and SURF.
- **AKAZE**: KAZE is updated to Accelerated-KAZE (AKAZE) [3]. It is similar to KAZE based on non-linear diffusion filtering. However, it solves the computational burden of creating a non-linear scale-space by using a mathematically well-organised framework Fast Explicit Diffusion (FED), embedded in a pyramidal framework to accelerate feature detection in non-linear scale-spaces significantly. The AKAZE detector contains a determinant of the Hessian Matrix. The rotation invariance quality has improved by Scharf filters [3]. The maxima of detector

responses in spatial locations are chosen as feature points. The feature descriptor of AKAZE is based on a highly efficient Modified Local Difference Binary (MLDB). By experiments, although authors in [3] have proved that AKAZE is computationally less expensive than KAZE, SIFT, and SURF, it is more expensive than BRISK and ORB.

#### D. Feature Matching

Feature matching is an essential process of many problems in computer vision such as object recognition, 3D reconstruction from multiple images, image registration and motion tracking [10] [11]. There are two types of matching strategies: Brute Force and K-Nearest Neighbours (KNN). Brute Force is classified based on the feature descriptors' type. The floating-point descriptor is used in SIFT, SURF and KAZE, whereas the binary descriptor is used in ORB, BRISK and AKAZE.

- Brute Force algorithm specifies two parameters between two features: distance metric and crosscheck boolean. Crosscheck boolean which is used to validate if the two features are matched. This method classifies the algorithms based on the descriptor types. For example, the L1-norm function from the OpenCV library (also called Least Absolute Deviations) is used for float-point descriptors such as SIFT, SURF and KAZE use and the second type of descriptors is Hamming distance is used for ORB, BRISK and AKAZE [12].
- KNN with a pre-defined value K is an approach, which may produce a large number of matching points. Therefore, there is a need to identify those suitable matches by using (NNDR), which is a threshold introduced by Lowe to reduce the number of false matches (Outlier) [15]. In this paper, we have applied Brute force and KNN then set NNDR threshold value as 0.7, which is a set to remove 90% of the false matches while discarding less than 5% of the correct matches between descriptors in reference and sensed images. Equation (3) describes NNDR [22]

$$\frac{\|D_R - D_{S1}\|}{\|D_R - D_{S2}\|} < T_{ratio} \quad (3)$$

where  $D_R$  is the reference image feature descriptor,  $D_{S1}$  and  $D_{S2}$  are the first and second closest descriptors to  $D_R$  in the sensed image respectively.  $\|D_R - D_{S1}\|$  is the distance of  $D_R$  to  $D_{S1}$ ,  $\|D_R - D_{S2}\|$  is the distance of  $D_R$  to  $D_{S2}$ , and  $T_{ratio}$  is threshold, set as 0.7. The matching value after NNDR is called putative matching. After applying NNDR, the RANSAC algorithm is applied to remove the rest of the outliers.

#### E. Outliers Rejection

Random Sample Consensus (RANSAC) is a famous algorithm used to reject the outlier points (it is also called false matching points) [23]. RANSAC is introduced as general and straightforward applicable to many different problems. It is used to reject a high number of outliers and estimate the homography matrix by using good matches. Homography

matrix transfers the second image (sensed image) to the first one (reference image).

### III. EXPERIMENTAL RESULTS WITH DISCUSSIONS

Experiments were conducted in two sets of data, which will be introduced in detail in Section III A. Python 3.7 with OpenCV 3.4 were used to perform the experiments. The platform of the experiments was Google Colaboratory environment with a Tesla GPU P100-PCIE-16 GB RAM, 147.15 GB Disk.

#### A. Datasets

Two datasets are used for the experiments to compare the six algorithms. Dataset A has three image pairs selected from different remote sensing databases, as demonstrated in Table I. Dataset A is used to evaluate the image matching precision and computational time both are described in Section III B. Moreover, it is used to demonstrate image stitching via image registration by SIFT, as in figure 2. Whereas dataset B has the same three reference images used in dataset A with different sizes that are used to examine the rotation and scaling scenarios. Table II illustrates the dataset B specifications.

TABLE I: Dataset A specifications

P	RI	SI	IT	RI size	SI size
1	WKC2019	WKC2019P	Aerial RGB	1345*983	1427*1165
2	lake	lake1	UAV RGB	767*1165	767*1165
3	Suez21	Suez25	Radar	1140*1165	1140*1165

P: pair, RI: Reference Image, SI: Sensed Image, IT: Image Type

TABLE II: Dataset B specifications

Image	Image name	Type of Image	Image Size (H*L)
1	lake [24]	UAV RGB image	672*1159
2	WKC2019 [25]	Aerial RGB image	672*446
3	Suez21 [26]	Radar image	570*478

#### B. Evaluation metrics

There are two types of experiments based on each dataset. Dataset A is used to examine the image matching for the six algorithms using three different evaluation metrics; Precision [3] [19] [5], the number of inlier points (Good matches) and computational cost [9].

- The number of Inlier points (Good matches) is the number of matching points after applying RANSAC algorithm.
- Computational cost is vital to examine which algorithm can be used for real-time applications.
- Precision (the inlier ratio) defines the number of correct matches out of the putative matches as shown in Figure 1, and it is expressed in Equation (4).

$$Precision = \frac{Good\ Matches}{Putative\ Matches} \quad (4)$$

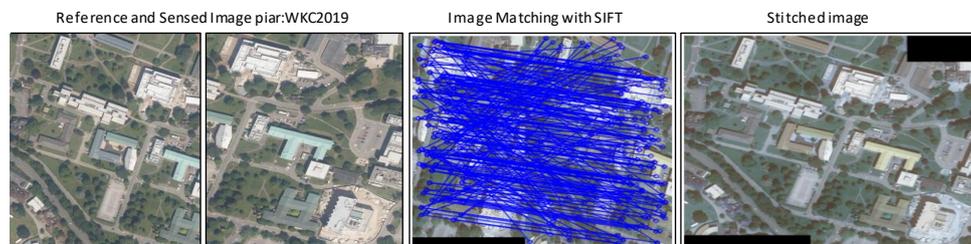


Fig. 2: Dataset A: WKC2019 and WKC2019P to demonstrate image stitching via image registration by using SIFT

### C. Results and Discussion

Table III presents the image matching results of image pairs in dataset A. The highest precision and lowest computational cost are highlighted. The mean value and standard deviation for all three image pairs show that SIFT provides the highest precision of  $94.11\% \pm 4.14$ , then BRISK with  $93.79\% \pm 6.86$  and then AKAZE with  $93.32\% \pm 6.63$ . The experimental results shown in Figure 3 for dataset B show that SIFT is the only stable algorithm among the six algorithms in rotation and scaled scenarios. It provides the highest mean value of precision for the three image pairs of dataset B in the rotation scenario and the image scaling range of (25% to 600%) in the scaling scenario. It has been noticed in figure 3 that the matching precision of AKAZE, KAZE, BRISK, ORB and SURF have dramatically dropped after scaling change to 400% of the original image. However, SIFT has a high computational cost compared to ORB and AKAZE. It provides a relatively low computational cost with Suez (radar image) and WKC2019 (aerial image). In the rotation scenario, AKAZE and KAZE provide high precision with fluctuation in some angles. Most algorithms provide high precision over 80% in a scaling scenario, but this precision drops dramatically after scaling the sensed image to 400% of the original, whereas SIFT matching precision gradually drops after scaling the sensed image to 500% of the original one. For the computational cost, ORB is always the fastest algorithm in all experiments with both datasets, and AKAZE is the second fastest algorithm.

### IV. CONCLUSION AND FUTURE WORK

This research has compared the six popular feature detection and description algorithms used for remotely sensed image matching. It also covers the different matching strategies and methods to remove outliers. Performance evaluation to compare these algorithms are conducted in two datasets with different scenarios. To summarise, SIFT precision outperforms all other algorithms in all experiments with Datasets A and B. However, it has a higher computational cost than ORB and AKAZE. BRISK and AKAZE have the second-best precision in experiment dataset A and scaling scenario in dataset B, respectively. AKAZE has slightly lower precision in the experiments of dataset A but is stabler than BRISK in rotation

scenario and much faster in all scenarios. Therefore, it can be used as an alternative to SIFT. Furthermore, AKAZE has the second-best precision in the rotation scenario and the third-best one in experiment dataset A. ORB is the fastest algorithm with reasonable precision, usually over 80%. Therefore, it can be recommended for real-time remote sensing applications, whilst AKAZE is the second-fastest one. This finding could further investigate comparing image matching precision and computational cost of SIFT with deep neural networks using our remote sensing datasets.

### REFERENCES

- [1] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE Features," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7577 LNCS, no. PART 6, 2012, pp. 214–227.
- [2] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [3] P. F. Alcantarilla, J. Nuevo, and A. Bartoli, "Fast explicit diffusion for accelerated features in nonlinear scale spaces," in *BMVC 2013 - Electronic Proceedings of the British Machine Vision Conference 2013*, 2013.
- [4] E. Rublee, W. Garage, and M. Park, "ORB : an efficient alternative to SIFT or SURF," *IEEE*, 2011, pp. 2564–2571.
- [5] M. Hassaballah, H. A. Alsharby, and A. A. Ali, "Analysis and evaluation of keypoint descriptors for image matching," in *Recent Advances in Computer Vision*, 2018, vol. 804, no. January, pp. 113–140.
- [6] Z. Saleem and S. Tareen, "A Comparative Analysis of SIFT , SURF , KAZE , AKAZE , ORB , and BRISK," in *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. IEEE, 2018, pp. 1–10.
- [7] S. Gauglitz, T. Höllerer, and M. Turk, "Evaluation of interest point detectors and feature descriptors for visual tracking," *International Journal of Computer Vision*, vol. 94, no. 3, pp. 335–360, 2011.
- [8] Z. Pusztai and L. Hajder, "Quantitative Comparison of Feature Matchers Implemented in OpenCV3," *21 st. Computer Vision Winter Workshop, Rimske Toplice*, 2016.
- [9] S. K. Sharma and K. Jain, "Image Stching using AKAZE Features," *Journal of the Indian Society of Remote Sensing*, vol. 48, no. 10, pp. 1389–1401, 2020.
- [10] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.
- [11] J. L. Moigne and N. S. Netanyahu, *Image Registration for remote sensing*, 2011.
- [12] K. Kuppala, S. Banda, and T. R. Barige, "An overview of deep learning methods for image registration with focus on feature-based approaches," *International Journal of Image and Data Fusion*, vol. 11, no. 2, pp. 113–135, 2020.
- [13] A. A. Goshtasby, *Theory And Applications of Image Registration*. John Wiley Sons, 2017.

TABLE III: Quantitative Comparison of the Six Algorithms

Algorithms	FD 1	FD 2	Feature Matches	Putative Matches	Inliers Points	Outliers Points	Precision %	Image Matching Time (s)
Image Pair 1: WKC2019								
SIFT	9091	10393	9091	3372	3332	40	98.81	9.21
SURF	10043	12908	10043	4053	3898	155	96.18	10.73
ORB	500	500	500	169	166	3	98.22	1.45
BRISK	10392	9765	10392	3026	3015	11	99.63	6.23
KAZE	4746	4977	4746	2071	2058	13	99.23	6.23
AKAZE	4023	4286	4023	1675	1667	8	99.52	2.37
Image Pair 2: Lake								
SIFT	6634	6880	6634	1206	1116	90	92.54	5.21
SURF	7054	7144	7054	998	921	77	92.28	5.49
ORB	500	500	500	128	107	21	83.59	1.37
BRISK	9067	8215	9067	1451	1386	65	95.52	2.81
KAZE	4750	4470	4750	1487	1328	159	89.30	4.57
AKAZE	3913	3402	3913	1266	1093	173	86.33	1.73
Image Pair 3: Suez Canal								
SIFT	7605	7958	7605	366	333	33	90.98	6.13
SURF	11683	14043	11683	377	335	42	88.86	10.82
ORB	500	500	500	18	13	5	72.22	1.74
BRISK	35861	38253	35861	1423	1227	196	86.23	27.23
KAZE	4896	4490	4896	1141	966	175	84.66	4.90
AKAZE	4347	4291	4347	811	763	48	94.10	2.39
Mean Value and Standard Deviation for all three image pair								
Metrics	SIFT	SURF	ORB	BRISK	KAZE	AKAZE		
Precision %	<b>94.11 ±4.14</b>	92.44 ±3.66	84.68 ±13.03	<b>93.79 ±6.86</b>	91.06 ±7.44	<b>93.32 ±6.63</b>		
IMT(s)	6.85 ±2.09	9.01 ±4.84	<b>1.52 ±0.19</b>	11.31 ±13.80	<b>5.23 ±0.88</b>	<b>2.16 ±0.38</b>		

The **Red** colour is the highest value, the **Blue** one is the second-highest value, and the **Green** is the third one. FD 1: Number of features Detected in Image 1 and FD 2: Number of features Detected in Image 2. IMT: Image Matching Time (s).

- [14] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image Matching from Handcrafted to Deep Features: A Survey," *International Journal of Computer Vision*, vol. 129, no. 1, pp. 23–79, 2021.
- [15] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [16] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*. Springer, 2006, pp. 404–417.
- [17] D. G. Viswanathan, "Features from Accelerated Segment Test (FAST) Deepak Geetha Viswanathan 1." 2009.
- [18] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6314 LNCS, no. PART 4, pp. 778–792, 2010.
- [19] J. Heinly, E. Dunn, and J.-M. Frahm, "Comparative evaluation of binary features," in *European Conference on Computer Vision*. Springer, 2012, pp. 759–773.
- [20] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *2011 International conference on computer vision*. Ieee, 2011, pp. 2548–2555.
- [21] E. Mair, G. D. Hager, D. Burschka, M. Suppa, and G. Hirzinger, "Adaptive and generic corner detection based on the accelerated segment test," in *European conference on computer vision*. Springer, 2010, pp. 183–196.
- [22] S. BAŞ, "Point-based matching of oblique images acquired from airplane and uav platforms," Master's Thesis, Hacettepe University, Available:openaccess.hacettepe.edu.tr, 2020.
- [23] M. A. Fischler and R. C. Bolles, "Random Sample Paradigm for Model Consensus: A Applications to Image Fitting with Analysis and Automated Cartography," *Graphics and Image Processing*, vol. 24, no. 6, pp. 381–395, 1981.
- [24] SenseFly, "Thammasat university campus in bangkok dataset," 2018, available:sensefly.com, Accessed on 18 Dec 2020.
- [25] Digimap, "High resolution 25cm vertical aerial imagery," 2019, [JPG geospatial data], Scale 1:500. Tiles: su7371, Updated: 11 October 2013, Getmapping, Using: EDINA Aerial Digimap Service, digimap.edina.ac.uk, Accessed on: 16 Feb 2021.
- [26] ESA, "Suez canal traffic jam seen from space," 2021, available: esa.int, Accessed on 30 Mar 2021.

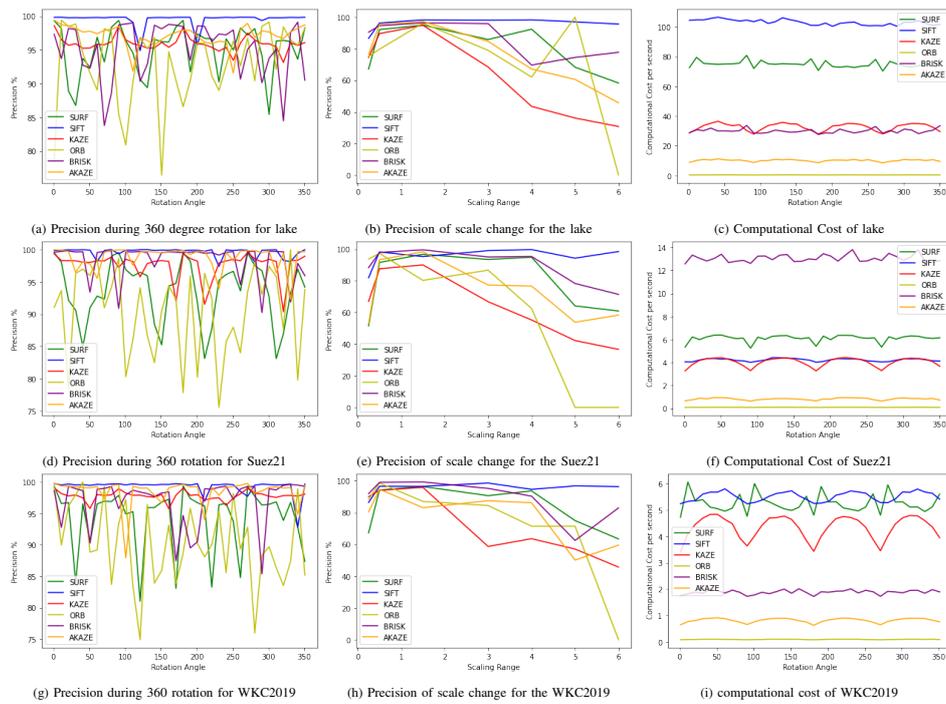


Fig. 3: Results of rotation and scaling scenarios, left column rotation precision and second column is scaling precision and last one is computational cost

## Deep Despeckling of SAR Images to Improve Change Detection Performance

Mohamed Ihmeida<sup>1</sup>[0009-0009-8728-6868] and Muhammad Shahzad<sup>1</sup>[0009-0002-9394-343X]

University of Reading, Reading RG6 6DH, UK  
m.a.g.ihmeida@pgr.reading.ac.uk, m.shahzad2@reading.ac.uk

**Abstract.** Synthetic aperture radar (SAR) image change detection (CD) focuses on identifying the change between two images at different times for the same geographical region. SAR offers advantages over optical sensors for disaster-related change detection in remote sensing due to its all-weather capability and ability to penetrate clouds and darkness. The performance of change detection methods is affected by several challenges. Deep learning methods, such as convolutional neural networks (CNNs), have shown promising performance in dealing with these challenges. However, CNN methods still suffer from speckle noise, adversely impacting the change detection performance F1 score. To tackle this challenge, we propose a CNN model that despeckles the noise prior to applying change detection methods. We extensively evaluate the performance of our method on three SAR datasets, and the results of our proposed method demonstrate superior performance compared to state-of-the-art methods such as DDNet and LANTNet performance. Our method significantly increased the change detection accuracy from a baseline of 86.65% up to 90.79% for DDNet and from 87.16% to 91.1% for LANTNet in the Yellow River dataset.

**Keywords:** Unsupervised Learning · SAR change detection · Despeckling noise.

### 1 Introduction

Remote sensing change detection (CD) is an essential technique for identifying changes in multi-temporal images of the same geographical region [10] [16]. It provides valuable information for various applications, including deforestation monitoring, target detection, and agricultural advancement [2] [23]. Additionally, CD algorithms support decision-making during natural disasters, enabling timely actions to prevent material losses and save lives [13]. Change detection in remote sensing involves distinguishing changed and unchanged pixels in multi-temporal Earth Observation (EO) images for the same geographical region. These multi-temporal EO images are required to be co-registered. This step is important in aligning EO images to the same coordinate system, which is useful for obtaining consistent radiometric characteristics, such as brightness and contrast. This process enhances the change detection performance [14] [19]. Key point extraction

techniques like SIFT, SURF, and CNNs are often used for image registration [6]. Classical change detection can be easily obtained by computing the intensity difference between images. The result of this process is called a change map (CM). However, challenges such as co-registration errors, illumination variations, and speckle noise affect the accuracy of change detection algorithms.

Synthetic aperture radar offers advantages over optical sensors for change detection in remote sensing due to its all-weather capability, penetration through clouds and vegetation, and sensitivity to small changes. SAR change detection methods primarily rely on unsupervised learning due to the lack of annotated SAR datasets. Various unsupervised CD methods use clustering algorithms, such as principal component analysis, fuzzy clustering algorithms (FCM) [12] and fuzzy local information C-mean (FLICM) [17]. Researchers make an effort to reduce the impact of speckle noise on CD methods. Qu et al. [22] introduced a dual domain neural network (DDNet) incorporating spatial and frequency domains to reduce speckle noise. Gao et al. [10] proposed a Siamese adaptive fusion network for SAR image change detection, which extracts semantic features from multi-temporal SAR images and suppresses speckle noise. Meng et al. [20] presented a noise-tolerant network called LANTNet that utilises feature correlations among multiple convolutional layers and employs a robust loss function to mitigate the impact of noisy labels. While these deep learning-based approaches show some robustness against speckle noise, they still struggle to eliminate it and reduce its effectiveness in change detection methods. Furthermore, the presence of speckle noise varies between single-look (pre-change) and multi-look (post-change) SAR imaging processes, further degrading the performance of change detection algorithms when considering different instances in time.

To address the issues with degrading CD performance, we propose a robust despeckling model (DM) architecture that effectively suppresses speckle noise in SAR CD datasets. This approach leads to significant improvements in change detection performance. Experimental evaluations on public SAR CD datasets provide compelling evidence of the superiority of our proposed method when compared to existing approaches.

## 2 Related Work

SAR change detection is widely used in various applications, including urban extension [16], agricultural monitoring [23], target detection [21], and disaster assessment [2]. Due to the lack of annotated SAR datasets, most researchers rely on unsupervised methods for SAR change detection. However, the presence of speckle noise poses a significant challenge and reduces the accuracy of change detection. Image pre-processing, including despeckling and image registration, is a crucial step in SAR change detection to enhance image quality and align multi-temporal images [19].

Generating a difference image (DI) is important in SAR change detection. Various methods, such as image differencing, log ratio, and neighbourhood-based ratio, have been proposed to generate the DI [5] [30]. The classification of the

DI typically involves thresholding and clustering. Some approaches use the preclassification result to train a classifier model and combine the preclassification and classifier results to generate a change map. These methods aim to improve change detection performance by leveraging preclassification and classifier information [8].

Recent approaches in SAR change detection focus on explicitly suppressing speckle noise to improve accuracy. Methods such as DDNet [22], Siamese adaptive fusion networks [10], and LANTNet [20] have been proposed to mitigate the impact of speckle noise and extract high-level features from multi-temporal SAR images. However, these approaches have limitations in effectively handling different speckle noise characteristics in images prior and after the change, especially when the number of looks varies. To address this challenge, we propose a despeckling model to suppress speckle noise and achieve effective SAR change detection for different numbers of looks in pre- and post-change images.

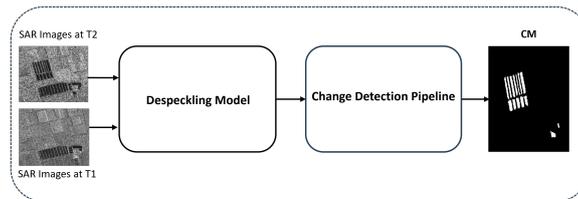


Fig. 1: An overview of the proposed modules

### 3 Methodology

The despeckling module applies a sequence of convolutional layers to reduce speckle noise in input SAR images. The resulting image with reduced noise is then passed to the subsequent CD methods. Figure 1 presents the DM and CD methods overview. The following sections explain the proposed despeckling model architecture and the change detection methods.

#### 3.1 Despeckling Model Architecture

The proposed despeckling architecture aims to learn a mapping from the input SAR image using convolutional layers to generate a residual image containing only speckle noise. The resulting speckle-only image can be combined with the original image through either subtraction [4] or division [27] operations to produce the despeckled image. The division operation is preferred as it avoids an additional logarithmic transformation step and allows for end-to-end learning. However, training such a network requires reference despeckled images, which are

4 M. Ihmeida and M Shahzad

Table 1: Proposed Despeckling Model Configuration. Where L1 and L10 refer to a series of Conv-ReLU layers, while the layers between L2 and L9 consist of Conv-BN and ReLU layers as illustrated in Figure 1.

-	layer	Filter Size	Filters	Output size
L1	Conv + ReLU	3*3*1	64	256*256*64
L2-L9	Conv + BN + ReLU	3*3*64	64	256 *256* 64
L10	Conv + ReLU	3*3*64	1	256 *256*1

typically unavailable for SAR images. To address this, researchers use synthetic reference images generated using multiplicative noise models [29] [4] [27]. This study also employs synthetic SAR reference images to train the proposed despeckling network architecture, consisting of ten convolutional layers with batch normalisation, ReLU activation functions, and a hyperbolic tangent as the final nonlinear function. The proposed architecture is similar to [27] [29] [4], but with additional convolutional layers and improved loss function presented in Figure 2. Moreover, the details on hyperparameters are also provided in Table 1 for clarity.

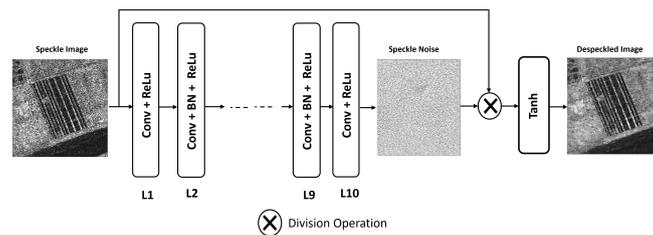


Fig. 2: Proposed despeckling model architecture

### 3.2 Proposed Loss Function

A common approach to training the despeckling network is to use the per-pixel Euclidean loss function  $LE(\theta)$ , computed by comparing the predicted despeckled image with the noise-free SAR image. The  $LE(\theta)$  calculates the squared Euclidean distance between corresponding pixels. While effective in various image restoration tasks, such as super-resolution, semantic segmentation, change detection, and style transfer, it often results in artifacts and visual abnormalities in the estimated image. Researchers have incorporated a total variation (TV) loss and an Euclidean loss function  $LE(\theta)$  as supplementary measures. The TV loss reduces artifacts but may lead to oversmoothing and information loss, thus impacting change detection performance. To overcome this, we design a loss

function which combines the  $LE(\theta)$  and a structural similarity index (SSIM), initially proposed for image quality assessment, which offers a better trade-off by removing artifacts while preserving essential information, ultimately enhancing change detection performance.

$$L_E(\theta) = \frac{1}{W \cdot H} \sum_{w=1}^W \sum_{h=1}^H \|X^{(w,h)} - \hat{X}^{(w,h)}\|^2 \quad (1)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1) \cdot (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1) \cdot (\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2)$$

The total loss is thus calculated as follows:

$$L_T = L_E(\theta) + \lambda_{SSIM} \cdot SSIM \quad (3)$$

Where  $X$  and  $\hat{X}$  are the reference (noise-free) and despeckled images, respectively,  $\mu_X$  and  $\mu_{\hat{X}}$  are the mean values of  $X$  and  $\hat{X}$  respectively. Similarly,  $\sigma_X$  and  $\sigma_{\hat{X}}$  are the standard deviations of  $X$  and  $\hat{X}$  respectively. While  $\sigma_{X\hat{X}}$  is the covariance between  $X$  and  $\hat{X}$ . Finally,  $C_1$  and  $C_2$  are constants set to be 0.01 and 0.03 respectively [28].

### 3.3 Change Detection

It is critical to suppress speckle noise in our proposed method to enhance CD performance. To evaluate the performance of the proposed despeckling model, we incorporated state-of-the-art CD methods, including DDNet [22] and LANTNet [20]. PCA- $k$ -means [3] is an unsupervised change detection method that utilises principal component analysis and  $k$ -means clustering to identify changes by splitting the feature vector space into two clusters. NR-ELM [9] employs a neighbourhood-based ratio to create a difference image and subsequently utilises an extreme learning machine to model high-probability pixels in the difference image. This information is then combined with the initial change map to produce the final change detection result. DDNet [22] combines spatial and frequency domain techniques to reduce speckle noise, while LANTNet [20] leverages feature correlations across multiple convolutional layers and incorporates a robust loss function to mitigate the impact of noisy labels.

## 4 Experimental Results & Evaluation

In this section, we introduced the datasets and evaluation metrics. Subsequently, we presented and evaluated the results by comparing them with those obtained from state-of-the-art CD methods.

6 M. Ihmeida and M Shahzad

#### 4.1 Datasets and Evaluation Metrics

Two types of datasets were used in this paper. The first is the Berkeley Segmentation Dataset 500, widely employed to generate synthetic SAR images for training the despeckling model. Real SAR images were used for testing, specifically for change detection purposes, to assess the model's performance. Detailed descriptions of both datasets can be found in the following subsections:

##### – Synthetic SAR Images

The Berkeley Segmentation Dataset 500 (BSD-500) was originally developed to evaluate the segmentation of natural edges, including object contours, object interior and background boundaries [1]. It included 500 natural images with carefully manually annotated boundaries and edges of natural objects collected from multiple users. This dataset has been widely used to generate synthetic SAR images for the purpose of despeckling [25] [18] [15]. Inspired by these studies, we have used it to train our despeckling model.

##### – Real SAR Images

For the purpose of change detection, we employed three real SAR image datasets that are multi-temporal and have been co-registered and corrected geometrically.

- *Farmland and Yellow River Datasets:* The images for both datasets were captured by RADARSAT-2 in the region of the Yellow River Estuary in China on 18th June 2008 (pre-change) and 19th June 2009 (post-change). The pre-change images are single-look, whereas the post-change images have been acquired via a multi-look (four) imaging process. The single-look pre-change image is significantly influenced by speckle noise compared to the four-look post-change image [10]. The disparity between the single and four looks in these two SAR datasets poses a significant challenge for change detection methods.
- *Ottawa Dataset:* The images for this dataset were also captured by RADARSAT-2 in May 1997 (pre-change) and August 1997 (post-change) in the areas affected by floods [22] [26] [11]. Because of the single imaging process, the pre- and post-change images are less affected by noise in this dataset.

The synthetic SAR images were utilised to train the proposed DM, as depicted in Figure 1. In contrast, the real SAR images were despeckled for the purpose of change detection (CD datasets). Figure 3 presents the real SAR datasets.

To evaluate the results, we used two common evaluation metrics, including Overall Accuracy and F1 score. The F1 score is usually used to evaluate the change detection accuracy [24] [7].

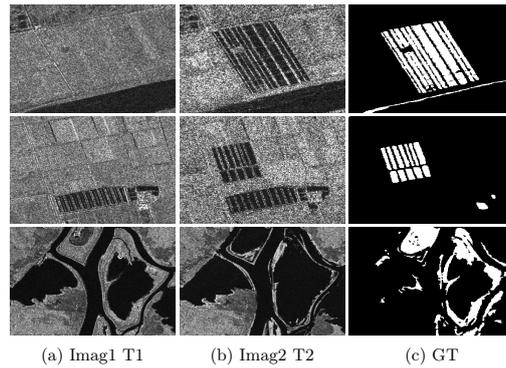


Fig. 3: The real SAR datasets. (a) Image acquired in T1. (b) Image acquired in T2. (c) Ground truth image(GT).

Table 2: Quantitative evaluation on three CD datasets based on despeckling model. Here, w/o means it is the original method without despeckling, and DM is our proposed despeckling model.

Methods	Metrics	Yellow River		Farmland		Ottawa	
		w/o	DM	w/o	DM	w/o	DM
PCKA [3]	OA $\uparrow$	89.80	95.82	88.22	94.44	97.55	98.31
	F1-Score $\uparrow$	72.66	<b>87.72</b>	47.52	<b>65.90</b>	91.93	<b>94.47</b>
NR-ELM [9]	OA $\uparrow$	94.11	95.73	97.86	98.42	98.17	95.82
	F1-Score $\uparrow$	81.59	<b>87.04</b>	78.28	<b>84.96</b>	94.15	84.84
DDNet [22]	OA $\uparrow$	95.35	96.83	98.50	98.87	98.09	98.43
	F1-Score $\uparrow$	86.65	<b>90.79</b>	86.67	<b>89.70</b>	93.90	<b>94.87</b>
LANTNet [20]	OA $\uparrow$	95.61	96.91	98.77	98.84	98.3	98.44
	F1-Score $\uparrow$	87.16	<b>91.1</b>	88.69	<b>89.20</b>	94.46	<b>94.88</b>

## 4.2 Experimental Results & Discussion

To evaluate the effectiveness of the despeckling model, we compared the results of change detection methods (namely PCA- $k$ -means (PCKA) [3], NR-ELM [9], DDNet [22] and LANTNet [20]) with and without the despeckling model using three real SAR datasets. Figures 5, 6 and 7 demonstrate the proposed despeckling model performance on Yellow River, Farmland and Ottawa datasets. DM has considerably enhanced the F1 score for existing (including state-of-the-art)

8 M. Ihmeida and M Shahzad

change detection methods. In all these experiments, we empirically set the  $\lambda_{SSIM}$  to be 5 in the loss objective (3) as a trade-off between despeckling and change detection performance. Table 2 presents the OA and F1 score on three real SAR datasets for four CD methods. However, in Figure 4, the NR-ELM algorithm with despeckling model achieved a lower F1 score because the Ottawa dataset is less affected by speckle noise. This is why we observe a higher F1 score for all other methods without DM. Additionally, compared to other methods, NR-ELM exhibits more resistance to speckle noise due to its built-in despeckling process within its architecture. Therefore, the decrease in the F1 score when incorporating the DM module is attributed to the extra despeckling process, which over-smooths the input image and subsequently reduces the F1 score.

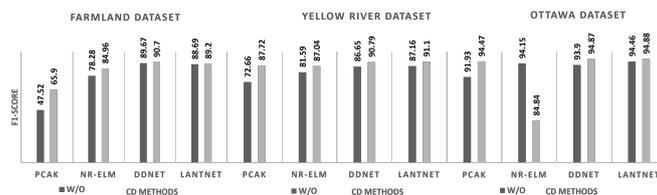


Fig. 4: The correlation between DM and the F1 score for SAR CD datasets

It can be observed that in Yellow River and Farmland datasets, the proposed DM achieves a superior F1 score for CD methods compared to without DM (W/O) results due to the ability to efficiently cope with the single-look pre-change and multi-look post-change SAR images via robust loss function. It should be noted that CD methods without the despeckling model perform well on Ottawa dataset because the dataset is slightly affected by speckle noise. Nevertheless, the performance of CD methods was further improved with the proposed DM as presented in Table 2 and Figure 4.

### 4.3 Hardware & Running Times

The experiments were conducted using three datasets (described in section 4.1) on a Tesla GPU P100 with 16 GB of RAM and 147.15 GB of disk space, resulting in a training duration of approximately 11 hours. The framework used to train the proposed despeckling model was TensorFlow 2.0.

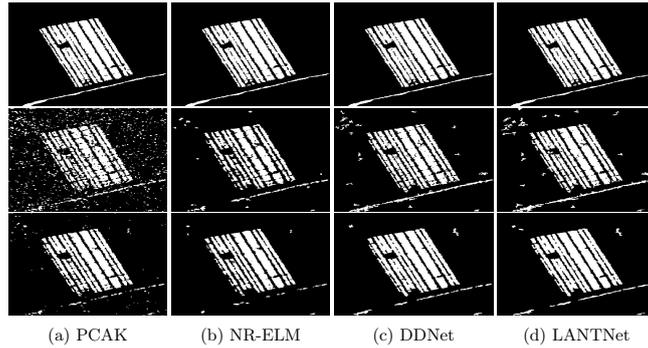


Fig. 5: Change detection results on Yellow River dataset. Rows: (1st row) Yellow River ground truth(GT), (2nd row) CD methods results without despeckling, (3rd row) the CD methods results with the proposed DM. Columns: (a) PCAk [3], (b) NR-ELM [9], (c) DDNet [22], and (d) LANTNet [20].

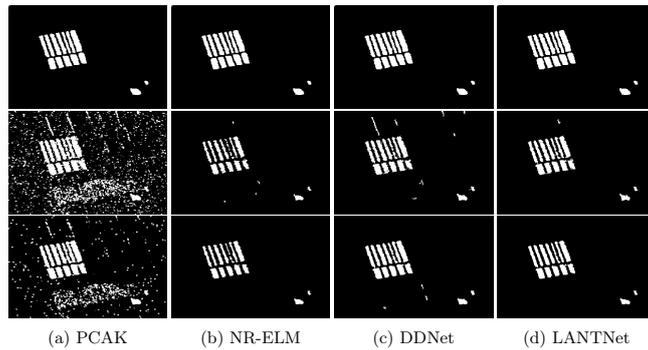


Fig. 6: Change detection results on Farmland dataset. Rows: (1st row) Yellow River ground truth(GT), (2nd row) CD methods results without despeckling, (3rd row) the CD methods results with the proposed DM. Columns: (a) PCAk [3], (b) NR-ELM [9], (c) DDNet [22], and (d) LANTNet [20].

## 5 Conclusion

In recent years, deep-learning architectures have shown promise in improving SAR change detection performance. However, the challenge of speckle noise per-

10 M. Ihmeida and M Shahzad

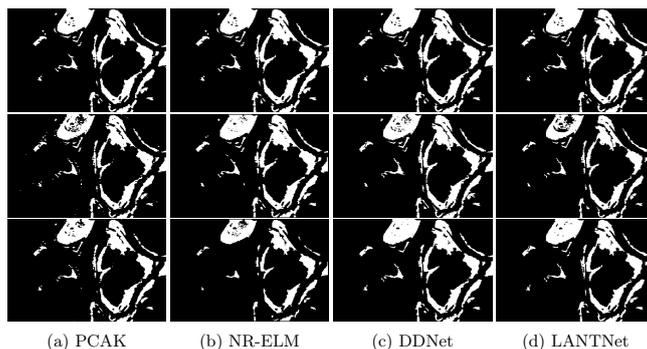


Fig. 7: Change detection results on Ottawa dataset. Rows: (1st row) Yellow River ground truth(GT), (2nd row) CD methods results without despeckling, (3rd row) the CD methods results with the proposed DM. Columns: (a) PCAk [3], (b) NR-ELM [9], (c) DDNet [22], and (d) LANTNet [20].

sists in these methods. To overcome this challenge, we propose a despeckling model that effectively suppresses speckle noise and enhances the performance of existing change detection methods. Extensive evaluations and comparisons with state-of-the-art methods demonstrate the superior performance of our proposed despeckling model. It should be noted that our current approach focuses solely on a single-imaging modality. Future work of this work could explore the domain of multi-modal change detection, incorporating both optical and SAR data.

## References

1. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. *IEEE transactions on pattern analysis and machine intelligence* **33**(5), 898–916 (2010)
2. Bai, X., Zhou, F.: Analysis of new top-hat transformation and the application for infrared dim small target detection. *Pattern Recognition* **43**(6), 2145–2156 (2010)
3. Celik, T.: Unsupervised change detection in satellite images using principal component analysis and  $k$ -means clustering. *IEEE geoscience and remote sensing letters* **6**(4), 772–776 (2009)
4. Chierchia, G., Cozzolino, D., Poggi, G., Verdoliva, L.: Sar image despeckling through convolutional neural networks. In: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 5438–5441. IEEE (2017)
5. Dekker, R.: Speckle filtering in satellite sar change detection imagery. *International Journal of Remote Sensing* **19**(6), 1133–1146 (1998)
6. Fan, R., Hou, B., Liu, J., Yang, J., Hong, Z.: Registration of multiresolution remote sensing images based on l2-siamese model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **14**, 237–248 (2020)

7. Foody, G.M.: Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment* **239**, 111630 (2020)
8. Gao, F., Dong, J., Li, B., Xu, Q.: Automatic change detection in synthetic aperture radar images based on peanet. *IEEE Geoscience and Remote Sensing Letters* **13**(12), 1792–1796 (2016). <https://doi.org/10.1109/LGRS.2016.2611001>
9. Gao, F., Dong, J., Li, B., Xu, Q., Xie, C.: Change detection from synthetic aperture radar images based on neighborhood-based ratio and extreme learning machine. *Journal of Applied Remote Sensing* **10**(4), 046019 (2016)
10. Gao, Y., Gao, F., Dong, J., Du, Q., Li, H.C.: Synthetic aperture radar image change detection via siamese adaptive fusion network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **14**, 10748–10760 (2021)
11. Gao, Y., Gao, F., Dong, J., Li, H.C.: Sar image change detection based on multi-scale capsule network. *IEEE Geoscience and Remote Sensing Letters* **18**(3), 484–488 (2020)
12. Gong, M., Zhou, Z., Ma, J.: Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering. *IEEE Transactions on Image Processing* **21**(4), 2141–2151 (2012). <https://doi.org/10.1109/TIP.2011.2170702>
13. Ihmeida, M., Shahzad, M.: Enhanced change detection performance based on deep despeckling of synthetic aperture radar images. *IEEE Access* (2023)
14. Ihmeida, M., Wei, H.: Image registration techniques and applications: comparative study on remote sensing imagery. In: 2021 14th International Conference on Developments in eSystems Engineering (DeSE). pp. 142–148. IEEE (2021)
15. Intajag, S., Chitwong, S.: Speckle noise estimation with generalized gamma distribution. In: 2006 SICE-ICASE International Joint Conference. pp. 1164–1167. IEEE (2006)
16. Khelifi, L., Mignotte, M.: Deep learning for change detection in remote sensing images: Comprehensive review and meta-analysis. *IEEE Access* **8**, 126385–126400 (2020)
17. Krinidis, S., Chatzis, V.: A robust fuzzy local information c-means clustering algorithm. *IEEE transactions on image processing* **19**(5), 1328–1337 (2010)
18. Li, H.C., Hong, W., Wu, Y.R., Fan, P.Z.: Bayesian wavelet shrinkage with heterogeneity-adaptive threshold for sar image despeckling based on generalized gamma distribution. *IEEE Transactions on Geoscience and Remote Sensing* **51**(4), 2388–2402 (2012)
19. Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A.: Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS journal of photogrammetry and remote sensing* **152**, 166–177 (2019)
20. Meng, D., Gao, F., Dong, J., Du, Q., Li, H.C.: Synthetic aperture radar image change detection via layer attention-based noise-tolerant network. *IEEE Geoscience and Remote Sensing Letters* **19**, 1–5 (2022)
21. Moulton, J., Kassam, S., Ahmad, F., Amin, M., Yemelyanov, K.: Target and change detection in synthetic aperture radar sensing of urban structures. In: 2008 IEEE Radar Conference. pp. 1–6. IEEE (2008)
22. Qu, X., Gao, F., Dong, J., Du, Q., Li, H.C.: Change detection in synthetic aperture radar images using a dual-domain network. *IEEE Geoscience and Remote Sensing Letters* **19**, 1–5 (2021)
23. Radke, R.J., Andra, S., Al-Kofahi, O., Roysam, B.: Image change detection algorithms: a systematic survey. *IEEE transactions on image processing* **14**(3), 294–307 (2005)

12 M. Ihmeida and M Shahzad

24. Tang, X., Zhang, H., Mou, L., Liu, F., Zhang, X., Zhu, X.X., Jiao, L.: An unsupervised remote sensing change detection method based on multiscale graph convolutional network and metric learning. *IEEE Transactions on Geoscience and Remote Sensing* (2021)
25. Ulaby, F., Dobson, M.C., Álvarez-Pérez, J.L.: *Handbook of radar scattering statistics for terrain*. Artech House (2019)
26. Wang, J., Gao, F., Dong, J., Zhang, S., Du, Q.: Change detection from synthetic aperture radar images via graph-based knowledge supplement network. *arXiv preprint arXiv:2201.08954* (2022)
27. Wang, P., Zhang, H., Patel, V.M.: Sar image despeckling using a convolutional neural network. *IEEE Signal Processing Letters* **24**(12), 1763–1767 (2017)
28. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing* **13**(4), 600–612 (2004)
29. Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L.: Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing* **26**(7), 3142–3155 (2017)
30. Zhang, W., Jiao, L., Liu, F., Yang, S., Liu, J.: Adaptive contourlet fusion clustering for sar image change detection. *IEEE Transactions on Image Processing* **31**, 2295–2308 (2022). <https://doi.org/10.1109/TIP.2022.3154922>

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2023.0322000

# Enhanced Change Detection Performance Based on Deep Despeckling of Synthetic Aperture Radar Images

MOHAMED IHMEIDA<sup>1</sup>, MUHAMMAD SHAHZAD<sup>1</sup>

<sup>1</sup>Department of Computer Science, University of Reading, Reading, RG6 6DH, UK (e-mail:m.a.g.ihmeida@pgr.reading.ac.uk, m.shahzad2@reading.ac.uk)

Corresponding author: Mohamed Ihmeida (e-mail: m.a.g.ihmeida@pgr.reading.ac.uk)

## ABSTRACT

Synthetic aperture radar (SAR) image change detection (CD) focuses on identifying changes between two images at different times for the same geographical region. Recently, several deep learning methods have been proposed for performing SAR based CD. However, speckle noise remains a major challenge for these methods. To address this, we propose a despeckling model (DM) that effectively suppresses speckle noise and enhances the performance of the existing CD methods. The proposed despeckling architecture is not only resilient to multi-temporal SAR acquired from one SAR imaging process (i.e., the same number of SAR images looks before and after the change) but also deals with any combination of single or multi-look images acquired prior and after the change. Moreover, as a second contribution, we have also proposed a loss function that effectively suppresses speckle noise, thereby improving the change detection accuracy. Both the despeckling model and the proposed tolerant noise loss function are evaluated extensively on three public real SAR datasets, achieving superior performance compared to existing state-of-the-art SAR CD methods in all datasets.

**INDEX TERMS** Change Detection, Convolutional Neural Network, Despeckling Noise, Synthetic Aperture Radar, Unsupervised Learning.

## I. INTRODUCTION

REMOTE sensing (RS) change detection (CD) aims to identify the change between two multi-temporal images for the same geographical region at different times [1] [2] [3] [4]. It offers valuable information for numerous applications, including deforestation monitoring [2], target detection [5], and agriculture investigation [6]. Moreover, the CD algorithms help to extract vital information to assess the change, especially in case of natural disasters (e.g., earthquakes, floods, droughts, and hurricanes [7] [8]), which in turn supports the local governments to make an effective and timely decision to prevent or mitigate material losses and lives.

In remote sensing, change detection endeavours to distinguish the changed and unchanged pixels of multi-temporal remote sensing images, this is Earth Observation (EO) images acquired for the same geographical region, but at different times [9] [10]. Typically, these multi-temporal images are co-registered (i.e., transformed into the same coordinate system) to obtain consistent radiometric characteristics such as brightness and contrast [11]. This enhances the change detection

performance by aligning the correct position for each pixel in both multi-temporal images prior to feeding them as input to the subsequent change detection process [12] [13] [14] [15]. Most image registration algorithms rely on robust extraction of key points either using shallow extraction methods such as Scale-Invariant Feature Transform (SIFT) [16], Speeded-Up Robust Features (SURF) [17] or deep methods including convolutional neural networks (CNNs) [18], Siamese networks [19], and spatial transformer networks [20].

Once co-registered, the change map (a result of the change detection algorithm) can be easily obtained using classical change detection methods by computing a difference image (DI), simply the intensity difference between the two images. However, change detection in EO is nontrivial owing to inherent challenges such as errors in co-registration, variations in illumination, viewpoint, shadows, atmospheric effects (e.g., presence of clouds, fog, etc.), and varying sensor characteristics. Moreover, surface reflectance from incoherent objects (such as vegetation) can adversely affect the performance of optical CD algorithms.

Synthetic aperture radar (SAR) offers distinct advantages

over optical sensors for CD in EO because it is not affected by weather conditions, provides penetration through clouds and vegetation, and offers sensitivity to small changes, making it capable of detecting changes that may be missed by optical CD methods. This technique allows us to remotely map the reflectivity of objects or environments with high spatial resolution through the emission and reception of electromagnetic signals in the microwave spectrum, which enables ease of penetration through clouds and provides all-weather day/night sensing capability, making it suitable for applications related to disaster assessment (such as flooding and earthquake) [21].

Typically, optical CD methods rely mainly on supervised machine learning approaches [22] [23] [24]. However, owing to the lack of annotated SAR datasets, the majority of SAR CD approaches primarily rely on unsupervised learning [25] [26] [27]. Several methods for unsupervised SAR CD have been proposed in literature. For instance, Celik [28] proposed a simple unsupervised CD method using principal component analysis and  $k$ -means where change detection was achieved by partitioning the feature vector space into two clusters. Krinidis *et al.* [29] proposed fuzzy local information C-means (FLICM) to improve the clustering quality and aim to be robust to noise and preserve the image details. Gong *et al.* [30] also proposed fuzzy c-means (FCM), a reformulated FLICM to cluster image pixels into changed and unchanged. The aforementioned methods are performed under speckle-free images. These approaches perform fairly well. However, SAR data suffer from speckle noise, which arises owing to the coherent nature of SAR imaging, which causes interference patterns in the received signals. This speckle noise makes information extraction from SAR images challenging and, consequently, adversely affects change detection accuracy [31] [32] [33].

Several approaches have been proposed to address speckle noise. For instance, pioneering work in the despeckling of SAR images was proposed by Lee [34]. Later, Lee [35] refined [36] to remove noisy edge boundaries in SAR images by enhancing the edge representation using local statistics (average and variance) within a  $7 \times 7$  window. However, a drawback of this approach is its reliance on a fixed mask size [21]. Kuan *et al.* [37] proposed an adaptive speckle-noise smoothing filter that can handle different noise types without prior knowledge of the original statistics of the image. However, it tends to over smooth image details and has high computational complexity. Lope *et al.* [38] then proposed an Enhanced Lee filter and comprehensively analysed well-known filters by experimenting with varying the local coefficients of despeckled SAR images. Their approach allows the preservation of fine details, such as texture and edge information, in the heterogeneous regions of the observed SAR image. Zhu *et al.* [39] further improved despeckling performance by combining an enhanced Lee filter with a median filter.

In the context of change detection, several recent approaches have tackled the despeckling problem using deep

neural networks. For instance, Zhang *et al.* [40] proposed unsupervised change detection using deep learning methods that employ multi-scale superpixel reconstruction method to suppress the speckle noise and generate a difference image. Subsequently, two-stage centre-constrained fuzzy c-means clustering algorithm is executed to classify the DI pixel into changed, unchanged and intermediate classes. Image patches belonging to changed and unchanged pixels are used as pseudo-label training samples, whereas the image patches belonging to the intermediate class are utilised as testing samples. The final stage is to train a convolutional wavelet neural network on the image patches belonging to changed and unchanged pixels to classify the intermediate classes. Wang *et al.* [41] introduced a sparse model that exploits structural features of changed regions in noisy DIs generated from multi-temporal SAR images. Wenhua *et al.* [42] introduced a multi-objective sparse feature learning mode. In this model, the sparsity of representation is dynamically learned to enhance robustness against various noise levels. The network is further fine-tuned using correctly labelled samples chosen from coarse results, allowing for learning semantic information related to changed and unchanged pixels. Liu *et al.* [43] presented a local restricted CNN for SAR change detection in which the original CNN was improved by incorporating a local spatial constraint. Qu *et al.* [44] also presented a dual domain neural network (DDNet) to obtain features from spatial and frequency domains to minimise the speckle noise. Gao *et al.* [3] proposed a Siamese adaptive fusion network for SAR image change detection to extract high-level semantic features from multi-temporal SAR images and suppress speckle noise. Meng *et al.* [45] proposed a robust loss function and a layer attention-based noise-tolerant network (LANTNet) that benefits from feature correlations among multi-convolutional layers and suppresses the impact of noisy labels.

Although these state-of-the-art deep learning-based approaches provide some robustness against different noise types, they still fail to fully suppress speckle noise, which hinders their effective change detection ability. Moreover, the amount of speckle noise varies between single-look or multi-look SAR imaging processes [6]. This is considered at different times (e.g., single-look at time instance  $t_1$  and multi-look at time instance  $t_2$ ) and consequently further degrades the performance of various change detection algorithms. To this end, in this paper, we propose a robust despeckling architecture that is not only resilient to multi-temporal SAR acquired from one SAR imaging process (i.e., the same number of SAR images looks before and after the change) but also deals with any combination of single or multi-look images acquired prior and after the change. To achieve this, the following are the significant contributions of this study:

- We propose a deep convolutional neural network-based Despeckling Model (DM) that can suppress speckle noise and improve the performance of state-of-the-art SAR CD methods.

- We develop a new speckle noise tolerant loss function, inspired by the works of [45], that is more resistant to speckle noise and significantly improves the baseline change detection accuracy.
- Both the despeckling model and the proposed tolerant noise loss function are evaluated on three public real SAR datasets and achieved superior performance compared with existing state-of-the-art SAR CD methods.

## II. RELATED WORK

SAR change detection has been widely used in many applications such as urban extension [46], agricultural monitoring [47], target detection [48] disaster monitoring [49] and assessment [50]. Typically owing to the lack of annotated SAR datasets, most researchers rely on unsupervised methods [3] [51] [52] [53] to address SAR CD. However, the problem is highly challenging owing to the presence of speckle noise, which negatively impacts SAR images and reduces the change detection accuracy [31] [32] [54]. For this purpose, many researchers have formulated SAR CD in three sequential steps image pre-processing, difference image generation, and classification [55]. The image pre-processing stage includes despeckling (denoising) and image registration. Image despeckling aims to reduce the impact of speckle noise and enhance SAR image quality. However, oversmoothing usually occurs in doing so, which may result in the loss of geometric details. After despeckling, the latter image registration aids in aligning multi-temporal images in the same reference coordinate system, enabling accurate change detection [12] [13]. To generate difference image, various methods have been proposed in the literature, including image differencing (also known as subtracting) [2], log ratio [56], neighbourhood-based ratio [57], Gauss-ratio operator [58] and mean- and log-ratio difference [59]. Finally, the classification of DI typically includes thresholding and clustering [60].

Some approaches use the clustered DI image (preclassification result) to subsequently train a classifier model and then combine the information from the preclassification and classifier results to generate a change map. For instance, Gao *et al.* [61] computed the preclassification result by computing a DI via log-ratio and fuzzy c-means clustering and later trained the PCANet model (classifier) to obtain the initial classification, which was fused with the preclassification results to obtain the final change map. Similarly, Gao *et al.* [62] proposed an approach that employs a neighbourhood-based ratio to generate the difference image and then adopts an extreme learning machine (ELM) to model the high probability pixel based on the difference image, which is later used with the initial change map to yield the final change map. Wang *et al.* [63] employed a semi-supervised Laplacian support vector machine (SVM) to differentiate between changed and unchanged regions. To initialise the SVM, a pseudo-training set is generated using saliency similarity detection. This pseudo-training set consists of labelled changed and unchanged pixels. The Laplacian SVM effectively utilises

the prior information from the available labelled samples and incorporates unlabelled samples to improve its discriminatory capabilities. Lv *et al.* [64] presented feature learning utilising a stacked contractive autoencoder to extract temporal change features from superpixels while effectively suppressing noise. Li *et al.* [65] proposed a Gamma correction and fuzzy local information c-means clustering model to reduce the impact of speckle noise and improve the performance. Liu *et al.* [43] introduced a locally restricted CNN for SAR change detection. They enhanced the original CNN architecture by incorporating a local spatial constraint, thereby improving CD performance.

Recently, a few approaches have aimed to explicitly suppress the inherent speckle noise to improve the SAR CD performance. For example, Qu *et al.* [44] proposed DDNet, a method that leverages features extracted from both the spatial and frequency domains to mitigate the impact of speckle noise. Gao *et al.* [3] also presented a Siamese adaptive fusion network for SAR image change detection, which focused on extracting high-level semantic features from multi-temporal SAR images while effectively suppressing speckle noise. Meng *et al.* [45] introduced a layer attention module that leverages the correlation among multiple convolutional layers and designed a loss function that minimises the influence of speckle noise, thereby enhancing the change detection performance. A limitation of these approaches is their inability to effectively tackle different speckle noises in images prior and after the change, for example, single-look prior image and multi-look post-change image, which makes it difficult for SAR CD methods to perform well due to varying speckle-noise characteristics [6]. In the following, we present a denoising framework that enables us to effectively tackle the SAR CD problem for both the same or different numbers of looks in the pre- and post-change images.

## III. METHODOLOGY

The proposed methodology consists of two modules where the first despeckling modules where the first despeckling module passes the input SAR image through a series of convolutional layers to suppress speckle noise and later feeds the resulting noise-reduced image to the subsequent change detection module. For change detection, we adapt [45], which first performs a preclassification step and then employs a layer attention module that exploits the correlations among the multi-layer convolutions and produces robust cascaded feature representations learned by the network. Furthermore, we propose a noise-tolerant loss function that is resilient to speckle noise and significantly improves baseline change detection accuracy. In the following, we discuss them in detail, where we first present the proposed despeckling architecture, despeckling loss function, and adaptations that we have made to the baseline change detection approach by proposing a noise-resilient loss function.

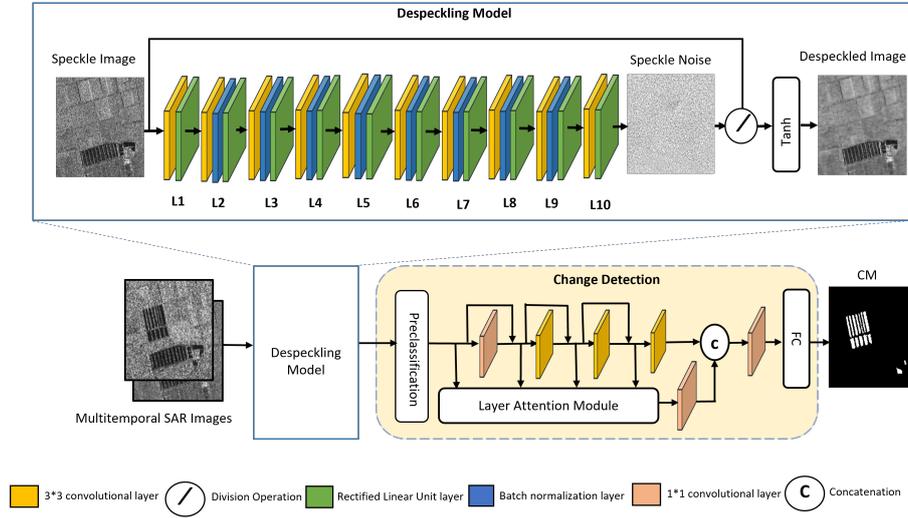


FIGURE 1: Proposed method for despeckling of SAR images for improved change detection task.

#### A. PROPOSED DESPECKLING ARCHITECTURE

The proposed despeckling architecture aims to learn a mapping from the input SAR image through a series of convolutional layers to yield a residual image [66] that contains speckle noise only. The learned residual (i.e., speckle-only in our case) image can then be passed along with the original image through either a subtraction [67] or division [68] operation to produce the resulting despeckled image. However, the division operation is preferable [68] because it avoids an additional logarithmic transformation step and enables end-to-end learning.

Practically, training such a network design requires ground truth or reference despeckled images, which is usually not the case for SAR images. To cope with this, several researchers synthetically generate reference noise-free SAR images using multiplicative noise models [69]. For our purpose, we also rely on synthetically generated SAR reference images and use them to train our proposed despeckling network architecture as depicted in Figure 1. It consists of ten convolutional layers, each incorporating batch normalisation and ReLU activation functions. Each layer has 64 filters with a stride of one, and zero padding is used to ensure that the output of each layer has the same dimensions as the input image, except for the last one, which has only one filter. At the end of the network, a hyperbolic tangent is employed as a nonlinear function [68]. For clarity, we also provide the architecture details of the proposed model along with the hyperparameter details in Table 1.

TABLE 1: Proposed Despeckling Model Configuration, where L1 and L10 refer to a series of Conv-ReLU layers, while the layers between L2 and L9 consist Conv-BN and ReLU layers as illustrated in Figure 1.

-	layer	Filter Size	Filters	Output size
L1	Conv + ReLU	3*3*1	64	256*256*64
L2-L9	Conv + BN + ReLU	3*3*64	64	256 *256* 64
L10	Conv + ReLU	3*3*64	1	256 *256*1

#### B. DESPECKLING LOSS FUNCTION

Let us assume that  $F \in \mathbb{R}^{W \times H}$  denotes the observed SAR image intensity with speckle,  $X \in \mathbb{R}^{W \times H}$  represent the noise-free SAR image, and  $N \in \mathbb{R}^{W \times H}$  represents the multiplicative speckle noise. Then we can describe the relation between the observed and noise-free SAR images as

$$F = N \odot X \quad (1)$$

Where  $\odot$  denotes the Hadamard product (i.e., the element-wise multiplication) between  $N$  and  $X$ . As mentioned earlier,  $X$  is synthetically generated by multiplicative noise using the procedure explained in [69] [70] [71].

One straightforward approach to train the despeckling network with learning parameters  $\theta$  is to simply use the predicted despeckled image and noise-free SAR image to compute the per-pixel Euclidean loss function  $L_E$  as follows:

$$L_E(\theta) = \frac{1}{W \cdot H} \sum_{w=1}^W \sum_{h=1}^H \|X^{(w,h)} - \hat{X}^{(w,h)}\|^2 \quad (2)$$

Where  $X$  is the reference image, and  $\hat{X}$  is the despeckled image.  $W$  and  $H$  represent the width and height of an image. Although this simple Euclidean loss  $L_E$  has been effective

in solving numerous image restoration problems such as super-resolution [72], semantic segmentation [73], change detection [74], and style transfer [75], it usually produces several artifacts (e.g., unwanted distortions such as irregular patterns, pixelation, blurring, or other visual abnormalities) in the resulting estimated image [76]. To address this problem, Wang et al. [68] integrated a supplementary total variation (TV) loss into  $L_E$ , which somewhat removes the artifacts but oversmooths the images, causing loss of information and consequently degrading the performance of change detection [41]. To overcome these issues, we utilised a structural similarity index (SSIM), originally proposed for image quality assessment [77], as an auxiliary to  $L_E$  to achieve a better trade-off performance by removing artifacts while maintaining the necessary information, which improves the change detection performance.

$$SSIM(X, \hat{X}) = \frac{(2\mu_X\mu_{\hat{X}} + C_1) \cdot (2\sigma_{X\hat{X}} + C_2)}{(\mu_X^2 + \mu_{\hat{X}}^2 + C_1) \cdot (\sigma_X^2 + \sigma_{\hat{X}}^2 + C_2)} \quad (3)$$

Where  $X$  and  $\hat{X}$  are the reference (noise-free) and despeckled images, respectively,  $\mu_X$  and  $\mu_{\hat{X}}$  are the mean values of  $X$  and  $\hat{X}$  respectively. Similarly,  $\sigma_X$  and  $\sigma_{\hat{X}}$  are the standard deviations of  $X$  and  $\hat{X}$  respectively. While  $\sigma_{X\hat{X}}$  is the covariance between  $X$  and  $\hat{X}$ . Finally,  $C_1$  and  $C_2$  are constants set to be 0.01 and 0.03 respectively [77].

The total loss is thus calculated as follows:

$$L_T = L_E(\theta) + \lambda_{SSIM} \cdot SSIM \quad (4)$$

Where  $L_T$  is the total loss and  $\lambda_{SSIM}$  represents the weight of the auxiliary SSIM in the loss.

### C. PROPOSED CHANGE DETECTION LOSS FUNCTION

Existing unsupervised change detection methods utilise clustering algorithms such as hierarchical Fuzzy C-Means [78] and Fuzzy C-Means (FCM) [79] to generate pseudo-labels with a high probability for network training. While this method solves the need for label data, errors commonly affect network performance. In addition to this, the attention mechanism is utilised to emphasise the essential parts of the input while disregarding irrelevant information, but it often neglects the correlations among multiple convolution layers. To address this limitation, Meng et al. [45] proposed a layer attention module to weigh features from different layers based on the learned correlation matrix. This module effectively combines spatial information from low-level layers with semantic information from high-level layers, emphasising informative layers and suppressing redundant ones. The process involves matrix multiplication to assign adaptive weights to the input feature groups, followed by calculating the attention matrix using a softmax operation. The weighted feature matrix is then multiplied by the attention matrix, reshaped, and combined with the original input to produce the final output. The change map is generated through a series of convolution and fully connected layers. The trained network

can classify all pixels from the multitemporal SAR images to obtain the final change map. We adapt the training strategy and propose a loss function that is more noise resistant to speckle noise. However, this loss function does not provide satisfactory performance. To this end, we designed a robust loss function that is more resistant to speckle noise. The loss function combines MSE and Kullback-Leibler Divergence (KL). The loss function is expressed as follows:

$$L_{MSE}(X, \hat{X}) = \|X - \hat{X}\|^2 \quad (5)$$

$$L_{KL}(X, \hat{X}) = \hat{X} \cdot (\log \hat{X} - X) \quad (6)$$

$$L_T = \alpha L_{MSE} + \beta L_{KL} \quad (7)$$

where  $\alpha$  and  $\beta$  are two weighting hyperparameters.

In our empirical study,  $\alpha$  and  $\beta$  were set to 0.9 and 0.1 to trade-off noise robustness and convergence efficiency. The KL acts similarly to CE with the difference that CE penalises the network based on its predictions, whereas KL mainly evaluates the disparity between the probability distribution predicted by the network and the distribution of the reference ground truth. Therefore, we argue that combining MSE and KL can provide a better change detection performance and suppress speckle noise (see Section IV-C2). In the following section, we present the results of our proposed methodology along with the training details.

## IV. EXPERIMENTAL RESULTS & EVALUATION

In this section, we first introduced the datasets and employed evaluation metrics. Subsequently, we investigated the effectiveness of the proposed despeckling model coupled with the CD loss function to improve the change detection accuracy. Finally, the results were presented and evaluated by comparing them with those of state-of-the-art methods.

### A. DATASETS

Two types of datasets were used in this paper. The first is the Berkeley Segmentation Dataset 500, widely employed to generate synthetic SAR images. In addition, real SAR images (for the purpose of change detection purpose) were employed to assess the model's performance. Both datasets are described in detail in the following subsections:

#### 1) Synthetic SAR Images

The Berkeley Segmentation Dataset 500 (BSD-500) was originally developed to evaluate the segmentation of natural edges, including object contours, object interior and background boundaries [80]. It included 500 natural images with carefully manually annotated boundaries and edges of natural objects collected from multiple users. This dataset has been widely used to generate synthetic SAR images for the purpose of despeckling [69] [70] [71]. Inspired by these studies, we have used it to train our despeckling model.

## 2) Real SAR Images

For the purpose of change detection, we employed three real SAR image datasets that are multi-temporal and have been co-registered and corrected geometrically.

- *Farmland and Yellow River Datasets*: The images for both datasets were captured by RADARSAT-2 in the region of the Yellow River Estuary in China on 18th June 2008 (pre-change) and 19th June 2009 (post-change). The pre-change images are single-look, whereas the post-change images have been acquired via a multi-looks (four) imaging process. The single-look pre-change image is significantly influenced by speckle noise compared to the four-look post-change image [3]. The disparity between the single and four looks in these two SAR datasets poses a significant challenge for change detection methods.
- *Ottawa Dataset*: The images for this dataset were also captured by RADARSAT-2 in May 1997 (pre-change) and August 1997 (post-change) in the areas affected by floods [44] [53] [81]. Because of the single imaging process, both the pre- and post-change images are less affected by noise in this dataset.

As mentioned above, synthetic SAR images were utilised to train the proposed DM, as depicted in Figure 1. While the real SAR images were despeckled for the purpose of change detection.

## B. EVALUATION METRICS

Quantitative evaluation indices, including precision (P), recall (R), overall accuracy (OA) and F1 score (F1) [82] [83] [84] were used in this study to evaluate the change detection process. These metrics were computed as follows:

$$R = \frac{TP}{(TP + FN)} \quad (8)$$

$$P = \frac{TP}{(TP + FP)} \quad (9)$$

$$OA = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (10)$$

$$F1 = \frac{(2 \cdot P \cdot R)}{(P + R)} \quad (11)$$

Here  $TP$ ,  $FP$ ,  $TN$ , and  $FN$  represent the true positives, false positives, true negatives, and false negatives, respectively.

A higher P value indicates a decrease in the occurrence of false alarms, while a greater R value indicates a reduced rate of incorrect detections. OA measures the proportion of accurately detected pixels in the image. However, relying solely on these three metrics can lead to overestimating the outcome when the number of altered pixels is only a small part of the entire image. The F1 score is used to address this, which considers the limitations of P and R, providing a more comprehensive evaluation of performance. It is important to note that larger F1 values indicate better overall performance [85].

## C. ABLATION STUDY

In this section, we initially investigate the performance of the proposed DM on F1 score using three real SAR CD datasets that are discussed in Section IV-A2. We then feed the despeckled SAR images by the proposed DM to five change detection methods PCA- $k$ -means (PCAK) [28], NR-ELM [62], DDNet [44], LANTNet [45] and the proposed CD method where the DDNet and LANTNet are the current state-of-the-art CD methods. Furthermore, we investigate the performance of the proposed CD loss function on F1 score by comparing it with different loss functions.

### 1) Performance Investigation of Despeckling Model

To validate the effectiveness of the despeckling model, we compared the results of change detection methods with and without the despeckling model using three real SAR datasets. Figure 2, 3 and 4 demonstrates that the proposed despeckling model considerably enhanced the F1 score for existing (including state-of-the-art) change detection methods. In all these experiments, we empirically set the  $\lambda_{SSIM}$  to be 5 in the loss objective (4) as a tradeoff between despeckling and change detection performance. It is evident that the performance of the CD methods improves once we passed them through the proposed despeckling model in three SAR datasets. However, in Figure 4, the NR-ELM algorithm with DM obtained a lower F1 because Ottawa dataset is less affected by the speckle noise. This is why we see a higher F1 score even with all other methods without DM. Secondly, Compared to other methods, the NR-ELM is more resistant to speckle noise because of the inherent despeckling process encoded within its architecture. Therefore, the decline in the F1 score when we include the DM module, is due to the fact that an additional despeckling process oversmooths the input image, which subsequently decreases the F1 score. These results will be explained in more detail in section IV-D.

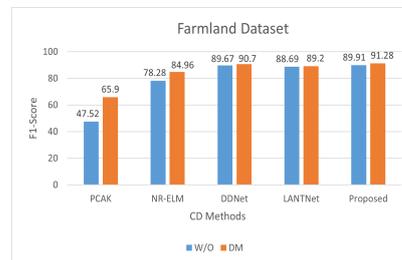


FIGURE 2: Relationship between DM and F1 score for Farmland dataset

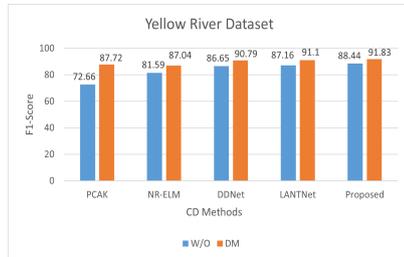


FIGURE 3: Relationship between DM and F1 score for Yellow River dataset

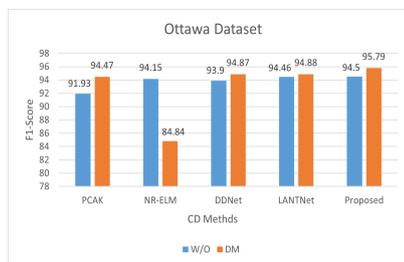


FIGURE 4: Relationship between DM and F1 score for Ottawa dataset

2) Performance Investigation of Proposed CD Loss Function  
 Furthermore, we compared various loss functions and analysed their performance over the baseline change detection methods. Table 2 shows that the loss function combining MSE and KL delivers the best performance, indicating its greater resilience to speckle noise.

TABLE 2: Relationship between loss functions and F1-Score.

Loss function	F1-Score		
	Farmland	Yellow River	Ottawa
MAE	89.1	87.24	94.24
MAE+CE [45]	88.69	87.16	94.46
MSE	86.91	86.22	94.75
MSE+CE(our)	89.60	88.23	94.54
MSE+KL(our)	<b>89.91</b>	<b>88.44</b>	<b>95.35</b>

#### D. EXPERIMENTAL RESULTS & DISCUSSION

To evaluate the impact of the proposed despeckling model on change detection purpose, we compare the effectiveness of the proposed DM with other existing despeckling methods such as Lee [34], Enhanced Lee [38], SAR2SAR [86] and IDCNN [68] on three real SAR datasets. Subsequently, we feed the despeckled SAR images to four aforementioned change detection methods, namely PCA-*k*-means (PCAK) [28], NR-ELM [62], DDNet [44] and LANTNet [45]. PCAK employs principle component analysis for feature extraction and utilises the *k*-means clustering algorithm for classification.

NR-ELM incorporates the neighbourhood ratio for feature extraction using the difference image, followed by classification using an extreme learning machine. DDNet is a dual-domain network that exploits spatial and frequency domain features to mitigate speckle noise. LANTNet is a layer attention-based noise-tolerant network that leverages the correlation between convolutional layers. Both DDNet and LANTNet are currently state-of-the-art change detection methods.

Figures 5, 6 and 7 present the visual results of the change maps obtained from the aforementioned change detection methods using various despeckling filters on Farmland, Yellow River and Ottawa datasets respectively. The corresponding quantitative evaluations are provided in Tables 3, 4, and 5. In the tables, the w/o means it is the original method without despeckling. The DM is our proposed despeckling model, while the Proposed in Methods column refers to the baseline CD with the proposed objective loss function. Figures 5, 6 and 7 are organised as following; the first two columns represent Farmland dataset at time T1 and time T2, while column three represents the reference change detection images used as ground truth (GT). Columns four to eight display the results of PCAK, NR-ELM, DDNet, LANTNet and the proposed method, respectively. The row one presents the results of the above-mentioned methods without despeckling, while rows two to six represent the despeckling method results with Lee [34], Enhanced Lee [38], SAR2SAR [86], IDCNN [68] and DM (our) respectively. In the following section, we discuss the details of the achieved results for individual datasets.

##### 1) Results of Farmland dataset

From Figure 5, it can be observed that the change map generated by PCAK misclassifies many unchanged pixels compared to GT. The Enhanced Lee filter significantly improves the results for PCAK, increasing the accuracy from 47.44% to 79.44%, while the proposed DM achieves 65.90%. It is worth mentioning that, Farmland dataset is heavily influenced by speckle noise, and change detection algorithms usually perform poorly compared to Ottawa dataset, which is less affected by speckle noise. Simply applying PCAK, which is a simple CD method, without despeckling, results in poor performance, as shown in Table 3. Another reason for this poor performance, in addition to the speckle noise, is because the pre- and post-change images in Farmland dataset are different looks, i.e., single and multi-looks before and after the change with varying noise levels. Using despeckling process somewhat takes this into account and improves the performance, as seen in Table 3, where all despeckling methods consistently improve the results with PCAK. Specifically, the Enhanced Lee performs the best here because it is well suited for stronger speckle noise and helps PCAK to significantly smooth the image, while DM is designed to support and generically enhance the overall CD performance. NR-ELM produces better results with less noise but misses some changed pixels. The DM filter improves NR-ELM's performance from 78.28% to 84.96%.

Furthermore, DDNet performed better than PCAK and

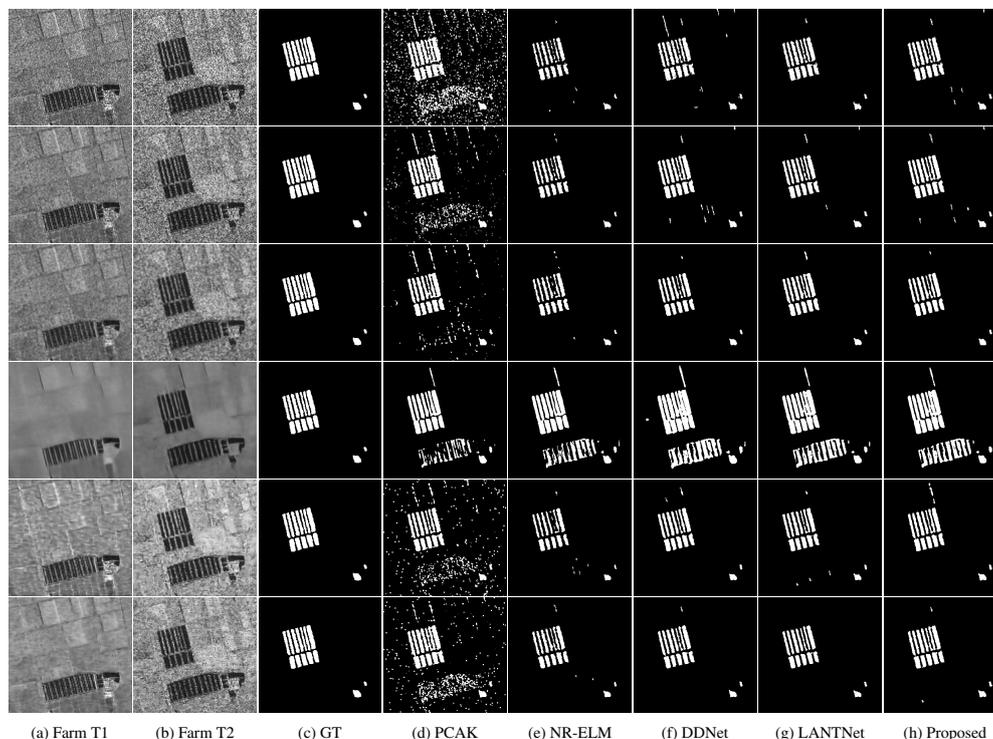


FIGURE 5: Visualised results of Farmland dataset with different despeckling methods. Rows: (1st row) Farmland without despeckling (w/o), (2nd row) Farmland despeckled with lee, (3rd row) Farmland despeckled with enhanced lee, (4th row) Farmland despeckled with SAR2SAR, (5th row) Farmland despeckled with IDCNN, (6th row) Farmland despeckled with proposed DM. Columns: Farmland image captured at (a) t1 and (b) t2. (c) refers to the ground truth (GT) image. Results obtained by methods (d) PCAK [28], (e) NR-ELM [62], (f) DDNet [44], (g) LANTNet [45], and (h) Proposed loss objective function.

NR-ELM did. The DM enhances the F1 score for DDNet from 86.67% to 89.70%, i.e., it demonstrates higher accuracy than PCAK and NR-ELM, although slightly lower than DDNet, while DM improves the accuracy of LANTNet from 88.69% to 89.20%. The proposed method improved performance after incorporating the DM module, increasing accuracy from 89.91% to 91.28%. Notably, the despeckled data using the SAR2SAR filter performed poorly and yielded lower results than the original methods without the despeckling model. It is evident that the DM outperforms other despeckling methods in terms of the F1 score for the purpose of change detection. Moreover, it consistently outperforms other change detection methods without a DM. It is primarily due to the fact that the proposed loss function is more resistant to speckle noise. In other words, the DM suppresses speckle noise even when two Farmland image pairs have different looks, such as single-look (pre-change) and four-look (post-change). This type of suppression is reflected positively in the performance of the change detection methods as shown in Table 3.

## 2) Results of the Yellow River dataset

In Figure 6, it is noticeable that the change map generated by PCAK misclassifies many unchanged pixels as changed ones compared with the GT. The Lee filter reduces speckle noise and improves the CM. The DM performs as the best filter, effectively suppressing noise and significantly improving the F1 score from 72.66% to 87.7% for the PCAK method. NR-ELM produces better results with less noise but misses some changed pixels, whereas the DM filter enhances NR-ELM's performance from 81.59% to 87.04%.

Furthermore, DDNet outperformed PCAK and NR-ELM results. The DM considerably enhance the F1 score from DDNet from 86.65% to 90.79%. LANTNet achieves higher accuracy than PCAK and NR-ELM. DM has enhanced the F1-score for LANTNet from 88.44% to 91.1%. After applying the proposed DM, the proposed method's performance has improved from 88.44% to 91.83%. Finally, the proposed method consistently outperforms all other change detection methods even without despeckling. With DM filtering, the

TABLE 3: Quantitative evaluation on Farmland change detection based on different despeckling filters. Here w/o means it is the original method without despeckling, DM is our proposed despeckling model while the Proposed in Methods column refers to the baseline CD with the proposed objective loss function.

Methods	Metrics	w/o	Lee [34]	Enhanced Lee [38]	SAR2SAR [86]	ID-CNN [68]	DM
PCAK [28]	Recall $\uparrow$	90.04	90.87	67.32	95.69	85.45	90.76
	Precision $\uparrow$	32.27	57.73	96.89	66.51	54.35	51.74
	OA $\uparrow$	88.22	95.53	97.94	96.89	94.89	94.44
	F1-Score $\uparrow$	47.52	70.60	<b>79.44</b>	78.48	66.44	65.90
NR-ELM [62]	Recall $\uparrow$	65.20	68.82	66.52	97.50	66.64	75.39
	Precision $\uparrow$	97.92	99.13	98.51	57.17	97.23	97.33
	OA $\uparrow$	97.86	98.12	97.96	95.52	97.91	98.42
	F1-Score $\uparrow$	78.28	81.24	79.42	72.08	79.08	<b>84.96</b>
DDNet [44]	Recall $\uparrow$	82.26	86.58	78.25	99.26	81.52	82.81
	Precision $\uparrow$	91.59	92.76	98.21	48.66	97.57	97.85
	OA $\uparrow$	98.50	98.81	98.63	93.76	98.79	98.87
	F1-Score $\uparrow$	86.67	89.57	87.11	65.30	88.82	<b>89.70</b>
LANTNet [45]	Recall $\uparrow$	81.35	80.51	81.76	98.46	79.87	81.18
	Precision $\uparrow$	97.50	96.14	96.27	52.74	96.98	98.98
	OA $\uparrow$	98.77	98.66	98.73	94.69	98.66	98.84
	F1-Score $\uparrow$	88.69	87.64	88.42	68.69	87.60	<b>89.20</b>
Proposed	Recall $\uparrow$	84.08	85.48	79.09	97.59	97.59	86.45
	Precision $\uparrow$	97.50	93.78	97.22	53.08	53.08	96.67
	OA $\uparrow$	98.89	98.81	98.63	94.75	94.75	99.02
	F1-Score $\uparrow$	<b>89.91</b>	89.44	87.22	68.76	68.76	<b>91.28</b>

results even further improve.

It is worth mentioning that the despeckled data using the SAR2SAR filter does not perform well and yields lower results compared to the original methods without despeckling, such as DDNet and LANTNet. It is evident that DM achieves a superior F1 score for change detection methods compared to other despeckling methods due to the ability to efficiently cope with the single-look pre-change and multi-look post-change SAR images via robust loss function.

### 3) Results of the Ottawa dataset

Compared to previous datasets, the Ottawa dataset is less affected by speckle noise. This is evident from the achieved better change detection results of 91.93% using the PCAK method without any despeckling process on Ottawa dataset compared to the previous two datasets. Including the proposed DM further improves the F1 score value from 91.93% to 94.47%. NR-ELM provides better results compared to PCAK, Lee slightly improves the F1 score from 94.15% to 94.77%, whereas DM reduces the performance to 84.84% as shown in Figure 7 and Table 5. The proposed DM with the NR-ELM degrades the performance because of oversmoothing. This is because NR-ELM has an inherent despeckling process encoded within its architecture. Moreover, this is also the case for other despeckling methods except the Lee method, which does not degrade (but slightly improve) the performance. A possible reason for this could be because, in comparison, Lee [34] is the least strong despeckling method and therefore does not result in much oversmoothing, which degrades the performance.

DDNet performed better than PCAK and NR-ELM, and the proposed DM improves the F1 score for DDNet from 93.90% to 94.87%. LANTNet produces better accuracy than

PCAK, NR-ELM and DDNet. Its accuracy has further improved by the proposed DM from 94.46% to 94.88%. With the proposed loss objective, the performance slightly improves from 94.46% to 94.50%, which is further enhanced from 94.50% to 95.79% when used in conjunction with the DM as shown in Figure 7 and Table 3. It can be observed from the Ottawa dataset results that the CD methods without despeckling already perform well because the data is less affected by noise. Nevertheless, with DM, the performance of these CD methods was further improved.

### E. TRAINING SETUP

All the experiments were conducted on three data sets detailed in section IV-A where Python 3.7 with OpenCV version 3.4.2.17 was used. The hardware specifications include a Tesla GPU P100-PCIE-16 GB RAM 147.15 GB Disk.

### V. CONCLUSION & OUTLOOK

In recent years, many deep-learning architectures have been employed for SAR change detection problems, leading to enhancements in the change detection performance. However, speckle noise remains a major challenge for these methods. To address this, we propose which are two-fold: 1) First, we have proposed a despeckling model which effectively suppresses the speckle noise and enhances the performance of existing CD methods; 2) Secondly, we have proposed a robust loss function that is able to take the performance of CD methods even further. The proposed solutions have been extensively examined and compared to the state-of-art SAR change detection methods. The achieved results with the proposed despeckling model and the noise tolerant loss function demonstrate superior performance compared to the current change detection methods. The proposed approach

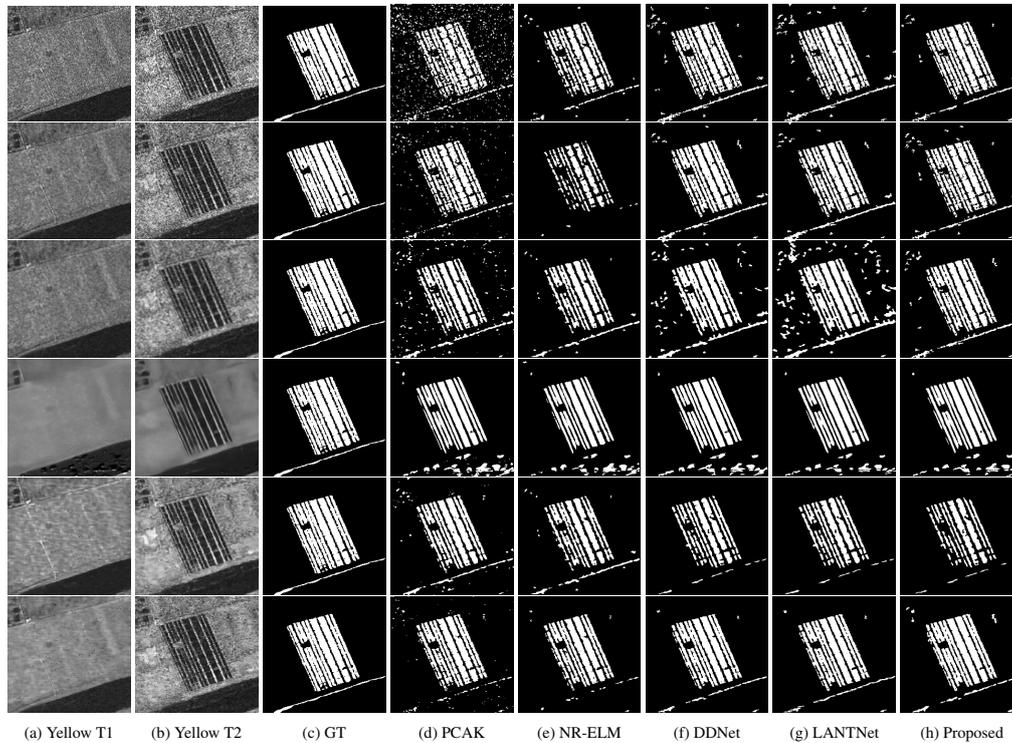


FIGURE 6: Visualised results of Yellow River dataset with different despeckling methods. Rows: (1st row) Farmland without despeckling (w/o), (2nd row) Farmland despeckled with lee, (3rd row) Farmland despeckled with enhanced lee, (4th row) Farmland despeckled with SAR2SAR, (5th row) Farmland despeckled with IDCNN, (6th row) Farmland despeckled with proposed DM. Columns: Farmland image captured at (a) t1 and (b) t2. (c) refers to the ground truth (GT) image. Results obtained by methods (d) PCAK [28], (e) NR-ELM [62], (f) DDNet [44], (g) LANTNet [45], and (h) Proposed loss objective function.

TABLE 4: Quantitative evaluation on Yellow River change detection based on different despeckling filters.

Methods	Metrics	w/o	Lee [34]	Enhanced Lee [38]	SAR2SAR [86]	ID-CNN [68]	DM
PCAK [28]	Recall $\uparrow$	74.96	78.40	74.52	81.80	81.37	82.59
	Precision $\uparrow$	70.50	87.80	82.75	83.31	92.79	93.53
	OA $\uparrow$	89.80	94.12	92.58	93.74	95.49	95.82
	F1-Score $\uparrow$	72.66	82.84	78.42	82.55	86.70	<b>87.72</b>
NR-ELM [62]	Recall $\uparrow$	72.18	48.35	70.19	78.30	79.76	79.32
	Precision $\uparrow$	93.83	99.72	92.22	85.53	95.08	96.42
	OA $\uparrow$	94.11	90.63	93.54	93.68	95.59	95.73
	F1-Score $\uparrow$	81.59	65.13	79.71	81.76	86.75	<b>87.04</b>
DDNet [44]	Recall $\uparrow$	83.46	86.32	82.86	80.46	64.06	86.58
	Precision $\uparrow$	90.09	91.41	81.89	85.00	90.40	95.44
	OA $\uparrow$	95.35	96.06	93.59	93.90	93.43	96.83
	F1-Score $\uparrow$	86.65	88.79	82.37	82.67	77.91	<b>90.79</b>
LANTNet [45]	Recall $\uparrow$	82.44	84.00	83.03	79.84	65.93	87.51
	Precision $\uparrow$	92.45	91.18	71.49	87.83	99.04	94.99
	OA $\uparrow$	95.61	95.64	90.94	94.35	93.72	96.91
	F1-Score $\uparrow$	87.16	87.44	76.83	83.64	79.16	<b>91.1</b>
Proposed	Recall $\uparrow$	84.08	85.93	81.51	79.08	61.89	89.53
	Precision $\uparrow$	93.28	87.05	82.01	86.44	97.75	94.25
	OA $\uparrow$	96.03	95.14	93.42	93.97	92.85	97.12
	F1-Score $\uparrow$	<b>88.44</b>	86.49	81.76	82.59	75.79	<b>91.83</b>

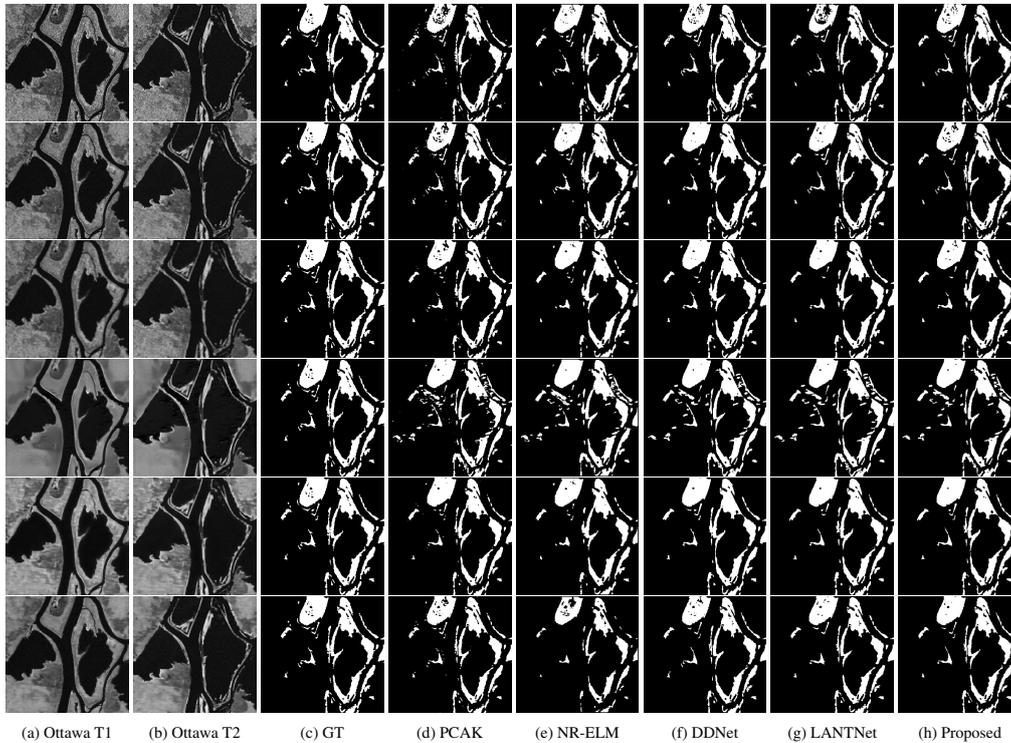


FIGURE 7: Visualised results of Ottawa dataset with different despeckling methods. Rows: (1st row) Farmland without despeckling (w/o), (2nd row) Farmland despeckled with lee, (3rd row) Farmland despeckled with enhanced lee, (4th row) Farmland despeckled with SAR2SAR, (5th row) Farmland despeckled with IDCNN, (6th row) Farmland despeckled with proposed DM. Columns: Farmland image captured at (a) t1 and (b) t2. (c) refers to the ground truth (GT) image. Results obtained by methods (d) PCAK [28], (e) NR-ELM [62], (f) DDNet [44], (g) LANTNet [45], and (h) Proposed loss objective function.

TABLE 5: Quantitative evaluation on Ottawa change detection based on different despeckling filters

Methods	Metrics	w/o	Lee [34]	Enhanced Lee [38]	SAR2SAR [86]	ID-CNN [68]	DM
PCAK [28]	Recall $\uparrow$	88.16	91.58	88.74	88.85	92.01	91.00
	Precision $\uparrow$	96.05	96.28	97.74	82.14	96.85	98.20
	OA $\uparrow$	97.55	98.11	97.89	95.18	98.26	98.31
	F1-Score $\uparrow$	91.93	93.87	93.02	85.36	94.37	<b>94.47</b>
NR-ELM [62]	Recall $\uparrow$	93.14	94.79	87.67	88.68	92.07	73.92
	Precision $\uparrow$	95.19	94.74	94.59	80.56	92.65	99.53
	OA $\uparrow$	98.17	98.34	97.25	94.82	97.59	95.82
	F1-Score $\uparrow$	94.15	<b>94.77</b>	91.00	84.42	92.36	84.84
DDNet [44]	Recall $\uparrow$	92.70	93.66	93.66	90.78	94.51	91.71
	Precision $\uparrow$	95.12	96.06	96.06	82.91	94.73	98.26
	OA $\uparrow$	98.09	98.39	98.39	95.58	98.30	98.43
	F1-Score $\uparrow$	93.90	94.84	94.85	86.67	94.62	<b>94.87</b>
LANTNet [45]	Recall $\uparrow$	91.8	94.67	90.73	89.91	92.62	91.66
	Precision $\uparrow$	97.30	94.48	95.11	82.49	95.23	98.33
	OA $\uparrow$	98.3	98.28	97.80	95.39	98.1	98.44
	F1-Score $\uparrow$	94.46	94.57	92.87	86.04	93.90	<b>94.88</b>
Proposed	Recall $\uparrow$	91.70	92.74	88.80	91.88	93.11	93.73
	Precision $\uparrow$	97.47	97.47	95.96	87.06	95.35	97.96
	OA $\uparrow$	98.31	98.47	97.64	98.07	98.19	96.56
	F1-Score $\uparrow$	<b>94.50</b>	95.04	92.24	89.41	94.22	<b>95.79</b>

so far only focuses on single-imaging modality. In future, an extension of the work could be in the domain of multi-modal (optical and SAR) change detection.

## REFERENCES

- [1] G. Quin, B. Pinel-Puysségur, J.-M. Nicolas, and P. Loreaux, "Mimoso: An automatic change detection method for sar time series," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5349–5363, 2013.
- [2] L. Khelifi and M. Mignotte, "Deep learning for change detection in remote sensing images: Comprehensive review and meta-analysis," *IEEE Access*, vol. 8, pp. 126385–126400, 2020.
- [3] Y. Gao, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Synthetic aperture radar image change detection via siamese adaptive fusion network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 10748–10760, 2021.
- [4] C. Chen, H. Ma, G. Yao, N. Lv, H. Yang, C. Li, and S. Wan, "Remote sensing image augmentation based on text description for waterside change detection," *Remote Sensing*, vol. 13, no. 10, p. 1894, 2021.
- [5] X. Bai and F. Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognition*, vol. 43, no. 6, pp. 2145–2156, 2010.
- [6] W. Shi, M. Zhang, R. Zhang, S. Chen, and Z. Zhan, "Change detection based on artificial intelligence: State-of-the-art and challenges," *Remote Sensing*, vol. 12, no. 10, p. 1688, 2020.
- [7] Z. Zheng, Y. Zhong, J. Wang, A. Ma, and L. Zhang, "Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters," *Remote Sensing of Environment*, vol. 265, p. 112636, 2021.
- [8] D. Qin, X. Zhou, W. Zhou, G. Huang, Y. Ren, B. Horan, J. He, and N. Kito, "Msim: A change detection framework for damage assessment in natural disasters," *Expert Systems with Applications*, vol. 97, pp. 372–383, 2018.
- [9] Q. Shi, M. Liu, S. Li, X. Liu, F. Wang, and L. Zhang, "A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection," *IEEE transactions on geoscience and remote sensing*, vol. 60, pp. 1–16, 2021.
- [10] X. Zhang, S. Cheng, L. Wang, and H. Li, "Asymmetric cross-attention hierarchical network based on cnn and transformer for bitemporal remote sensing images change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–15, 2023.
- [11] H. Noh, J. Ju, M. Seo, J. Park, and D.-G. Choi, "Unsupervised change detection based on image reconstruction loss," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1352–1361, 2022.
- [12] M. Ihmeida and H. Wei, "Image registration techniques and applications: comparative study on remote sensing imagery," in *2021 14th International Conference on Developments in eSystems Engineering (DeSE)*, pp. 142–148, IEEE, 2021.
- [13] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS journal of photogrammetry and remote sensing*, vol. 152, pp. 166–177, 2019.
- [14] J. Zhang, W. Ma, Y. Wu, and L. Jiao, "Multimodal remote sensing image registration based on image transfer and local features," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 8, pp. 1210–1214, 2019.
- [15] X. Dai and S. Khorram, "The effects of image misregistration on the accuracy of remotely sensed change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 5, pp. 1566–1577, 1998.
- [16] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [17] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*, pp. 404–417, Springer, 2006.
- [18] H. Zhang, W. Ni, W. Yan, D. Xiang, J. Wu, X. Yang, and H. Bian, "Registration of multimodal remote sensing image based on deep fully convolutional neural network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 3028–3042, 2019.
- [19] R. Fan, B. Hou, J. Liu, J. Yang, and Z. Hong, "Registration of multiresolution remote sensing images based on 12-siamese model," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 237–248, 2020.
- [20] L. Li, L. Han, M. Ding, and H. Cao, "Multimodal image fusion framework for end-to-end remote sensing image registration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–14, 2023.
- [21] P. Singh, M. Diwakar, A. Shankar, R. Shree, and M. Kumar, "A review on sar image and its despeckling," *Archives of Computational Methods in Engineering*, vol. 28, pp. 4633–4653, 2021.
- [22] L. Xu, W. Jing, H. Song, and G. Chen, "High-resolution remote sensing image change detection combined with pixel-level and object-level," *IEEE Access*, vol. 7, pp. 78909–78918, 2019.
- [23] G. M. Gandhi, B. Parthiban, N. Thummalu, and A. Christy, "Ndvi: Vegetation change detection using remote sensing and gis—a case study of vellore district," *Procedia computer science*, vol. 57, pp. 1199–1210, 2015.
- [24] H. Chen and Z. Shi, "A spatial-temporal attention-based method and a new dataset for remote sensing image change detection," *Remote Sensing*, vol. 12, no. 10, p. 1662, 2020.
- [25] F. Gao, X. Wang, Y. Gao, J. Dong, and S. Wang, "Sea ice change detection in sar images based on convolutional-wavelet neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 8, pp. 1240–1244, 2019.
- [26] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized gaussian model to automatic change detection in multitemporal sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 874–887, 2005.
- [27] H.-C. Li, T. Celik, N. Longbotham, and W. J. Emery, "Gabor feature based unsupervised change detection of multitemporal sar images based on two-level clustering," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 12, pp. 2458–2462, 2015.
- [28] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and  $k$ -means clustering," *IEEE geoscience and remote sensing letters*, vol. 6, no. 4, pp. 772–776, 2009.
- [29] S. Krinidis and V. Chatzis, "A robust fuzzy local information c-means clustering algorithm," *IEEE transactions on image processing*, vol. 19, no. 5, pp. 1328–1337, 2010.
- [30] M. Gong, Z. Zhou, and J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2141–2151, 2012.
- [31] H. Dong, W. Ma, Y. Wu, M. Gong, and L. Jiao, "Local descriptor learning for change detection in synthetic aperture radar images via convolutional neural networks," *IEEE Access*, vol. 7, pp. 15389–15403, 2019.
- [32] N. Lv, C. Chen, T. Qiu, and A. K. Sangaiah, "Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in sar images," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 12, pp. 5530–5538, 2018.
- [33] S. Saha, M. Shahzad, P. Ebel, and X. X. Zhu, "Supervised change detection using prechange optical-sar and postchange sar data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 8170–8178, 2022.
- [34] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE transactions on pattern analysis and machine intelligence*, no. 2, pp. 165–168, 1980.
- [35] J.-S. Lee, "Speckle analysis and smoothing of synthetic aperture radar images," *Computer graphics and image processing*, vol. 17, no. 1, pp. 24–32, 1981.
- [36] J.-S. Lee, "Refined filtering of image noise using local statistics," *Computer graphics and image processing*, vol. 15, no. 4, pp. 380–389, 1981.
- [37] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE transactions on pattern analysis and machine intelligence*, no. 2, pp. 165–177, 1985.
- [38] A. Lopes, R. Touzi, and E. Nezry, "Adaptive speckle filters and scene heterogeneity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 6, pp. 992–1000, 1990.
- [39] J. Zhu, J. Wen, and Y. Zhang, "A new algorithm for sar image despeckling using an enhanced lee filter and median filter," in *2013 6th International congress on image and signal processing (CISP)*, vol. 1, pp. 224–228, IEEE, 2013.
- [40] X. Zhang, H. Su, C. Zhang, X. Gu, X. Tan, and P. M. Atkinson, "Robust unsupervised small area change detection from sar imagery using deep learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 173, pp. 79–94, 2021.
- [41] R. Wang, J.-W. Chen, L. Jiao, and M. Wang, "How can despeckling and structural features benefit to change detection on bitemporal sar images?," *Remote Sensing*, vol. 11, no. 4, p. 421, 2019.

- [42] W. Zhang, L. Jiao, F. Liu, S. Yang, W. Song, and J. Liu, "Sparse feature clustering network for unsupervised sar image change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.
- [43] F. Liu, L. Jiao, X. Tang, S. Yang, W. Ma, and B. Hou, "Local restricted convolutional neural network for change detection in polarimetric sar images," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 3, pp. 818–833, 2019.
- [44] X. Qu, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Change detection in synthetic aperture radar images using a dual-domain network," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.
- [45] D. Meng, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Synthetic aperture radar image change detection via layer attention-based noise-tolerant network," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.
- [46] Q. Shi, M. Liu, S. Li, X. Liu, F. Wang, and L. Zhang, "A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2022.
- [47] R. E. Kennedy, P. A. Townsend, J. E. Gross, W. B. Cohen, P. Bolstad, Y. Wang, and P. Adams, "Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects," *Remote sensing of environment*, vol. 113, no. 7, pp. 1382–1396, 2009.
- [48] F. Hao, Z.-F. Ma, H.-P. Tian, H. Wang, and D. Wu, "Semi-supervised label propagation for multi-source remote sensing image change detection," *Computers & Geosciences*, vol. 170, p. 105249, 2023.
- [49] E. Hamidi, B. G. Peter, D. F. Muñoz, H. Moftakhari, and H. Moradkhani, "Fast flood extent monitoring with sar change detection using google earth engine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–19, 2023.
- [50] F. Wang and Y. J. Xu, "Comparison of remote sensing change detection techniques for assessing hurricane damage to forests," *Environmental monitoring and assessment*, vol. 162, pp. 311–326, 2010.
- [51] Y. Li, C. Peng, Y. Chen, L. Jiao, L. Zhou, and R. Shang, "A deep learning method for change detection in synthetic aperture radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 5751–5763, 2019.
- [52] D. Xue, T. Lei, X. Jia, X. Wang, T. Chen, and A. K. Nandi, "Unsupervised change detection using multiscale and multiresolution gaussian-mixture-model guided by saliency enhancement," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 1796–1809, 2020.
- [53] J. Wang, F. Gao, J. Dong, S. Zhang, and Q. Du, "Change detection from synthetic aperture radar images via graph-based knowledge supplement network," *arXiv preprint arXiv:2201.08954*, 2022.
- [54] Y. Li, C. Peng, Y. Chen, L. Jiao, L. Zhou, and R. Shang, "A deep learning method for change detection in synthetic aperture radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 5751–5763, 2019.
- [55] L. Bruzzone and D. F. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images," *IEEE Transactions on image processing*, vol. 11, no. 4, pp. 452–466, 2002.
- [56] R. Dekker, "Speckle filtering in satellite sar change detection imagery," *International Journal of Remote Sensing*, vol. 19, no. 6, pp. 1133–1146, 1998.
- [57] M. Majidi, S. Ahmadi, and R. Shah-Hosseini, "A saliency-guided neighbourhood ratio model for automatic change detection of sar images," *International Journal of Remote Sensing*, vol. 41, no. 24, pp. 9606–9627, 2020.
- [58] B. Hou, Q. Wei, Y. Zheng, and S. Wang, "Unsupervised change detection in sar image based on gauss-log ratio image fusion and compressed projection," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 7, no. 8, pp. 3297–3317, 2014.
- [59] W. Zhang, L. Jiao, F. Liu, S. Yang, and J. Liu, "Adaptive contourlet fusion clustering for sar image change detection," *IEEE Transactions on Image Processing*, vol. 31, pp. 2295–2308, 2022.
- [60] Y. Gao, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Synthetic aperture radar image change detection via siamese adaptive fusion network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 10748–10760, 2021.
- [61] F. Gao, J. Dong, B. Li, and Q. Xu, "Automatic change detection in synthetic aperture radar images based on pcanet," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1792–1796, 2016.
- [62] F. Gao, J. Dong, B. Li, Q. Xu, and C. Xie, "Change detection from synthetic aperture radar images based on neighborhood-based ratio and extreme learning machine," *Journal of Applied Remote Sensing*, vol. 10, no. 4, p. 046019, 2016.
- [63] S. Wang, S. Yang, and L. Jiao, "Saliency-guided change detection for sar imagery using a semi-supervised laplacian svm," *Remote sensing letters*, vol. 7, no. 11, pp. 1043–1052, 2016.
- [64] N. Lv, C. Chen, T. Qiu, and A. K. Sangaiah, "Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in sar images," *IEEE transactions on industrial informatics*, vol. 14, no. 12, pp. 5530–5538, 2018.
- [65] L. Li, H. Ma, and Z. Jia, "Gamma correction-based automatic unsupervised change detection in sar images via flicm model," *Journal of the Indian Society of Remote Sensing*, pp. 1–12, 2023.
- [66] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE transactions on image processing*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [67] G. Chierchia, D. Cozzolino, G. Poggi, and L. Verdoliva, "Sar image despeckling through convolutional neural networks," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 5438–5441, IEEE, 2017.
- [68] P. Wang, H. Zhang, and V. M. Patel, "Sar image despeckling using a convolutional neural network," *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1763–1767, 2017.
- [69] F. Ulaby, M. C. Dobson, and J. L. Álvarez-Pérez, *Handbook of radar scattering statistics for terrain*. Artech House, 2019.
- [70] H.-C. Li, W. Hong, Y.-R. Wu, and P.-Z. Fan, "Bayesian wavelet shrinkage with heterogeneity-adaptive threshold for sar image despeckling based on generalized gamma distribution," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 2388–2402, 2012.
- [71] S. Intajag and S. Chitwong, "Speckle noise estimation with generalized gamma distribution," in *2006 SICE-ICASE International Joint Conference*, pp. 1164–1167, IEEE, 2006.
- [72] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part IV 13*, pp. 184–199, Springer, 2014.
- [73] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431–3440, 2015.
- [74] Y. Liu, C. Pang, Z. Zhan, X. Zhang, and X. Yang, "Building change detection for remote sensing images using a dual-task constrained deep siamese convolutional network model," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 5, pp. 811–815, 2021.
- [75] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2414–2423, 2016.
- [76] M. Derezhinski and M. K. K. Warmuth, "The limits of squared euclidean distance regularization," in *Advances in Neural Information Processing Systems (Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, eds.)*, vol. 27, Curran Associates, Inc., 2014.
- [77] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [78] Y. Zheng, B. Jeon, D. Xu, Q. Wu, and H. Zhang, "Image segmentation by generalized hierarchical fuzzy c-means algorithm," *Journal of Intelligent & Fuzzy Systems*, vol. 28, no. 2, pp. 961–973, 2015.
- [79] J. C. Bezdek, R. Ehrlich, and W. Full, "Fcm: The fuzzy c-means clustering algorithm," *Computers & geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.
- [80] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 5, pp. 898–916, 2010.
- [81] Y. Gao, F. Gao, J. Dong, and H.-C. Li, "Sar image change detection based on multiscale capsule network," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 3, pp. 484–488, 2020.
- [82] A. Varghese, J. Gubbi, A. Ramaswamy, and P. Balamuralidhar, "Changenet: A deep learning architecture for visual change detection," in *Proceedings of the European conference on computer vision (ECCV) workshops*, pp. 0–0, 2018.
- [83] X. Tang, H. Zhang, L. Mou, F. Liu, X. Zhang, X. X. Zhu, and L. Jiao, "An unsupervised remote sensing change detection method based on multiscale graph convolutional network and metric learning," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

- [84] G. M. Foody, "Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification," *Remote Sensing of Environment*, vol. 239, p. 111630, 2020.
- [85] Y. Chen and L. Bruzzone, "Self-supervised change detection in multi-view remote sensing images. arxiv 2021." *arXiv preprint arXiv:2103.05969*.
- [86] E. Dalsasso, L. Denis, and F. Tupin, "Sar2sar: A semi-supervised despeckling algorithm for sar images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4321–4329, 2021.



**MOHAMED IHMEIDA** received the B.S.c degree in electrical and electronic engineering from University of Benghazi, Libya, in 2010, the M.S.c degree in Advanced Computer Network from University of Derby, UK, in 2018. In 2019 started his PhD journey at University of Reading in Synthetic Radar Images change detection.



**MUHAMMAD SHAHZAD** is currently working as a Lecturer in the Department of Computer Science, University of Reading, Reading, United Kingdom. His research interests include deep learning for remote sensing and computer vision-related applications.

...

# **Appendix B**

## **Source code**

The source code for the experiments is available at the link below, last updated on 31.1.2023.

<https://github.com/Mohamed-DL/Experiments-code>

