Effective Segmentation of Arctic Sea Ice Floes from High-Resolution Optical Images

Siyuan Chen

In the fulfilment of the requirement for the degree of

Master of Philosophy

Centre for excellence in Signal and Image Processing
Department of Electronic and Electrical Engineering
University of Strathclyde

Supervised by

Professor Jinchang Ren
Professor Stephen Marshall
Doctor Yijun Yan
Doctor Jaime Zabalza

Nov 25, 2022

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for the examination which has led to the award of a degree.

Siyuan Chen

Acknowledgements

2021 may not be a good year for most people due to the pandemic of COVID-19. As a student who started the research life with the working at home restriction, I have to say this is definitely the year I will never forget throughout my life.

Firstly, I would like to express my special thanks of gratitude to Professor Jinchang Ren for his long-time guidance. I will not have such a great improvement in the research area and personal skills without Jinchang's supervision. I am also grateful to my co-supervisor, Professor Steve Marshall and Dr J. Zabalza for their help and support to my work.

I am also appreciative of all the co-supervisors, especially Doctor Yijun, who has given me a lot of technical support and valuable advice regarding the career selection. Also, I would like to thank my roommate Yinhe, who has given me a lot of inspiration and emotional support during the remote working periods.

Finally, I would like to thank my parents and all my friends for their encouragement and understanding along the way.

Abstract

Arctic sea ice, as an essential environmental component that is closely connected to the Arctic ecosystem, plays an important role in the global weather and climate system. In addition, the Marginal Ice Zone (MIZ) which is the region close to the open water, plays a significant role in studying the physical and dynamic processes in Arctic area as it reflects the intense interactions between the atmosphere, open water, and sea ice. In addition, the MIZ can provide a physical buffer to protect the inner larger pack ice from being breakup by the ocean waves, which accordingly prevents the sea ice retreat that may occur due to the increasing Arctic open water area. Floe Size Distribution (FSD), a keystone indicator of the MIZ which has great impact on multiple sea ice processes such as the lateral melt rate and the propagation of waves underneath the sea ice, is particularly beneficial for the weather prediction, and the management of the Arctic region.

Until now, remote sensing data is one of the most popular and very often the only sources of information regarding sea ice conditions in the Arctic. During the last few decades, many efforts have been made for sea ice segmentation from the Synthetic Aperture Radar (SAR) images and optical images. However, separating the touching floes is still a main obstacle for accurate FSD retrieval. In recent years, the High-Resolution Optical (HRO) images with less speckle noise compared to the SAR images have provided an alternative solution to accurately delineate the floe boundaries and extract FSD. Nevertheless, traditional floe separation approaches require manual interactions. Meanwhile, the data annotation of sea ice images requires domain knowledge and can be labor extensive, resulting in the deep learning based methods hard to be applied.

In this thesis, a multi-stage segmentation and floe separation model is proposed to effectively investigate ice pixels and separate the touching ice floes automatically from the HRO images. For ice pixels investigation, a novel segmentation framework is proposed, where a combination of superpixel K-means clustering is employed for identifying the ice, water-ice mixed, and open water regions. Afterwards, the contrast enhancement technique is applied on the water-ice mixed regions to improve the performance of the subsequent thresholding process. For floe separation, the marker controlled watershed transformation based method is proposed, where different strategies are employed to generate the markers for preventing the floes from being

over-segmented. The robustness of the segmentation framework is validated on the image dataset and compared to two state-of-the-art methods. The result shows that the proposed framework has yielded the highest performance in terms of the accuracy, Matthews correlation coefficient (MCC) and F-1 score. For FSD retrieval, the proposed method outperforms the traditional distance transformation based watershed segmentation with the closer power law exponents and less mean square error compared to the ground truth. In addition, the experiments for optimising the model by evaluating the scenarios using different algorithms and parameter settings are conducted. As a result, for the proposed model, a combination of bilateral filter for pre-processing, Simple Linear Iterative Clustering (SLIC) for superpixel generation, and Top-bottom-hat transformation for contrast enhancement is recommended.

Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Motivation and objectives

The Arctic sea ice cover composed of discrete pieces of sea ice, known as floes, is a keystone indicator of ocean and climate research. The sea ice cover not only has a significant effect on regulating the heat, water, and momentum balance between the ocean and atmosphere [1], but also greatly affects human activities such as navigation and resource development in the polar regions [2]. In the past few decades, the Arctic sea ice extent has seen a rapid declining trend in the context of the anthropogenic climate change [3] and Arctic Amplification [4]. As can be seen in Figure 1. 1, the minimum sea ice area obtained in September each year has diminished from about 7 million km^2 before 2000 to less than 5 million km^2, with a record minimum value of 3.57 million km^2 in 2012. [5]. Such rapid sea ice extent reduction was beyond the expectation of the scientific community [6]. Hence, to both understand the trend and predictive forecast the sea ice future evolution, a better understanding of the region where ocean processes such as waves and wind can significantly affect the dynamics of the sea ice, so called the Marginal Ice Zone (MIZ), is required [7].

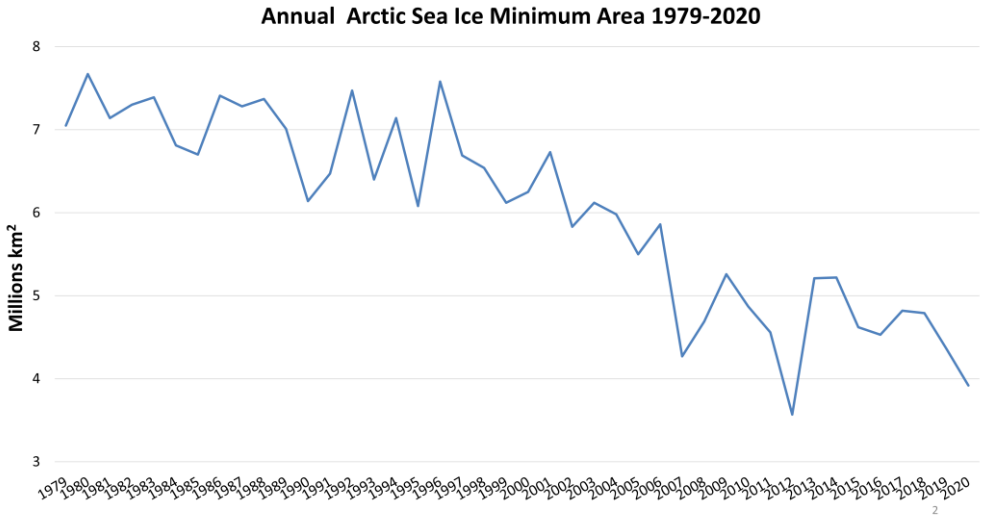

Figure 1. 1: Arctic sea ice minimum 1979-2020

The MIZ, defined as the region with 15 %–80 % sea ice concentration, is a dynamic and biologically active band of the sea ice cover close to the open ocean [8] and has been widely studied in the disciplines such as oceanography and marine biology [9]. It is not only the sea ice cover that makes the MIZ valuable, but rather the

biological effects that benefits the species that are already under pressure due to the climate change by acting as a living food container. Moreover, the physical and dynamic processes that reflect the intense interactions between the atmosphere, open water, and sea ice are significantly important. The MIZ can provide a physical buffer to protect the inner larger pack ice from being broken by the ocean waves [10], which accordingly prevents the sea ice retreat that may occur due to the increasing Arctic open water area [11]. In addition, the MIZ variation also has a great impact on the human accessibility to the Arctic, as the broken small floes allow easier navigation than the dense pack ice. Therefore, accurate modelling of MIZ would be beneficial for understanding how the environments inside and around the Arctic may develop. This can contribute to Arctic communities to plan for the knowledge-based management of the Arctic region, such as the effective measures for protecting the vulnerable species [12], and the projection of the potential viability of new routes for the ship navigation in the next few decades [13].

It is significantly challengeable to model the MIZ due to its complexity and variability. The floes in the MIZ can have varying sizes from meters to kilometers and vulnerable to ocean processes. Therefore, for the description of the evolution in the MIZ, floe size is an important criterion [14]. The smaller floes can cause the increment of the perimeter of the sea ice cover, leading to the enhancement of the lateral melt rate [15]. The increased melt rate will further expand the area of the open water and accordingly allow the solar insolation to transmit more heat to the ocean, which is known as the Albedo feedback [16]. As a result, the heated upper layer of the ocean causes the ocean to be re-stratified. In addition, with the expansion of MIZ, lateral ice melting is expected to be a significant factor of the seasonal ice loss [17].

Floe Size Distribution (FSD) is a probability function that characterizes the number of floes in different sizes categories in a unit area, and the FSD data are generally considered to follow a power law [14]. Accordingly, the FSD observations are usually reported by the power law exponent. However, current climate models seldomly take the floe size or the FSD into account [18], or assume the exponent of FSD as a constant value [17] due mainly to the difficulty of accurate FSD retrieval. Since floe size can have great impact on multiple sea ice processes such as the lateral melt rate, and the propagation of waves underneath the sea ice [19], such assumption of the fixed floe size prevents the climate model from accurately reacting to the

processes of the sea ice evolution induced by floes, and influence the accuracy of thermodynamics in the model as the ocean processes are highly coupled [20].

Remote sensing is usually considered as a suitable tool for comprehensively understanding and characterising the FSD across the Arctic region, as the satellite acquired images can provide good spatial and temporal coverage. A variety of remote sensing data acquired by different sensors have been employed for sea ice detection, such as the images collected by Synthetic Aperture Rader (SAR) and optical sensors. With the advancement of satellite technology, in recent years, the High-Resolution Optical (HRO) images containing less speckle noise compared to the SAR images have provided an alternative solution to accurately delineate the floe boundaries and extract FSD. Whether the image is "high-resolution" depends on if its features of interest are much larger than the pixel resolution [21]. The sea ice HRO imagery with a resolution at meter level conforms the definition in the context of the feature of ice floes, which thus enables the accurate detection of the weak boundaries in order to derive the FSD [22].

Both SAR data and optical data contain rich earth observation information but have their own advantages and disadvantages. SAR images, by recording the reflected echo of the emitted microwave radiation to the earth surface, have the advantage of high resistance to the meteorological influences. Yet the speckle noise that causes non-linear data distribution [23] may affect small floe investigation and weak boundary detection. On the other hand, the HRO images with less speckle noise can provide detailed information such as edges, shapes and texture of the floes, which can be beneficial for boundary detection and touching floe separation. However, HRO images also complicate the background, leading to the difficulty in cleanly discriminating the ice pixels from the melt ponds and submerged ice in the water-ice mixed regions.

Despite the existence of several HRO sea ice image datasets, records of the FSD retrieval from satellite observations are rare and the relevant approaches for floe segmentation and touching floe separation remain under-developed. The most challenging task is to automatically separate the ice floes that are seemingly connected to each other, as the weak boundaries in the conjunction have very similar features to the floe itself.

Currently, to achieve FSD retrieval from the optical images, the popularly used approach is to firstly discriminate the ice pixels from background, namely the water, for example using the thresholding [24, 25], graph cut [26], watershed segmentation

[27-29], etc. The binarized image is then proceeded to the post-processing stage to reduce erroneous floe splitting or merging, such as using the rule-based boundary revalidation [26]. And finally, manual correction is required [26, 30]. It is mentioned in the study [26] that for a single 10×10 km sea ice image, it can take more than three days to generate the floe separated ground truth data for FSD retrieval, as each touching boundary needs to be manually delineated by the sea ice expert, which is labor extensive. As a result, the records of the floe separated data are rare.

Machine learning approaches have also been widely used for processing the satellite observations [21, 31, 32]. For the sea ice HRO images, the emergence of melt pond causes the scattering characteristics of ice surface to change. By learning the pattern of this change, the melt pond can therefore be retrieved [31]. Similarly, by learning the characteristics of the ice surface, different types of the ice surface can be discriminated [21]. In addition, deep learning methods have also made a great progress in segmenting the remote sensing images, such as the Fully Convolutional Network (FCN) [33] and DeepLab [34]. However, for FSD retrieval, deep learning methods with insufficient labelled data can hardly to meet the needs. To tackle with this issue, semi-supervised methods could be beneficial, as a relatively small amount of labeled data is required compared to the deep learning [35].

In addition to the abovementioned approaches, in recent years, a growing interest in pixel grouping can be observed (i.e. object-based segmentation). Object-based approaches bottom-up cluster the pixels into groups according to their inherent characteristics (i.e. intensity, texture, and spatial correlation), which provides the potentiality of better preserving the shape and size of an individual floe [36]. Furthermore, the individual pixel schemes are likely to produce “speckled” errors that are incorrectly scattered across the image, leading to the error in FSD calculation.

Despite the existence of several approaches developed for sea ice detection, precise water-ice segmentation and automated floe separation method for FSD retrieval from the HRO imagery are still challenging. The main issues can be summarised as 1) Accurate delineation of the floe boundaries; 2) Identification of the floes in the water-ice mixed regions; 3) Separation of the touching floes. For identifying the ice pixels, the traditional pixel-based approaches could miss the small floes in water-ice mixed regions due to the low contrast, whilst producing holes inside the floes due to the working mechanism that solely relying on the pixel intensity. In terms of the deep learning based methods, it could be difficult to separate the touching

floes to achieve accurate FSD retrieval, due mainly to the lack of the labelled floe separated data [22]. In addition, although a growing interest in object-based methods can be observed such as the method based on gradient vector flow (GVF) [36], manual interactions are still inevitable and the effectiveness on high spatial resolution images containing size varying floes has not been validated. Therefore, an effective automated sea ice detection and floe separation method is beneficial for the MIZ study using HRO imagery. And such a method can generate labelled FSD data for a semi-supervised learning approach.

Since the contrast can be low in the water-ice mixed regions, contrast enhancement methods can be applied to enhance the segmentation performance. However, a global contrast enhancement may decrease the dark ice pixels inside the floes and cause over-segmentation. Therefore, local contrast enhancement is employed by only applying the contrast enhancement methods to the selected water-ice mixed regions. To effectively extract the water-ice mixed regions from the HRO image, superpixel based image segmentation techniques can be beneficial. Superpixel algorithms group pixels with visual coherence to create perceptually meaningful over-segmented small regions [37]. It has the foremost advantage that well adherence to the real object boundaries. Moreover, the superpixel segmentation can help to provide an intermediate representation between pixels and objects, which is a convenient basis for extracting image features and reducing the computational complexity in subsequent tasks [38]. For example, superpixels can be complementary to the classification methods such as ensemble learning and incremental learning [39] and conditional random fields based models [40] thus improve the performance and speed up the computation. By applying the superpixels to the sea ice image, the ice floe regions and water-ice mixed regions can be identified according to the features of the extracted superpixels, such as the pixel intensity and variance. As a result, the contrast enhancement methods and segmentation methods can be applied locally to enhance the segmentation quality in such regions without affecting the other regions. For floe separation, as the general watershed transformation will inevitably cause over-segmentation, it is not preferred in the model. Nevertheless, the floe regions formed by the superpixels can be employed as the marker to guide the watershed and prevent the floes from being wrongly separated.

Therefore, in this thesis, the overarching aim is to develop an effective semi-automated workflow to derive the sea ice floe size distribution from the HRO images. To achieve this goal, several objectives can be summarised as follows:

1) Due to the lack of the labelled data of separated floes , machine learning based methods are difficult to be applied for FSD applications. Therefore, one of the objectives here is to develop an unsupervised method that can reduce the labour of manual correction and be complementary to supervised learning methods.
2) For the precise determination of the floe size and the FSD, accurate segmentation is required to correctly identify the floes with the correct boundary adherence. Therefore, another objective is to enhance the source image to achieve higher distinguishability of floes especially in water-ice mixed regions, where superpixel techniques is used as the superpixels can provide the intermediate representation by aggregating the pixels with similar features for separating among ice, water-ice mixed, and open water regions. Additional modules are then needed to tackle further the challenges cases of water-ice mixed regions for refined segmentation.
3) The main obstacle in accurate FSD retrieval is to separate the touching floes that have weak boundaries, as those regions usually have similar features to the floe itself. Accordingly, another objective here is to develop an algorithm for automated separation of touching floes.
4) The performance of a multi-stage model can be affected significantly by the algorithm selection and parameter settings. Therefore, another objective is to find out the best performed scenario and experimental conditions, including the recommended parameter settings.

1.2 Thesis Organisation

The rest of this thesis is organized as follows:

Chapter 2 presents an overview of the background knowledge of the satellite remote sensing images and the related works for sea ice segmentation and FSD retrieval, including the superpixel-based segmentation methods. In addition, a brief review of the existing superpixel generation methos is provided.

Chapter 3 firstly introduces the technical background including the algorithms selection and their working principles. Then, the description of the proposed multi-stage model is introduced in detail, including the segmentation framework that integrates the superpixel, clustering, contrast enhancement, and thresholding

techniques, and floe separation method based on marker controlled watershed transformation with the refinement made in marker generation.

Chapter 4 first introduces the experimental image dataset and the settings corresponding to the model optimisation and validation. Various scenarios of the algorithm usage and parameter settings are evaluated. The optimsed model is then applied for sea ice segmentation and FSD retrieval. The quantitative evaluation and analysis are presented.

Chapter 5 briefly concludes the contributions of the thesis and discusses the future works.

1.3 Contributions

For tackling with the aforementioned research problems, a multi-stage model is proposed in this thesis. The major contributions can be summarised as follows:

1) A novel semi-automated multi-stage workflow is developed to effectively derive the FSD from HRO sea ice images.
2) For sea ice segmentation, by identifying the water-ice mixed regions and applying the contrast enhancement technique in such regions, the segmentation performance has been much improved. As a result, the proposed workflow has yielded the highest performances in terms of the overall accuracy, Matthews correlation coefficient and F-1 score as well as the computational efficiency compared to the other two state-of-the-art algorithms
3) By applying distance transform based watershed in water-ice mixed regions and morphological erosion in other regions, the floes can be successfully separated whilst suppressing the over-segmentation issue with the guidance of superpixels. As a result, the proposed method outperforms the baseline watershed algorithm with much improved FSD results in comparison to the ground truth.

Chapter 2. Literature Review

2.1 Introduction

In this chapter, according to the aforementioned motivation and objectives, the background and related works in remote sensing image processing, including sea ice segmentation and FSD retrieval are firstly introduced. Afterwards, the superpixel-based approaches for remote sensing and sea ice detection are described in Section 2.2.3. Section 2.2.4 surveys the superpixel generation algorithms and discusses their advantage and disadvantages. Finally, a summary is given in Section 2.3.

2.2 Development of Image Processing in Sea Ice Imagery

2.2.1 Background of SAR and Optical Imagery

As mentioned, SAR and optical images are the two main data sources in the sea ice study. The SAR images are derived by recording the reflected echoes of the emitted numerous microwave radiations to the earth surface, and then merging the echoes to form a "synthetic aperture". The SAR images over sea ice were collected since the launch of the satellite SEASAT (USA) in 1978. In the following decades, other satellites carrying the SAR instruments such as Kosmos-1870 (1987) and Almaz-1 (1991) have fully confirmed the applicability of the acquired images for disciplines including oceanology, cartography, geology, etc. [41]. In 2002, with the development of satellite technology, the ENVISAT (Environmental Satellite) with multiple polarisation ability for acquiring Advanced SAR (ASAR) data was launched. The improvement of the data source resulted in the feasibility of identifying the polarimetric scattering signatures of different types of sea ice and open water from different angles [41]. From the studies of the sea ice signatures [42, 43], the multi-polarised SAR data shows the great potential for the operational mapping of sea ice types and concentrations. SAR, as an active microwave imaging radar, has the ability of strong penetration, which can provide stable image observations at all times with rich texture information regardless of severe weather and nonuniform illumination conditions [44]. However, due to the multiple signal reflection, the prominent signatures in SAR images tend to be sparse and often not correspond to the physical structure of the object [45].

On the other hand, another important data source for sea ice detection is the optical imagery. The optical imagery with resolution at meter level can provide detailed information such as edges, shapes and texture of the floes which offers the potentiality to precisely investigate ice pixels in water-ice mixed regions for accurate FSD retrieval and weak boundary identification [22]. In recent years, there has been growing interest in the optical image datasets such as the MODIS optical data [46] and the Landset dataset [47]. For SAR images, it is difficult to explicitly identify the open water due to the varying backscatter signatures [48]. The factors such as incidence angle and wind speed can cause the backscatter signatures of open water to vary and span a range that overlaps with the signatures of the surrounding ice [49]. Optical images, by contrast, has better discrimination in ice and water due to the less speckle noise and better-preserved structural information. In addition, the HRO image datasets are becoming more and more accessible, thus providing an alternative solution that enables accurate ice floe boundary detection and FSD retrieval.

2.2.2 Sea Ice Detection and FSD Retrieval

For processing sea ice images, a variety of approaches have been developed. It is worth noting that most of the segmentation methods can be used for both SAR and optical images, although they are originally designed for only one of them. This is because after removing of the speckle noise SAR images can have similar characteristics as that of the optical images. Therefore, the image segmentation techniques for both SAR and optical images need to be reviewed.

The SAR images have been studied for a long time due to its advantage of collecting data regardless of weather conditions. The early attempts, for example in the study [50], used the local dynamic thresholding to segment the sea ice image. Pixels are divided into different classes through a set of threshold values derived according to the overall image intensities and the relative intensities within the images. In addition, to achieve the floe separation, the study [50] post-processed the segmented image with a morphological operation-based restricted shrinking/growing algorithm [51]. The key idea is to iteratively "grow" the binary object that can fill the holes within the object, and then "shrink" the object to break thin connections between the touching floes. However, such an algorithm contains much redundancy, as the iterations vary from floe to floe, resulting in the extended processing time for the large floes. Moreover, the morphological operations can deform the floe and affect the boundary

adherence of the segmentation result. In the study [52], the functional-based Markov Random Field (MRF) model was used to achieve the sea ice segmentation, in which the Gamma distribution was used instead of the generally used Gaussian distribution. Another MRF based method was applied in the study [53], where the edge penalty was incorporated to ensure the stability of model parameters estimation, however, with a relatively low efficiency.

More recently, high-resolution SAR imagery has been acquired with the advancement of satellite technology, the corresponding methods, therefore, have been further developed. In the study [26], the Kernel Graph Cuts (KGC) algorithm [54] has been applied. The KGC algorithm has the advantage of well handling the ununiformly illuminated image [55]. In addition, a combination of median, bilateral and Gaussian filters was used in [26] to refine the image data distribution to a Gaussian distribution for enhancing the performance of the KGC algorithm. Furthermore, to achieve FSD retrieval, the segmented results were further processed by the distance transformation controlled watershed transformation. Distance transformation is a technique to calculate the distance from each pixel within the object to the object edge. By selecting the regions with maximum distance to the edges as the local minima, so called the "basins", in the watershed transformation, the weak boundaries can be detected. This approach can work well if the ice floes are close to circle. However, the floes usually have irregular shapes in practical (i.e. two hump-like), which causes multiple local minima within one floe and accordingly leads to the over-segmentation. Therefore, to mitigate this issue, rule-based boundary revalidation [26] was then applied to remove the erroneous separation. Although the method has yielded satisfying results in the delineation of individual floes and avoiding the over-segmentation issue, considerable manual interactions including the manual correction for the floe boundary delineation and KGC parameter specification are required. Therefore, it can hardly be applied readily to the unknown dataset.

For optical images, some early attempts include, for instance, the study [56] used the thresholding method to detect ice floes where the threshold value was determined by analysing the histogram of the image, as the histogram of the grayscale of the sea ice optical images tend to be bimodal. The connected floes were separated by manually eliminating the weak boundaries. In the study [25], to determine the floe number from the digital images acquired by the optical camera integrated on an Unmanned Aerial Vehicle (UAV), two edge detection algorithms using gradient and morphology were

applied. However, since the traditional edge detection methods have limited effect in separating the touching floes, the number of detected floes greatly varied when using different edge detection methods. More recently, to investigate the weak boundaries of the seemingly connected floes, in the study [29], the original grayscale image was firstly binarised with the global Otsu thresholding algorithm. Afterwards, local minima regions were generated based on the segmented binary image to be complementary to the watershed transformation. In addition, to correct numerous over-segmentation errors, the junction lines produced by watershed are recovered according to the concavity of the boundary touching the lines, depending on whether the concavity is larger than the pre-defined threshold. Consequently, by eliminating the over-segmenting lines, the floe separated results can be refined. Nevertheless, since the shapes of floes tend to be irregular, the over-segmentation issue cannot be fully addressed by solely relying on the boundary concavity. Furthermore, the correct separation line could also be removed in some exceptional circumstances (i.e. touching square shape floes).

With the rapid development of the artificial intelligence. Various Machine Learning (ML) techniques have been applied to both SAR and optical images. For example, the clustering techniques are widely used in SAR sea ice images, and the efforts have been made to improve the object function of the clustering method (i.e. Fuzzy C-Means clustering (FCM) [57]) for achieving higher accuracy and efficiency. In the study [58], the energy measures of the wavelet decomposition were used to trace the texture information, and kernel distance was adopted for measuring the feature similarity, which can significantly reduce the influence caused by the speckle noise. Similarly, the study [59] incorporated a non-local spatial filtering to suppress the speckle noise. Subsequently, the fuzzy between-cluster variation term was adopted in the objective function, which enabled the regulation of the distance between the cluster centroids.

To tackle with the computation complexity of the common methods for SAR image segmentation, Shang et.al [60] proposed a segmentation method using region smoothing and label correction (RSLC). In this method, the input image was firstly smoothed to remove the speckle noise whilst retaining the edge information, so that the later segmentation process can be accelerated. In the smoothening process, different templates were applied for the edge regions and the homogeneous regions, in which the edges were detected and enhanced based on the convolution result of the

direction template with the maximum response, and the homogeneous regions were smoothened by the Gaussian filters that took the values of the direction-difference map as the variances of the kernel functions. This process enables the homogeneous regions to be smoothed efficiently, moreover, the smoothed image accelerates subsequent processes. Then, the two smoothed results are fused with K-means clustering, and the majority voting algorithm was consequently applied to correct the wrongly classified labels. Object boundaries were well adhered, and the structure was maintained from the experiments, but the robustness to the small objects remains unsatisfactory. The algorithm spent 5.43s for processing a 1001 × 779 SAR image, which is 170s faster than the algorithm for comparison.

Deep Learning (DL) approaches have also been applied for sea ice detection and classification. With multiple layers, DL models have strong capabilities of well extracting the features and autonomously learning, which are therefore suitable for object detection. In the study [61], an unsupervised neural network based on the Learning Vector Quantisation is used to classify the sea ices from the SAR images acquired by the airborne platform, yielding a total accuracy of 77.8%. In [62], an Artificial Neural Network (ANN) is used for the sea ice classification by taking the texture features of the image as input, where the features are extracted by using the Gray-Level Co-occurrence Matrix (GLCM). This algorithm was tested in the sea ice images with a size of 2200 × 3000 and yielded the classification accuracies of 79.4% for the open water, 89.3% for the dark ice, and 94.5% for the medium gray ice. Nevertheless, the computation time for a single image was nearly 30min. When applying deep learning on the remote sensing images, the lack of labeled data has always been a major hurdle. In this context, the U-Net [63] can be beneficial, as it was specifically designed for learning with a small number of training samples. In [64], five U-Net classifiers are created to tackle with sea ice classification under different conditions. These classifiers are then fused to derive an aggregate output by applying the stacked generalization [65], achieving an overall accuracy of 96.1%.

The ML and DL methods have also been applied in the optical remote sensing images [66-68]. For the sea ice detection, the watershed transformation and Random Forest (RF) [69] were combined for the adaptively detecting sea ice and melting ponds in the HRO images from different sources and data owners [21]. To produce the local minima for the watershed transformation, the gradient image was firstly created by applying the Sobel filter. The watershed segmented results were manually classified

to create the training dataset for the RF classifier. As a result, the algorithm has yielded an average accuracy of 96%. In another study [31], ANN was implemented for the retrieval of melting ponds in the optical images with a pixel resolution of 2 m, achieving an average accuracy of 85.5%. For the sea ice detection from close-range optical images (i.e. the images taken from the boat), the study [70] employed the Segmentation Net (SegNet) [71] and Pyramid Scene Parsing Network (PSPNet) [72] to classify ice, water, sky and boat. With the aid of employing the transfer learning [73] techniques for generating the training dataset, the PSPNet outperformed the SegNet with an average accuracy of 97.8%.

Despite the existing ML and DL approaches for optical image sea ice detection, none of them is designed for the FSD retrieval. Since the lack of the floe separated data can be one of the main issues, in this context, an automated approach for producing the floe separated data can be used to generate the training samples, i.e. in a semi-supervised learning process.

To achieve this goal, object-based methods can be beneficial, as object-based methods aggregate the homogenous pixels into objects and then classify the objects individually. In contrast to the pixel-based classification that classifies the pixels directly, object-based methods consider the spatial patterns that characterise the shape [74], so as to potentially distinguish the weak boundaries between the touching floes. In addition, the properties extracted from the objects such as spatial, textural, and contextual patterns can be fed as input into ML models for classification. In study [75], to segment the optical image with high resolution (pixel size of ~ 20cm), the fast edge-based segmentation [76], one of the object-based methods was employed. And then a total of 13 spectral, textural, and spatial features were extracted and fed into RF model for classification, achieving an overall accuracy of 95.5%.

For accurate sea ice detection and FSD retrieval, Zhang et.al in the study [36], applied the gradient vector flow (GVF) snake algorithm to the optical image acquired by UAV. The snake algorithm (also known as the deformable contour or active contour [77]), is a bottom-up region growing technique that works on the principle of controlling the continuous closed curve to iteratively evolve the shape and position until the desired image features are achieved. To apply GVF to the sea ice images, the method firstly differentiated the ice pixels that are normally illuminated and insufficiently illuminated, namely the “light” ice and “dark ice”. The light ice pixels are identified through Otsu thresholding. And the dark pixels, by contrast, are derived

from the K-means segmented image by subtracting the light ice image. According to the binary images of the light and dark ice, the seeds of the GVF algorithm are generated via distance transformation. The floe boundaries are then detected by GVF in the light and dark images separately based on the seeds. The segmentation result is achieved by fusing the two results together. Finally, morphological cleaning is applied to enhance the segmented image by filling the holes and eliminating the small ice pieces.

Although the approach proposed in [36] can separate the touching floes automatically, there are some limitations. Firstly, there is a trade-off between weak boundary detection capability and GVF capture range. For better floe separation effect, specific parameter settings are required for each image. In addition, with the increment of the floe size, more iterations are needed to effectively identify the boundaries, resulting in extended computation time. In the tested images acquired by UAV, no extremely large floe is contained. As mentioned, the floe sizes in the MIZ can vary from meters to kilometers. As a result, region growing may have difficulty in adhering object boundaries efficiently in the satellite acquired images with high spatial resolution. In this context, superpixel segmentation can be introduced.

Since the main purpose of this thesis is to effectively detect ice floes and extract FSD from HRO images. The obstacle is to accurately delineate floe boundaries in the water-ice mixed regions. Therefore, superpixel algorithms that group a set of pixels with visual coherence into meaningful over-segmenting regions can be used for replacing the regular pixel grid, which can provide a simplified form that is beneficial for computing image features and reducing the complexity [78]. Superpixels with similar properties can be merged to effectively categorise the pixels whilst well adhering object boundaries. Thus, the superpixel has the potentiality to detect the water-ice mixed regions efficiently, because within-class spectral variation can be significantly reduced by changing the classification unit from pixels to image objects [79]. In the following section, the superpixel-based segmentation methods in remote sensing, including sea ice detection are reviewed.

2.2.3 Superpixel-based Methods in Remote Sensing and Sea Ice Detection

A variety of computer vision approaches have considered the superpixel as a basic building block. This section surveys the applications of the superpixel-based methods in remote sensing, including the sea ice detection. The superpixels are generally used

to produce a first segmentation. The study [39] applied the grid-graph-based superpixel segmentation [80] for automatic iceberg detection from SAR images (pixel size of ~ 100m). The features were then extracted from the superpixels and fed into the RF classifier. With the aid of superpixel, the low-variability image regions, ice regions, can be well preserved, whilst removing the high-variability background regions. An accuracy of 98% was achieved.

In the study [40], Zhang et al. introduced a conditional random fields classifier for sea ice classification for SAR images in the Arctic during the melting season. The superpixels were obtained using the mean shift algorithm [81]. The features of backscatters of the generated superpixels were extracted and modelled for classifying the ice and water, which enhanced the performance in distinguishing the speckle noise and wind roughened open water.

For optical images, several superpixel based approaches have also been proposed. In the study [32]. The Simple Linear Iterative Clustering (SLIC) [37] algorithm was foremostly used to produce small segments, namely the superpixels. Afterwards, the superpixels were merged to determine the optimal segmentation scale through the minimum spanning tree. Usually, prior knowledge is required to determine the segmentation scale to obtain meaningful objects. With the employment of the minimum spanning tree, different proposals can be simulated and evaluated. Based on the local variance and the changing rate of the local variance in terms of the number of segments, the proper segmentation scale was then determined.

Similarly, the study [82] also adopted the SLIC algorithm, and further refined the superpixels results produced by SLIC through purifying the superpixels that contain different classes based on colour quantisation. And to improve the computational speed, a classification scheme is employed in which only the edge superpixels were purified. Therefore, the boundary accuracy can be improved compared to the original SLIC superpixels. In the study [22], Chai et al. improved the original SLIC algorithm by adding the texture descriptor term to the objective function. The local directional ZigZag pattern (LDZP) [83] was selected for extracting the texture descriptor due to its superior texture representation capability. The texture distance for every two pixels was calculated from the Euclidean distance of the surrounding elements of the targeted pixels in the 3 by 3 local window. The texture distance was then added to the SLIC objective function. Afterwards, Chai et al. applied two-stage thresholding to refine the superpixels. According to the local mean and variance of the superpixels, they

classified the superpixels into three categories: ice, water, and ice-water mixed, then the ice-water mixed regions are further segmented via support vector machine (SVM) [84], where 50% labelled data was used for training and rest for testing. The algorithm has achieved an accuracy of 95.51%.

In the study [85], to evaluate the effectiveness of different superpixel generation algorithms in processing the HRO sea ice images, four algorithms of which the effectiveness has been validated on other datasets were employed with a multi-stage segmentation model consists of 1) preprocessing for contrast enhancement, 2) superpixel generation, 3) superpixel grouping using K-means clustering, 4) and post-processing for ice shape enhancement. The experiments show that the segmentation results have yielded superior segmentation accuracy, which validates the feasibility of applying the superpixel techniques for the sea ice detection. However, the performance can be significantly affected with different superpixel number specifications. In addition, the size distribution has not been demonstrated.

2.2.4 Superpixel Generation Methods

A wide range of the superpixel generation algorithms have been developed and have been increasingly popularly used in the computer vision applications such as visual tracking [86], image segmentation [87], and image retrieval [88]. By taking the superpixels as the base unit rather than hundreds of thousands of image pixels, the computation efficiency can be particularly improved [89]. A brief review of the state-of-the-art superpixel algorithms is presented in this section. Basically, the existing superpixel algorithms can be categorized to the graph cut-based and seeding-based approaches according to their working principles.

For graph cut-based methods, the Normalised Cuts (NC) [90] is a classic algorithm and widely considered as the pioneer of the superpixel segmentation. The algorithm produces the sueperpixels in a top-down way by firstly charaterising the image by a graph. And minimize the costs of cutting over the graph. The cost is calculated according to the similarity between two pixels. Yet, NC has the disadvantage of high computational complexity which is reported to be $O(N^{1.5})$, where N is the total pixel number in the image. Similar to the NC algorithm, the methods based on graph cut usually have high complexity, such as the Entropy Rate Superpixels (ERS) [91] and

Efficient graph-based superpixel [80], which respectively have the complexity of $O(N^2 \log N)$ and $O(N \log N)$.

For the seeding-based methods, the universal concept is to firstly identify a number of seeds (also known as the centers). Afterwards, by bottom-up growing the seeds until meeting some certain conditions, the superpixels can therefore be generated. Quick Shift (QS) [92] and Mean Shift (MS) [81] are the two representative attempts based on the mode shifting. The superpixel generation is achieved by iteratively moving/shifting the seeds to the regions where the pixels have maximum probability density. However, the superpixel number cannot be explicitly controlled when using QS and MS. Manual parameter tuning with prior knowledge is required for changing the produced superpixel number. In addition, the computational complexity for both QS and MS are reported to be $O(N^2)$, which is relatively high and can cause extended running time when processing the large images. For improving the superpixel number controllability, the Lazy Random Walk (LRW) was proposed by Shen et al. [93]. The seeds are initialised by using the random walk technique to ensure the even distribution of the seeds in the whole image. The seeds are then iteratively shifted by energy optimization. The generated superpixels are refined by using the LRW for enhancing the compactness. Nevertheless, the computational complexity is still $O(N^2)$ as reported in [93]. The high computational complexity of these methods is a huge drawback when dealing with the HRO data.

In addition to the mode shifting methods, clustering is also widely used in the seeding-based superpixel algorithms. For instance, K-means clustering algorithm can be beneficial due to the capability of minimizing the local colour variance in the image. However, the computational cost can be high for deriving the distance between seeds and the rest pixels in the whole image. In this context, the Simple Linear Iterative Clustering (SLIC) [37] was proposed to solve the issue. In the SLIC algorithm, the searching window of the K-means clustering is restricted by considering the superpixel size. As a result, the computation complexity is relaxed to a level of $O(N)$. Texture Sensitive SLIC (TS-SLIC), an extended method from SLIC, was proposed in [22] which is specifically designed for sea ice detection. As mentioned in the previous section, TS-SLIC not only calculates the distance of spatial and colour information, but also measures the texture distance for better handling the ice surface. Another method clustering-based method by employing the Density-Based Spatial Clustering

of Applications with Noise (DBSCAN) is proposed in [94]. The algorithm firstly groups pixels into small seeds through using the DBSCAN, and secondly merge the seeds into superpixels, achieving a computation complexity of $O(N)$. However, the DBSCAN method have some shortcomings, such as the training is required for selecting appropriate threshold for distance measurement and appropriate description of the relation between colour and spatial information, leading to the algorithm hard to be implemented.

Other two well-known seeding-based methods are respectively the morphological-based and the watershed transformation based. For the morphological-based method, a popular work is the TurboPixel (TP) [95]. The algorithm creates evenly distributed seeds and then gradually dilates the seeds by means of geometric flows, achieving highly regular superpixel results. However, the regular pixel shape results in the poor performance in boundary adherence. The computational complexity of TP is reported to be $O(N)$ in [95]. Nevertheless, it is also pointed out in the study [96] that TP is usually much slower than other algorithms with the same computational complexity (i.e. SLIC). For watershed transformation, due to the characteristic of over-segmenting the image, it is available to produce superpixels. A computational complexity of $O(N \log N)$ is reported [97]. Nevertheless, due to its mechanism, the size and number of the segments cannot be controlled. To tackle with the drawbacks of the superpixel generation directly using watershed transformation, Water Pixel (WP) [98], a variant of the watershed segmentation was proposed. A regular hexagonal grid is firstly employed for generating the seeds as the basins, and the watershed transformation is performed on the gradient image to achieve better boundary adherence. According to the [98], the complexity is $O(N)$.

In addition to the graph cut-based and seeding-based methods, the learning-based methods have been recently researched for the superpixel generation. Bayesian Adaptive Superpixel Segmentation (BASS) [99] is one of the learning-based algorithms that employs the Bayesian mixture model, and provide the size adaptive superpixel by iteratively merging or splitting the superpixels. As a result, the BASS algorithm not only respects topology, but also favors spatial coherence. The computational complexity is not given in the original paper, but it is pointed out in the study [100] that BASS has a higher computational cost compared to SLIC.

2.3 Summary

This chapter discusses the research background and related work in the following three aspects. The first is the developed techniques in sea ice detection and FSD retrieval from both SAR and optical sea ice imagery. For the sea ice imagery with a high resolution (features of interest much larger than the pixel resolution) that has complicated information in the background, traditional top-down methods such as the global thresholding may not adhere to the floe boundaries well in the water-ice mixed regions, resulting in holes in the floes due to the variation of the pixel intensity. For FSD retrieval from the HRO imagery, the existing approaches focus mainly on the watershed transformation, and graph cuts. However, manual inspection and correction are inevitable, including refining the results by delineating the floe boundaries and separating the touching floes. For the deep learning-based methods, the FSD retrieval tends to be less focused, due mainly to the lack of sufficient data for training the models. As a result, an effective automated sea ice detection and FSD retrieval method is needed to benefit the MIZ, as it can generate labelled FSD data and be complementary to the semi-supervised learning approach. To achieve accurate FSD retrieval, a good boundary adherence is required. Therefore, the object-based approaches can be utilised, as they can aggregate the homogenous pixels by considering the spatial patterns. Superpixel techniques that can provide intermediate representation can therefore be applied for achieving better boundary adherence and provide convenience for further discriminating the water- ice mixed regions.

Accordingly, in the second aspect, several superpixel-based methods for remote sensing are discussed, including the review of sea ice detection methods. The superpixels are generally categorised according to the extracted features. Therefore the water-ice regions can be retrieved and further processed, e.g. using the SVM [22].

For a superpixel based image segmentation approach, the superpixel generation methods can affect the performance. Therefore, in the third aspect, several superpixel generation algorithms are reviewed, including the graph cut-based, seeding-based approaches, and learning based-approach. The selection of the superpixel algorithms is detailed in Chapter 3.

Chapter 3 Technical Background and Proposed Methodology

3.1 Introduction

Due to the requirement of the accurate FSD retrieval from the HRO sea ice imagery, in this chapter, a multi-stage model is proposed for effectively investigating the ice pixels and separating the touching floes. In the segmentation stage, a novel method is proposed by employing the superpixel techniques for identifying the floe, water-ice mixed, and open water regions, followed by applying the thresholding on the contrast enhanced water-ice mixed region to achieve better segmentation performance. Meanwhile, in the post-processing stage, a novel marker controlled watershed transformation-based floe separation method is proposed. To tackle with the over-segmentation issue with the general watershed transformation, different marker generation strategies are employed for the floes within or without the floe regions.

The rest of the chapter is organized as follows: Section 3.2 firstly introduces the technical background of the image processing techniques used in the thesis. In addition, calculation of the FSD is also presented. In Section 3.3, the overall workflow of the multi-stage segmentation model is introduced, in which the detailed information of the proposed methods in terms of the sea ice detection and floe separation is presented. Finally, some concluding remarks are suumarised in Section 3.4.

3.2 Technical Background

For sea ice detection from the HRO satellite imagery, accurate delineation of the floe boundaries has always been the major obstacle, as the detailed information provided by the HRO images also complicates the background in the water-ice mixed regions. Until now, in the remote sensing fields, the traditional pixel-based algorithms such as watershed, is usually sensitive to the noise. Also, the graph-based methods are usually inefficient due to the computational complexity. Therefore, superpixel-based methods have the potential to obtain the over-segmentation results efficiently and adhere well to the real object boundaries from the HRO sea ice images.

3.2.1 Superpixel Algorithms Used in Thesis

A superpixel segmentation of a sample image is presented in Figure 3. 1 with different superpixel sizes. As can be seen, the resulted superpixels can simplify the exemplification of the image to homogeneous small regions that are significant and easier for analysis and further processing. In addition, the determination of the size and compactness parameters can significantly affect the performance of the segments for maintaining the boundaries in the original image. Consequently, the smaller superpixels are usually required for delineating the boundaries of the informative regions.

Figure 3. 1: (left) Sample image. (right) Superpixel segmentation results with 200 (lower right) and 1000 (upper left) superpixels and the boundaries of each superpixel marked in yellow.

There are many algorithms developed for generating the superpixels. The state-of-the-art superpixel algorithms are briefly reviewed in Section 2.2.4, and each of them has its certain advantages and shortcomings that might adapt better to some specific situations. Despite the various approaches to generating superpixel, there are some key characteristics that most of the research agree with. The foremost characteristic is that the superpixel algorithm and method should be time-efficient whilst reducing the memory usage; secondly, superpixels should adhere well to object boundaries; thirdly, superpixels should improve the quality of the segmentation results.

As mentioned, the main purpose of using the superpixel technique is to tackle with the obstacle of accurately delineating the floe boundaries and generating the masks for the ice floe region and water-ice mixed region. Considering the pixel quantity in the HRO image can be over 1×10^8 (i.e. Figure A. 4 presented in appendix), algorithms with a lower computational complexity (i.e. $O(N)$) can significantly reduce the processing time and hardware requirement. In this context, the SLIC and WP are preferred, as the complexity of both of them are $O(N)$, along with good boundaries adherence capability to benefit accurate sea ice segmentation. The TS-SLIC [22], an

extended method from SLIC, has the awareness of the texture features which could be beneficial for identifying the floes from the water-ice mixed regions. The BASS algorithm, one of the recently proposed superpixel methods, produces superpixels with adaptively determined sizes and tends to preserve the object connectivity [99], resulting in the potentiality in coping with large or small sea ice floes. In addition, in a previous study of mine [85], the four aforementioned algorithms have been evaluated with a segmentation model using the K-means clustering to categorise the generated superpixels into ice and water to form a binary segmented image. The scenarios incorporating the four algorithms have yielded a segmentation accuracy of 98.19% on average, which validates the feasibility of using the four algorithms for sea ice detection. And the TS-SLIC scenario that slightly outperforms the other three methods with an accuracy of 98.30% was recommended. However, when the model was test on a single HRO image, the results seem to be highly sensitive to the parameter settings. Since the model proposed in this thesis further employs contrast enhancement and thresholding techniques during the segmentation, the result could be different to that in [85]. Therefore, the four algorithms are selected for evaluating their effect in sea ice segmentation in the method proposed in this thesis. A brief introduction to these four superpixel algorithms is given as follows.

Simple Linear Iterative Clustering (SLIC) [37]: SLIC, one of the most popular superpixel algorithms, adapts the k-means clustering algorithm and generates superpixels with uniform size. To well adhere to the boundaries of the objects with visual coherence, while controlling the compactness and size of the superpixels over the whole image, the weighted distance has been introduced that considers spatial proximity and colour similarity. In addition, the k-means algorithm is implemented locally in terms of each superpixel centroid by limiting the search space, which significantly reduces the complexity of computation even the number of superpixels is extremely large.

When implementing the SLIC algorithm, the superpixel number, N_{SP}, and the compactness coefficient, m, are the two parameters that need to be specified. Afterwards, according to the pixel number, N, of the whole image, the sampling interval, S, can be derived: $S = \sqrt{N/N_{SP}}$, which is used for restricting the size of the search area to form each superpixel instead of computing the distance of each pixel to the centroid globally. The sampling interval is firstly used to generate the regular grid

where the centroids are initially determined at the low gradient position. Then, by denoting each pixel by a five-dimensional vector (x, y, l, a, b), where the former two are the spatial coordinate and the latter three are the colour components from CIELAB colour space, the Euclidean distances of spatial and colour components d_c and d_s and the combined distance, D, for measuring proximity of two pixels i and j can be derived:

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \tag{3.1}$$

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \tag{3.2}$$

$$D = \sqrt{(d_c)^2 + (\frac{d_s}{S} m)^2} \tag{3.3}$$

Texture-Sensitive SLIC (TS-SLIC) [22]: The TS-SLIC, as a modified SLIC, introduced the measurement of texture descriptor to the distance calculation. Hence, the distance of the texture feature between two pixels can be derived by introducing the texture descriptor, local directional zigzag pattern, LDZP [83]. For the two pixels i and j, the texture distance d_t can be determined as in the Equation 3.4, where N_r is the number of neighbouring pixels in a selected local window (i.e. $N_r = 8$ for a 3 by 3 window). Afterwards, the modified distance between two pixels, D' can be calculated as in the Equation 3.5.

As a result, the generated superpixel become more sensitive to the texture information.

$$d_t = \sqrt{\sum_{n=1}^{N_r} (LDZP_{i,n} - LDZP_{j,n})^2} \tag{3.4}$$

$$D' = \sqrt{(d_c)^2 + (\frac{d_s}{S})^2 m^2 + d_t} \tag{3.5}$$

Water Pixel (WP) [98]: WP is the approach based on the watershed transform. As mentioned, the main concept of the watershed segmentation is to create catchment basins which are usually the local minima of the distance transformation, then the ridges of the basins filled by water are the watersheds. Although the traditional watershed segmentation is considered an efficient approach, it is not able to control the amount and the compactness of the segments. To improve the controllability thus avoid superpixels with irregular shapes and sizes, WP firstly generates a gradient

image based on the input image to determine the initial seeds, which are the minimum gradient regions in the cells of the regular hexagonal grid that is generated according to the predefined grid step. Afterwards, distance transformation is performed based on the seed regions and the result is added to the gradient image by multiplying a regularisation parameter that enforces the compactness of the latter superpixels generation. Consequently, the superpixels are determined by performing the watershed transformation on the regularised gradient image. Compared to the traditional watershed segmentation, the WP provide controllability over the superpixel amount and compactness, whilst maintaining the high efficiency.

Bayesian Adaptive Superpixel Segmentation (BASS) [99]. The BASS algorithm is a refinement of the Dirichlet-Process Gaussian Mixture Model (DPGMM), which is a Bayesian Non-Parametric (BNP) mixture model. By introducing the Potts term to the Bayesian estimation of the spatial covariances, the resulting superpixels respect more spatial coherence than the DPGMM. Moreover, BASS can produce size-adaptive superpixels without predefining the superpixel number by iteratively evaluating the split and merged superpixel proposals through the Hastings ratios, encouraging the superpixels to retain only the connected regions.

3.2.2 Noise Removal Algorithms Used in Thesis

Due to various internal and external factors, noise can be hardly avoided and may affect the subsequent processing stages in the remote sensing images. For example, in the optical imagery, white noise is more or less contained to degrade the performance of edge detection and the accuracy and efficiency of segmentation when using the methods based on texture features such as region growing [68]. Therefore, numerous denoising methods have been developed, which can be categorised to digital filtering-based methods and statistical-based methods [101]. The digital filtering-based methods suppress the noise by firstly transforming the spatial domain image to the frequency domain, then performing noise removal on the frequency domain. Despite the success of noise suppression, it is computationally expensive to accomplish the spatial-to-frequency transformation and the revert transformation. The statistical-based methods, on the other hand, suppress the noise by employing the statistical terms of the input images, such as the pixel mean or median value, and the histogram, which can avoid the high computation cost of the domain transformation.

In this context, two widely used statistical-based methods, Median filter, and Gaussian filter are firstly selected for evaluating the effect of the segmentation. In addition, in the study [22], Bilateral filter [102] and Robust Principal Component Analysis (Robust PCA) [103] are recommended for noise removal in the sea ice HRO image. However, the algorithms are compared using fixed parameter setting, and the effect of the noise removal to the segmentation quality is not evaluated. Therefore, these two methods are also adopted in this thesis for assessment. In addition, due to the limited training samples, the supervised noise removal methods such as the deep learning-based algorithms are not selected. A brief introduction to these four methods is given as follows.

Median filter: The median filter works on the principle that replacing the value of each pixel with the median value of the neighbouring pixels in a local window. In such a manner, the outliers that have unrepresentative values compared to the surrounding pixels at the local window can be eliminated. Depending on the window size and shape chosen, the performance can vary a lot.

Gaussian filter: Gaussian filter, aslo known as the Gaussian blur, is also a widely used technique in pre-processing that can reduce the noise. It works on the principle that convolving the image with a Gaussian function where the effect can be affected by the standard deviation of the Gaussian distribution. Similar to the median filter, the Gaussian filter may blur the image while reducing the noise. Therefore, the parameter selection is important for the filter.

Bilateral filter [102]: It is a non-linear noise removal approach that smooths the image whilst having the capability of preserving the edges. In contrast with the mean and median filter, the bilateral filter has introduced the weighted average instead of taking the average value directly from the nearby pixels. The weight considers not only the geometric closeness of the pixels, but also the photometric similarity(i.e. colour or intensity difference).

Robust PCA [103]: The Robust PCA is originated from the widely used dimension reduction technique of Principal Component Analysis (PCA). Although the traditional PCA has superior performance in many practical applications, it has the shortcoming against outliers or corrupted observations. To enhance the robustness of traditional PCA, the Robust PCA transforms the input image into the sum of a low-rank matrix and a sparse matrix, where the corruption component is contained in the sparse matrics. The decomposition of the original image brings benefits since the

corruption component can be arbitrary in magnitude, but are considered to be sparsely supported.

3.2.3 Contrast Enhancement Algorithms Used in Thesis

As the transition area between water and ice, the water-ice mixed regions can be observed to have relatively low contrast compared to that between the large floes and the open water. Therefore, as aforementioned, contrast enhancement techniques are introduced to improve the performance of the subsequent processes. Due to the histogram of the water and ice pixels in the transition area is usually compacted in a small range, a commonly employed contrast enhancement method, Histogram Equalisation (HE) can be utilised, as it is able to flatten and stretch the dynamic range of the image histogram. In addition, the Top-Bottom-hat transformation is selected due to is capability of maximised enhancing the bright regions and suppressing the dark regions, which can be beneficial for increasing the discrimination between floes and the background. Meanwhile, the Multiscale Retinex (MSR) [104] is also selected for comparison as it considers the characteristics of dynamic range compression and image contrast improvement. The three selected methods are briefly introduced as follows.

Histogram Equalisation (HE): As a widely used contrast enhancement approach, HE rearranges the histogram of the input image in which some of the pixel intensity values falls densely in a certain range to a uniform distributed histogram. HE is a nonlinear transformation that works on the principle of rearranging the histogram distribution based on the Cumulative Distribution Function (CDF). As denoted in Equation 3.6, suppose that $p_r(i)$ is the probability of the occurrence of the gray level i, where n_i is the number of the occurrence of i, and P_{total} is the total number of pixels in the image.

$$p_r(i) = \frac{n_i}{P_{total}}, 0 \leq i < 256 \tag{3.6}$$

$p_r(i)$ can also be seen as the normalised histogram of the original image. Then, the CDF function corresponds to the gray level i can be defined in Equation 3.7, which is namely the accumulated normalised probability.

$$CDF(i) = \sum_{j=0}^{i} p_r(j) \tag{3.7}$$

Given the accumulated normalised histogram, the gray levels can be rearranged and remapped back to the gray level range of the original image as shown in Equation 3.8. During the practical implementation, to meet the need of the data type, which is an unsigned integer in this case, the round operation is applied.

$$HE(k) = \text{round}(CDF(k) \times [max(i) - min(i)] + \min(i)), k \in [0, 255] \quad (3.8)$$

Top- Bottom-hat transformation: The Top-Hat and Bottom-Hat transformation is another contrast enhancement method which is widely used in medical image processing (i.e. retinal vessel investigation [105]). The transformation is based on the mathematical morphology. Dilation, $\oplus$ and erosion, $\ominus$ are the two fundamental operations in morphological filtering of images. As denoted in Equations 3.9 and 3.10, for each foreground component in a binary image, I, dilation and erosion will "thicken" or "thin" it by translating the Structure Element (SE) throughout the image [106]. SE here is defined by the shape, such as a disk, rectangle, diamond, etc., and the size is usually denoted by the radius in the image plane. The erosion operation thins the component by 1) setting the pixel on the origin (usually the center) of the SE to 0 when the SE overlaps the background, 2) retaining the pixel when the SE is entirely contained in the foreground component of the image. The dilation operation, in contrast, sets the pixel on the origin of the SE to 1 when any of the SE pixels overlaps the foreground of the original image. Opening and closing are the another two morphological operations based on the dilation and erosion which can be respectively denoted as in Equation 3.11 and 3.12.

$$I \ominus \text{SE} = \{i | (SE + i) \in I\}, \text{i} \in [0, size(I)] \quad (3.9)$$

$$I \oplus SE = \{i | (SE + i) \cup I \neq 0\}, \text{i} \in [0, size(I)] \quad (3.10)$$

$$I_{opening} = (I \ominus \text{SE}) \oplus \text{SE} \quad (3.11)$$

$$I_{closing} = (I \oplus SE) \ominus SE \quad (3.12)$$

Then, the Top-hat and Bottom-hat transforms can be denoted as in Equation 3.13, 3.14, and 3.15. The Top-hat transform retrieves the bright objects or elements that are smaller than the SE, by contrast, the Bottom-hat transform produces the dark ones. By adding the bright elements and suppressing the dark areas, the contrast of the image can be enhanced.

$$I_{top} = I - (I_{opening}) \quad (3.13)$$

$$I_{bot} = (I_{closing}) - I \quad (3.14)$$

$$I_{enhanced} = I + I_{top} - I_{bot} \quad (3.15)$$

Multiscale Retinex (MSR) [104]: The specific expression of the MSR for the grayscale image is shown in Equation 3.16. In the equation, $I(x,y)$ denotes the input image, N and ω_n are respectively the number and weight of the scale, and $*$ represents the convolution calculation.

$$MSR(i) = \sum_{n=1}^{N} \omega_n [\log I\,(x,y) - \log(F_n(x,y) * I(x,y))] \tag{3.16}$$

$$F(x,y) = Const \cdot e^{-(x^2+y^2)/\sigma_{MSR}} \tag{3.17}$$

In general, the N is set to 3 and ω_n is set to $1/3$ [104]. And the $F_n(x,y)$ represents a Gaussian function which can be defined in Equation 3.17, where the $Const$ is a normalisation factor used to fulfil the condition that the integral of $F(x,y)$ is always equal to 1. And σ_{MSR} is the standard deviation of the filter to control the amount of the spatial details to be retained. As suggested in the [104], $\sigma_{MSR} = [15,80,250]$ for the three different scales are the general settings.

3.2.4 Calculation of the Otsu threshold

The Otsu thresholding [107] is widely used in the image segmentation. The algorithm works on the principle that exhaustively assesses the threshold that can minimise the intra-class variance, σ_w^2, which can be denoted as in Equation 3.18. The σ_{c1}^2 and σ_{c2}^2 are the variances of the two classes separated by the threshold T_{Otsu}, where T_{Otsu} varies from 0 to 255 which is the range of the gray level in the uint8 type grayscale image. The weighting parameter, ω_{c1} and ω_{c2} are respectively the probability of the two classes, which can be denoted in Equation 3.19 and 3.20, where $P(i)$ is the percentage possibility that pixels with intensity equal to i appear in the image.

$$\sigma_w^2(T_{Otsu}) = \omega_{c1}(T_{Otsu})\sigma_{c1}^2(T_{Otsu}) + \omega_{c2}(T_{Otsu})\sigma_{c2}^2(T_{Otsu}), T_{Otsu} \in [0,255] \tag{3.18}$$

$$\omega_{c1}\,(T_{Otsu}) = \sum_{i=0}^{T_{Otsu}-1} P(i) \tag{3.19}$$

$$\omega_{c2}\,(T_{Otsu}) = \sum_{i=T_{Otsu}}^{255} P(i) \tag{3.20}$$

3.2.5 FSD Calculation

As mentioned in Section 1.1.1, the floe size and size distribution play important roles in the Arctic ecosystem. According to the Albedo feedback, the decreased floe size leads to the accelerated melting rate, and accordingly increase the heat transmission from the solar insolation to the upper layer of the ocean, which may cause the ocean re-stratification. According to the binary segmentation results that only contain water pixels and ice pixels (water = 0, ice = 1), the properties (i.e. area, caliper diameter) can be extracted at pixel-level. Afterwards, to practically apply the geoscience concepts into images, such properties can be converted to the SI units.

The observations of the FSD are usually presented in terms of the Floe Number Distribution (FND), which is frequency of the floes with different sizes in a unit domain area (typically in km^{-2}), and Cumulative Floe Number Distribution (CFND, hereinafter referred to as $N(d)$). The $N(d)$ is derived by cumulatively summing the FND from the floes with the largest size. And in the previous studies, $N(d)$ is usually defined as a power-law distribution as $N(d) \propto d^{-\alpha}$, where d is the floe size which is typically measured by the mean caliper diameter, and α is the power-law exponent [108]. Therefore, if logarithm is taken on $N(d)$, then the equation becomes $\log\big(N(d)\big) = C - \alpha \cdot \log(d)$ in which C is the constant value. Visually, the straight line can be observed on the log-log plot of $N(d)$ with the slope equal to $-\alpha$. Hence, for the FSD evaluation and climate model enhancement, finding the value of α that is as close to the true value as possible is the main issue.

Typically, the Least-Square Fit (LSF) is adopted for the slope estimation by employing the LSF on the log-log space. Usually the truncated range of the $N(d)$ is used for the slope fitting by artificially removing the small fraction of extremely small floes as those floes may not fit the power law and can affect the LSF fitting results [26]. Therefore, in this thesis, LSF is adopted for the α estimation. The results and comparison are presented in the Chapter 4.

3.3 Methodology

3.3.1 Overall Workflow of the Proposed Method

The proposed multi-stage segmentation model is centered around the superpixels derived from the input image. The overall workflow is presented in Figure 3. 2. First,

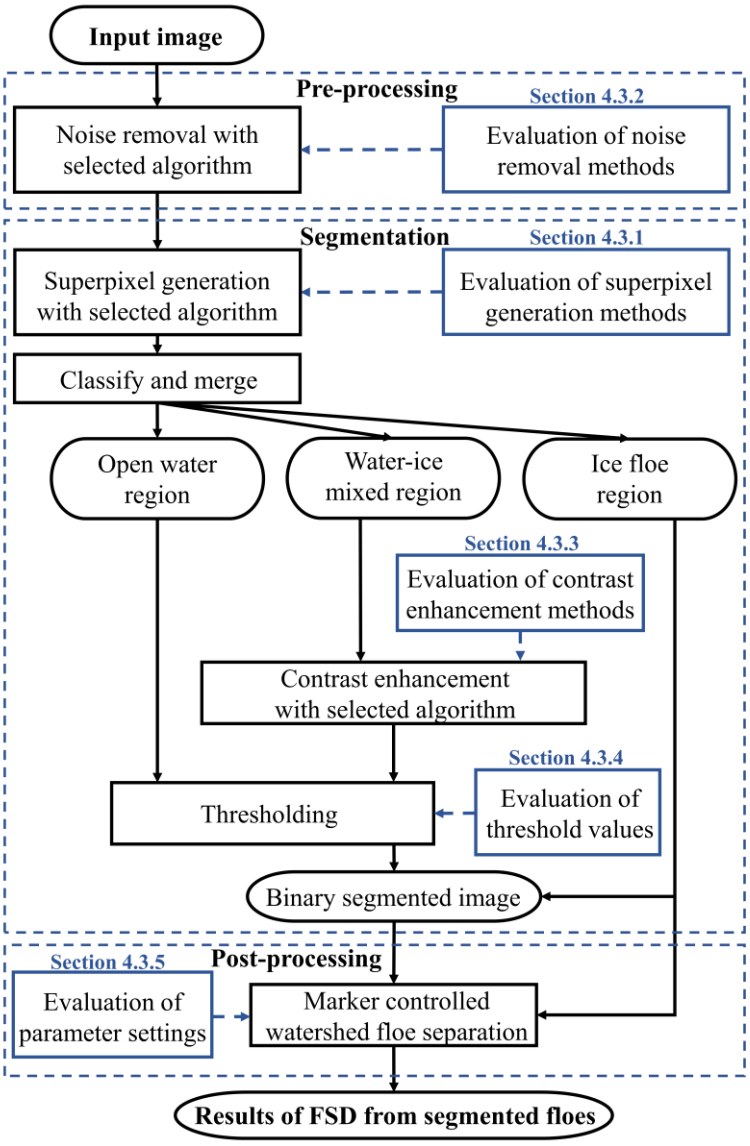

Figure 3. 2: Workflow of the proposed multi-stage segmentation model

in the pre-processing stage, noise removal algorithm is applied to enhance the image for producing better performance in the subsequent stages. Then, for precisely identifying the ice pixels from the HRO images, a novel Superpixel-based Segmentation with Contrast Enhancement in water-ice Mixed region (SS-CEM) is proposed. In SS-CEM, superpixels are firstly generated and categorised as ice (ICE), water-ice mixed (MIX), and open water (OW) according to the features extracted from the superpixels. The categorised superpixels are then merged respectively to form the ICE Region (ICE-Reg), MIX region (MIX-Reg), and OW region (OW-Reg). Afterwards, the contrast enhancement technique is applied in MIX-Reg for enhancing the performance of the subsequent thresholding process, where the threshold value is determined by combining the Otsu threshold and the standard deviation derived from

the MIX-Reg. Finally, the thresholded image is combined with the ICE-Reg to form the segmentation result. For the segmentation framework, the effects of the algorithm selection and parameter settings on the segmentation performance are evaluated for the model optimisation, which are presented in Chapter 4.

In the post-processing stage, a novel marker-controlled watershed segmentation based on distance transformation and morphological erosion is applied for floe separation. Compared to the general watershed segmentation, the proposed method applies different strategies to generate markers for the floes within or without ICE-Reg. In this manner, the mitigation of the over-segmentation issue in the general watershed segmentation is achieved. The framework of floe separation in detail is presented in Section 3.3.2.

3.3.2 Detailed Information of SS-CEM

As introduced in the overall workflow in Section 3.1, the segmentation framework, SS-CEM consists of four main steps. The intermediate results are presented in Figure 3. 3. In step 1, as shown in the Figure 3. 3 (a), the superpixels are firstly generated, where the boundaries of the superpixels are marked in yellow. Next, K-means clustering is applied to categorise the superpixels. The results are shown in Figure 3. 3 (b) (red: OW, green: ICE, blue: MIX). The K-means algorithm extract useful insights from data by iteratively assigning the cluster centroid and grouping the surrounding data points with the smallest distance until convergence. From the visual inspection, the superpixels in the ice floes usually have high pixel intensities and low standard deviation, and those in the water-ice mixed regions contain not that high but variant pixel intensities. Therefore, K-means can be employed for the classification by taking the mean intensity and standard deviation values extracted from each superpixel as inputs.

The classification results are plotted and shown in Figure 3. 3 (d). It is worthwhile noting that due to the Euclidean distance is used for calculating the distance between two superpixels in K-means, the input values measured on different scales may lead to the unbalanced weight for each feature. Therefore, the two statistical characteristics are normalised to the range of [0,1] during the implementation. Assuming the superpixel number is N_{sp}, the determined mean intensity and standard deviation values

are respectively $Intensity_n$ and Std_n, $n = 1,2, \dots, N_{sp}$, the normolised values are given by:

$$NormolisedIntensity_n = \frac{Intensity_n - \min\left(Intensity_n\right)}{\max(Intensity_n) - \min\left(Intensity_n\right)} \tag{3.21}$$

$$NormolisedStd_n = \frac{Std_n - \min\left(Std_n\right)}{\max(Std_n) - \min\left(Std_n\right)} \tag{3.22}$$

The classified superpixels are then merged to form the OW-Reg, ICE-Reg, and the MIX-Reg. Afterwards, in step 3, the contrast enhancement techniques are applied for a producing a better segmentation performance in the subsequent processing. The visual comparison can be seen from the Figure 3. 3 (e) and Figure 3. 3 (f), where the

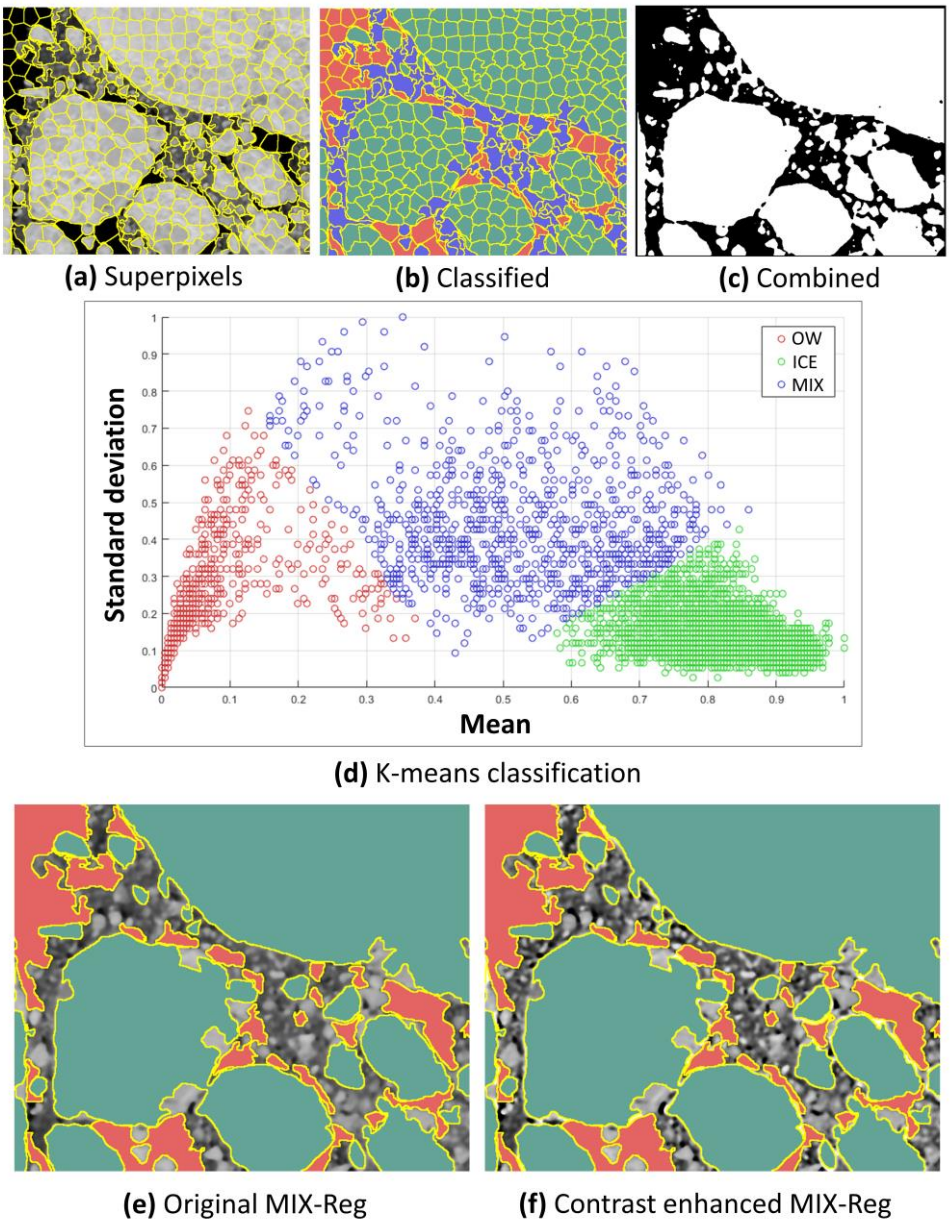

Figure 3. 3: Visualisation of the intermediate results in SS-CEM

ICE-Reg and OW-Reg are masked for better visual effects. From visual inspection, the floe boundaries in the MIX-Reg can be more clearly identified. The quantitative comparisons of the effect of contrast enhancement algorithms with different parameter settings are detailed in the Chapter 4.

In step 4, the thresholding technique is applied for investigating the ice pixels in the MIX-Reg and identifying the small floes that may contained in the OW-Reg. The histograms for the entire input image and the MIX-Reg are presented in Figure 3. 4. As can be seen from the Figure 3. 4 (a), a bimodal distribution can be observed, however, it is difficult to find an appropriate threshold for the investigation of the ice pixels with not that high intensity (i.e. [50,100]). And in Figure 3. 4 (b), the frequencies of the pixel intensity are more discriminative in the histogram of MIX-Reg by removing a large amount of the pixels with very high and very low intensities. Thus, the Otsu thresholding can be applied. Nevertheless, although the Otsu thresholding can produce the threshold automatically by minimising the intra-class variance, the problem is however, the submerged ice contained in such water-ice transition regions

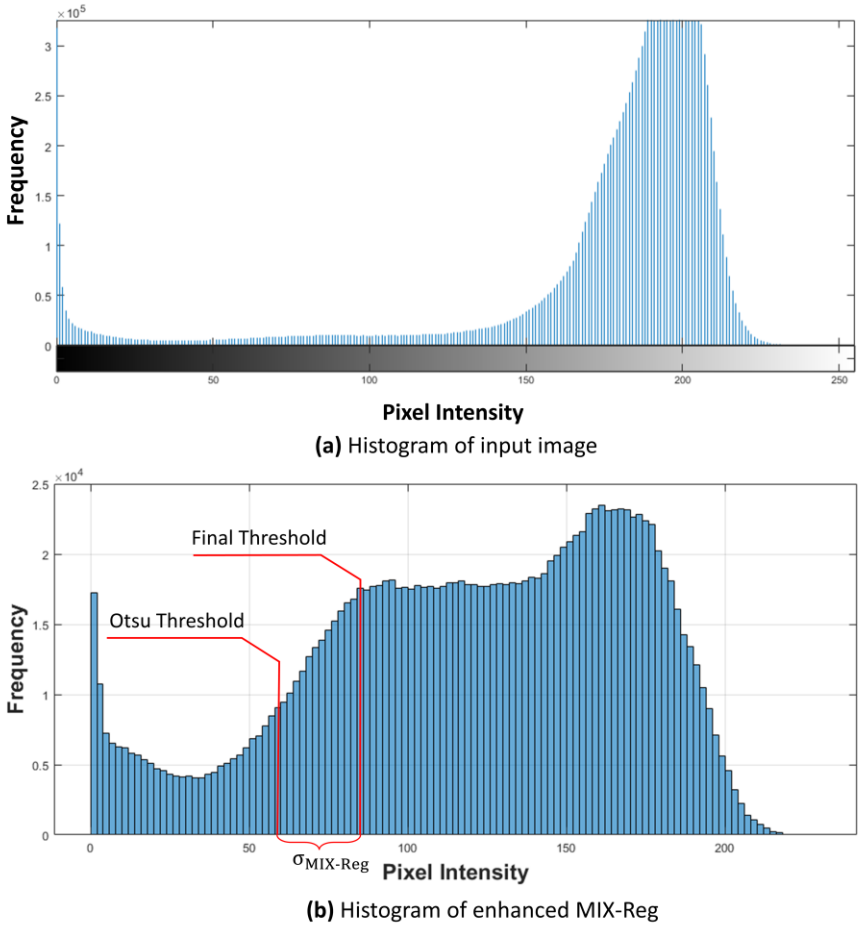

Figure 3. 4: Histogram of (a) the whole input image; (b) the enhanced MIX-Reg

can affect the performance of the Otsu thresholding. In this context, the local standard deviation derived from the MIX-Reg can be introduced. The final threshold is determined as in Equation 3.23, where the $\sigma_{MIX-Reg}$ is the standard deviation of MIX-Reg, and C_{thresh} is a constant for adjusting the threshold value. From trial and error, the C_{thresh} is suggested to set to 1. A brief comparison is presented in the section of model optimisation in Chapter 4. Finally, the thresholded results are combined with the ICE-Reg to generate the segmentation result as can be seen in Figure 3. 3 (c).

$$T_{final} = T_{Otsu} + C_{thresh} \cdot \sigma_{MIX-Reg} \tag{3.23}$$

3.3.3 Erosion and Distance transformation-based Marker Controlled Floe Separation (ED-MCFS)

In the post-processing stage, the watershed algorithm has the potential to separate the connected floes with a main obstacle of the over-segmentation. To tackle with the issue of over-segmentation, a novel Erosion and Distance Transformation-based Marker-Controlled Floe Separation (ED-MCFS) framework is proposed. The workflow of the framework is presented in Figure 3. 5.

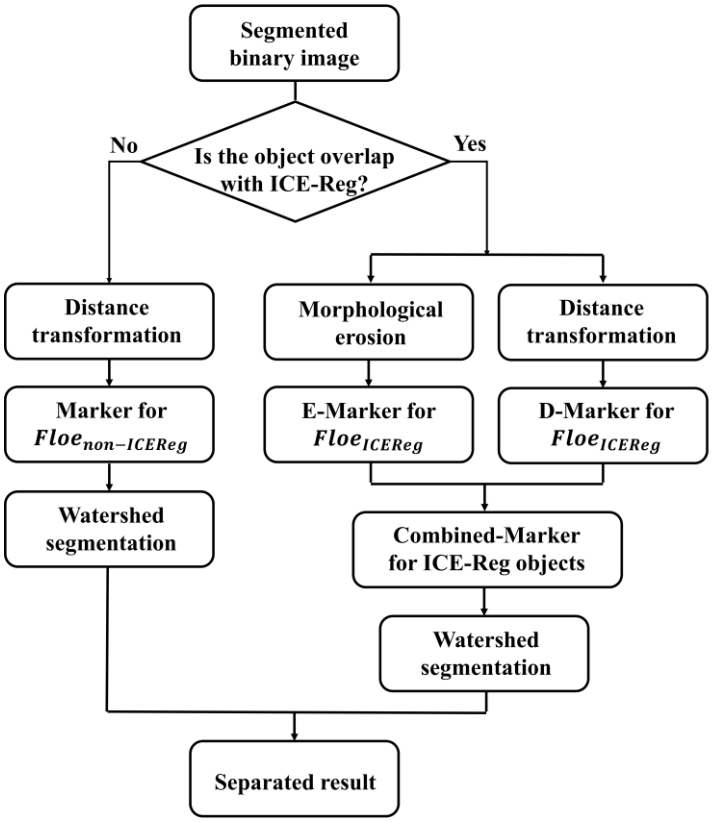

Figure 3. 5: Workflow of the floe separation

First, according to the binary segmented result and the ICE-Reg identified in the segmentation stage, the objects with 8-connectivity are categorised into $Floe_{ICEReg}$, and $Floe_{non-ICEReg}$. The two categories are defined as in the Equation 3.24 and Equation 3.25, where N_{obj} is the total number of the object with 8-connectivity in the segmented binary image, and the $Floe(k)$ is accordingly the k^{th} object. $Floe_{ICEReg}$ are namely the objects that have overlapping regions with the ICE-Reg, and $Floe_{non-ICEReg}$ are the resting objects.

$$Floe_{ICEReg} = \{Floe(k) | Floe(k) \cap ICEReg \neq 0\}, k \in [1, N_{obj}] \quad (3.24)$$

$$Floe_{non-ICEReg} = \{Floe(k) | Floe(k) \cap ICEReg = 0\}, k \in [1, N_{obj}] \quad (3.25)$$

Different strategies are employed for the two categories mainly due to difference in size. The intermediate results of the floe separation are presented in Figure 3. 6. Firstly, for the $Floe_{ICEReg}$ as shown in Figure 3. 6 (c), morphological erosion and

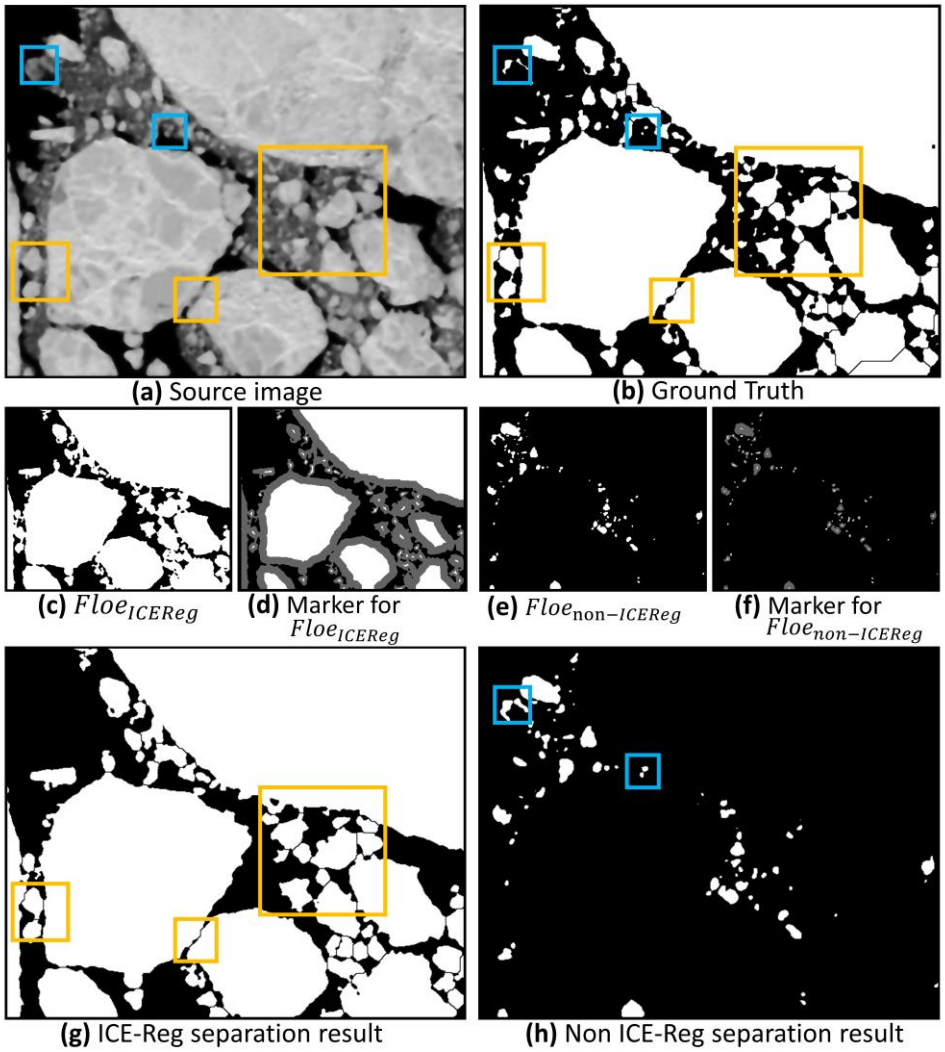

Figure 3. 6: Floe separation intermediate results visualisation

distance transformation are applied to generate the E-Marker and D-Marker. The combined result of the two markers is marked in white in the Figure 3. 6 (d), where the $Floe_{ICEReg}$ is marked in gray. Afterwards, the combined marker is used as the local minima for guiding the watershed stransformation. The result is presented in Figure 3. 6 (g). As highlighted in the yellow bounding box, the floe separation is achieved without obvious over-segmentation compared to the segmented result.

As can be seen in Figure 3. 6 (e), for $Floe_{non-ICEReg}$, the floes that do not have overlap with the ICE-Reg, distance transformation is applied, as such floes usually have small size and relatively regular shape. The generated marker for $Floe_{non-ICEReg}$ is presented in Figure 3. 6 (f). Similarly, the watershed segmentation is applied with the guidance of the marker, and the result is presented in Figure 3. 6 (h).

3.4 Summary

In this chapter, the technical background is firstly introduced, including the working principle of the algorithms used in different stages of the proposed model, followed by the calculation of the FSD.

Afterwards, a multi-stage model consists of the pre-processing, segmentation, and post-processing stages for ice pixels detection and floe separation is introduced. In the segmentation stage, a novel framework, SS-CEM is proposed, where the combination of the superpixel technique and K-means clustering is firstly employed for identifying the ice floe regions (ICE-Reg), water-ice mixed regions (MIX-Reg), and open water regions (OW-Reg). Next, contrast enhancement techniques are employed to enhance the discrimination of water and ice pixels in the MIX-Reg. Subsequently, thresholding is applied in MIX-Reg and OW-Reg. In such a divide and conquer manner, the segmentation performance in the water-ice mixed region can be improved, whilst preserving the floe surface. Finally, by combining the thresholded result and the ICE-Reg, the segmented result is achieved.

In the post-processing stage, a novel marker controlled watershed-based floe separation framework, ED-MCFS is proposed. In the framework, for the floes that that have overlap with the ICE-Reg, morphological erosion and distance transformation are applied for marker generation. The created markers are then combined for guiding the watershed transformation for the floes in ICE-Reg. For the resting floes that tend to have relatively smaller sizes, the marker is generated by solely applying the distance

transformation. By employing the markers for floes in different regions, it is aimed to tackle with the over-segmentation issue which is hard to be avoided in the general watershed transformation. The effect of the parameter settings and the evaluation of the performance applied to the HRO image dataset are detailed in Chapter 4.

Chapter 4. Superpixel-based Multi-stage Segmentation of Sea Ice Images

4.1 Introduction

Since several algorithms are selected in each stage of the proposed multi-stage, the best performed scenarios can be selected by quantitatively evaluating the effect to the segmentation quality and FSD exponent estimation. A dataset consists of 8 HRO images are used for validation. The experiments in this chapter were all conducted on a computer with a 4.1 GHz CPU and 16 GB RAM.

The rest of the chapter is organized as follows. In Section 4.2, the image dataset and ground truth generation are firstly introduced, followed by the experimental design and the key parameter settings for the model optimisation. Afterwards, in Section 4.3, the comparison of the results produced by different scenarios is detailed. In Section 4.4, the assessments of the optimised model on the whole image dataset against the other methods are provided. Finally, a summary of the recommended algorithm and parameter settings, and the optimised model performance is given in Section 4.5.

4.2 Experimental Design and Key Parameter Settings

4.2.1 Image Dataset and Ground Truth Generation

Brief introduction to the images used in this thesis is presented in Table 4. 1, where the acquisition date and image size for each image is also presented. The original HRO images are accessible on the Global Fiducials Library (GFL) [109], a long-term archive of the images from U.S. National Imagery Systems released by the U.S. Geological Survey (USGS) in 2009.

Among the large amount of the HRO data in the GFL, the images with relatively high qualities (i.e. no heavy cloud and severe variant illumination) are selected from Chukchi Sea, East Siberian Sea, and Arctic Sea Ice Buoys datasets for creating the image dataset for testing. The source images are presented in the Appendix A. It is worthwhile noting that the images used in the thesis are not exactly the original image in the GFL but have been pre-processed. The interested region was firstly cropped out

Table 4. 1: Brief introduction to the image dataset

Name	Image Size	Acquisition date
Chukchi_2013	2548×6368	31 May 2013
Chukchi_2010_1	7632×8544	30 May 2010
Chukchi_2010_2	7080×10944	30 May 2010
Chukchi_2014	13504×12672	02 May 2014
Esiber_2001	3684×7056	16 Jun 2001
MIZ_2014_1	6720×9168	18 Jul 2014
MIZ_2014_2	10352×9392	30 Jul 2014
MIZ_2014_3	10944×7728	14 Aug 2014

from the rotated original HRO images to remove the background pixels and the distorted regions. And the images are then downsized to half resolution (2 meter). The generated dataset is then used to evaluate the effectiveness of the proposed multi-stage model. Last but not least, the data type of all the images are unsigned 8-bit integer.

To quantitatively evaluate the segmentation performance, Ground Truth (GT) data is necessary. In this thesis, following the GT generation methodology adopted in the study [26], The GT data is then generated via firstly producing the automated segmentation results through the environment for visualizing images (ENVI) software, then the manual correction is performed by domain experts to trace the floe boundaries and separate the touching floes. An edge-preserving and merging algorithm are used in the ENVI to produce the segmentation result. However, the threshold need to be specified manually to obtain satisfactory outcomes [26].

4.2.2 Evaluation Metrics

For the quantitative assessment of the segmentation results, several popularly used metrics are adopted in the experiments, including the pixel-level Accuracy (ACC), Matthews correlation coefficient (MCC), and F1 score [110-112]. These three metrics are derived from the confusion matrix that consists of the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The positive or negative labels the pixels that are foreground or background in the segmented binary image, and the true or false indicates if the pixel is correctly predicted compared to the GT data. ACC, as denoted in Equations 4.1, is the metric that quantifies the how many foreground pixels are correctly classified compared to GT. On the other hand, MCC

as in Equations 4.2 can comprehensively assess the performance of the image within the four quadrants of the confusion matrix. Similarly, F-1 score describes the overall performance in both Precision and Recall, as can be seen in Equations 4.3 to 4.5. The Precision and Recall are attractive as the criteria of segmentation quality because they are sensitive to the over and under segmentation. Over-segmentation decreases the Precision and under-segmentation leads to low Recall.

$$ACC = (TP + TN)/(TP + TN + FP + FN) \tag{4.1}$$

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \tag{4.2}$$

$$F1 = (2 \times Precision \times Recall)/(Precision + Recall) \tag{4.3}$$

$$Precision = TP/(TP + FP) \tag{4.4}$$

$$Recall = TP/(TP + FN) \tag{4.5}$$

In addition to the ACC, MCC and F1 score, the peak signal to noise ratio (PSNR) is also adopted. The PSNR that considers image type and Mean Square Error (MSE) is defined as in Equation 4.6, where the $PeakVal$ is determined by the maximum value of the input image. In the context of the image used in this thesis, the data type is uint8, the $PeakVal$ therefore is 255. On the other hand, the MSE is defined in Equation 4.7. MSE is the expectation of the square difference between the pixel values of the denoised image, $f'(i,j)$ and those of the original image, $f(i,j)$, where M and N are namely the length and width of the image. In general, the higher the PSNR, the less the image is distorted.

$$PSNR = 10\,log_{10}(\frac{PeakVal^2}{MSE}) \tag{4.6}$$

$$MSE = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}(f'(i,j) - f(i,j))^2 \tag{4.7}$$

For quantitative assessment of the proposed floe separation method, the floes in GT and separated images with 8-connectivity are quantified into nine categories, $\mathrm{CAT(c)}$ according to the size. The definition of $\mathrm{CAT(c)}$ is as in Equation 4.8, where $Obj(i)$ is the floe with 8-connectivity and $Area\big(Obj(i)\big)$ denotes how many pixels is contained in the $Obj(i)$. N_{Obj} is the total number of the floes with 8-connectivity in the binary image.

$$\mathrm{CAT(c)} = \{Obj(i) \big| Area\big(Obj(i)\big) \in [10^{c-1}, 10^{c}]\}, c \in [1,9], i \in [1, N_{Obj}] \tag{4.8}$$

For the produced separated result and GT image, $CAT_{SEP}(c)$ and $CAT_{GT}(c)$ are calculated respectively. Afterwards, for evaluating the similarity of the pixel level floe size distribution in the two images, the mean square error can be derived as denoted in Equation 4.9.

$$MSE_{CAT} = \frac{1}{9}\sum_{c=1}^{9}(CAT_{SEP}(c) - CAT_{GT}(c))^2 \tag{4.9}$$

Finally, for evaluating $N(d)$, the cumulative floe number distribution that is proportional to $d^{-\alpha}$ (introduced in Section 3.2.5), the absolute difference between $\alpha_{Predicted}$ and α_{GT}, the exponent values estimated from the separated result and GT image, is compared as shown in Equation 4.10.

$$Diff = \left|\frac{\alpha_{GT} - \alpha_{Predicted}}{\alpha_{GT}}\right| \times 100\% \tag{4.10}$$

4.2.3 Parameter settings

Since several algorithms are selected for comparison in each stage of the proposed multi-stage model, an optimisation process is necessary for the determination of the best performed combination and key parameter settings. Therefore, the experiments for model optimisation are designed and detailed in the following subsections. In addition, a sample image is required for the experiments. From visual inspection, Chukchi_2013 that is less affected by the environmental factors such as the cloud and illumination, is therefore selected as the sample image.

4.2.3.1 Settings for the Superpixel Algorithms

For the evaluation and comparison of the four selected superpixel generation algorithms, SLIC, WP, BASS, TS-SLIC, a simplified model is applied by removing the pre-processing stage and the contrast enhancement in MIX-Reg. The effects of the noise removal and contrast enhancement algorithms with different parameters are evaluated after the superpixel algorithm selection.

The performance of the SLIC, WP, and TS-SLIC can vary as the superpixel number, N_{SP} changes. Since the generated superpixels are usually compact in size, the superpixels therefore consist of approximately the same number of pixels. Therefore, the effect of N_{SP} need to be evaluated. However, according to Table 4. 1, the selected

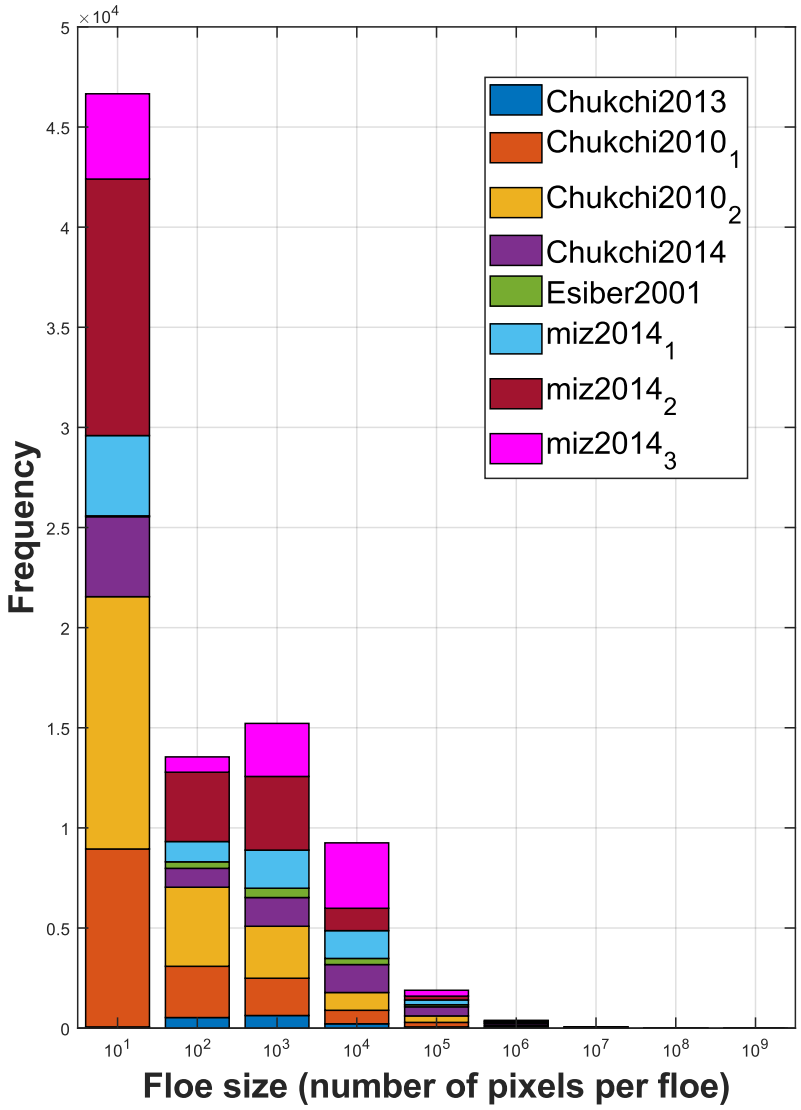

Figure 4. 1: Pixel level floe distribution of the image dataset observed from GT

images have different size. Therefore, a fixed superpixel number setting will lead to the varying superpixel size for different images and may affect the segmentation performance. For instance, too large superpixel size could cause the superpixel containing ice regions and water-ice mixed regions simultaneously. And if too small superpixels are used, the features contained could be unstable and accordingly affect the classification performance by K-means.

Figure 4. 1 presents the pixel level floe size distribution of all the images in the dataset, where the floes are categorised as $\mathrm{CAT}(c), c \in [1,9]$. As can be seen, floes within 10 pixels appear most frequently in the image dataset. However, it would be less meaningful to derive and compare the mean intensity and standard deviation from

the superpixels that only contain roundly 10 pixels, as the features could be significantly affected by the disturbance such as melt ponds, cloud, and shadow. In addition, in the proposed segmentation model, such small floes can be investigated in the thresholding stage. Therefore, the floe sizes with the second and third highest frequency are selected to be the superpixel sizes. Considering the size of the sample image is 2548×6368, the superpixel number, N_{SP} is then set to the range of 1×10^4 to 10×10^4(roundly 150 to 1500 pixels for each superpixel) with an increment of 1×10^4 for SLIC and TS-SLIC algorithms. For WP, the superpixel number is controlled by specifying the grid step, σ_{WP}, of the initial regular superpixel grid. Meanwhile, the superpixel number is capped at 65536 in the implementation provided by the author of WP. Thus, the results are produced with WP until its superpixel number reaches the highest value.

For implementing the superpixel algorithms, some other parameters need to be specified. In this study, the compactness coefficient $m = 10$ is adopted for both SLIC and TS-SLIC, and the regularisation parameter k_{wp} is set to 8 for WP as suggested in the original papers [22, 37, 98]. The gradient image used for WP is produced by performing a basic morphological gradient using the SE with the radius $R = 3$. In addition, to ensure consistency of the algorithm performance to the original papers, WP and BASS are implemented based on the codes provided by the authors. The WP superpixel results were produced in Ubuntu 20.04.2, while the BASS superpixels were generated via Python 3.8.5 with Pytorch 1.3.0. The generated results are then imported to the MATLAB.

4.2.3.2 Settings for Pre-processing

In the experiments for pre-processing, the segmentation model incorporating SLIC and removing the contrast enhancement in MIX-Reg is employed. The effect of contrast enhancement stage is evaluated in later experiments. To evaluate the selected

Table 4. 2: Parameter settings for pre-processing

Method	Parameters	Values
Median Filter	Window Size, m	1:1:20
Gaussian Filter	Standard deviation, σ_g	0.5:0.25:5
Robust PCA	Regularisation parameter, λ	$[\lambda_0 : \lambda_0 : 10\lambda_0]$
Bilateral Filter	Degree of Smoothing, DoS	1000:2000:50000
	Domain parameter for spatial kernel, σ_d	[1, 3, 5]

algorithms for pre-processing, namely the median filter, Gaussian filter, Robust PCA, and bilateral filter, for each algorithm, the effect of using different parameter settings is evaluated. The parameter settings for the four algorithms are presented in Table 4. 2. It is worth noting that for the Robust PCA, a regularisation parameter, λ is required to relax the computation. The suggested λ can be derived from Equation 4.11, where M and N are respectively the length and width of the input image.

$$\lambda_0 = \frac{1}{\sqrt{max\ (M, N)}} \tag{4.11}$$

For example, for the sample image Chukchi_2013, $M = 2548$ while $N = 6368$. Hence, $max(M, N) = 6368$. The value of the suggested λ, can be calculated accordingly to be $\lambda_0 \approx 0.0125$. To evaluate how the parameter can affect the performance in the HRO image, the experiments that using λ from λ_0 to $10\lambda_0$ with an increment of λ_0 are conducted.

For the bilateral filter, the two keystone parameters respectively are the Degree of Smoothing, DoS, which is the range parameter for colour kernel, and the domain parameter for spatial kernel, σ_d. As illustrated in the study [102], there is a trade-off between these two parameters. Basically, larger range parameters can produce smoother results. The weight of colour of each pixel becomes the same as the value of range parameter approaches infinity, resulting in the filter becoming equivalent to the mean filter. Therefore, the range parameter cannot be too large. Certainly, the value of the range parameter cannot be too small as well, because the extremely small DoS can significantly increase the weight of the center point, resulting in the suppressed denoising effect. Accordingly, when the DoS is set to 0, the output image is exactly equivalent to the input image. On the other hand, the domain parameter for spatial kernel, σ_d constraints the effect of the smoothing parameter and has the capability of controlling the edge preservability. As a result, the blurring of the edge can be prevented by setting a small σ_d, however, the denoising and smoothening effect can be suppressed.

4.2.3.3 Settings for Contrast Enhancement

For the experiments on the contrast enhancement algorithms applied in MIX-Reg, the best performed pre-processing algorithm and the segmentation model incorporating SLIC is adopted to ensure the consistency with the pre-processing experiments.

Table 4. 3: Parameter settings for contrast enhancement

Method	Parameters	Values
MSR	Percentage of saturation pixel, PoS	[1, 5, 10]
Top-Bottom Hat	Radius of SE, R_{SE}	[1, 5, 10]

The parameters for each method are specified in Table 4. 3. For the MSR method, the Percentage of Saturation (PoS) pixels refers to the percentage of the saturation on each side of the histogram. The PoS value as shown in the table has a range from [1, 5, 10]. Generally, the PoS can be different on either the dark or light side, but for the convenience of computation, the same values are applied to both sides in this study.

For the Top-Bottom Hat method, as mentioned in Section 3.2.3, the radius of the structure element, R_{SE} for the morphological operations can significantly affect the result. Therefore, the Top-Bottom Hat transformation is evaluated through three radius setting scenarios.

4.2.3.4 Settings for Floe Separation

In the floe separation stage, to apply the erosion for the marker generation, the radius of the structure element, R_{ERO}, is the main parameter that need to be specified. As concluded in Table 4. 4, the range of 5 to 30 with an increment of 5 is selected for evaluation.

Table 4. 4: Parameter settings for Floe Separation

Parameter	Values
R_{ERO}	5:5:30

4.3 Results and Analysis for Model Optimisation

4.3.1 Evaluation of Superpixel Algorithms and Effect of Superpixel Number

To comprehensively evaluate the effect of different superpixel number settings on the proposed segmentation pipeline, the three superpixel number adjustable algorithms were tested with the superpixel number varying from 1×10^4 to 10×10^4 with an increment of 1×10^4. The results are presented in Figure 4. 2.

As can be seen in the Figure 4. 2 (a) and (c), the pixel level accuracy and F1 score values have a strong resistance within 1% variation against superpixel number variation for SLIC and WP. In terms of MCC in Figure 4. 2 (b), a relatively more obvious increasing or decreasing trend can be observed as the superpixel number varies because the MCC is more sensitive to the false attributes in the confusion matrix. In the previous study of mine [85], the superpixels are directly merged to form the segmentation result for Chukchi_2013. As a result, the MCC decreased rapidly to 80% as the superpixel number decreased to 1×10^4. By contrast, the proposed SS-CEM yields the variations of MCC values within 5%, which indicates the framework has strong stability against the superpixel number variation. This is mainly owing to

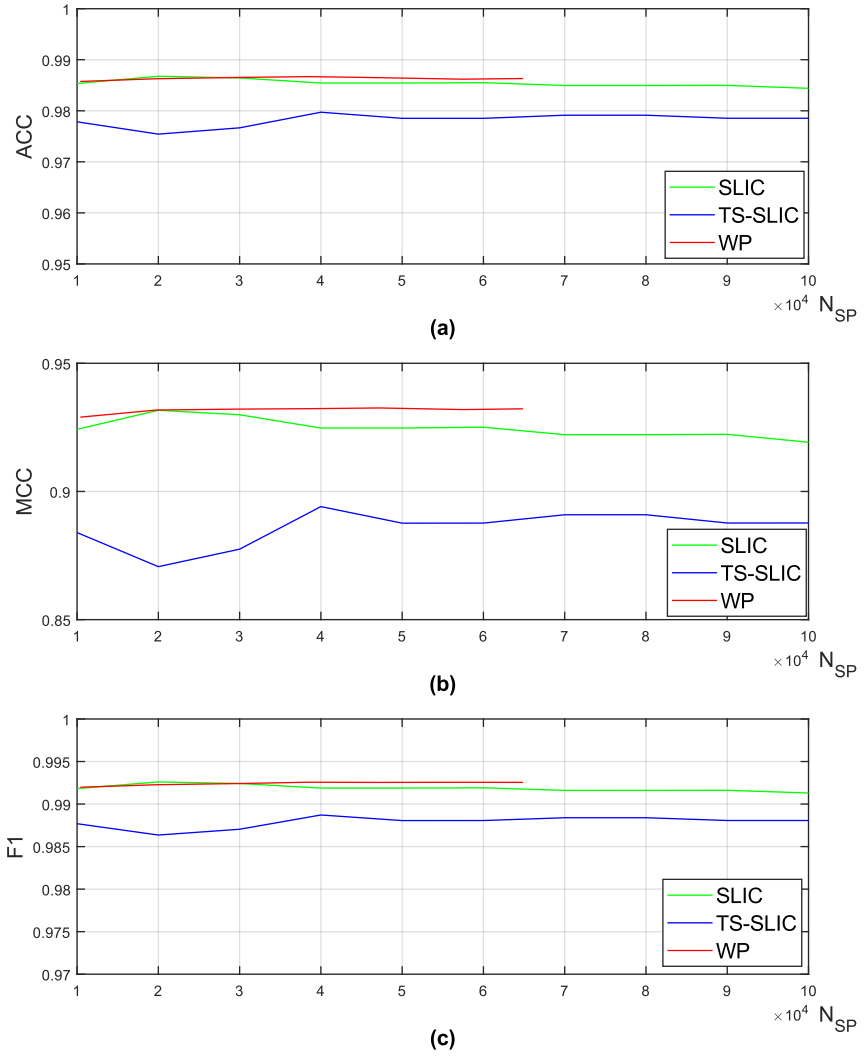

Figure 4. 2: ACC (a), MCC (b), and F1 score (c) of the three superpixel algorithms with different superpixel number settings

Table 4. 5: Quantitative comparison among four selected superpixel algorithms

Methods	ACC/%	MCC/%	F-1 score/%	Running time/s
SLIC	98.68	93.17	99.26	4.11
TS-SLIC	97.97	89.41	98.87	327.15
WP	98.70	93.21	99.26	8.71
BASS	91.73	72.52	94.83	510

the segmentation scheme that investigates ice pixels in the water-ice mixed regions by thresholding. Hence, the superpixels with under-segmentation issues can be refined.

Table 4. 5 summarises the quantitative results of the four superpixel methods under the best performed scenarios in the simplified model where pre-processing and contrast enhancement are not used. As can be seen, the WP and SLIC outperform the other two algorithms in terms of the ACC, MCC and F-1 scores, also the WP can further improve the ACC and MCC by 0.02% and 0.04% compared to the SLIC. However, the WP method suffers from certain limitations, especially the uncontrollable number of superpixels, which may affect the subsequent segmentation and FSD analysis. With the increased images size, the superpixel number can vary significantly as the grid step varies, in a non-linear way, which is not desired for the proposed framework in implementation. The SLIC, on the contrary, has a similar performance compared to the WP method, but has a very good controllability to the number of the superpixels. Therefore, the SLIC is chosen and recommended in my implementation.

In addition, the running time for each superpixel generation algorithm is also reported in Table 4. 5. As can be seen, the SLIC algorithm is of the highest efficiency. The BASS and TS-SLIC take much longer running time yet the results are still unsatisfactory. It is worth noting that these algorithms are implemented on different platforms, leading to the time comparison a bit unfair. Therefore, the running time is just provided for reference.

According to the experimental results, the SLIC method is chosen and recommended for the proposed framework due to the high segmentation accuracy and good size controllability. For the SLIC method, the best performed result can be observed when the superpixel number is set to 2×10^4, where each superpixel consists

of approximately 800 pixels. Therefore, for the experiments on the remaining images of the dataset, the superpixel size of 800 is adopted.

4.3.2 Evaluation of Noise Removal Algorithms

The Effect of the preprocessing stage is then evaluated by using the model incorporating the SLIC method for superpixel generation whilst removing the contrast enhancement stage. Figure 4. 3 compares the segmentation accuracy when using different parameter settings for each method, where the red horizontal line denotes the

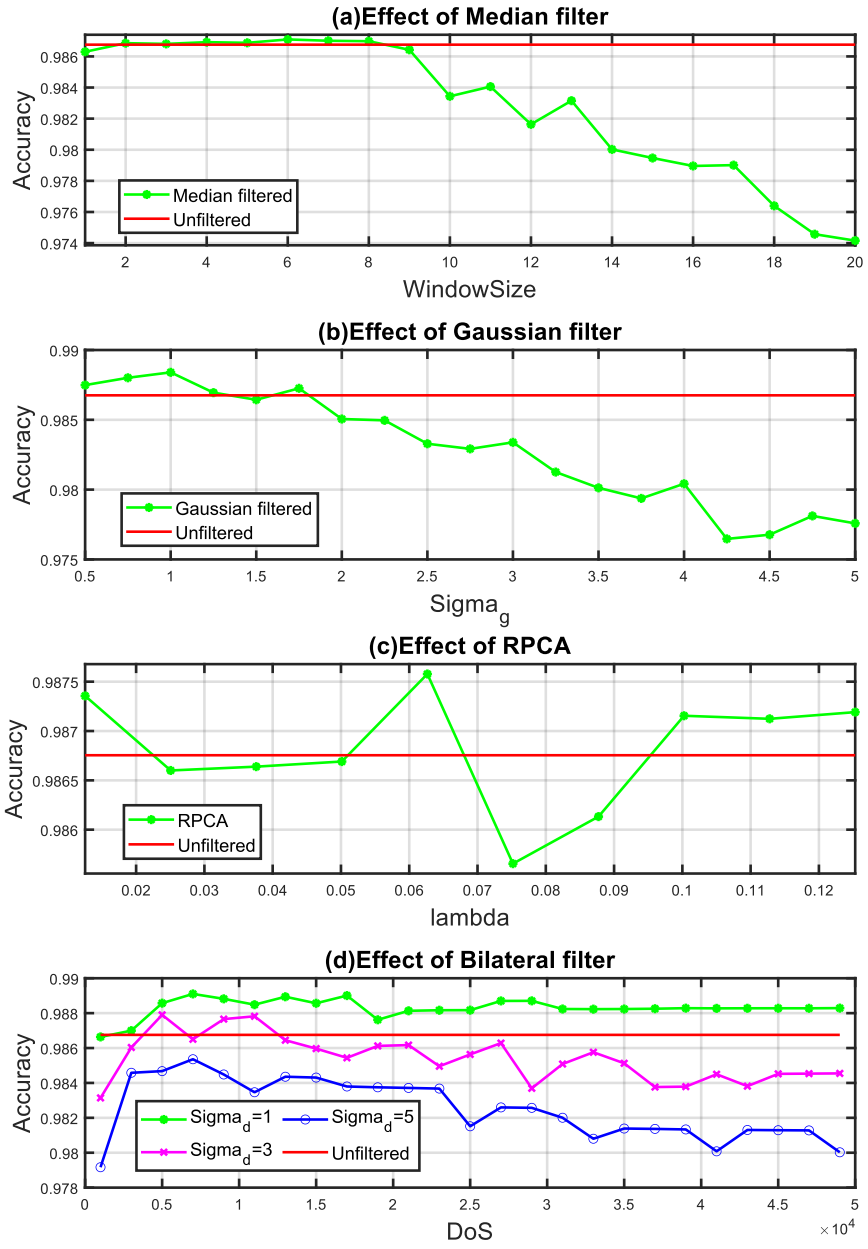

Figure 4. 3: Comparison of different parameter settings in the pre-processing stage

segmentation results conducted on the unfiltered image. As can be seen from Figure 4. 3 (a), the median filter can slightly increase the accuracy when the window size is lower than 8. However, the accuracy rapidly decreases as the window size become larger. Similarly, for the gaussian filter, as shown in Figure 4. 3 (b), the best performed scenario can be observed when the Standard deviation, σ_g is set to 3. Then a decreasing trend can be observed as the σ_g increases.

As can be seen from Figure 4. 3 (c), for the RPCA method, a peak accuracy can be observed when λ is set to $5\lambda_0$. However, the performance tends to be unstable as the λ varies, where the accuracies in several scenarios are lower than that of the model using unfiltered image.

For bilateral filter, as presented in Figure 4. 3 (d), there is an obvious trend that the segmentation accuracy increases as the smaller standard deviation σ_d is selected, which means the edge preservability provided by the small σ_d value do help the performance of the proposed segmentation model. In addition, the bilateral filter has yielded stable results with higher accuracy than the unfiltered result when σ_d is set to 1, and the highest accuracy is achieved when DoS is set to 1.7×10^4.

To further evaluate the performance of the four methods, the best performed scenarios are presented in Table 4. 6, where the running time for each scenario is recorded by measuring the whole processing time of the model. As can be seen, the Bilateral filtered image can increase the ACC, MCC, and F-1 score by 0.22%, 1.18% and 0.1% respectively. Meanwhile, PSNR is derived by taking the original image as reference. Therefore, a high PSNR value can be observed with the small MSE between the filtered and original image. Although the denoised image produced by median filter has the lowest PSNR value and the highest mean square error compared to the

Table 4. 6: Results of the best performed scenarios in the pre-processing experiments

Methods	ACC/%	MCC/%	F-1/%	PSNR/dB	Running time/s
Unfiltered	98.68	93.17	99.26	-	7.82
Median Filter	98.71	93.35	99.28	41.57	8.01
Gaussian Filter	98.84	94.03	99.35	45.87	7.96
RPCA	98.76	93.60	99.31	52.94	85.02
Bilateral Filter	98.90	94.35	99.36	46.71	9.01

unfiltered image, the accuracy of the segmentation result is inferior to the other methods. On the other hand, bilateral filer produces the image with similar PSNR, but a higher accuracy compared to that of the gaussian filter, indicating the edge preservability is beneficial in the proposed model. In addition, the running time within 3 seconds can be observed when using median, gaussian, and bilateral filters. By contrast, the RPCA method is less efficient. In this context, the bilateral filter with the standard deviation, σ_d set to 1 is recommended in the proposed model.

4.3.3 Evaluation of Contrast Enhancement Algorithms

To evaluate the effect of contrast enhancement in MIX-Reg to the segmentation accuracy, different scenarios are tested using the model incorporating SLIC. The experimental results are provided in Table 4. 7, where the result produced by original image without contrast enhancement is provided for comparison. The running time measured for the whole processes is recorded. From the table, the MSR yields the highest performance when 1% pixels are allowed to be saturated on both side of the histogram. And a decreasing trend can be observed as the PoS increases. And for the Top-Bot hat, the scenario $R_{SE} = 1$ has a slightly enhanced performance compared to the scenario that contrast enhancement is not applied. And the highest performance is

Table 4. 7: Effect of the contrast enhancement methods

Methods	ACC/%	MCC/%	F-1/%	Running time/s
Original	98.68	93.17	99.26	7.82
Original + HE	98.87	94.17	99.36	8.59
Original + MSR(PoS=1)	98.87	94.24	99.37	12.89
Original + MSR(PoS=5)	98.51	92.67	99.16	12.67
Original + MSR(PoS=10)	96.68	85.55	98.10	12.71
Original + TopBot (R_{SE}=1)	98.72	93.38	99.28	8.05
Original + TopBot(R_{SE}=5)	98.88	94.26	99.37	8.03
Original + TopBot(R_{SE}=10)	98.27	91.64	99.02	8.33
Bilateral + HE	99.03	95.11	99.43	10.20
Bilateral + MSR(PoS=1)	99.00	94.97	99.44	13.90
Bilateral + TopBot($\boldsymbol{R_{SE}}$=5)	99.05	95.16	99.47	10.17

achieved in R_{SE}=5 scenario. Among the scenarios that taking the original image as input, the Top-Bot hat scenarios are most efficient, and the scenario of R_{SE} set to 5 yielded the highest ACC, MCC, and F-1 performance.

In addition, the performances of the complete model that combines the bilateral filter for pre-processing and the best performed contrast enhancement scenarios are evaluated. As can be seen in the last three rows of Table 4. 7, the combination of bilateral filter and Top-Bot hat has yielded the highest ACC, MCC, and F-1 of 99.05%, 95.16%, and 99.47% respectively. Therefore, for the proposed sea ice segmentation method, SS-CEM, the Top-Bot hat with R_{SE} set to 5 is recommended.

4.3.4 Effect of the Threshold Value

Furthermore, to evaluate the effect of the threshold value, the segmentation results produced by different threshold values in a complete segmentation model (bilateral + SLIC + Top-Bot hat) are compared. The results are presented in Table 4. 8. As can be seen from the table, by introducing the standard deviation of the MIX-Reg to the Otsu threshold, the performance can be improved. And when the regularisation constant, C_{thresh} is set to 1, the best performance can be achieved.

Table 4. 8: Effect of the threshold value

	ACC/%	MCC/%	F-1/%
T_{Otsu}	96.78	82.83	98.22
$T_{Otsu} + 0.5 \times \sigma_{MIX-Reg}$	98.15	90.35	98.97
$T_{Otsu} + 1 \times \sigma_{MIX-Reg}$	99.05	95.16	99.47
$T_{Otsu} + 1.5 \times \sigma_{MIX-Reg}$	98.24	91.72	99.00

4.3.5 Effect of the Structure Element Radius in Floe Separation

Table 4. 9 evaluates the effect of structure element radius, R_{ERO} used for morphological erosion in the proposed ED-MCFS, where [50,5000] is the truncation range of the floe size (Unit: meter) used for LSF fitting. As can be seen, compared to the unseparated image, the mean square error of the pixel level size distribution, MSE_{CAT} has been suppressed through applying the floe separation. Meanwhile, with

Table 4. 9: Effect of the structure element radius in floe separation

	α	MSE_{CAT}	$Diff/\%$
GT [50, 5000]	1.1502	-	-
Not separated	1.3285	51.65	15.50
Separated ($R_{ERO} = 5$)	1.1863	16.21	3.14
Separated ($R_{ERO} = 10$)	1.1539	10.35	0.32
Separated ($R_{ERO} = 15$)	1.1569	10.08	0.58
Separated ($R_{ERO} = 20$)	1.1482	10.18	0.17
Separated ($R_{ERO} = 25$)	1.1437	10.14	0.57
Separated ($R_{ERO} = 30$)	1.1407	10.34	0.83

the increment of the radius, the difference between the α values estimated from the separated image and the GT can be reduced to within 1%. The lowest $Diff$ can be observed when R_{ERO} is set to 20, which can be used in the following experiments.

Figure 4. 4 further compares the pixel level floe size distribution of GT and different scenarios. As can be seen, the scenarios of the proposed floe separation method produce similar distribution with GT compared to the unseparated image.

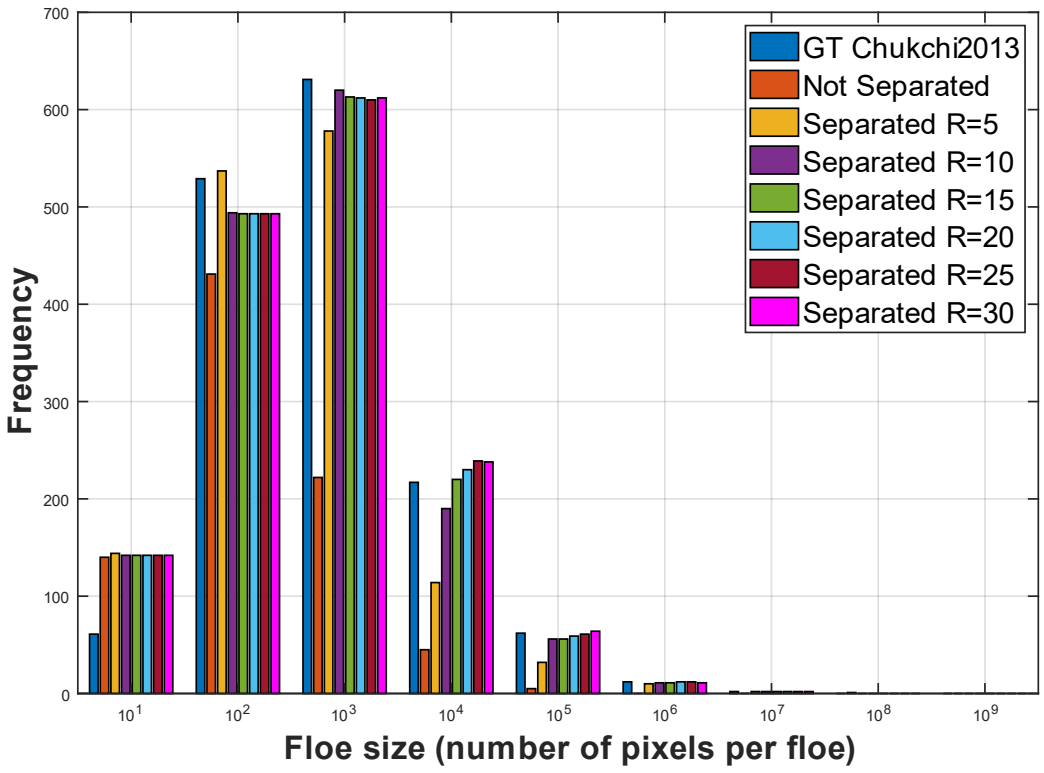

Figure 4. 4: Effects of floe separation and different scenarios to the pixel level size distribution

4.4 Results and Analysis

4.4.1 Assessment of Sea Ice Segmentation Results

As aforementioned, segmentation and floe separation results of the whole dataset are generated using the best performed scenarios after model optimisation. The selected methods and corresponding parameter settings in the optimised model are: 1) bilateral filtering with the standard deviation set to 1 and DoS set to 1.7×10^4; 2) SLIC for superpixel generation, where each superpixel has about 800 pixels; 3) Top-Bottom Hat transforms with the structure element radius set to 5; 4) The threshold is determined by adding the Otus threshold with the standard deviation calculated from the water ice mixed regions; and 5) Radius of the structure element in the floe separation stage is set to 20.

To evaluate the performance of the proposed method, against two state-of-the-art methods that have been briefly introduced in Chapter 2, the Open Source Sea-ice

Table 4. 10: Quantitative comparison of SS-CEM to the other two algorithms

Image	Method	ACC/%	MCC/%	F-1/%	Running time/s
Chukchi_2013	OSSP	98.28	90.04	98.04	41.49
	RSLC	98.44	94.70	98.96	113.45
	Proposed	**99.05**	**95.16**	**99.47**	**10.17**
Chukchi_2010_1	OSSP	95.24	88.40	96.73	284.32
	RSLC	97.40	93.35	98.24	429.50
	Proposed	**97.67**	**94.00**	**98.43**	**41.84**
Chukchi_2010_2	OSSP	93.79	80.97	96.15	341.80
	RSLC	96.49	88.19	97.86	577.01
	Proposed	**96.98**	**90.04**	**98.15**	**53.54**
Chukchi_2014	OSSP	96.01	90.97	97.13	628.14
	RSLC	95.95	90.86	97.07	1316.62
	Proposed	**98.49**	**96.34**	**98.94**	**136.08**
Esiber_2001	OSSP	98.26	90.94	99.03	75.81
	RSLC	98.31	91.23	99.06	170.24
	Proposed	**98.71**	**95.91**	**99.20**	**14.51**
MIZ_2014_1	OSSP	87.50	74.39	90.90	301.42
	RSLC	86.82	73.37	90.36	580.22
	Proposed	**93.20**	**84.37**	**95.25**	**40.14**
MIZ_2014_2	OSSP	92.71	85.93	93.65	386.62
	RSLC	92.53	85.63	93.48	694.32
	Proposed	**96.55**	**92.97**	**97.11**	**67.54**
MIZ_2014_3	OSSP	89.81	81.32	88.91	274.29
	RSLC	90.62	82.70	89.89	619.78
	Proposed	**95.14**	**90.67**	**95.01**	**66.08**

Processing (OSSP) [21] and Region Smoothing and Label Correction (RSLC) [60] were employed for comparison. OSSP and RSLC are two methods designed for tackling with the remote sensing images, where OSSP is designed specifically for providing a universal processing protocol for sea ice detection in the HRO images from different data sources and data owners, whilst the RSLC is designed for processing the SAR images that has the advantage of high object boundary adherence and the structure preservability. Table 4. 10 presents the quantitative results of the proposed model applied to the whole image dataset against OSSP and RSLC. The running time of the whole segmentation model is also provided. As discussed in the previous sections, the segmentation model used here consists of the bilateral filtering for pre-processing, SLIC for superpixel generation, and Top-Bot hat for contrast enhancement in the MIX-Reg.

Quantitatively, as can be seen in Table 4. 10, the proposed method consistently produces the highest ACC, MCC and F-1 score on the HRO image dataset in comparison to the other two algorithms. Furthermore, for the three seasonal MIZ images, $\mathrm{MIZ\_2014}_1$, $\mathrm{MIZ\_2014}_2$, and $\mathrm{MIZ\_2014}_3$, the proposed model outperforms the second-best algorithm 9.98%, 7.04% and 7.79% in terms of the MCC. A visual comparison of the segmentation result in the region of interest cropped from the

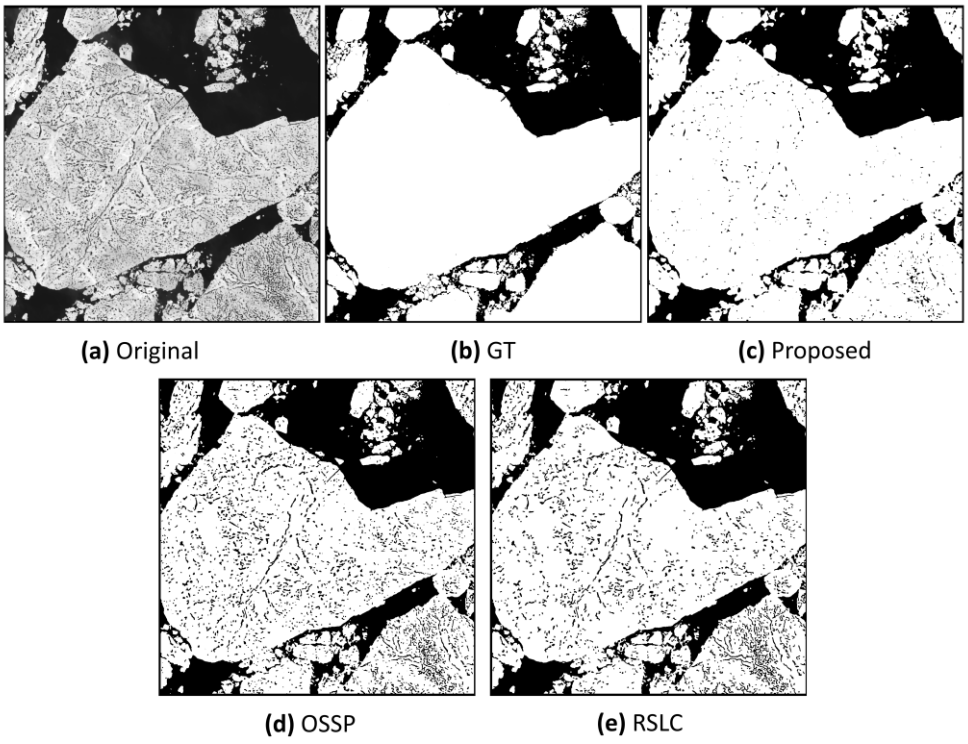

Figure 4. 5: Visual comparison of the segmentation results in MIZ_2014_2

MIZ_2014$_2$ is provided in Figure 4. 5. As can be seen, a large amount of the melt ponds can be found in the image, resulting in the existence of the small holes in the OSSP and RSLC results (Figure 4. 5 (d) and Figure 4. 5 (e)). By contrast, the proposed method can produce the result closer to the manual corrected GT data with less holes in the floes, as shown in Figure 4. 5 (c). Meanwhile, from the running time provided in Table 4. 10, for the proposed method and RSLC that are both implemented on MATLAB, the proposed method has yielded much higher efficiency. Nevertheless, since the OSSP was implemented on Python, the running time for OSSP is just provided for a general comparison.

4.4.2 Assessment of FSD Retrieval

Similar to the experiments conducted in Section 4.3.4, to evaluate the proposed floe separation method, ED-MCFS is applied on the segmentation results. Table 4. 11 presents the exponent α, the percentage difference of the α values compared to GT,

Table 4. 11: Quantitative comparison of floe separation

Image	Method	α	$Diff/\%$	MSE_{CAT}	Running time/s
Chukchi_2013	GT [50,5000]	1.1502	-	-	-
	Watershed	0.8819	23.33	49.31	3.59
	ED-MCFS	1.1482	0.17	10.18	28.84
Chukchi_2010$_1$	GT [50,5000]	1.2271	-	-	-
	Watershed	1.0492	14.50	567.89	33.13
	ED-MCFS	1.3241	7.90	503.93	103.51
Chukchi_2010$_2$	GT [20,3000]	1.0848	-	-	-
	Watershed	0.9439	12.99	376.04	41.45
	ED-MCFS	1.1003	1.43	321.24	165.29
Chukchi_2014	GT [50,5000]	1.3759	-	-	-
	Watershed	1.2590	8.50	466.02	63.28
	ED-MCFS	1.4481	5.25	451.84	121.83
Esiber_2001	GT [50,4000]	1.0970	-	-	-
	Watershed	1.1349	3.45	45.68	6.56
	ED-MCFS	1.1007	0.34	29.64	35.61
MIZ_2014$_1$	GT [50,3000]	1.4121	-	-	-
	Watershed	1.6443	16.44	579.43	34.49
	ED-MCFS	1.5189	7.56	432.29	83.13
MIZ_2014$_2$	GT [50,4000]	1.4852	-	-	-
	Watershed	1.9466	31.10	621.64	30.83
	ED-MCFS	1.3411	9.70	441.83	81.12
MIZ_2014$_3$	GT [80,3000]	2.0974	-	-	-
	Watershed	2.2876	9.07	472.00	38.23
	ED-MCFS	2.1048	0.35	386.20	114.29

and the mean square error between the pixel level size distribution, where the results produced by the general distance transformation based watershed with the radius set to 20 is provided for comparison. It is worth noting in the table that the range followed by GT indicates the truncated floe size in meters used for the α determination. As can be seen, the proposed method outperforms the general distance transformation based watershed in the α estimation and pixel level size distribution similarity. In general, the proposed method has yielded a good performance where the α values are within 10% compared to GT.

To further demonstrate the efficacy of the proposed approach and for more comprehensive performance assessment, Figure 4. 6 and Figure 4. 7 show the $N(d)$ plots before and after using the proposed floe separation method, in comparison to the GT derived for different images. The $N(d)$ plots produced from the baseline watershed method are also given for comparison. As aforementioned, LSF is applied to determine the exponent coefficient in FSD. As can be seen from these two figures, the distribution of the proposed results has been much enhanced compared to those without the floe separation scheme. Actually, the results align close to the distribution of the GT compared to the baseline watershed indicating a more accurate exponent α which has derived from the proposed approach. For the N(d) in the Esiber_2001 and MIZ_2014_3 images as shown in Figure 4. 7 (a) and Figure 4. 7 (g), the difference between the derived α values are within 1% from those of the GT, which has fully validated the effectiveness of the proposed method. By contrast, the results from the baseline watershed appear to be convex and non-linear in the truncation range due to the over-segmentation.

Nevertheless, in Chukchi_2010_1, MIZ_2014_1, and MIZ_2014_2, the estimated α values have a relatively high error of 7.90%, 7.56%, and 9.70%, respectively, due to various reasons as explained below. In Chukchi_2010_1, this is mainly due to numerous large touching floes, which are in general difficult to be separated using the erosion. As seen in Figure 4. 6 (c), the floes with a diameter more than 1×10^4 m remain after applying the floe separation. Such unseparated large floes have inevitably led to a higher exponent value than that of the GT. For MIZ_2014_1 and MIZ_2014_2 with melt ponds contained, the effects of floe separation on the determined exponent value are not as good as in other scenarios mainly due to the over-segmentation. The over-segmented floes with different the sizes will result in the increased or decreased

α values. For example, as shown in Figure 4. 7 (c), the floes with a diameter ranges from 10 ~ 300m appear more frequently than those in the GT, and the over-segmentation has resulted in a higher α. By contrast, as shown in Figure 4. 7 (e), the high frequency of the floes with a diameter ranges from 100 ~ 1000m leads to a lower exponent value compared to the GT. Therefore, further refinement of the proposed method to tackle images containing large touching floes and melt ponds is needed as my future work.

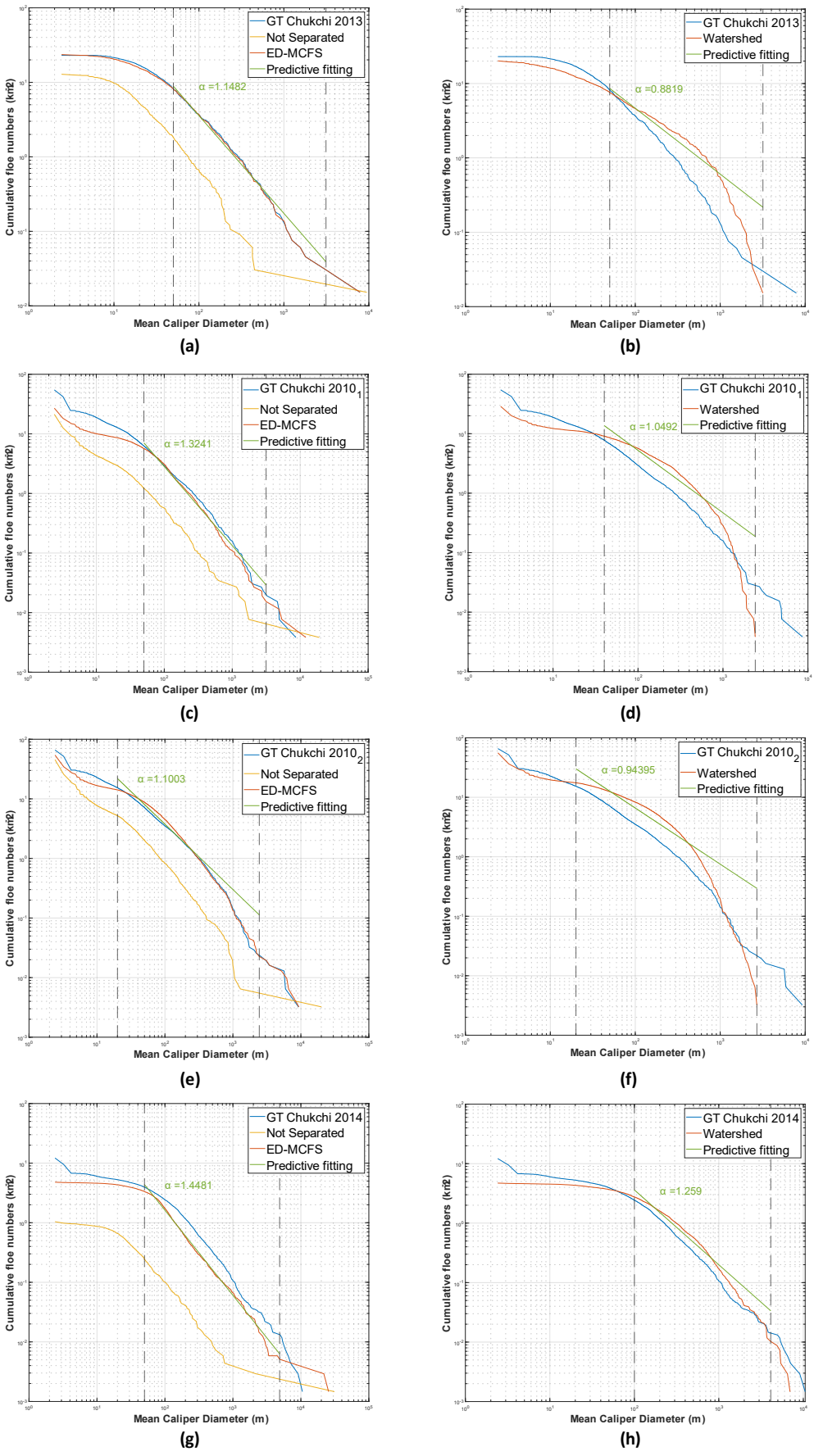

Figure 4. 6: Comparison of N(d) in the Chukchi images

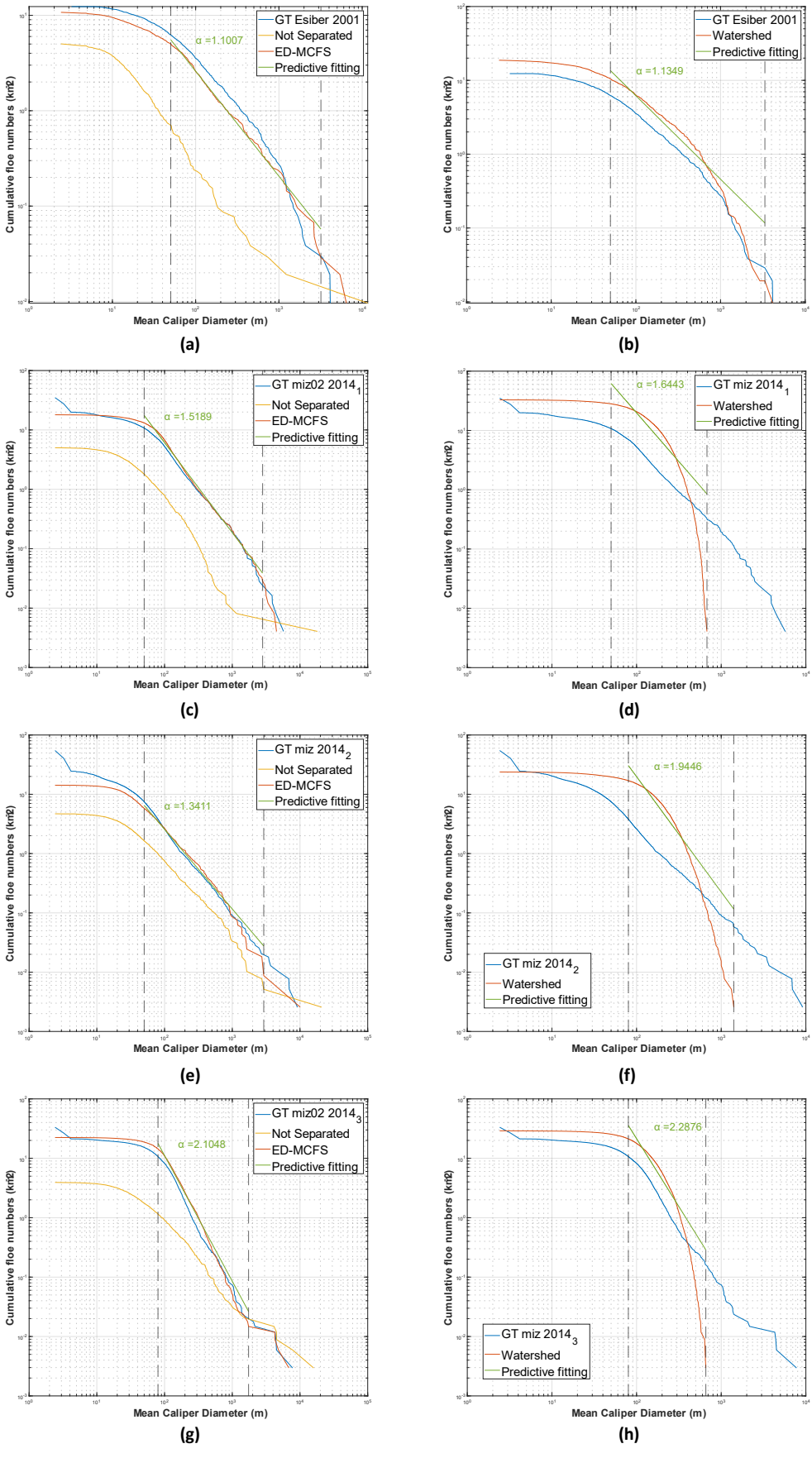

Figure 4. 7: Comparison of N(d) in the Esiber and MIZ images

4.5 Summary

In this chapter, the model proposed in Chapter 3 is optimised by experimentally comparing the effect of using different algorithms with various parameter settings. As a result, SLIC method can be selected for superpixel generation, as it has yielded relatively high accuracy compared to TS-SLIC and BASS with the highest efficiency. The WP method although has yielded highest accuracy, is not recommended. Currently, the implementation caps the superpixel number, which makes it difficult to be applied to the image with large size. For the pre-processing stage, bilateral filter with the standard deviation set to 1 is recommended due to the edge preservability. For the segmentation stage, bilateral filtered image combined with the SLIC method for superpixel generation and Top-Bottom hat transformation for contrast enhancement is the best performed scenario, where each superpixel consists of roundly 800 pixels and the radius of the structure element can be set to 5. In addition, by introducing the standard deviation of the MIX-Reg to the threshold value, a better segmentation result can be observed compared to solely using the Otsu method. In the post-processing stage, the floe separation is proposed with the key concept of marker-controlled watershed segmentation, where the markers for floes within or without the ICE-Reg are produced separately. From the experiments, the recommended radius is 20 for the structure element.

Afterwards, the results of the SS-CEM with preprocessing were validated against two state-of-the-art methods. In the experiment, the model achieved the highest ACC, MCC, and F-1 score in less time. Meanwhile, for the proposed floe separation method, ED-MCFS, the FSD observations compared with the general distance transformation based watershed segmentation. From the comparison of the estimated α values, it has been validated that ED-MCFS can produce closer results compared to the manually corrected GT data with a minimum difference of 0.17% and maximum difference of 9.7%.

In the future work, the refinement of the superpixel generation will be studied in depth. Meanwhile, a deep learning model can be applied by taking the results produced by the proposed model as training data for achieving the semi-supervised learning.

Chapter 5. Conclusions and future works

5.1 Conclusions

The thesis mainly emphasised on the sea ice investigation and FSD retrieval from the HRO image using a superpixel based multi-stage model. The main contributions include the proposal of the novel sea ice detection and floe separation methods. In addition, various scenarios of using different algorithms with the corresponding parameter settings were evaluated for the model optimisation. The recommendation of the algorithm selection and parameter settings are discussed in the previous chapter.

In Chapter 3, for effectively investigating the ice pixels from HRO images, a novel segmentation framework is proposed. For tackling the segmentation in the transition area of water and ice, a combination of superpixel and k-means clustering is employed for discriminating the ice floe regions and water-ice mixed regions according to the features that superpixels contained. The contrast enhancement technique is then applied to the water-ice mixed regions, which helps the later thresholding process to better investigate the ice pixels from such regions. For accurate FSD retrieval, a novel floe separation framework is proposed as the post-processing stage to separate the touching floes. With the employment of the ice floe regions identified in the segmentation stage, different strategies are applied to generate the makers for controlling the watershed segmentation. As a result, the touching floes can be separated whilst avoiding the over-segmentation issue.

In Chapter 4, a comprehensive evaluation of the model optimisation is firstly presented by taking one of the HRO images as the sample. The proposed segmentation method has shown a stable performance in terms of the accuracy to the variation of the superpixel number. The optimised model is then applied to the whole HRO image dataset and has yielded the highest segmentation performance compared to the other two state-of-the-art methods. However, in some images that contain numerous melt ponds, holes still can be observed in the segmented image which can affect the performance of the subsequent floe separation. For floe separation, the proposed method outperforms the general distance transformation based watershed with closer α values compared to GT images. Overall, the difference within 10% in α estimation has been achieved by the proposed method. For the segmentation results with less holes such as Chukchi_2013, Esiber_2001 and MIZ_2014_3, the difference can be kept

within 1% compared to GT. Nevertheless, when coping with the segmented images that contain large touching floes or holes inside floes, the percentage difference increased to 7-10%.

The code of the proposed workflow is available at https://github.com/SiYC/Sea-Ice.git, and the implementation of FSD calculation is available at https://github.com/SiYC/FSD-calculation.

5.2 Future works

Despite the progress which has been shown in the thesis, there are still some limitations of the proposed approaches for possible refinement in the future. In the pre-processing stage, the employed methods are not able to remove the environmental factors such as the cloud. Therefore, a deep-learning based approach (i.e. U-Net for cloud removal [113]) could be beneficial as the U-Net has a relatively small demand for the training data.

Regarding the suitability of the proposed segmentation and separation methods working on SAR images, this will be investigated as future work, along with proper pre-processing to remove the speckle noise from the SAR images. Meanwhile, the results produced by the proposed model can be complementary to the semi-supervised learning methods. In terms of the FSD estimation, although the power law expression is adopted in this thesis, its universality has not been widely demonstrated [114]. Also, the utilised truncated ranges are determined by visual assessment. Therefore, improvement can be made by employing other methods for CFND regression, such as Singular Spectrum Analysis (SSA) [115], and even forecasting the trend with auto-regressive moving average (ARMA) model [116] and prophet model [117].

Also, the refinement regarding the superpixel generation is another important topic for future work. In this study, although four superpixel algorithms were evaluated, none of them is perfectly suitable for the ice floe segmentation, which is one of the main reasons to apply the thresholding for refining the superpixel segments. Therefore, the feature of the sea ice will be learned more in depth for developing a more appropriate superpixel generating method. In addition, the classification stage can be further explored by considering other algorithms. Although K-means is powerful and widely used, it still has some limitations due to the unsupervised classification mechanism. Therefore, some supervised methods, such as Support Vector Machine (SVM) and even deep learning, especially reinforcement learning

[118] and zero-shot learning [119] can be employed for categorising the superpixels or image segmentation when there are insufficient training samples. In addition, the Generative Adversarial Network (GAN) can also be employed for producing simulated samples.

Appendix

Appendix A: Experimental image dataset

The information of the four original HRO images is summarised in Table A. 1. Source images used throughout the processing stages in this thesis are presented in Figure A. 1 to Figure A. 8. Note that the images used in this thesis are downsized to half resolution compared to the original HRO images that has the resolution of 1 meter. Therefore, the resolution in both X and Y axis are accordingly 2 meters.

Table A. 1: Source Image Information

Attribute	Acquisition Date	Image Size	Comments
Chukchi_2013	31/May/2013	2548×6368	Large floe in corner
Chukchi_2010_1	30/May/2010	7632×8544	Contain thin cloud and shadow
Chukchi_2010_2	30/May/2010	7080×10944	Roughened open water
Chukchi_2014	02/May/2014	13504×12672	Insufficiently illuminated
Esiber_2001	16/June/2001	3684×7056	Contain cloud
MIZ_2014_1	18/Jul/2014	6720×9168	Contain melt ponds
MIZ_2014_2	30/Jul/2014	10352×9392	Contain melt ponds
MIZ_2014_3	14/Aug/2014	10944×7728	Contain melt ponds

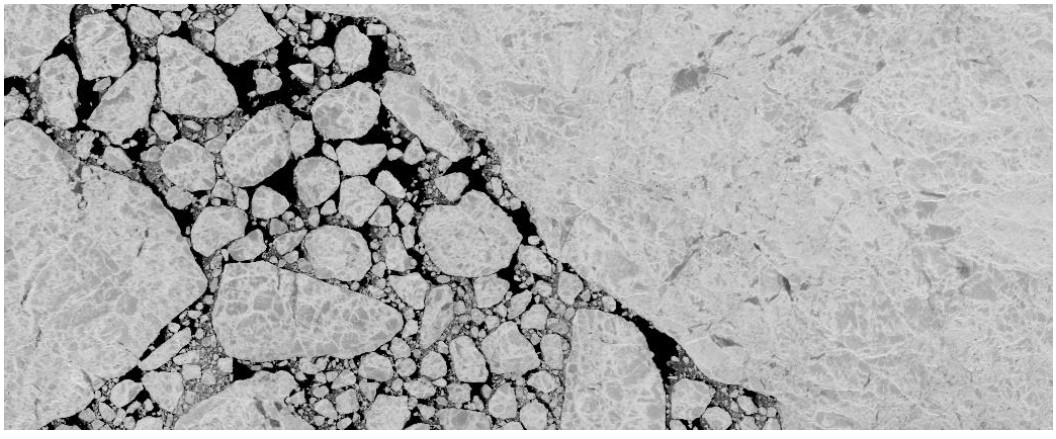

Figure A. 1: Source image of Chukchi_2013

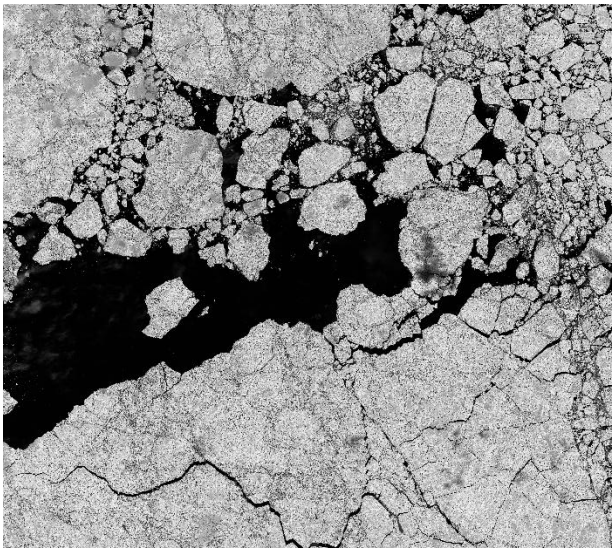

Figure A. 2: Source image of $Chukchi\_2010_1$

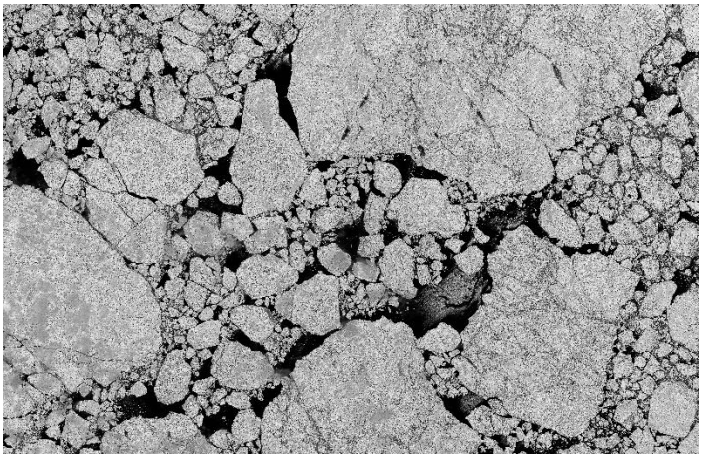

Figure A. 3: Source image of $Chukchi\_2010_2$

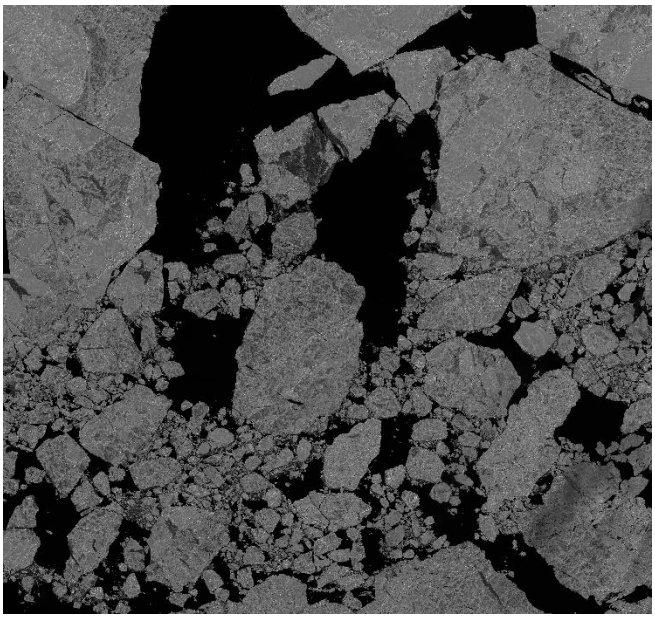

Figure A. 4: Source image of Chukdhi_2014

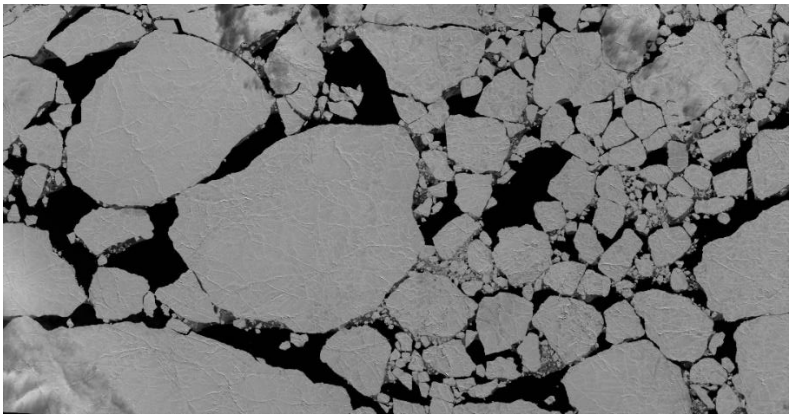

Figure A. 5: Source image of Esiber_2001

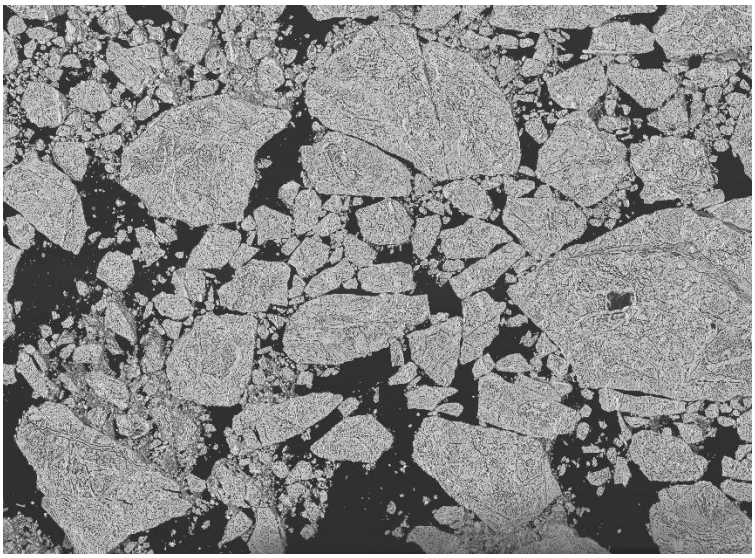

Figure A. 6: Source image of miz_2014_1

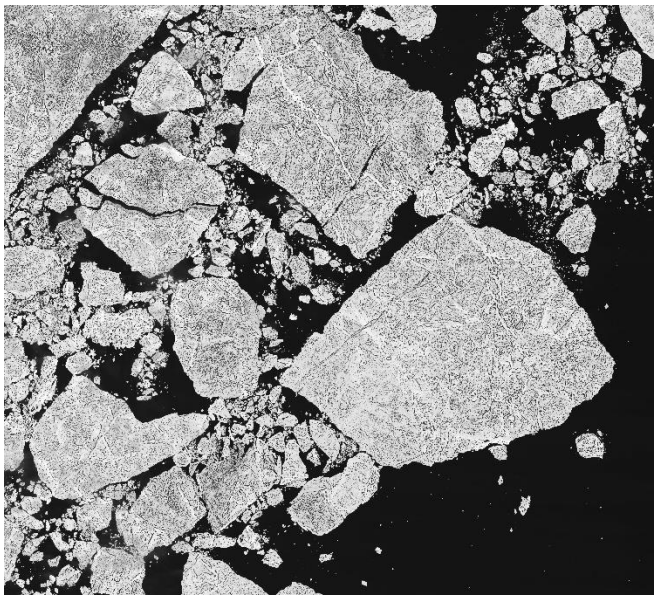

Figure A. 7: Source image of miz_2014_2

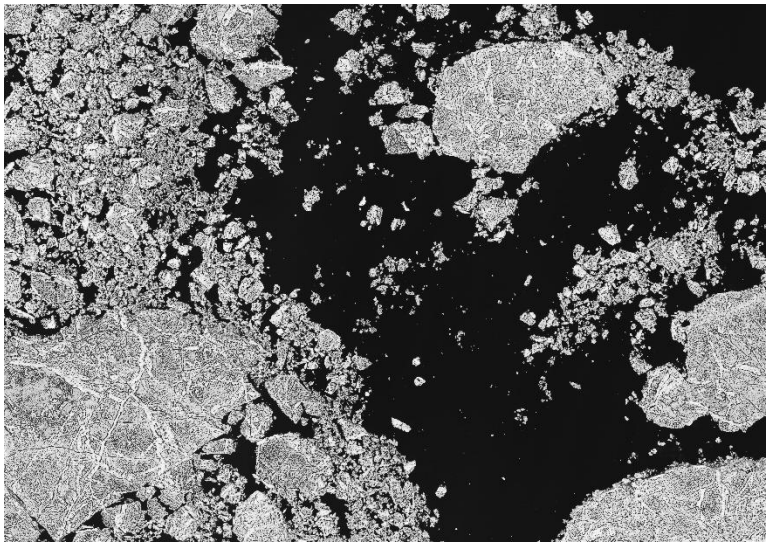

Figure A. 8: Source image of miz_2014$_3$

Appendix B: Publications During the MPhil Study

There are two publications during my MPhil study as detailed below.

1) S. Chen, Y.Y., J. Ren, Phil Hwang, S. Marshall, T. Durrani, *Superpixel Based Sea Ice Segmentation with High-Resolution Optical Images: Analysis and Evaluation*, in *Communications, Signal Processing, and Systems*. 2022, Springer

In this paper [85], as discussed in Section 2.2.3, different superpixel generation methods applied on the segmentation of high-resolution sea ice image are analysed and evaluated. In the segmentation model, superpixels are generated from the pre-processed image, followed by the k-means classification to group the superpixels to form the binary segmentation result. Afterwards, the segmentation result is post-processed with morphological opening for the shape refinement and floe separation. Although the model has yielded a segmentation accuracy of 98.19% on average, the performance highly relies on the proper parameter specification. The specification of superpixel number significantly affects the performance. In addition, the effect of the floe separation with morphological opening is not ideal. To tackle these limitations, a two-stage segmentation and hierarchal floe separation methods are proposed. The efficacy and efficiency have been validated in the previous sections.

2) S. Chen, J.R., Y. Yan, Z. Wang, H. Zhao, *Multi-sourced sensing and machine learning for effective detection of fire hazard in early stage whilst minimising the false alarms.* Computers and Electrical Engineering, in press, 2022

This paper is based on the extension of my undergraduate final year project, which presents a fast and cost-effective indoor fire alarm system that integrating multimodal sensors for tackling the early fire detection issue. The collected data is analysed and classified by machine learning techniques for suppressing the false alarm rate, where SVM is found to outperform the random forest, K-means, and Artificial Neural Networks (ANN). The experiments were designed under different scenarios to simulate the non-fire, smouldering, flaming, and the false stimuli situations. The documented data is also used for the model validation. The proposed system has yielded the accuracy, recall, and F1 of 99.8%, 99.6%, and 99.7%, respectively.

References

1. Stroeve, J. and D. Notz, *Changing state of Arctic sea ice across all seasons.* Environmental Research Letters, 2018. **13**(10): p. 103001.
2. Haarpaintner and J., *Arctic-wide operational sea ice drift from enhanced-resolution QuikScat/SeaWinds scatterometry and its validation.* IEEE Transactions on Geoscience and Remote Sensing, 2005. **44**(1): p. 102-107.
3. Bobylev, L.P. and M.W. Miles, *Sea Ice in the Arctic Paleoenvironments*, in *Sea Ice in the Arctic*. 2020, Springer. p. 9-56.
4. Serreze, M.C. and R.G. Barry, *Processes and impacts of Arctic amplification: A research synthesis.* Global and planetary change, 2011. **77**(1-2): p. 85-96.
5. NASA. *Arctic Sea Ice Minimum from 1979 to 2020*. Available from: https://climate.nasa.gov/vital-signs/arctic-sea-ice/.
6. Schröder, D., et al., *September Arctic sea-ice minimum predicted by spring melt-pond fraction.* Nature Climate Change, 2014. **4**(5): p. 353-357.
7. Meylan, M.H., et al., *Dispersion Relations, Power Laws, and Energy Loss for Waves in the Marginal Ice Zone.* Journal of Geophysical Research: Oceans, 2018. **123**(5): p. 3322-3335.
8. Strong, C. and I.G. Rigor, *Arctic marginal ice zone trending wider in summer and narrower in winter.* Geophysical Research Letters, 2013. **40**(18): p. 4864-4868.
9. Palma, D., et al., *Cruising the marginal ice zone: climate change and Arctic tourism.* Polar Geography, 2019. **42**(4): p. 215-235.
10. Squire, V.A., *Of ocean waves and sea-ice revisited.* Cold Reg. Sci. Technol., 2007. **49**: p. 110-133.
11. Thomson, J. and W.E. Rogers, *Swell and sea in the emerging Arctic Ocean.* Geophys. Res. Lett., 2014. **41**: p. 3136-3140.
12. Hauser, D.D.W., et al., *Decadal shifts in autumn migration timing by Pacific Arctic beluga whales are related to delayed annual sea ice formation.* Global Change Biology, 2017. **23**.
13. Aksenov, Y., et al., *On the future navigability of Arctic sea routes: High-resolution projections of the Arctic Ocean and sea ice.* Marine Policy, 2017. **75**: p. 300-317.
14. Rothrock, D. and A. Thorndike, *Measuring the sea ice floe size distribution.* Journal of Geophysical Research: Oceans, 1984. **89**(C4): p. 6477-6486.
15. Steele, M., *Sea ice melting and floe geometry in a simple ice-ocean model.* Journal of Geophysical Research: Oceans, 1992. **97**(C11): p. 17729-17738.
16. Curry, J.A., J.L. Schramm, and E.E. Ebert, *Sea ice-albedo climate feedback mechanism.* Journal of Climate, 1995. **8**(2): p. 240-247.
17. Bateson, A.W., et al., *Impact of sea ice floe size distribution on seasonal fragmentation and melt of Arctic sea ice.* The Cryosphere, 2020. **14**(2): p. 403-428.
18. Roach, L.A., et al., *An Emergent Sea Ice Floe Size Distribution in a Global Coupled Ocean-Sea Ice Model.* Journal of Geophysical Research: Oceans, 2018.
19. Boutin, G., et al., *Floe Size Effect on Wave-Ice Interactions: Possible Effects, Implementation in Wave Model, and Evaluation.* Journal of Geophysical Research: Oceans, 2018.
20. Zhang, J., et al. *Modeling the seasonal evolution of the Arctic sea ice floe size distribution*. 2016.
21. Wright, N.C. and C.M. Polashenski, *Open-source algorithm for detecting sea ice surface features in high-resolution optical imagery.* The Cryosphere, 2018. **12**(4): p. 1307-1329.

22. Chai, Y., et al., *Texture-Sensitive Superpixeling and Adaptive Thresholding for Effective Segmentation of Sea Ice Floes in High-Resolution Optical Images.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021. **14**: p. 577-586.
23. Xu, C., et al. *Unsupervised Classification of High-Resolution SAR Images Using Multilayer Level Set Method*. in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. 2019. IEEE.
24. Pringle, D., et al., *Pore space percolation in sea ice single crystals.* Journal of Geophysical Research: Oceans, 2009. **114**(C12).
25. Zhang, Q., et al. *Digital image processing for sea ice observations in support to Arctic DP operations*. in *International Conference on Offshore Mechanics and Arctic Engineering*. 2012. American Society of Mechanical Engineers.
26. Hwang, B., et al., *A practical algorithm for the retrieval of floe size distribution of Arctic sea ice from high-resolution satellite Synthetic Aperture Radar imagery.* Elementa: Science of the Anthropocene, 2017. **5**.
27. Blunt, J., et al. *Image analysis techniques for high Arctic, deepwater operation support*. in *OTC Arctic Technology Conference*. 2012. OnePetro.
28. Ijitona, T.B., J. Ren, and P.B. Hwang. *SAR sea ice image segmentation using watershed with intensity-based region merging*. in *2014 IEEE International Conference on Computer and Information Technology*. 2014. IEEE.
29. Zhang, Q., R. Skjetne, and B. Su. *Automatic image segmentation for boundary detection of apparently connected sea-ice floes*. in *The proceedings of the 22nd International Conference on Port and Ocean Engineering under Arctic Conditions*. 2013. Port and Ocean Engineering under Arctic Conditions.
30. Herman, A., K.U. Evers, and N. Reimer, *Floe-size distributions in laboratory ice broken by waves.* The Cryosphere, 2017. **12**: p. 685-699.
31. Lee, S., et al., *Machine learning approaches to retrieve pan-Arctic melt ponds from visible satellite imagery.* Remote Sensing of Environment, 2020.
32. Wang, M., et al., *Optimal Segmentation of High-Resolution Remote Sensing Image by Combining Superpixels With the Minimum Spanning Tree.* IEEE Transactions on Geoscience and Remote Sensing, 2018. **56**(1): p. 228-238.
33. Shelhamer, E., J. Long, and T. Darrell, *Fully Convolutional Networks for Semantic Segmentation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017. **39**: p. 640-651.
34. Chen, L.-C., et al., *Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs.* CoRR, 2015. **abs/1412.7062**.
35. Wang, J., et al., *Semi-Supervised Remote Sensing Image Semantic Segmentation via Consistency Regularization and Average Update of Pseudo-Label.* Remote. Sens., 2020. **12**: p. 3603.
36. Zhang, Q. and R. Skjetne, *Image processing for identification of sea-ice floes and the floe size distributions.* IEEE Transactions on geoscience and remote sensing, 2014. **53**(5): p. 2913-2924.
37. Achanta, R., et al., *SLIC superpixels compared to state-of-the-art superpixel methods.* IEEE transactions on pattern analysis and machine intelligence, 2012. **34**(11): p. 2274-2282.
38. Zhang, Q., et al., *Salient object detection based on super-pixel clustering and unified low-rank representation.* Comput. Vis. Image Underst., 2017. **161**: p. 51-64.
39. Barbat, M.M., et al., *An adaptive machine learning approach to improve automatic iceberg detection from SAR images.* ISPRS Journal of Photogrammetry and Remote Sensing, 2019.

40. Zhang, Y., et al., *Sea ice and water classification on dual-polarized Sentinel-1 imagery during melting season.* The Cryosphere Discussions, 2021: p. 1-26.
41. Zakhvatkina, N., V. Smirnov, and I. Bychkova, *Satellite sar data-based sea ice classification: An overview.* Geosciences, 2019. **9**(4): p. 152.
42. Dierking, W., H. Skriver, and P. Gudmandsen, *On the improvement of sea ice classification by means of radar polarimetry.* Remote Sensing in Transition, 2004: p. 203-209.
43. Scheuchi, B., et al. *The potential of cross-polarization information for operational sea ice monitoring*. in *Envisat & ERS Symposium*. 2005.
44. Soh, L.-K. and C. Tsatsoulis, *Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices.* IEEE Transactions on geoscience and remote sensing, 1999. **37**(2): p. 780-795.
45. Auer, S., S. Hinz, and R. Bamler, *Ray-Tracing Simulation Techniques for Understanding High-Resolution SAR Images.* IEEE Transactions on Geoscience and Remote Sensing, 2010. **48**: p. 1445-1456.
46. Zeng, T., et al., *Sea ice thickness analyses for the Bohai Sea using MODIS thermal infrared imagery.* Acta Oceanologica Sinica, 2016. **35**(7): p. 96-104.
47. Rösel, A. and L. Kaleschke, *Comparison of different retrieval techniques for melt ponds on Arctic sea ice from Landsat and MODIS satellite data.* Annals of Glaciology, 2011. **52**(57): p. 185-191.
48. Kwok, R., *Declassified high-resolution visible imagery for Arctic sea ice investigations: An overview.* Remote sensing of environment, 2014. **142**: p. 44-56.
49. Kwok, R., et al., *Identification of sea ice types in spaceborne synthetic aperture radar data.* Journal of Geophysical Research, 1992. **97**: p. 2391-2402.
50. Holt, B. and S. Martin, *The effect of a storm on the 1992 summer sea ice cover of the Beaufort, Chukchi, and East Siberian Seas.* Journal of Geophysical Research: Oceans, 2001. **106**(C1): p. 1017-1032.
51. Soh, L.-K., C. Tsatsoulis, and B. Holt. *Identifying Ice Floes and Computing Ice Floe Distributions in SAR Images*. 1998.
52. Deng, H. and D.A. Clausi, *Unsupervised segmentation of synthetic aperture radar sea ice imagery using a novel Markov random field model.* IEEE Transactions on Geoscience and Remote Sensing, 2005. **43**(3): p. 528-538.
53. Yu, Q. and D.A. Clausi, *IRGS: Image segmentation using edge penalties and region growing.* IEEE transactions on pattern analysis and machine intelligence, 2008. **30**(12): p. 2126-2139.
54. Salah, M.B., A. Mitiche, and I.B. Ayed, *Multiregion Image Segmentation by Parametric Kernel Graph Cuts.* IEEE Transactions on Image Processing, 2011. **20**: p. 545-557.
55. Hwang, B., et al., *Intensified management of coffee forest in Southwest Ethiopia detected by landsat imagery.* Forests, 2020. **11**(4): p. 422.
56. Toyota, T., C. Haas, and T. Tamura, *Size distribution and shape properties of relatively small sea-ice floes in the Antarctic marginal ice zone in late winter.* Deep Sea Research Part II: Topical Studies in Oceanography, 2011. **58**(9-10): p. 1182-1193.
57. Bezdek, J.C., R. Ehrlich, and W.E. Full, *FCM: The fuzzy c-means clustering algorithm.* Computers & Geosciences, 1984. **10**: p. 191-203.
58. Xiang, D., et al., *A kernel clustering algorithm with fuzzy factor: Application to SAR image segmentation.* IEEE Geoscience and remote sensing letters, 2013. **11**(7): p. 1290-1294.
59. Ji, J. and K.-L. Wang, *A robust nonlocal fuzzy clustering algorithm with between-cluster separation measure for SAR image segmentation.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014. **7**(12): p. 4929-4936.

60. Shang, R., et al., *SAR Image Segmentation Using Region Smoothing and Label Correction.* Remote Sensing, 2020. **12**(5): p. 803.
61. Hara, Y., et al., *Application of neural networks for sea ice classification in polarimetric SAR images.* IEEE Trans. Geosci. Remote. Sens., 1995. **33**: p. 740-748.
62. Ressel, R., A. Frost, and S. Lehner, *A Neural Network-Based Classification for Sea Ice Types on X-Band SAR Images.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015. **8**: p. 3672-3680.
63. Ronneberger, O., P. Fischer, and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation.* ArXiv, 2015. **abs/1505.04597**.
64. Wang, Y.-R. and X. Li, *Arctic sea ice cover data from spaceborne synthetic aperture radar by deep learning.* Earth System Science Data, 2021. **13**: p. 2723-2742.
65. Wolpert, D.H., *Stacked generalization.* Neural Networks, 1992. **5**: p. 241-259.
66. Maggiori, E., et al. *Can semantic labeling methods generalize to any city? the inria aerial image labeling benchmark*. in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2017. IEEE.
67. Wang, S., et al., *Weakly supervised deep learning for segmentation of remote sensing imagery.* Remote Sensing, 2020. **12**(2): p. 207.
68. Yao, J., et al., *Cloud detection of GF‐7 satellite laser footprint image.* IET Image Processing, 2021.
69. Breiman, L., *Random forests.* Machine learning, 2001. **45**(1): p. 5-32.
70. Dowden, B., O. de Silva, and W. Huang, *Sea Ice Image Semantic Segmentation Using Deep Neural Networks.* Global Oceans 2020: Singapore – U.S. Gulf Coast, 2020: p. 1-5.
71. Badrinarayanan, V., A. Kendall, and R. Cipolla, *SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017. **39**: p. 2481-2495.
72. Zhao, H., et al., *Pyramid Scene Parsing Network.* 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: p. 6230-6239.
73. Perkins, D.N. and G. Salomon, *Transfer of learning.* International encyclopedia of education, 1992. **2**: p. 6452-6457.
74. Liu, D. and F. Xia, *Assessing object-based classification: advantages and limitations.* Remote Sensing Letters, 2009. **1**: p. 187 - 194.
75. Miao, X., et al., *Object-based detection of Arctic sea ice and melt ponds using high spatial resolution aerial photographs.* Cold Regions Science and Technology, 2015. **119**: p. 211-222.
76. Robinson, D.J., N.J. Redding, and D.J. Crisp. *Implementation of a Fast Algorithm for Segmenting SAR Imagery*. 2002.
77. Kass, M., A. Witkin, and D. Terzopoulos, *Snakes: Active contour models.* International journal of computer vision, 1988. **1**(4): p. 321-331.
78. Ibrahim, A. and E.-S.M. El-kenawy, *Image segmentation methods based on superpixel techniques: A survey.* Journal of Computer Science and Information Systems, 2020. **15**(3).
79. Guo, Q., et al., *An Object-Based Classification Approach in Mapping Tree Mortality Using High Spatial Resolution Imagery.* GIScience & Remote Sensing, 2007. **44**: p. 24 - 47.
80. Felzenszwalb, P.F. and D.P. Huttenlocher, *Efficient Graph-Based Image Segmentation.* International Journal of Computer Vision, 2004. **59**: p. 167-181.
81. Comaniciu, D. and P. Meer, *Mean Shift: A Robust Approach Toward Feature Space Analysis.* IEEE Trans. Pattern Anal. Mach. Intell., 2002. **24**: p. 603-619.
82. Tong, H., et al., *Purifying SLIC Superpixels to Optimize Superpixel-Based Classification of High Spatial Resolution Remote Sensing Image.* Remote. Sens., 2019. **11**: p. 2627.

83. Roy, S.K., et al., *Local directional ZigZag pattern: A rotation invariant descriptor for texture classification.* Pattern Recognition Letters, 2018. **108**: p. 23-30.
84. Chang, C.-C. and C.-J. Lin, *LIBSVM: A library for support vector machines.* ACM Trans. Intell. Syst. Technol., 2011. **2**: p. 27:1-27:27.
85. S. Chen, Y.Y., J. Ren, Phil Hwang, S. Marshall, T. Durrani, *Superpixel Based Sea Ice Segmentation with High-Resolution Optical Images: Analysis and Evaluation*, in *Communications, Signal Processing, and Systems*. 2022, Springer.
86. Yeo, D., et al. *Superpixel-based tracking-by-segmentation using markov chains*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017.
87. Vargas-Muñoz, J.E., et al., *An iterative spanning forest framework for superpixel segmentation.* IEEE Transactions on Image Processing, 2019. **28**(7): p. 3477-3489.
88. Wang, K., et al. *Finger vein image retrieval via coding scale-varied superpixel feature*. in *Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval*. 2017.
89. Ibrahim, A., M. Salem, and H.A. Ali, *Block-based illumination-invariant representation for color images.* Ain Shams Engineering Journal, 2018. **9**(4): p. 917-926.
90. Shi, J. and J. Malik, *Normalized cuts and image segmentation.* Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997: p. 731-737.
91. Liu, M.-Y., et al., *Entropy rate superpixel segmentation.* CVPR 2011, 2011: p. 2097-2104.
92. Vedaldi, A. and S. Soatto. *Quick Shift and Kernel Methods for Mode Seeking*. in *ECCV*. 2008.
93. Shen, J., et al., *Lazy Random Walks for Superpixel Segmentation.* IEEE Transactions on Image Processing, 2014. **23**: p. 1451-1462.
94. Shen, J., et al., *Real-Time Superpixel Segmentation by DBSCAN Clustering Algorithm.* IEEE Transactions on Image Processing, 2016. **25**: p. 5933-5942.
95. Levinshtein, A., et al., *TurboPixels: Fast Superpixels Using Geometric Flows.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009. **31**: p. 2290-2297.
96. Gong, Y.-j. and Y. Zhou, *Differential Evolutionary Superpixel Segmentation.* IEEE Transactions on Image Processing, 2018. **27**: p. 1390-1404.
97. Meyer, F. *Color image segmentation*. 1992.
98. Machairas, V., et al., *Waterpixels.* IEEE Transactions on Image Processing, 2015. **24**(11): p. 3707-3716.
99. Uziel, R., M. Ronen, and O. Freifeld. *Bayesian adaptive superpixel segmentation*. in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.
100. Simsar, E., et al., *Object-Aware Monocular Depth Prediction With Instance Convolutions.* IEEE Robotics and Automation Letters, 2022. **7**: p. 5389-5396.
101. Chang, Y., et al., *Remote sensing image stripe noise removal: From image decomposition perspective.* IEEE Transactions on Geoscience and Remote Sensing, 2016. **54**(12): p. 7018-7031.
102. Tomasi, C. and R. Manduchi. *Bilateral filtering for gray and color images*. in *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*. 1998. IEEE.
103. Wright, J., et al. *Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization*. in *NIPS*. 2009.
104. Petro, A.B., C. Sbert, and J.-M. Morel, *Multiscale retinex.* Image Processing On Line, 2014: p. 71-88.
105. Kushol, R., et al., *Contrast Enhancement by Top-Hat and Bottom-Hat Transform with Optimal Structuring Element: Application to Retinal Vessel Segmentation*. 2017, Springer International Publishing. p. 533-540.

106. Zhang, Q. and R. Skjetne. *Sea Ice Image Processing with MATLAB*. 2020.
107. Otsu, N., *A threshold selection method from gray-level histograms.* IEEE transactions on systems, man, and cybernetics, 1979. **9**(1): p. 62-66.
108. Rothrock, D.A. and A.S. Thorndike, *Measuring the sea ice floe size distribution.* Journal of Geophysical Research, 1984. **89**: p. 6477-6486.
109. U.S.G.S. *Chukchi_20130531_2*. 2013; Available from: https://earthexplorer.usgs.gov/metadata/4246/3874.
110. Chang, H.-H., et al., *Performance measure characterization for evaluating neuroimage segmentation algorithms.* NeuroImage, 2009. **47**(1): p. 122-135.
111. Chicco, D. and G. Jurman, *The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation.* BMC Genomics, 2020. **21**(1).
112. Taha, A.A. and A. Hanbury, *Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool.* BMC Medical Imaging, 2015. **15**(1).
113. Zheng, J., X.-Y. Liu, and X. Wang, *Single Image Cloud Removal Using U-Net and Generative Adversarial Networks.* IEEE Transactions on Geoscience and Remote Sensing, 2021. **59**: p. 6371-6385.
114. Horvat, C. and E. Tziperman, *The evolution of scaling laws in the sea ice floe size distribution.* Journal of Geophysical Research: Oceans, 2017. **122**(9): p. 7630-7650.
115. Hassani, H., *Singular Spectrum Analysis: Methodology and Comparison.* Journal of Data Science, 2021.
116. Singh, B. and D. Pozo, *A Guide to Solar Power Forecasting using ARMA Models.* 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), 2019: p. 1-4.
117. Toharudin, T., et al., *Employing long short-term memory and Facebook prophet model in air temperature forecasting.* Communications in Statistics - Simulation and Computation, 2021: p. 1-24.
118. Sutton, R.S. and A.G. Barto, *Reinforcement Learning: An Introduction.* IEEE Transactions on Neural Networks, 2005. **16**: p. 285-286.
119. Wang, W., et al., *A Survey of Zero-Shot Learning.* ACM Transactions on Intelligent Systems and Technology (TIST), 2019. **10**: p. 1 - 37.